Efficiently Determining the Appropriate Mix of Personal Interaction and Reputation Information in Partner Choice

Shulamit Reches[1] Philip Hendrix[2] Sarit Kraus[1] Barbara J. Grosz[2]

[1] Department of Computer Science, Bar-Ilan University, Ramat-Gan, 52900 Israel
reches@013.net, sarit@cs.biu.ac.il
[2] SEAS, Harvard University, Cambridge, MA 02138 USA
{phendrix, grosz}@eecs.harvard.edu

ABSTRACT

Many multi-agent settings require that agents identify appropriate partners or teammates with whom to work on tasks. When selecting potential partners, agents may benefit from obtaining information about alternative possibilities through gossip (i.e., by consulting others) or using a reputation system (a centralized repository of information about past behavior). This paper defines a statistical model, the "Information-Acquisition Source Utility model" (IASU) by which agents operating in an uncertain world can determine the amount of information to collect about potential partners before choosing one and which information sources they should consult (gossip, reputation system, or additional personal interaction with the agent). The IASU model explicitly represents the cost of information, which may vary by information source. To maximize the expected gain from a choice, it estimates the utility of choosing a partner by iteratively estimating the benefit of additional information. The paper reports empirical studies that compare the effectiveness of the IASU model with a baseline in which only prior experience with a potential partner is used as the basis of the decision and with a model that determines in advance both the amount of information and its allocation among the different sources. Two different application domains are used in these empirical studies, the Surrogate Venture Game model, which concerns choosing an optimal partner for a business venture, and a restaurant domain. The results of the experiments show that the use of the model significantly increases the agents' overall utility.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Multiagent systems

1. INTRODUCTION

In many multi-agent system (MAS) settings, agents may need to choose a partner for a long term collaboration [21]. For example, an agent may look for a possible partner with whom to cooperate in an online negotiation. Usually, agents have many possible partners to choose from, and thus they can benefit from obtaining information from other agents' experience with potential partners. Relying on personal information may be risky, because an agent may not have enough information to make a good decision.

This paper focuses on the problem of deciding how much information to collect about potential partners before choosing one of them, given a variety of information sources, with different costs associated with accessing each source. For expository purposes we use the term *candidate* to refer to a potential partner and the term *Chooser* to refer to the agent that needs to choose a partner.

To obtain an accurate picture of a candidate in an environment in which information is incomplete, the Chooser would have to interact many times before its personal experience was sufficient to accurately judge the candidate as a long term collaborator. In some MAS environments, agents may not be able to take part in enough such interactions before making a long term commitment. Gossip [6] and reputation systems [8] provide mechanisms which agents may use to gain knowledge about the behavior of other agents. This paper considers environments that have three information sources: (1) personal interaction, i.e. experience gained through direct interaction with a candidate; (2) gossip, information obtained from another agent about the behavior of a candidate; (3) reputation system, a centralized repository of information. Obtaining information about candidates from these information sources may improve a Chooser's decision of which candidate is most likely to make the best partner.

In most situations, there is a cost associated with obtaining information. For example, the cost of accessing information in a reputation system may be the time needed to search on-line, the time needed to check reliability, or the money that must be paid for information from a specific supplier. As a result, a resource-bounded Chooser must be able to decide how to allocate its resources among the different information sources to obtain a better picture of the candidates.

In this paper, we investigate analytic and empirical methods for deciding how much information to acquire from each source (personal, gossip or reputation system) given their associated costs, so that the agent will make the best decision in choosing a long term partner.

We define a statistical model, the Information-Acquisition Source Utility (IASU) to determine the maximum expected utility when choosing a partner for a long term collaboration. The IASU model combines prior information about available candidates with beliefs about the distribution of the candidates in the environment. Given this information, it estimates the amount of information a Chooser should obtain about each candidate and which information sources to use to maximize the Chooser's reward. We derive a utility function that estimates the expected reward from every possible allocation of the Chooser's resources and then finds its maxima. Locating the maxima of this utility function provides an estimation of the allocation that maximizes the Chooser's utility.

Extensive research has been devoted to the problem of determin-

ing how to allocate resources to select the best candidate among several available options [2, 3, 16, 18]. However, prior research has not taken into account the use of such resources as gossip information or information obtained through a centralized reputation system. In contrast, research in the reputation arena [11, 22, 10] has not addressed the question of appropriate allocation of the information obtained.

We have applied the IASU model to two test domains: (1) the *Venture Game* in which an agent has to choose an optimal candidate with which to play in order to obtain a maximum utility score; (2) a *Restaurant Domain* in which an agent must identify the best restaurant among several candidates to choose as a subcontractor for a long term catering contract. The Venture Game domain is a simulation domain similar to the Surrogate Venture game introduced by Hendrix [9] to capture the essence of the dependence among partners undertaking business ventures. In this game an agent has the ability to enter business ventures with other agents by investing its recourses in a venture. The agent working with the invested recourses will successfully complete the venture and return a high reward, or fail and produce no reward.

The restaurant test domain was created using information obtained through customer surveys and online restaurant reviews. In this domain, an employer seeks a restaurant from which it can purchase a quantity of meals for its employees at a reduced cost. The employer is committed to selecting a high quality restaurant to avoid having to pay for additional meals if employees choose to select a different restaurant. The IASU model is intended to help an employer pick the best restaurant to minimize the employer's costs when buying meals.

In the next section we describe the construction of the IASU model, describe a two-candidate case and then the k-candidate case. Section 3 contains an evaluation of the IASU model. In Section 4, we present related work in the field. Conclusions are given in Section 5.

2. THE IASU MODEL

The Information-Acquisition Source Utility (IASU) model, is a statistical model that represents the tradeoff between exploration and exploitation when choosing a candidate for long term collaboration. It specifies the allocation of the information among various information sources.

