On the Value of Commitment Flexibility in Dynamic Task Allocation via Second-Price Auctions

Tinglong Dai
Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213
dai@cmu.edu

Katia Sycara
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
katia@cs.cmu.edu

Guoming Lai
Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213
guomingl@cmu.edu

Robin Glinton
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
rglinton@andrew.cmu.edu

ABSTRACT

Motivated by multi-agent systems applications, we study a task allocation problem in a competitive environment with multiple self-interested autonomous agents. Tasks dynamically arrive to a contractor that oversees the process of task allocation. Tasks are auctioned to contractees, who submit prices they require to accept tasks. The agent with the lowest bid wins but is rewarded with the second-lowest price. Each agent, based on his own state, will decide whether to participate in the auction or not, and will decide the bidding price if he chooses to participate. If a busy agent wins a new task, he has to decommit from his current task and pay a decommitment fee.

We formulate the problem and derive structural properties of equilibrium strategies. We also provide heuristics that are practical for multiagent system designers. Issues related to system design are discussed in the context of numerical simulations. The contribution is that (a) we provide formal analysis of contractees' optimal strategies in a given dynamic task allocation system with commitment flexibility; (b) we study the value of commitment flexibility in the presence of different system parameters.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence; I.2.8 [**Problem Solving, Control Methods, and Search**]: Plan execution, formation, and generation

General Terms

Management, Economics, Experimentation

Keywords

Commitment Flexibility, Task Allocation, Multirobot Systems, Second-Price Auction

AAMAS 2009 Workshop on Multi-agent Sequential Decision-Making in Uncertain Domains, May 11, 2009, Budapest, Hungary.

1. INTRODUCTION

Task allocation, especially in competitive environments, is one of the most important issues in the field of multi-agent systems. In contrast to a cooperative environment in which agents work toward a common objective, a competitive environment is characterized by agents who seek to maximize their individual social utilities. In a dynamic task-allocation problem, agents not only compete directly with each other, but also compete indirectly with agents over time through the opportunity cost of potentially losing more lucrative tasks to others. A successful mechanism design, therefore, should induce the agents to contribute to the achievement of their common goals while maximizing their individual utilities.

Our work is motivated by the fast-developing multiagent resource/task allocation applications, including large-scale disaster relief (e.g., [10]), automated resource allocation of multiple distributed UAVs (e.g., [15]), and task allocation among multirobot coordination (e.g., [5]). These problems share the following characteristics: 1) significantly dynamic environments full of uncertainties, which makes it difficult to adopt centralized, static control schemes because tasks with different priorities/values/durations might arrive unpredictably. In a disaster relief system, for instance, agents face uncertain, and time stressed environments that require timely, flexible response. 2) The state of the agent network evolves over time such that it is possible that not the same set of agents are active in the agent society at each time. These features make full commitment contracts (i.e., a contract is binding once it is made), as assumed in most of the existing task allocation literature, less lucrative in terms of both system performance and agent utilities.

Prior research (e.g., [12], [2], [13]) has shown that it is indeed beneficial to the system to provide agents with a certain degree of commitment flexibility, i.e., an agent is allowed to walk away from its contract by paying a decommitment penalty. Such flexibility is especially favorable in cases with multiple job types, wherein agents' incentive to drop low-value and/or time-consuming tasks can enhance the system's capability of handling tasks. Nevertheless, aligning individual utilities and system-wide objectives cannot happen by magic. The agents need to be well coordinated by setting appropriate decommitment penalties: too great decommitment penalties will force agents to become timid and continue on less lucrative jobs; too small decommitment penalties would allow agents to decommit from their jobs too frequently and sacrifice system-wide performance.

Our study is based on the contract net (CNET) protocol ([16]). A CNET consists of four parts: (a) Problem recognition, (b) task announcement, (c) bidding, and (d) awarding the contract. While the formats of (a), (b), and (d) are usually given in a multi-agent system, an efficient and effective auction mechanism is crucial in (c). A second-price-sealed-bid auction scheme is widely adopted in various systems due to its ability to induce bidders to bid their true valuations. In a second-price auction, the bidder who places the highest bid is granted the contract, but will pay the second-highest bid. We adopt an "inverse" second-price auction, i.e., the agent who asks for the lowest reward is the winner and is paid the second lowest reward in the auction.

Our goal is to design an auction-based mechanism that effectively takes advantage of commitment flexibility so as to enhance the system's task handling capability. The first step, however, is to examine individual agents' equilibrium strategies given various environment settings, e.g., decommitment penalty, arrival pattern of incoming tasks, and reward mechanism. The paper is organized as follows: In Section 2, we present relevant research. Then in Sections 3 and 4, we describe the problem, formulate each agent's problem and validate structural properties of equilibrium strategies. In Section 5, we apply heuristics to compare system performance under different environmental variables and mechanism settings. Issues related to system design are discussed in light of the numerical simulations. Finally, we conclude and discuss future research possibilities.

2. RELEVANT LITERATURE

Several research papers in the area of task allocation are relevant to our research. Sandholm and Lesser [12] study a leveled-commitment game between a contractor and a contractee, each of which is reluctant to walk away from the contract first. The study focuses on contract design and does not consider dynamic arrivals of incoming tasks. Abdallah and Lesser [1] build a model that integrates different aspects of mediator decision-making into a Semi-MDP model. The model is essentially a centralized model wherein the *mediator* rather than the agents has the option of decommitting from an unfinished task in pursuit of a more lucrative one. Sarne et al. [14] study the problem where multiple self-interested, autonomous agents compete for dynamically arriving tasks through a Vickrey-type auction mechanism. Each agent exits from the system after he gets a task assignment. They introduce the model and formulate equations by which the agents determine their equilibrium strategies. An efficient algorithm is proposed to calculate the equilibria. However, their model does not allow for the possibility of decommitment because it assumes one-shot task assignment for each agent. Brandt et al. [4] explore an automated task allocation mechanism combining auctioning protocols and contracts with commitment flexibility. They do not provide a mathematically rigorous model, but instead use simulation experiments to show the benefits of commitment flexibility. In addition, they do not take the equilibrium behavior of each agent into consideration. Finally, their model does not capture the various trade-offs the agents face in a dynamic task allocation environment and thus provides only limited insights.

Our research enriches the literature in that it brings together dynamic task allocation and commitment flexibility, with mathematical formulations, structural results of equilibrium behaviors, and computational analysis. This has not been done in the existing literature. This study helps system designers achieve a better understanding of the way commitment flexibility helps enhance a decentralized system's performance through auction mechanisms among self-interested, autonomous agents.

3. PROBLEM DESCRIPTION

3.1 Assumptions

- There is one contractor and a number of homogeneous contractees denoted by $n = 1, 2, \cdots, N$. N can be a random number with known probabilistic distribution. The contractees are homogeneous in the sense that they have the same capabilities in performing the incoming tasks.
- Each contractee's activity is confidential to other agents. This implies $N \geq 3$.
- We assume an infinite planning horizon with intervals of equal length denoted by $t = 1, 2, 3, \ldots$.
- Each task has a constant maximum reward M.
- The duration of each task, denoted by L, follows a discrete probabilistic distribution with minimum length $L_{\min}$ and maximum length $L_{\max}$.
- There is an entry fee for participating in an auction, denoted by C.
- The probabilistic distribution of the duration of incoming tasks, the costs and rewards, are public to all the agents. However, each agent's task allocation information is confidential to other agents.
- A contractee who wins a task will receive his due reward in full immediately[1].
- The cost for each agent to perform a task is c per period, and will be incurred period by period.
- A contractee's total cost of each incoming task does not exceed the contractor's budget, i.e.,

$$c \cdot L_{\max} \leq M.$$

3.2 Sequence of Events

Each round of auction follows the subsequent procedure:

- A task arrives at the beginning of each period.
- The contractor observes the incoming task and makes an announcement about the task duration to all the contractees.

[1]We make this assumption primarily for conciseness of the subsequent POMDP formulation. This assumption is also made in [14]. Assuming this would not necessarily encourage contractees to walk away, since the decommitment penalty might be larger than total reward of a task. We will extend this assumption in Section 4.3

- Each contractee decides whether to participate in the subsequent auction or not. A bidding contractee submits his own bid in a sealed envelope to the contractor.

  An agent who has an incomplete task on hand is allowed to participate in the auction. However, if he wins an auction, he has to decommit from his current task and pay a decommitment penalty, denoted by d.
- The contractor compares contractees's bids and announces the outcome of the auction. The winner is the contractee whose bid is lowest; he is awarded with a reward given by the second-lowest bid. Both the submissions of bids and announcements of results are conducted in a confidential manner such that each contractee is only aware of his own bids/results.

4. MODELING AND STRUCTURAL ANALYSIS

Our goal is to design a decentralized system that effectively makes use of agents' commitment flexibility to enhance the overall system's ability to complete as many tasks as possible. As the initial step, we analyze each contractee's behavior in the face of the environment we described in the preceding section.

4.1 Formulation

A contractee's objective is to decide whether to participate in the auction or not, and, if yes, how much to bid, so as to maximize his total discounted reward over the infinite planning horizon. We formulate contractee i's problem as a POMDP (Partially observed Markov decision process, see [7]) characterized by a tuple $< \mathcal{S}, \mathcal{A}_i, \mathcal{T}, \mathcal{R}_i, \Omega_i, \mathcal{O}_i >$, the elements of which denote state space, action space, state-transition functions, the reward functions, observation space and observation functions, respectively. Here $\mathcal{S}$ and $\mathcal{T}$ are defined across all the contractees.

1) $\mathcal{S} = \mathbb{Z}_+^N \times \mathbb{Z}_+$ contains all the contractees' state (since one contractee's outcome is affected by other contractees' state) and the duration of the incoming task. Contractee i's state at the beginning of period t is defined as x_{it} such that $x_{it} = 1, 2, \ldots$ denotes time-to-go until the completion of the currently active job and $x_{it} = 0$ means that the contractee is idle. The duration of the task arriving to the system at time t is denoted by y_t.

2) $\mathcal{A}_i = \{0, 1\} \times \mathbb{R}$ involves a two-fold decision: whether to bid or not, and if yes, how much to offer.