We formally model the set of candidates as $c_1, ..., c_k \in C$. Each candidate c_i is assigned a *competence* value p_i ($0 \leq p_i < 1$), which is the probability the candidate will succeed in performing a task. The *competence* values of the candidates are initially unknown to the Chooser and are sampled from a Beta random variable [1]. Although the p_i's are unknown in advance, the Chooser has some prior beliefs about the candidates' mean and variance based on its knowledge of the world. These prior assumptions may be inaccurate. The agent can estimate the values of α_i and β_i, such as $p_i \sim B(\alpha_i, \beta_i)$.

In an ideal world, the Chooser aims to select a partner with the highest *competence* value, but since obtaining information is costly, it may not be able to do so. Therefore, the Chooser's goal is to determine the amount of information to be requested from the different information sources such that its choice of a partner will maximize its overall expected utility. The expected utility will take into consideration the benefit and the associated costs of obtaining the additional information.

Initially, the agent has very little prior information about any candidate. As a result, it needs to obtain additional units of information about the potential candidates. A unit of information about a candidate is the result of one interaction with that candidate. Each

Parameter	Description
$C = c_1, ..., c_k$	Set C is a set of k candidates.
p_i	The *competence* value of candidate c_i.
T_{B,c_i}	The total number of prior information units (i.e., baseline information) about candidate c_i.
S_{B,c_i}	The number of successful interactions with candidate c_i among T_{B,c_i}.
T_{P,c_i}	The total number of additional information units about c_i obtained through personal interaction.
S_{P,c_i}	The number of successful interactions with candidate c_i among T_{P,c_i}.
γ	The expected number of information units available about a candidate from gossip.
T_{g,c_i}	The number of gossip requests about c_i.
T_{g,c_i}	The number of additional information units about c_i obtained from gossip ($T_{g,c_i} \cdot \gamma$ on average).
S_{G,c_i}	The number of successful interactions reported about c_i among T_{g,c_i}.
ρ	The expected number of information units available about a candidate from the reputation system.
T_{r,c_i}	$T_{r,c_i} = \begin{cases} 1 & \text{if reputation system is used} \\ 0 & \text{otherwise} \end{cases}$
T_{r,c_i}	The number of additional information units about c_i obtained from the reputation system
S_{R,c_i}	The number of successful interactions reported about c_i among T_{r,c_i} ($T_{r,c_i} \cdot \rho$ on average).
T_{ν,c_i}	$T_{P,c_i} + T_{g,c_i} + T_{r,c_i}$, the total number of new information units about c_i.
S_{ν,c_i}	The number of successful interactions with candidate c_i among T_{ν,c_i}.
μ	Average candidate competence.
$cost_P$	The direct cost for each unit of information obtained by personal interaction.
$cost_G$	The direct cost of each request for gossip.
$cost_R$	The direct cost of obtaining information about a candidate from the reputation system.
r	The reward obtained from a successful performance.
$Pr(p_i = x_i)$	The posterior function for p_i, using the prior distribution, $p_i \sim B(\alpha_i, \beta_i)$ and the sample with T_{B,c_i} units and a value of S_{B,c_i}.

Table 1: The notation used for the theorems and proofs.

outcome provides one unit of information on the *competence* of the candidate. For simplicity, we assume that the outcome of an interaction is either a success or a failure.

Table 1 presents the notation we use in describing the IASU model. The parameters T_{B,c_i} and S_{B,c_i} refer respectively to the total amount of prior information units about candidate c_i and the number of successful interactions with candidate c_i among T_{B,c_i}. The remaining parameters represent additional information the Chooser can obtain through personal interaction, gossip or the reputation system. We describe the model in two stages, the simple case of only two candidates, and the general case of an arbitrary number of candidates.

2.1 The two candidate case

In the two candidate case, the Chooser chooses between two candidates, c_1 and c_2. With no loss of generality, we assume that the Chooser believes c_1 is the best choice given the prior infor-

mation (T_{B,c_1} and T_{B,c_2}). The information obtaining procedure is worthwhile only if consequently the agent changes its decision from alternative c_1 to alternative c_2. Otherwise, the agent needlessly spent its resources. To estimate the optimal amount of information to acquire, the Chooser must calculate the probability that this information would cause it to change its decision and select c_2. Additional information may change the Chooser's decision if the information is positive about c_2 or negative about c_1. We use $P\_change(c_1, c_2)$ to denote the probability that the Chooser changes its decision from the current best candidate c_1, to the second candidate c_2, due to the acquisition of additional information[1]. The amount of information collected for each candidate may not be identical, and thus the Chooser must decide how it should determine what information to gather and which candidate to select given that information. For example, suppose the agent has one information unit about alternative c_1 that is "success" and six units of information about alternative c_2, five of which are "success" and one is "failure" ($T_{B,c_1}=1, S_{B,c_1}=1$ and $T_{B,c_2}=6, S_{B,c_2}=5$). Although alternative c_2 has had more failures, the agent has more information about this alternative. Given this information, how should the Chooser determine which candidate would make the best partner?

The heuristic used in this paper to determine the best candidate given possibly unequal amounts of information is to calculate the average number of successful interactions among t interactions for each candidate, and then choose the candidate with the highest value. If the Chooser has $n < t$ units of information for a specific candidate, it will add to its prior information $t - n$ units of information with an average success rate of μ. We demonstrate late that the value of t does not matter.

We use $E\_gain(c_1, c_2)$ to denote the expected gain from obtaining T_{ν,c_i} ($i = 1, 2$) units of information about candidates c_1 and c_2. We use $E\_utility(c_1, c_2)$ to denote the total expected utility of the agent from obtaining T_{ν,c_i} ($i = 1, 2$) units of information about candidates c_1 and c_2, while taking the costs into consideration. $E\_utility(c_1, c_2)$ is the difference between the expected gain from obtaining the information , denoted $E\_gain(c_1, c_2)$, and the cost of obtaining this information.

To determine the optimal allocation of information, we need to calculate the utility function of the agent from allocation of the information on candidates c_1 and c_2, denoted by $E\_utility(c_1, c_2)$. Since the cost of obtaining information is a denoted value, we have to estimate the value of $E\_gain(c_1, c_2)$.