3) $\mathcal{T} : \mathcal{S} \times \mathcal{A} \to \Pi(\mathcal{S})$ defines state-transition function. Define $h_{it} = 0$ or 1 as a contractee i's bidding result at time t. We see that

$$\Pr[x_{i,t+1}|x_{it} \neq 0] = \begin{cases} 1 & \text{if } x_{i,t+1} = x_{it} - 1, h_{it} = 0 \\ 0 & \text{if } x_{i,t+1} \neq x_{it} - 1, h_{it} = 0 \\ 1 & \text{if } x_{i,t+1} = y_t - 1, h_{it} = 1 \\ 0 & \text{if } x_{i,t+1} \neq y_t - 1, h_{it} = 1 \end{cases}$$

and

$$\Pr[x_{i,t+1}|x_{it} = 0] = \begin{cases} 1 & \text{if } x_{i,t+1} = 0, h_{it} = 0 \\ 0 & \text{if } x_{i,t+1} \neq 0, h_{it} = 0 \\ 1 & \text{if } x_{i,t+1} = y_t - 1, h_{it} = 1 \\ 0 & \text{if } x_{i,t+1} \neq y_t - 1, h_{it} = 1 \end{cases}$$

4) $\mathcal{R}_i = \mathbb{R}$ specifies contractee i's reward given an action. Our definition follows that if contractee i places a bid, then his one-stage reward would be the second lowest bid minus decommitment penalty (if he has an unfinished task) if his wins, and 0 if he does not.

5) $\Omega_i = \{0, 1\}^\infty$ contains $h_{it}, t = 0, 1, \ldots,$ which is the only signal that contractee i uses to reason about other contractees' state.

6) $\mathcal{O}_i : \mathcal{S} \times \mathcal{A} \to \Pi(\Omega_i)$ is the observation function $\omega(a, s')$ specifing the transition probabilities of making each possible observation $o \in \Omega_i$ given that contractee's bidding decision and state in the next period. Solving the above-defined POMDP is by no means an easy task due to the existence of a great number of agents and requirement of large state space. However, given that the initial system state is known, our POMDP can be converted to the following MDP formulation. We study the contractee's decision at the beginning of period t. $V_t^i(x_{it}, y_t)$ is defined as the maximum total discounted reward discounted by δ from time t until infinity, given x_{it} and y_t.

We see that the transition probability of h_{it} is exactly contractee i's probability of winning if he bids at time t. Instead of using h_{it} as part of contractee i's state, we include it to estimate the probabilities of winning and thus expected total rewards in the future.

A contractee's bidding strategy $s_{it} = (\Delta_{it}(x_{it}, y_t), b_{it}(x_{it}, y_t))$ for an auction consists of two parts: $\Delta_{it}(x_{it}, y_t) = 1$ indicates contractee i will participate in the auction, and $\Delta_{it}(x_{it}, y_t) = 0$ otherwise. b_{it} is contractee i's bidding price if he chooses to participate in the auction.

Let s_{-it} denote the symmetric bidding strategy of all the contractees other than contractee i at period t. Given s_{-it}, and assuming the system state follows its stationary distribution, contractee i would be able to determine the probabilistic distribution of the number of participants in the auction, and his probability of winning for a given bid. His estimated expected maximum total discounted reward function can be written as

$$V_t^i(x_{it}, y_t; s_{-it}) \tag{4.1}$$
$$= \begin{cases} -c + \max_{b_{it}} \Big\{ \delta\tilde{\mathbb{E}}V_{t+1}^i(x_{it} - 1, y_{t+1}; s_{-it}) \ , \\ \quad -C + \omega(b_{it}, s_{-it}) \Big[\tilde{\mathbb{E}}[z_t|b_{it}, s_{-it}] - d + \\ \quad \delta\tilde{\mathbb{E}}V_{t+1}^i(y_t - 1, y_{t+1}; s_{-i,t+1}) \Big] \\ \quad +[1 - \omega(b_{it}, s_{-it})]\delta\tilde{\mathbb{E}}V_{t+1}^i(x_{it} - 1, y_{t+1}; s_{-i,t+1}) \Big\}, \\ \quad \text{if } x_{it} = 1, 2, 3, \ldots \\ \max_{b_{it}} \Big\{ \delta\tilde{\mathbb{E}}V_{t+1}^i(0, y_{t+1}; s_{-it}) \ , \\ \quad -C + \omega(b_{it}, s_{-it}) \Big[-c + \tilde{\mathbb{E}}[z_t|b_{it}, s_{-it}] \\ \quad +\delta\tilde{\mathbb{E}}V_{t+1}^i(y_t - 1, y_{t+1}; s_{-i,t+1}) \Big] \\ \quad + [1 - \omega(b_{it}, s_{-it})]\delta\tilde{\mathbb{E}}V_{t+1}^i(0, y_{t+1}; s_{-i,t+1}) \Big\}, \\ \quad \text{if } x_{it} = 0 \end{cases}$$

where $\omega(b_{it}, s_{-it})$ is the estimated probability of winning the auction given s_{-it} and b_{it}, when $x_n \geq 1$ and $x_n = 0$, respectively. $\tilde{\mathbb{E}}[z_t|b_{it}, s_{-it}]$ denotes the estimated expected second-lowest bid in the tth round auction given b_{it} and s_{-it}.