In order to calculate $E\_gain(c_1, c_2)$, we first need to find the value of $P\_change(c_1, c_2)$, which is the probability that the information would cause the agent to change its decision from c_1 to c_2.

Proposition 1 gives the value of $P\_change(c_1, c_2)$, which Proposition 2 uses to provide the value of the function $E\_gain(c_1, c_2)$.

To compute $P\_change(c_1, c_2)$, we need to calculate the number of successful performances for each candidate in t trials, where t is equal to the maximum total number of information units of a candidate which is the sum of the number of prior information units T_{B,c_i}, and the total number of additional information units T_{ν,c_i} (i.e. $\max(T_{B,c_1}+T_{\nu,c_1}, T_{B,c_2}+T_{\nu,c_2})$). For the candidate that has fewer information units c_i we add $t-(T_{B,c_l}+T_{\nu,c_l})$ so that the two candidates have an equal number of information units. For these additional units of information, we use μ as the success rate. The number of successful information units about candidate c_i, among the total additional information units T_{ν,c_i} is S_{ν,c_i} (for $i = 1, 2$). Therefore, the new average for c_1 after obtaining T_{ν,c_1} units of in-

[1]$P\_change(c_i, c_j)$ is used as an abbreviation with reduced parameters.

formation, about candidate c_1 is:
$(S_{B,c_1}+S_{\nu,c_1}+(t\text{-}T_{\nu,c_1}\text{-}T_{B,c_1})\cdot\mu)/t$,
and similarly the new average for c_2. The Chooser will change from candidate c_1 to candidate c_2, after T_{ν,c_1}, T_{ν,c_2} additional units of information if the following inequality holds:
$S_{B,c_2} + S_{\nu,c_2} + (t - T_{\nu,c_2} - T_{B,c_2}) \cdot \mu > S_{B,c_1} + S_{\nu,c_1} + (t - T_{\nu,c_1} - T_{B,c_1})\mu$
This inequality is equivalent to:
$S_{\nu,c_1}\text{-}S_{\nu,c_2} < S_{B,c_2}-S_{B,c_1}+\mu\cdot(T_{B,c_1}+T_{\nu,c_1}-(T_{B,c_2}+T_{\nu,c_2}))$.
We denote Δ_{c_1,c_2} to be
$S_{B,c_2}-S_{B,c_1}+\mu\cdot(T_{B,c_1}+T_{\nu,c_1}-(T_{B,c_2}+T_{\nu,c_2}))$. Δ_{c_1,c_2} represents the difference between the number of estimated successful information units of candidate c_1 and c_2. $P\_change(c_1, c_2)$ is the probability that $S_{\nu,c_1}\text{-}S_{\nu,c_2} < \Delta_{c_1,c_2}$.

The number of successful information units from each knowledge source [personal interactions, gossip and reputation system, which are respectively S_{P,c_i}, S_{G,c_i} and S_{R,c_i} for $i = (1, 2)$], are binomial random variables with probability of p_i for success, ($S_{P,c_i} \sim B(p_i, T_{P,c_i})$ and on average: $S_{G,c_i} \sim B(p_i, T_{G,c_i})$ and $S_{R,c_i} \sim B(p_i, T_{R,c_i})$). Since S_{P,c_i}, S_{G,c_i} and S_{R,c_i} are independent, the total number of successful information units, S_{ν,c_i} is also a binomial random variable: $S_{\nu,c_i} \sim B(p_i, T_{\nu,c_i})$, for $i = 1, 2$.

Proposition 1. The value of $P\_change(c_1, c_2)$ is

$$\Sigma_{i=-T_{\nu,c_2}}^{\lceil\Delta_{c_1,c_2}\rceil-1}\Sigma_{k=i}^{T_{\nu,c_1}} \quad \binom{T_{\nu,c_1}}{k}\binom{T_{\nu,c_2}}{k-i}p_1^k(1-p_1)^{T_{\nu,c_1}-k}\cdot p_2^{k-i}(1-p_2)^{T_{\nu,c_2}-k+i}$$

where
$$\Delta_{c_1,c_2}=S_{B,c_2}-S_{B,c_1}+\mu\cdot((T_{B,c_1}+T_{\nu,c_1})-(T_{B,c_2}+T_{\nu,c_2}))$$

PROOF. For any given integer y, the probability of $S_{\nu,c_1}-S_{\nu,c_2}=y$, is:
$P(S_{\nu,c_1}-S_{\nu,c_2}=y)=$
$\sum_{i=y}^{T_{\nu,c_1}}P(S_{\nu,c_1}=i, S_{\nu,c_2}=i-y)=$
$\sum_{i=y}^{T_{\nu,c_1}}P(S_{\nu,c_1}=i)\cdot P(S_{\nu,c_2}=i-y)=$
$\sum_{i=y}^{T_{\nu,c_1}}\binom{T_{\nu,c_1}}{i}p_1^i(1-p_1)^{T_{\nu,c_1}-i}\binom{T_{\nu,c_2}}{i-y}p_2^{i-y}(1-p_2)^{T_{\nu,c_2}-i+y}$
We then obtain:
$P\_change(c_1, c_2)=P(S_{\nu,c_1} - S_{\nu,c_2} < \Delta_{c_1,c_2})$
$$=\Sigma_{y=-T_{\nu,c_2}}^{\lceil\Delta_{c_1,c_2}\rceil-1} P(S_{\nu,c_1} - S_{\nu,c_2} = y)$$
$$=\Sigma_{y=-T_{\nu,c_2}}^{\lceil\Delta_{c_1,c_2}\rceil-1}\Sigma_{i=y}^{T_{\nu,c_1}} \binom{T_{\nu,c_1}}{i}p_1^i(1-p_1)^{T_{\nu,c_1}-i}\cdot$$
$$\binom{T_{\nu,c_2}}{i-y}p_2^{i-y}(1-p_2)^{T_{\nu,c_2}-i+y}$$

(We use the notation $\lceil\Delta_{c_1,c_2}\rceil$, since Δ_{c_1,c_2} may not always have an integer value) $\square$

The function $P\_change(c_1, c_2)$ is expressed in terms of p_1 and p_2 which are not known to the Chooser. As a result, the equations for calculating the reward and the gain from choosing a particular candidate integrate over all possible values of p_1 and p_2. In Proposition 2, we present the gain function $E\_gain(c_1, c_2)$ considering all the possible values of p_1 and p_2. The input to function $E\_gain(c_1, c_2)$ is a specific allocation of the $T_{\nu,c_i}, (i = 1, 2)$ units of information, and the prior units of information (T_{B,c_i}, and S_{B,c_i} for $i = 1, 2$), and the output value is the reward obtained from that information.