Since we are dealing with a discounted infinite-horizon problem, each contractee's decision is not time sensitive. Let L be the random variable denoting the length of an incoming task, the above equation could be rewritten as

$$
\begin{aligned}
&V^i(x_i, y; s_{-i}) \qquad (4.2)\\
&= \begin{cases}
-c + \max_{b_i} \Big\{ \delta\tilde{\mathbb{E}} V^i(x_i - 1, L; s_{-i}) \ , \\
\quad -C + \omega(b_i, s_{-i}) \Big[\tilde{\mathbb{E}}[z|b_i, s_{-i}] - d \\
\quad +\delta\tilde{\mathbb{E}} V^i(y - 1, L; s_{-i}) \Big] \\
\quad +[1 - \omega(b_i, s_{-i})]\delta\tilde{\mathbb{E}} V^i(x_i - 1, L; s_{-i}) \Big\} , \\
\quad \text{if } x_i = 1, 2, 3, \ldots \\
\max_{b_i} \Big\{ \delta\tilde{\mathbb{E}} V^i(0, L; s_{-i}) \ , \\
\quad -C + \omega(b_i, s_{-i}) \Big[-c + \tilde{\mathbb{E}}[z|b_i, s_{-i}] + \\
\quad \delta\tilde{\mathbb{E}} V^i(y - 1, L; s_{-i}) \Big] \\
\quad + [1 - \omega(b_i, s_{-i})]\delta\tilde{\mathbb{E}} V^i(0, L; s_{-i}) \Big\} , \\
\quad \text{if } x_i = 0
\end{cases}
\end{aligned}
$$

We have the following propositions that help determine the form of optimal equilibrium strategies:

PROPOSITION 1. *When $x \geq 1$, $V^i(x, y; s_{-i})$ is decreasing in x.*

PROOF. Intuitively, it is easy to see that $V(x, y; s_{-i})$ is decreasing in x since the agent is always better off with a shorter task duration. We can show this by induction. To begin with, we look at the finite-horizon problem defined in (4.2). Assume that there are N periods in total. Define the terminal equation as

$$V^i_{N+1}(x_{i,N+1}, y_{N+1}; s_{-i,N+1}) = -cx_{i,N+1}.$$

The above equation suggests that in the final periods, each contractee, if busy, will continue to work unless the task is finished.

Clearly $V^i_{N+1}(x_{i,N+1}, y_{N+1}; s_{-i,N+1})$ decreases in $x_{i,N+1}$. Now suppose that $V^i_k(x_{i,k}, y_k; s_{-i,k})$ decreases in $x_{i,k}$, we see

$$
\begin{aligned}
&V^i_{k-1}(x_{i,k-1}, y_{k-1}; s_{-i,k-1}) \\
&= -c + \max_{bi,k-1} \Big\{ \delta\tilde{\mathbb{E}} V^i_k(x^i_{k-1} - 1, y_k; s_{-i,k-1}) \ , \\
&- C + \omega(b_{i,k-1}, s_{-i,k-1}) \cdot [\tilde{\mathbb{E}}[z|b_{i,k-1}, s_{-i,k-1}] - d \\
&+ \delta\tilde{\mathbb{E}} V^i_k(y_{k-1} - 1, y_k; s_{-i,k})] \\
&+[1 - \omega(b_{i,k-1}|s_{-i,k-1})]\delta\tilde{\mathbb{E}} V^i_k(x_{i,k-1} - 1, y_k; s_{-i,k}) \Big\}
\end{aligned}
$$

is decreasing $x_{i,k-1}$ because, for any given $y_{k-1}, s_{-i,k-1}, b_{i,k-1}$, both of the following two functions are increasing in $x_{i,k-1}$:

$$
\begin{aligned}
&f_1(x_{i,k-1}; y_{k-1}, s_{-i,k-1}, bi, k - 1) \\
&= \delta\tilde{\mathbb{E}} V^i_k(x^i_{k-1} - 1, y_k; s_{-i,k-1}) \\
&f_2(x_{i,k-1}; y_{k-1}, s_{-i,k-1}, bi, k - 1) \\
&= -C + \omega(b_{i,k-1}, s_{-i,k-1}) \cdot \{\tilde{\mathbb{E}}[z|b_{i,k-1}, s_{-i,k-1}] \\
&- d + \delta\tilde{\mathbb{E}} V^i_k(y_{k-1} - 1, y_k; s_{-i,k})\} \\
&+ [1 - \omega(b_{i,k-1}|s_{-i,k-1})]\delta\tilde{\mathbb{E}} V^i_k(x_{i,k-1} - 1, y_k; s_{-i,k})
\end{aligned}
$$

We have shown that $V^i_k(x_{i,k}, y_k; s_{-i,k})$ decreases in $x_{i,k}$, $k = 1, \ldots, N + 1$ in the above finite-horizon MDP. Now let $N \to \infty$, we see that $V^i(x, y; s_{-i})$ is decreasing in x. □

PROPOSITION 2. *$V^i(x, y; s_{-i})$ is decreasing in y.*

PROOF. An intuitive explanation to this proposition is that a larger task duration increases a contractee's cost, i.e., total costs throughout the duration of the task, and possible decommitment fees incurred in order to switch to a shorter task. Strict proof can be done by induction and is similar to that of Proposition 1. □

4.2 Structural Forms of Equilibrium Strategies

In this section, we establish the contractees' optimal bidding strategies (to bid or not to bid, and the optimal bidding price). We also apply our analytical framework to study the special case in which commitment flexibility is not allowed.

Before we derive our optimal bidding strategy, we make the assumption that a contractee with a shorter remaining time to finish the current task would like to place a higher bid. This makes sense because the contractee, when facing a shorter x_i, has less incentive to switch to other tasks. Similarly, we assume that a contractee would like to place a higher bid for a task with longer duration. This makes sense since tasks with longer duration have higher opportunity cost for the contractees.