Proposition 2. $E\_gain(c_1, c_2)$ is the expected gain from obtaining T_{ν,c_1} units of information about c_1 and T_{ν,c_2} units of information about c_2, denoted

$$\int_0^1\int_0^1 P\_change(c_1, c_2)(x_2-x_1)r\cdot N\cdot Pr(p_1=x_1)Pr(p_2=x_2)dx_1dx_2$$

where r is the reward obtained from a successful performance and

N is the number of interactions the agent intends to perform with that candidate.

PROOF. Candidate c_1 is currently the prior-best candidate. If the Chooser does not obtain any additional information, it will choose c_1. The expected reward is

$$p_1 \cdot r \cdot N$$

The Chooser will change its decision from c_1 to c_2 with a probability of $P\_change(c_1, c_2)$. In this case the reward will be

$$P\_change(c_1, c_2) \cdot p_2 \cdot r \cdot N$$

The agent will continue with c_1 with a probability of $1 - P\_change(c_1, c_2)$. So, the expected reward will be

$$(1 - P\_change(c_1, c_2)) \cdot p_1 \cdot r \cdot N$$

The expected gain from obtaining T_{ν, c_1} and T_{ν, c_2} additional units of information is the difference between the reward obtained with the additional information and the reward obtained without the additional information, which is:
$P\_change(c_1, c_2) \cdot p_2 \cdot r \cdot N + (1 - P\_change(c_1, c_2) \cdot p_1 \cdot r \cdot N - p_1 \cdot r \cdot N = P\_change \cdot (c_1, c_2)(p_2 - p_1) \cdot r \cdot N$

To calculate the expected gain we integrate over all possible values of the parameters p_1 and p_2, and multiply the equation by their posterior functions, denoted $Pr(p_i = x_i)$ which is the probability for partner c_i to have a competence value p_i, given the prior belief of p_i being a Beta random variable such as: $p_i \sim Beta(\alpha_i, \beta_i)$, and the prior sample which is the T_{B, c_i} baseline units of information about each possible partner c_i and the value of S_{B, c_i}.

We calculate the parameters α_i, and β_i using the estimated values of the mean and the variance of p_i that can be obtained from our knowledge about the world [2]. Given the results of the prior sample with T_{B, c_i} units and that $p_i \sim Beta(\alpha_i, \beta_i)$, the posterior function for p_i is: $Pr(p_i = x_i) = \frac{\Gamma(\alpha_i + \beta_i + T_{B, c_i})}{\Gamma(\alpha_i + s_i)\Gamma(\beta_i + T_{B, c_i} - s_i)} p_i^{s_i + \alpha_i - 1}(1 - p_i)^{T_{B, c_i} - s_i + \beta_i - 1}$

$\square$

Finally, we consider the cost of obtaining information. The expected utility from obtaining T_{P, c_i}, T_{G, c_i}, T_{R, c_i} units of information about candidate c_i for $i = 1, 2$, is

$$E\_utility(c_1, c_2) =$$
$$E\_gain(c_1, c_2) - \sum_{z \in \{P, g, r\}} cost_x \cdot (T_{x, c_1} + T_{x, c_2})$$

2.2 The K-candidate case

We denote the probability $best_b$ as the probability that after obtaining T_{ν, c_i} additional units of information about each c_i, candidate c_b will be superior to all other candidates and therefore will be chosen by the Chooser. Since the information units are independent,

$$best_b = \prod_{\forall i \in C, i \neq b} P\_change(c_i, c_b) \qquad (1)$$

Without loss of generality, assume candidate c_1 is currently the best candidate among these k candidates. The Chooser's gain after obtaining any additional units of information consists of the following elements:

---
[2]This is due to the fact that in a Beta distribution $\mu_i = \frac{\alpha_i}{\alpha_i + \beta_i}, \sigma_i = \frac{\alpha_i \beta_i}{(\alpha_i + \beta_i)^2(\alpha_i + \beta_i + 1)}$

- Candidate c_1 is currently the prior-best candidate. If the Chooser does not obtain any additional information, it will choose c_1. The expected reward will be

$$p_1 \cdot r \cdot N$$

- The Chooser will change its decision from candidate c_1 to candidate c_i with a probability of $best_i$. In this case the expected reward will be

$$\sum_{i=2}^{k} best_i \cdot p_i \cdot r \cdot N$$

- The Chooser will continue with candidate c_1 with probability $best_1 = 1 - \sum_{i=2}^{k} best_i$. The expected reward will be

$$(1 - \sum_{i=2}^{k} best_i) \cdot p_1 \cdot r \cdot N$$

The expected gain from obtaining additional units of information is the difference between the reward of the process after obtaining the additional information, and the reward without obtaining the additional information which is:

$$\sum_{i=2}^{k} best_i \cdot (p_i - p_1) \cdot r \cdot N$$

To calculate the expected gain we integrate over all possible values of the parameters p_i and p_j, as was done in the two candidate case. Considering all possible values of p_i we attain the following proposition. The proof is immediate from the above explanation.

Proposition 3. The expected reward from obtaining T_{ν, c_i} additional units of information about each candidate c_i is $E\_gain(1, .., k) =$

$$\int_0^1 \ldots \int_0^1 \sum_{i=2}^{k} best_i \cdot (x_i - x_1) \cdot r \cdot N \cdot \prod_{i=1}^{k} Pr(p_i = x_i) dx_1 \ldots dx_k$$

We consider the various costs involved in obtaining $\sum_{i=1}^{k} T_{\nu, c_i}$ units of information and present the agent's utility. the expected utility of the Chooser after collecting T_{ν, c_i} units of information for every candidate c_i given the direct costs $cost_P$, $cost_G$ and $cost_R$ is,

$$E\_utility(c_1, ..., c_k) = E\_gain(c_1, ..., c_k) - $$
$$\sum_{z \in \{P, G, R\}} \sum_{j=1}^{k} (cost_z \cdot T_{z, c_j})$$

The function $E\_utility(c_1, .., c_k)$ estimates the expected utility of the agent from obtaining $T_{\nu, c_1}, ..., T_{\nu, c_k}$ units of information from the different information sources. The function determines an allocation of information for the different candidates and information sources, and finds the expected reward from that allocation. As a result, its maxima is actually the information allocation: $T_{\nu, c_1}^m, ..., T_{\nu, c_k}^m$ that maximizes the Chooser's expected utility, Since the Chooser's resources are limited an integer $M > 0$ exists such that M is the maximum number of information units the Chooser can obtain ($\sum_{i=1}^{k} T_{\nu, c_i} \leq M$). According to this upper bound, we are able to find the maximum of the utility function using the Nelder-Mead (Simplex) method [13]. This maxima estimates the expected optimal allocation of the information units that will maximize the Chooser's gain. Then, the final decision about

the best candidate is made using the information collected according to this optimal allocation.

3. EVALUATION AND EMPIRICAL RESULTS

We compared the results of IASU with baseline results that took the best prior candidate without using additional information. We also compared IASU with the results of the *fixed number of experiments* model (the FNE model) [19] which decides in advance how many additional units of information a Chooser should obtain and divides the units among the candidates in proportion to their quality, as determined by the baseline results. Since the FNE model does not differentiate between the additional sources of information, we allocated the units of information in such a way that more information is gathered about the candidate with the highest estimated competence and the least amount of information is gathered about the candidate with the lowest estimated competence according to prior information.

To emphasize the importance of IASU, we describe two domains in which obtaining information is necessary, the Surrogate Venture Game and a Restaurant domain. In both domains the Chooser needs to select a candidate for a long term commitment. In the Surrogate Venture domain, using personal interactions reduces the number of the future N interactions, whereas in the restaurant domain personal interactions do not reduce N.

3.1 Surrogate Venture Domain

In the Surrogate Venture domain [9] there are two types of agents: investment agents and candidates. The investment agent plays many games of the Surrogate Venture and every game consists of a sequence of rounds. In each game it chooses a candidate with whom to make an investment in a single venture. Each candidate is assigned a competence at the beginning of the game. A venture can end in either a success or a failure. The investor agent offers an investment for the venture and if the venture is successful, the agent receives a reward, otherwise, the venture fails and the agent loses its investment. After any rewards have been received a new round begins.

3.1.1 Scoring

The success or failure of a venture is based on the *competence* of the candidate. The probability of success of the venture is the same as the candidate's *competence*. Formally, $P(success) = p_i$ where p_i is the *competence* of candidate i. The outcome of a venture is a Boolean value determined by

$$outcome = \begin{cases} success & \text{if } x \leq P(success) \\ failure & \text{otherwise} \end{cases}$$

where x is sampled from a uniformly distributed random value in the range $[0, 1)$. The scoring function may be likened to an employer utilizing the work of an employee. When the employer assigns an employee to a task the employer must pay the employee (the investment), and the employee may or may not succeed at their assigned task. The employer receives a reward for an employee completing a task. If the employee fails, then the venture investment is lost, and there is no reward.

3.1.2 The Utility Function

We hypothesized that the IASU model would lead to higher rewards than the baseline and FNE models because of its ability to acquire precise amounts of additional information.

In this game any interaction with a candidate c_i produces a successful venture with a probability of p_i. Assuming that from prior

According to	Set A	Set B	Set C	Set D
baseline	74.21	184.08	184.08	184.08
FNE model	84.09	209.18	205.41	206.23
IASU	89.41	219.01	215.77	218.82

Table 2: Results for the 2-candidate case in the surrogate venture domain depicting the average reward (in points) for each of the four parameter settings. The higher the better.

personal information candidate c_1 is the best choice, then obtaining T_{ν,c_i} additional units of information for each candidate c_i in this domain is as follows:

1. Since c_1 is currently the best candidate, if the Chooser does not obtain any additional information it will perform all N interactions using c_1. Thus, the mean number of points in this case will be $N \cdot (r \cdot p_1 - cost_P)$.

2. The average reward, from the total number of additional information units about c_i obtained through personal interaction $\sum_{l=1}^{k} T_{P,c_l}$, is: $\sum_{i=1}^{k} T_{P,c_i} \cdot r \cdot p_i$.

3. The Chooser will change from candidate c_1 to candidate c_i with a probability of $best_i$. In this case, the total reward will be: $\sum_{i=2}^{k} best_i \cdot (N - \sum_{l=1}^{k} T_{P,c_l}) \cdot (r \cdot p_i - cost_P)$.

4. The agent will continue with candidate c_1 with a probability of $best_1 = (1 - \sum_{i=2}^{k} best_i)$. In this case the total reward will be: $(1 - \sum_{i=2}^{k} best_i) \cdot (N - \sum_{l=1}^{k} T_{P,c_l}) \cdot (r \cdot p_1 - cost_P)$

5. Finally, we need to subtract the cost of asking for gossip ($\sum_{i=1}^{k} T_{g,c_i}$ times) and the reputation system ($\sum_{i=1}^{k} T_{c_j,r}$ times) and the cost of the total number of personal interactions $\sum_{i=1}^{k} T_{P,c_i}$, which is: $\sum_{i=1}^{k} T_{g,c_i} \cdot cost_G + \sum_{i=1}^{k} T_{c_j,r} \cdot cost_R + \sum_{i=1}^{k} T_{P,c_i} cost_P$.