To find out a symmetric bidding strategy, each contractee assumes that the policy function of other contractees is stationary, and makes its decisions conditioned on the stationary distributions formed when all the contractees choose symmetric strategies. This assumption is usually referred to as "fixed-strategy assumption" and made primarily for simplicity of modeling. Athey and Segal [3] give a review of dynamic mechanism design literature, and find out it is a common practice to assume that information is independent across periods and each agent has little access to information over time about other agent's type. Hu and Wellman [6] show that it is reasonable to make such an assumption when studying contractee's learning behaviors in that it is easy to model and implement, and, in some cases, such a model outperforms more sophisticated models.

THEOREM 1. ***(Optimal Bidding Strategy)***

(i) A participant (denoted by i) with time-to-go x in the auction for a task of duration y, given the equilibrium strategies of other contractees, will place a bid in the amount $b^*(x, y; s_{-i})$ such that

$$
\begin{aligned}
&b^*_i(x_i, y; s_{-i}) \\
&= \begin{cases}
\delta\tilde{\mathbb{E}}\left[V^i(x_i - 1, L) - V^i(y - 1, L)\right] + d \\
\quad + \frac{C}{\omega(b^*_i(x_i,y;s_{-i}),s_{-i})}, \text{ if } x_i = 1, 2, \ldots \\
\delta\tilde{\mathbb{E}}\left[V^i(0, L) - V^i(y - 1, L)\right] \\
\quad + \frac{C}{\omega(b^*_i(0,y;s_{-i}),s_{-i})} + c, \text{ if } x_i = 0
\end{cases} \qquad (4.3)
\end{aligned}
$$

(ii) If $x_i \geq 1$, contractee i's willingness to bid is increasing in x_i.

(iii) A contractee's willingness to bid is decreasing in y.

(iv) A busy contractee will choose to participate in the auction if and only if $y \leq \ell(x)$, where $\ell(x)$ is increasing in x; An idle contractee will choose to participate in the auction if and only if $y \leq \ell_0$, where ℓ_0 is a constant.

PROOF. (i) Our model can be viewed as a correlated private-value second-price auction in which each contractee's remaining duration is a private signal. Facing the same set of environmental variables (e.g., y, C,

c, d), the only factor that separates one contractee's decision from another contractee's is only their own remaining duration. Applying existing auction theoretic results (see, for example, [8] and [9]), contractee i's optimal bidding price makes him indifferent as to whether to participate in the auction or not. In addition, contractee i bids his own true valuation, i.e., we might replace $E[\text{second}|b_i, s_{-i}]$ with b_i^*.

When $x_i = 1, 2, 3\ldots$ The optimal bidding price b_i^* must satisfy:

$$\begin{aligned}
&\delta\tilde{\mathbb{E}}V^i(x_i-1, L; s_{-i}) \\
&= -C + \omega(b_i^*, s_{-i})\left[b_i^* - d + \delta\tilde{\mathbb{E}}V^i(y-1, L; s_{-i})\right] \\
&+ [1-\omega(b_i^*, s_{-i})]\delta\tilde{\mathbb{E}}V^i(x_i-1, y; s_{-i}) \qquad (4.4)
\end{aligned}$$

which gives

$$\begin{aligned}
b_i^*(x_i, y; s_{-i}) &= \delta\tilde{\mathbb{E}}\left[V^i(x_i-1, L) - V^i(y-1, L)\right] + d \\
&+ \frac{C}{\omega(b_i^*(x_i, y; s_{-i}), s_{-i})}. \qquad (4.5)
\end{aligned}$$

Similarly, we see that

$$\begin{aligned}
b_i^*(0, y; s_{-i}) &= \delta\tilde{\mathbb{E}}\left[V^i(0, L) - V^i(y-1, L)\right] \\
&+ \frac{C}{\omega(b_i^*(0, y; s_{-i}), s_{-i})} + c.
\end{aligned}$$

(ii) We have previously assumed that $b_i^*(x_i, y; s_{-i})$ decreases in x_i. What we need to do now is to verify that this actually holds.

Examine the right-hand side of (4.5), given s_{-i} when x_i increases, it follows from Proposition 1 that

$$\tilde{\mathbb{E}}\left[V^i(x_i-1, L) - V^i(y-1, L)\right]$$

decreases. Since $b_i^*(x_i, y; s_{-i})$ decreases in x_i, $\omega(b_i^*, s_{-i})$ will increase (a larger bid implies a lower chance of winning), which yields a decreasing $\frac{C}{\omega(b_i^*(x_i, y; s_{-i}), s_{-i})}$. We then see that both the right-hand and left-hand sides of (4.5) are decreasing in x_i.

We have assumed that the highest reward that a contractee could receive from accomplishing a task is fixed, which means a large $b_i^*(x_i, y; s_{-i})$ would put a contractee in a disadvantaged position. Hence contractee i's willingness to bid is decreasing in x_i.

(iii) Similar to the proof of (ii).

(iv) Combining (ii) and (iii), we see that contractee i's willingness to bid is increasing in x_i but decreasing in y_i, which is exactly what we need to prove. Since all the contractees are homogeneous, all the contractees must follow the same strategies, which completes the proof.

(i)–(iv) give the stationary, symmetric optimal equilibrium strategies for the repeated second-price auction game. □

A Special Case: No Commitment Flexibility

We have provided analysis for the general case. Now we study the special case in which commitment flexibility is disabled. We will see that its MDP formulation as well as optimal bidding strategies are in simplified forms.