6. The expected utility gain from all the new information units $\sum_{i=1}^{k} T_{\nu,c_i}$, is the difference between the expected reward using the additional information and the expected reward obtained using only the prior information. As a result, the utility function is:

$$E\_utility(c_1, .., c_k) = \sum_{i=2}^{k} best_i \cdot (N - \sum_{l=1}^{k} T_{P,c_l}) \cdot r \cdot (p_i - p_1) +$$

$$\sum_{i=1}^{k} T_{P,c_i} \cdot r \cdot (p_i - p_1) - (\sum_{i=1}^{k} T_{g,c_i} \cdot cost_G + \sum_{i=1}^{k} T_{c_j,r} \cdot cost_R)$$

3.1.3 Empirical Results

Simulations were run for the two- and three-candidate cases of Surrogate Venture. We set the number of prior information units, T_{B,c_i}, to 6; the length of collaboration, N, to 100; the average competence of the candidates to 0.5; the average number of units of gossip and reputation information received upon requesting the information to 10 and 25 respectively; and the cost of personal interaction $cost_P$ was set to 1. The values that were varied were the reward for a successful venture, the cost of gossip information $cost_G$, and the cost of reputation information $cost_R$. The IASU, baseline, and FNE models were tested for each value setting.

For the two-candidate case, we randomly clustered eight candidates into four pairs. The competence values for the four pairs were (0.155,0.539), (0.5781,0.7236), (0.357,0.597), and (0.3961,0.7302). Each competence value was randomly sampled from a Beta distribution having parameters Beta(2,2). The remaining parameters were varied across four experiments. We use the tuple $(r, cost_G, cost_R)$ to represent the values of the parameters. Our aim in choosing the values for these tuples was to investigate how the changing costs and rewards in the game affect the rewards
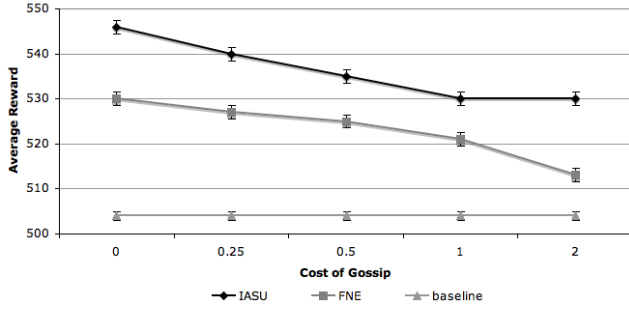
Figure 1: Rewards for varying costs of gossip.

of the Chooser. For these tuples we used values

- Set A: (3.0, 0.5, 1.25)

- Set B: (5.0, 0.5, 1.25)

- Set C: (5.0, 3.0, 5.0)

- Set D: (5.0, 0.5, 5.0)

Each of the candidate clusters was run with each of these sets of parameters.

Table 2 summarizes the average reward the agent obtained by using the baseline alternative, the FNE model, and by applying IASU. The agent's gain was higher (t-test, PV<0.05) when using IASU, compared with the FNE and the baseline models.

For the three-candidate case, we used four clusters of candidates and the average competence of the candidates was 0.5. Each one of the clusters contained three candidates. The competence values for the four triples were (0.155,0.539,0.578), (0.155, 0.539, 0.724), (0.539, 0.578, 0.724), and (0.357, 0.539, 0.724). As in the 2-candidate case, each of the candidate clusters was run with each of these varying parameters.

- Set A: (5, 3, 5)

- Set B: (3, 0.5, 1.25)

- Set C: (5, 0.5, 1.25)

Table 3 summarizes the average reward for each experiment. In most of these cases, the agent's gain was significantly higher (t-test, PV <0.05) when using IASU, compared with the FNE and the prior best models.

In a second round of experiments we varied the cost of gossip while fixing all other parameters. The goal of this set of experiments was to determine how the cost of gossip affects the average reward of the Chooser. The number of prior information units was set to 5, and the Chooser received one unit of information for each gossip request. This was done so that one gossip request and one personal interaction would yield the same amount of information. As in the previous experiment, we set the length of collaboration, N, to 100; the average competence of the candidates to 0.5; and the cost of personal interaction $cost_P$ was set to 1. Six candidates were clustered into three pairs, with the competence values for the three pairs set as (0.538, 0.408), (0.823, 0.225), and (0.646, 0.285). The models were run with various initial successes of prior information units, dependent on the worker agents' competence values, and for each input value the model was run 1000 times.

Figure 1 illustrates the effect of the cost of gossip on the average reward of the IASU model. The baseline model is not affected by the cost of gossip, because it does not take gossip information into

According to	set A	set B	set C
baseline	173.82	65.25	175.42
FNE model	198.6	87.78	214.98
IASU	210.86	96.77	228.668

Table 3: The three-candidate case in the surrogate venture domain-the average reward for each parameter setting, according to the different approaches. The higher the better.

account. The FNE model acquires a fixed amount of personal and gossip information, meaning that as the cost of gossip increases, the average reward for the FNE model will decrease. As expected, the average reward obtained by the IASU model was significantly higher than the FNE or baseline model (t-test, PV<0.05). The Chooser requested a lot of gossip information when the cost of gossip was low, in turn helping to increase the Chooser's average reward. The average reward decreased as the cost of gossip increased because the Chooser sought less gossip information and the cost of the gossip information that it did seek reduced the reward. The average reward of the IASU model remains greater than that of the FNE model even when the cost of gossip is high. This difference is in part due to IASU's ability to dynamically determine how much information (personal or gossip) to obtain, whereas the FNE model seeks fixed amount of information.

3.2 Restaurant Domain

To model the restaurant domain we denote a set of k possible restaurants: $r_1...r_k$; each restaurant r_i has an unknown parameter p_i which is its satisfaction rate; the agent employer has to pay the subcontractor for N meals in advance. According to the agent's prior knowledge, r_1 is currently the best choice. This models a minimum problem, where the agent has to choose the best restaurant at a minimal cost. We evaluate the utility function as follows:

- Candidate r_1 is currently the best candidate, so the agent does not obtain additional knowledge, r_1 will be chosen. Since the agent pays for N meals in advance, if an employee is not satisfied with the subcontractor, the agent has to pay him twice; se denote sub for the unused meal and exp for the meal the employee used. In this case the expected cost will be:

$$N \cdot sub + (1 - p_1) \cdot N \cdot exp$$

because $1 - p_1$ is the unsatisfactory rate which estimates the percentage of employees that will choose another restaurant.

- The employer will change the decision from restaurant r_1 to restaurant r_i, with a probability $best_i$. If the agent decides to get information, the cost will be:

$$\sum_{i=2}^{k} best_i \cdot (N \cdot sub + (1 - p_i) \cdot N \cdot exp) +$$

$$(1 - \sum_{i=2}^{k} best_i) \cdot (N \cdot sub + (1 - p_1) \cdot N \cdot exp) =$$

$$\sum_{i=2}^{k} best_i \cdot N \cdot exp \cdot (p_1 - p_i) + N \cdot sub + (1 - p_1) \cdot N \cdot exp$$

- The utility of obtaining additional units of information is the difference between the utility received from choosing c_1 and

	N=200	N=500	N=1000
baseline	1793.88	4484.71	8969.42
FNE model	1653.35	4106.06	8193.93
IASU	1606.55	4014.79	8020.63

Table 4: The average costs (in dollars) for each value of N according to the different approaches. (Two-candidate case in the restaurant domain.) The lower the better.

choosing a candidate with a higher competence than c_1, i.e.,

$$\sum_{i=2}^{k} best_i \cdot N \cdot exp \cdot (p_1 - p_i)$$

- The utility function is defined as
  $E\_utility(r_1, ..., r_k) =$

$$\int_0^1 ... \int_0^1 \sum_{i=2}^{k} best_i \cdot N \cdot exp \cdot (p_1 - p_i) dp_1...dp_k$$

$$+ \sum_{i=1}^{k} T_{P,c_i} cost_P + \sum_{i=1}^{k} T_{g,c_i} cost_G + \sum_{i=1}^{k} T_{r,c_j} cost_R$$

3.2.1 Empirical Results

For the restaurant domain we, used a database derived from a survey conducted on employees at several high-tech companies. The employees were asked to judge restaurants they had visited recently and to rate their satisfaction level after each visit. Each impression of an employee from one visit at a specific restaurant contributed a unit of information about that restaurant. This data yields both a gossip database and a personal interaction database. In addition, any person who visited a specific restaurant is a reputable source of gossip. We refer to the information obtained from one of the employees as the personal experience of the Chooser. The reputation system for the restaurants, was taken from on-line websites www.2eat.co.il, and www.rest.co.il. We set the cost of personal interaction, $cost_P$ at \$1, the cost of obtaining information through gossip, $cost_G$ at \$2, and the cost of obtaining information through the reputation system, $cost_R$ at \$3. We set $\mu = 0.5$; sub, the cost of a meal according to the agreement with the subcontractor at \$6; and exp, the cost of a meal in another restaurant at \$10.

We investigated four pairs of restaurants and ran experiments for each of them, varying the value of N, the number of meals the agent intends to pay for in advance, between 200, 500, and 1000. We investigated three groups, each containing three different restaurants. For this three-candidate case we set N to 200. We compared the results that were obtained with the results of the baseline model and the FNE model.

The IASU succeeded in reducing the agent's expenses. The agent's costs in the two candidate case were statistically significantly lower when using the IASU in contrast to the baseline model (t-test, PV<0.005 for all the values of N) and to the FNE model(t-test, PV=0.01 for N=200, PV=0.05 for N=500 and PV=0.06 for N=1000). Table 4 summarizes the average costs of the agent according to IASU, the baseline model and the FNE model for the different values of N according to the two-candidate case. In the three-candidate case the IASU improved the costs incurred by the agent in all three experiments. Table 5 describes the costs incurred by the agent according to the baseline, the FNE model and the IASU. The results are statistically significant in comparison to the two other models.(t-test, PV=0.004 compared to the baseline model, PV=0.02 compared to the FNE model.)

	exp. 1	exp. 2	exp. 3
baseline	1699.88	1991.29	1664.44
FNE model	1438.24	1706.84	1714.34
IASU	1331.86	1680.73	1507.82

Table 5: The average costs (in dollars) for each group in the three-candidate case of the restaurant domain. The lower the better.

4. RELATED WORK

Extensive research has been conducted on the problem of choosing an alternative among k available candidates. This research has determined how to allocate personal interaction information among different alternatives in order to maximize the expected gain from the choice.

Research on the Max K-armed bandit problem [3, 18] which is concerned with finding the trials' allocation that maximizes the expected maximum payoff obtained from one alternative.

Azoulay-Schwartz and Kraus [2] constructed a formal statistical model to find the optimal additional units of information to reveal the best alternative. Talman *et al.* [19] generalized the model of Azoulay-Schwartz & Kraus [2] to suit domains that involve choosing between heuristics or strategies. The two-candidate model was adapted to these domains and experimental results were presented. For the general k-candidate case, $k > 2$, the paper presented the fixed number of experiments model (FNE model), which decides in advance how many additional units of information the agent should obtain, denoted by N and divides it between the candidates in proportion to their quality according to the agent's baseline knowledge (initial units of information). Experimental results show that the gain from utilizing this approach was very small. In contrast to the FNE model, the EURICA model [16] presents a statistical model for finding the best heuristic among several candidates using information from personal interactions.

Cicirello & Smith [3] repeatedly ran trials on different candidate agents, each time trying to improve the current maximum reward. The same approach can be found in heuristics-related works, such as Selman *et al.* [17] who conducted a large number of experiments in order to identify the best heuristic.

Grass & Zilberstein [7] developed a decision theoretic approach that uses an explicit representation of the user's decision model to plan and execute information gathering actions. Their system is based on information sources that return perfect information about the query.