THEOREM 2. *If decommitment is forbidden, at the beginning of period t, an idle agent's strategy is to not to bid unless $y \le \hat{\ell}$, where $\hat{\ell}$ is a constant.*

PROOF. When full commitment is assumed, each contractee's problem can be formulated as the following MDP (we inherit most of the notation in (4.1) except that we have to define a state variable for each contractee, since a contractee must be idle to be eligible for participating in the auction):

$$\begin{aligned}
V^i(y; s_{-i}) = \max_{b_i} \Big\{ &\delta\tilde{\mathbb{E}}V^i(L; s_{-i}), -C + \omega(b_i, s_{-i})\cdot \\
&\Big\{\tilde{\mathbb{E}}[z|s_i, b_{-i}] - yc + \delta^y\tilde{\mathbb{E}}V^i(L; s_{-i})\Big\} \\
&+[1-\omega(b_i, s_{-i})]\delta\tilde{\mathbb{E}}V^i(L; s_{-i})\Big\}
\end{aligned}$$

where $\omega(b_i, s_{-i})$ denotes the contractee's probability of winning the auction given b_i, s_{-i}.

We can then use similar argument as in the proof of Part (i) of Theorem 1 to show that it is optimal for the contractee not to bid unless y is below a certain threshold. □

Theorems 1 and 2 establish the basis for our simulation-based study in the following section. Theorem 2 implies that, when only full commitments are assumed, the equilibrium strategy for each idle agent is to bid when the duration of the incoming task is equal to or lower than a fixed threshold, and reject all the tasks with durations longer than the threshold. In the contrast, Theorem 1 says that in an environment where commitment flexibility is enabled, agents have different thresholds of task durations as to whether to participate in the auction or not, depending on each agent's state. Such commitment flexibility can make the system more "friendly" to tasks with longer durations and thus can accommodate more tasks in a given period.

4.3 Extension: Different Reward Mechanisms

The model we presented above does not consider comparison of different reward mechanisms. Consider, for instance, that a contractee is not given his full reward immediately after he wins an auction. Instead, he has to finish his task to get the corresponding reward. Another possible mechanism is that the contractee might receive partial rewards period by period according to his progress.

Our POMDP model can be revised to accommodate such differences by adding a new variable r_i denoting the remaining reward from the currently active task (if any) to a contractee's state. Threshold policy in terms of x_i, y and r_i can be shown in similar fashion and is omitted here.

5. NUMERICAL STUDY

5.1 Experiment Design

We have previously established the structure of optimal equilibrium strategies. To have a concrete understanding of the benefits of commitment flexibility under different settings, we design a set of comparative experiments.

We resort to heuristics in our numeric study, recognizing that it is difficult to calculate the exact optimal solutions implied by our MDP model due to the following two difficulties:

- The number of bidders in each auction is uncertain. In a traditional task-allocation problem with an auction protocol, only idle agents would bid for the incoming task. Our framework, in contrast, allows every contractee, busy or idle, to participate in the auction as long as it is profitable for him to do so. What makes our problem even more challenging is the existence of entry fee.

- Each contractee's valuation of the same new task depends on his current state. We do not have a straightforward expression of the probability distribution of agents' valuation of an incoming task, as opposed to well-studied auction problems in the Economics literature (e.g., [11]). Furthermore, the distribution of each agent's progress is closely related to the equilibrium strategies of the agent. This further separates our paper from [14], in which an agent's valuation of the incoming task solely depends on his own capabilities and the state of the world.

Without commitment flexibility, we assume that each contractee will bid at the beginning of period t if and only if

$$y_t \leq \rho \cdot (L_{\max} - L_{\min}) + L_{\min}$$

where $0 < \rho \leq 1$ is a parameter that determines each agent's bidding strategy. A bidding agent's bid, depending on the duration of the incoming task, will be

$$b(y_t) = (\delta - \delta^{y_t})\mathbf{E}_{y_{t+1}} V(y_{t+1}) + y_t c + \frac{C}{\omega'},$$

$\delta \approx 1$ further simplifies the bid as

$$b(y_t) = y_t c + \frac{C}{\omega'},$$

where ω' is only partial known but can be estimated as follows: each contractee uses his own prior probability of winning to determine his bidding prices, and then use the bidding history over a period of time to update $\hat{\omega}'$ iteratively.

With commitment flexibility, we assume that contractee i will use the following strategy: to bid if and only if

$$y_t \leq \rho_1 \cdot (L_{\max} - L_{\min}) + L_{\min}$$

when the agent is idle; to bid if and only if

$$y_t \leq \rho_2 \cdot x_{it}$$

when the agent is busy. An idle contractee will place a bid in the amount of

$$b(0, y_t) = y_t c + \frac{C}{\omega_0},$$

while a busy contractee will place a bid in the amount of

$$b(x_{it}, y_t) = (x_{it} - y_t)c + d + \frac{C}{\omega(x_{it})}.$$

Both the values of ω_0 and $\omega(x_t)$ can be obtained through an iterative update approach (similar to the way we learn ω').

5.2 Simulation Experiments

Our model fits into the situation marked by dynamically incoming tasks and relatively limited number of agents. Our default experimental parameters reflect such a feature with $N = 7$, $C = 10$, $M = 100$, $c = 5$ and $d = 45$. We further assume that L, the duration of each incoming task, is uniformly distributed between $L_{\min} = 2$ and $L_{\max} = 20$. We observe 1000 periods and assume that the discount factor $\delta \approx 1$. In each set of experiment, we change one parameter and keep all the other parameters unchanged.