Teacy *et al.* and Huynh *et al.* [20, 11] provide quantitative results with respect to reputation systems that vary the amount of interaction between agents in the system. The IASU differs from these models in that the IASU does not have a population of agents choosing to interact and create reputation information. Instead, only one agent has makes decisions, i.e. the interaction is one-sided.

Huynh, Maes, Pujol, Teacy [11, 22, 14, 20] use agents that collect reputation information from many individual agents, rather than a central reputation authority. These papers do not address the cost of continuously asking many agents for reputation information about multiple agents. The current research extends Hendrix's [10] use of cost and tackles this problem by charging a defined cost for the reputation information.

A variety of ways of combining personal information and reputation information have been introduced by Ramchurn *et al.* [15] and Teacy *et al.* [20], including a model named TRAVOS that uses beta distributions to help model other agents. Neal and Hinton [12]

use the EM algorithm to estimate values of unobserved variables. In particular, estimating values when sample data is sparse is a necessary component of estimating another agent's competency after only a few interactions.

Fullam and Barber [5] learned how to appropriately mix previously acquired information from personal interaction and gossip sources. Our research learns how to seek appropriate amounts of such information, not how to combine them.

Dearden et al. [4] determined the value of information when choosing actions (physical actions in the environment) based on the need to explore vs. exploit. We add to this action space information seeking options (gossip and reputation information). Dearden et al. learned an optimal policy over many iterations while we focused on partner selection in only one iteration. These differences in objectives lead to significant differences in models and formalization.

5. CONCLUSIONS

This paper presents the Information-Acquisition Source Utility model (IASU), which is a statistical model which aims to improve the automated decision-making process of choosing the best candidate for a partner by incorporating gossip and reputation information. The IASU model computes the amount of additional information the agent should obtain about each candidate and from which sources (personal interaction, gossip and reputation system). It takes into account the cost associated in acquiring such information. IASU is a generalization of a model developed by Azoulay-Schwartz and Kraus [2] and implemented by Reches,Talman and Krause [16], which considered the problem of choosing between k candidates when an automated agent has a small amount of prior knowledge about each candidate. We evaluated the IASU in the Venture Game domain and the Restaurants domain, both of which require the Chooser to determine how many units of information to obtain about each candidate and through which information source. These experiments determine that the IASU model is significantly better than FNE.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1964.

[2] R. Azoulay-Schwartz and S. Kraus. Acquiring an optimal amount of information for choosing from alternatives. In *CIA VI*, pages 123–137, 2002.

[3] V. Cicirello and S. Smith. The max k-armed bandit: A new model for exploration applied to search heuristic selection. In *AAAI-05*, July 2005. Best Paper Award.

[4] R. Dearden, N. Friedman, and S. J. Russell. Bayesian q-learning. In *AAAI-98*, pages 761–768, 1998.

[5] K. Fullam and S. Barber. Dynamically learning sources of trust information: Experience vs. reputation. *Autonomous Agents and Multi-Agent Systems*, pages 1055–1062, 2007.

[6] K. K. Fullam and K. S. Barber. Learning trust strategies in reputation exchange networks. In *AAMAS '06*, pages 1241–1248, New York, NY, USA, 2006. ACM Press.

[7] J. Grass and S. Zilberstein. A value-driven system for autonomous information gathering. *Journal of Intelligent Information Systems*, 14:5–27, 2000.

[8] R. Gupta and A. K. Somani. Reputation management framework and its use as currency in large-scale peer-to-peer networks. In *P2P '04: Proceedings of the Fourth International Conference on Peer-to-Peer Computing (P2P'04)*, pages 124–132, Washington, DC, USA, 2004. IEEE Computer Society.

[9] P. Hendrix and B. J. Grosz. Reputation in the joint venture game. In *AAMAS '07*, New York, NY, USA, 2007. ACM Press.

[10] P. Hendrix and B. J. Grosz. Reputation in the venture games. In *AAAI-07*, 2007.

[11] T. D. Huynh, N. Jennings, and N. Shadbolt. Fire: An integrated trust and reputation model for open multi-agent systems. *Proceeding of 16th European Conference on Artificial Intelligence*, pages 18–22, 2004.

[12] R. Neal and G. Hinton. A view of the em algorithm that justifies incremental, sparse, and other variants. In M. I. Jordan, editor, *Learning in Graphical Models*. Kluwer, 1998.

[13] J. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7:308–313, 1964.

[14] M. Pujol, R. Sanguesa, and J. Delgado. Extracting reputation in multi agent systems by means of social network topology. *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Ssytems: Part 1*, pages 467–474, 2002.

[15] S. D. Ramchurn, C. Sierra, L. Godo, and N. Jennings. A computational trust model for multi-agent interactions based on confidence and reputation. *Proceedings of 6th International Workshop of Deception, Fraud and Trust in Agent Societies*, pages 69–75, 2003.

[16] S. Reches, S. Talman, and S. Kraus. A statistical decision-making model for choosing between multiple alternatives. In *AAMAS '07*, New York, NY, USA, 2007. ACM Press.

[17] B. Selman, H. A. Kautz, and B. Cohen. Local search strategies for satisfiability testing. In *the 2nd DIMACS Challenge on Cliques, Coloring and Satis.*, 1993.

[18] M. J. Streeter and S. Smith. An asymptotically optimal algorithm for the max k-armed bandit problem. In *AAAI-06*, July 2006.

[19] S. Talman, R. Toester, and S. Kraus. Choosing between heuristics and strategies: an enhanced model for decision-making. In *Proceedings of the International Joint Conference of Artificial Intelligence*, pages 324–330, 2005.

[20] W. Teacy, J. Patel, N. Jennings, and M. Luck. Coping with inaccurate reputation sources: Experimental analysis of a probabilistic trust model. *AAMAS '05*, pages 997–1004, 2005.

[21] M. M. Vanzin and K. S. Barber. Decentralized partner finding in multi-agent systems. In *Coordination of Large-Scale Multiagent Systems*, pages 75–98. Springer, March 2006.

[22] G. Zacharia and P. Maes. Trust through reputation mechansims. *Applied Artificial Intelligence*, 14:881–907, 2000.