Define THC (Task handling capability) as

$$THC = \frac{\text{Number of completed tasks}}{\text{The maximum possible number of completed tasks}}$$

THC measures a system's ability to handle tasks.

5.2.1 *Different Decommitment Penalties*

We want to find out the behavior of the system over different values of decommitment penalties. The results are shown in Figures 1 and 3. From our experimental design, we see that $0 \leq \rho_1^* \leq 1$ measures an idle contractee's willingness to bid: the larger ρ_1^* is, the contractee is more willing to bid for an incoming task. Similarly, $0 \leq \rho_2^* \leq 1$ measures a busy contractee's willingness to participate in the auction for the incoming task, which makes it possible for him to switch to a more attractive task.

We observe from Figure 1 that when the decommitment penalty is set too low, contractees have the incentive to bid for most of the incoming tasks and always to decommit from the current task, which leads to low productivity (see Figure 3). When the decommitment penalty is set too high, contractees are discouraged to decommit from long tasks, which make the system's ability of handling tasks relatively low.

The results also enable us to look at the benefits of commitment flexibility. When commitment flexibility is not allowed, each contractee would choose $\rho^* = 0.7$, resulting in $THC = 73.72\%$ and a total discounted reward of 972. In contrast, when commitment flexibility is enabled, contractees would choose $\rho_1^* = 0.95, \rho_2^* = 0.65$, which leads to $THC = 85.13\%$ and a total discounted reward of 1194, or a 15% improvement in THC.

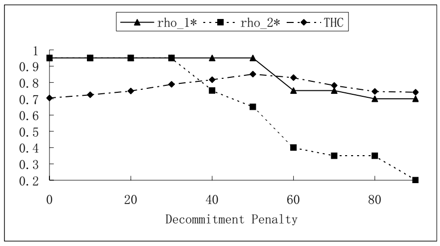

Figure 1: ρ_1^*, ρ_2^* and THC for different decommitment penalties ($C = 10, c = 5, M = 100, L \sim [2, 20], N = 7$)

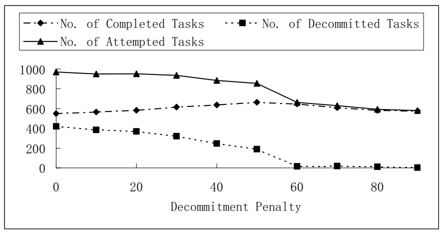

Figure 2: Numbers of attempted, completed and decommitted tasks for different commitment penalties ($C = 10, c = 5, M = 100, L \sim [2, 20], N = 7$)

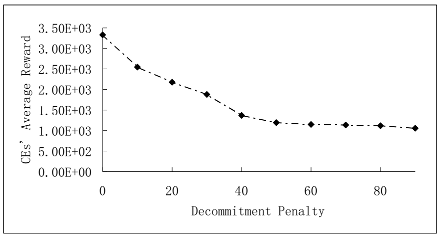

Figure 3: Contractee's average reward for different decommitment penalties ($C = 10, c = 5, M = 100, L \sim [2, 20], N = 7$)

5.2.2 *Different Entry Fees*

Choosing appropriate entry fee is also important in designing the auction mechanism. Our simulation results, as shown in Figures 4 and 5, reveal that a low entry fee will lead to low system productivity (as indicated by THC) because it encourages busy contractees to participate in auctions more frequently. When the entry fee is set to be too high, the system's task handling ability is inhibited due to contractees' limited participation.

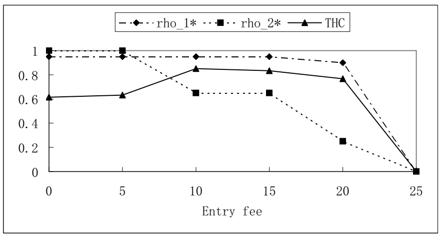

Figure 4: ρ_1^*, ρ_2^* and THC for different entry fees ($d = 50, c = 5, M = 100, L \sim [2, 20], N = 7$)

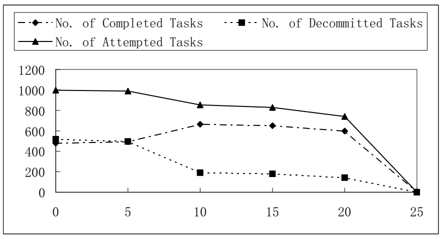

Figure 5: Numbers of attempted, completed and decommitted tasks for different entry fees ($d = 50, c = 5, M = 100, L \sim [2, 20], N = 7$)

5.2.3 *Different Maximum Rewards Per Task*

We observe from the experiment results (Figure 6) that a larger M in general leads to high system productivity. If, however, the contractor's objective is not just to improve task handling ability, but also to reduce the total payments to contractees, the choice will be between cost and performance.

5.2.4 *Different Task Duration Distributions*

We fix $L_{\min} = 2$ and observe the changes brought by different $L_{\max}$'s. Our results (Figure 7) indicate that the system's ability of handling tasks decreases when $L_{\max}$ increases.

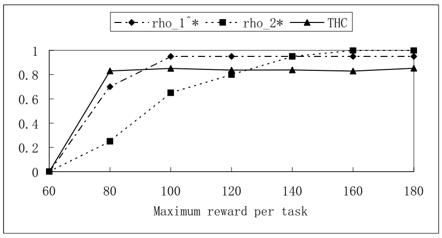

Figure 6: ρ_1^*, ρ_2^* and THC for different M($d = 50, C = 10, c = 5, L \sim [2, 20], N = 7$)

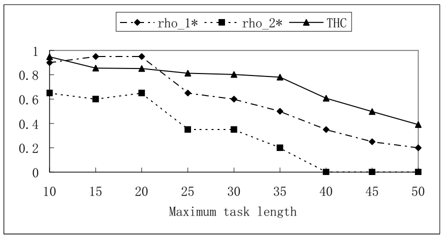

Figure 7: ρ_1^*, ρ_2^* and THC for different $L_{\max}$($d = 50, C = 10, c = 5, M = 100, L \sim [2, L_{\max}], N = 7$)

5.3 Discussions

Our experiments reveal the power of commitment flexibility. We might view ρ or ρ_1 (as used in Section 5.1) as the threshold of an idle contractee when deciding whether to bid for the incoming task or not: a lower ρ or ρ_1 means that the contractee is more "picky" about the duration of tasks, while a higher ρ or ρ_1 makes it possible for relatively longer tasks to be put on the auction. When there is no commitment flexibility, the contractee would choose a relatively low ρ, after accounting for the opportunity cost brought by a long task. When a contractee is allowed to decommit from his current task, however, he would take a more relaxed manner when making an initial bid for a task, i.e., he can tolerate long tasks, because he still has a "second chance" made possible by commitment flexibility.

6. DISCUSSION AND FUTURE RESEARCH

In this paper, we study a dynamic task-allocation problem using repeated second-price auctions. We provide mathematical formulations and analysis, which lead to structural results of the equilibrium strategies. Our model is helpful for multi-agent system designers in revealing the benefits of commitment flexibility. In addition to the existing literature and our current work, more sophisticated models are needed. Below are several potential extensions of the current model that we plan to pursue:

1. Different values. Our model assumes a constant maximum reward for any incoming task. What would contractees' behavior vary if the incoming tasks possess different value? Such an extension can give us a more comprehensive understanding of the benefits of commitment flexibility.

2. Bid on decommitted tasks. Under our current problem settings, a decommitted task is discarded and cannot be completed. Allowing contractees to bid on decommitted tasks would further improve the flexibility of the system. This is aligned with the notion of resale in the auction literature.

3. Team work. It would be interesting to consider the case when contractees can form teams to handle an incoming task. More interestingly, how does the team dynamics among contractees affect the contractor's decision of optimal bidding mechanism that induces contractees to complete as many tasks as possible? We leave this extension for future research.

Acknowledgement

This research was funded by AFOSR grants FA95500810356, FA95500710039 and ARO grant W911NF0810301. We appreciate Alan Scheller-Wolf, Isa Hafalir and Praveen Paruchuri for their helpful comments in the preparation of this paper.

7. REFERENCES

[1] S. Abdallah and V. Lesser. Modeling task allocation using a decision theoretic model. Utrecht, Netherlands, 2005. AAMAS'05.

[2] M. Andersson and T. Sandholm. Leveled commitment contracts with myopic and strategic agents. *Journal of Economic Dynamics and Control*, 25:615–640, 2001.

[3] S. Athey and I. Segal. An efficient dynamic mechanism. Technical report, UCLA Department of Economics, 2007.

[4] F. Brandt, W. Brauer, and G. Weiss. Task assignment in multiagent systems based on vickrey-type auctioning and leveled commitment contracting. *Cooperative Information Agents IV, Lecture Notes in Artificial Intelligence*, 1860:95–106, 2000.

[5] M. Dias and A. Stentz. A comparative study between centralized, market-based, and behavioral multirobot coordination approaches. In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '03)*, volume 3, pages 2279–2284, October 2003.

[6] J. Hu and M. P. Wellman. Online learning about other agents in a dynamic multiagent system. In *Proceedings of the second international conference on Autonomous agents*, pages 239–246, Minneapolis, Minnesota, 1998.

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

[8] V. Krishna. *Auction Theory*. Academic Press, 2002.

[9] P. Milgrom. *Putting Auction Theory to Work*. Cambridge University Press, 2004.

[10] I. Nourbakhsh, M. Lewis, K. Sycara, M. Koes, M. Yong, and S. Burion. Human-robot teaming for search and rescue. *IEEE Pervasive Computing*, 4(1):72–78, Jan-Mar 2005.

[11] J. G. Riley. Expected revenue from open and sealed bid auctions. *Journal of Economic Perspectives*, 3:41–50, 1989.

[12] T. Sandholm and V. Lesser. Leveled commitment contracts and strategic breach. *Games and Economic Behavior*, 35:212–270, 2001.

[13] T. Sandholm and V. Lesser. Leveled-commitment contraction. *AI Magazine*, 23:89–100, 2002.

[14] D. Sarne, M. Hadad, and S. Kraus. Auction equilibrium strategies for task allocation in uncertain environments. *Lecture Notes in Computer Science*, 3191:271–285, 2004.

[15] P. Scerri, T. Von Goten, J. Fudge, S. Owens, and K. Sycara. Transitioning multiagent technology to uav applications. In *Proceedings of AAMAS 2008 Industry Track*, 2008. to appear.

[16] R. G. Smith. The contract net protocol high-level communication and control in a distributed problem solver. *IEEE Transactions on Computers*, 29(12):1104–1113, 1980.