Probability in the Engineering and Informational Sciences, **27**, 2013, 277–296.
doi:10.1017/S0269964813000028

STOCHASTIC SEQUENTIAL ASSIGNMENT PROBLEM WITH THRESHOLD CRITERIA

Golshid Baharian
Industrial and Systems Engineering
University of Illinois at Urbana-Champaign
Urbana, IL
E-mail: gbahari2@illinois.edu

Sheldon H. Jacobson
Computer Science
University of Illinois at Urbana-Champaign
Urbana, IL
E-mail: shj@illinois.edu

The stochastic sequential assignment problem (SSAP) allocates distinct workers to sequentially arriving tasks with stochastic parameters to maximize the expected total reward. In this paper, the assignment of tasks is performed under the *threshold criterion*, which seeks a policy that minimizes the probability of the total reward failing to achieve a target value. A Markov-decision-process approach is employed to model the problem, and sufficient conditions for the existence of a deterministic Markov optimal policy are derived, along with fundamental properties of the optimal value function. An algorithm to approximate the optimal value function is presented, and convergence results are established.

1. INTRODUCTION

Consider the stochastic sequential assignment problem (SSAP) introduced by Derman, Lieberman, and Ross [4]: There are n workers available to perform n *i.i.d.* sequentially-arriving tasks, where the random variable X_j denotes the value of the jth task that arrives during time period j, and a fixed value (or success rate) p_i is associated with worker i. Whenever the ith worker is assigned to the jth task, the worker becomes unavailable for future assignments, with the expected reward associated with this assignment given by $p_i x_j$, where x_j is the observed value of the jth task.

Several extensions to the stochastic sequential assignment problem have been discussed in the literature. Albright [1], Albright [2], and Righter [12] study the SSAP with various task–arrival-time distributions. Nikolaev and Jacobson [6] consider a variation of SSAP in which the number of tasks is unknown until after the final arrival and follows a given probability distribution. An application of SSAP in kidney allocation to patients is addressed by Su and Zenios [15]. McLay, Jacobson, and Nikolaev [5] and Nikolaev, Jacobson, and McLay [7] address applications of SSAP in aviation security. Existing SSAP literature

 0269-9648/13 $25.00

focuses on a risk-neutral objective function, seeking a policy that maximizes the expected total reward obtained from the sequential assignment of tasks to workers. However, a risk-neutral policy is not always desirable since the probability distribution function (*p.d.f.*) of the total reward may carry with it a high probability of low unaccepted values; therefore, there are instances that a decision-maker is interested in a stable reward and looks for a risk-sensitive optimal assignment policy.

The work presented here is distinct from the existing literature in two ways. First, it considers the SSAP under a different objective function, termed the *threshold criterion*, which seeks to find a policy that minimizes the *threshold probability*: the probability (or risk) of the total reward failing to achieve a specified value (target or *threshold*). Specifically, let

$$R_n^\phi = \sum_{i=1}^{n} p_{\phi(\tilde{d}_i)} X_i$$

denote the total reward obtained after assigning all n tasks to available workers under policy ϕ, where $\phi(\tilde{d}_i)$ is the index of the worker assigned to the ith task under ϕ. For a given target value τ, the goal is to find an optimal policy ϕ^* that achieves the following infimum:

$$\inf_{\phi\in\Phi} P_\phi\left\{R_n^\phi \le \tau\right\},$$

where Φ is the set of all admissible policies. For simplicity, the target-dependent stochastic sequential assignment problem is denoted here as the TSSAP, and a n-stage TSSAP refers to a TSSAP with n tasks and n workers.

The second distinction between the work presented here and the existing literature is that the problem is modelled as a Markov decision process (MDP) and results in an uncountable-state-space MDP, while the countability of the state space is a basic assumption in the existing target-dependent, risk measure literature. Hence, the present work extends the threshold criteria literature to uncountable-state-space MDPs and obtains sufficient conditions for the existence of a deterministic Markov optimal policy. Fundamental characteristics of the optimal value function and the optimal policy are also presented. Finally, the algorithm proposed by Boda et al. [3] to approximate the optimal value function is adapted to the uncountable-state-space model, and convergence of the approximate value function to the optimal value function is established under certain conditions.

Several authors have studied Markov decision processes with the threshold criterion. As mentioned before, the focus of these papers is on Markov decision processes over a countable state space. White [16] considers a finite state space MDP with a bounded reward set, and characterizes the optimal value function by an optimality equation. Wu and Lin [17] show that the optimal value function is a distribution function of the target value and prove the existence of an optimal deterministic Markov policy. Sufficient conditions for the existence of an optimal policy for an infinite horizon MDP over a countable state space are provided by Ohtsubo and Toyonaga [10]. An algorithm is proposed by Boda et al. [3] to approximate the optimal value function, which decreases the computation time significantly. Sakaguchi and Ohtsubo [13] consider undiscounted semi-Markov decision processes with countable state and action spaces, with the objective of minimizing the threshold probability. The existence of an optimal stationary policy is proven, and value iteration methods and a policy improvement method are proposed. Other variations and applications of the threshold problem are discussed by Ohtsubo [8], Ohtsubo [9], and Ohtsubo and Toyonaga [11].

The paper is organized as follows. Section 2 mentions potential examples and applications of the TSSAP. Section 3 studies the model of a n-stage TSSAP with discrete task values, describes it as a MDP with a countable state space, and presents optimality equations so as to find a policy that minimizes the threshold probability. Section 4 discusses exact and approximate methods to solve the optimality equations given in Section 3. Section 5 extends the model of a n-stage TSSAP to the case where the *p.d.f.* of task values has uncountable support, which results in a MDP with an uncountable state space. Furthermore, sufficient conditions for the existence of an optimal policy under the threshold criterion and optimality equations to derive the optimal policy are presented. The approximate algorithm discussed in Section 4 is adapted to the generalized TSSAP, and its behavior is studied. Section 6 presents numerical results, and finally, Section 7 provides concluding comments and future directions of the research.

2. ILLUSTRATIVE EXAMPLES

This section provides an example for the TSSAP which demonstrates the application of the threshold criteria. Consider a SSAP which allows sequentially arriving passengers to be assigned to available aviation security resources as they check in at an airport. The time interval for screening passengers is divided into n slots (stages), where passenger j arrives during stage j. Upon the arrival of each passenger, a pre-screening system determines their threat (risk) value, classifying them as non-selectees (i.e., the passengers who have been cleared of posing a risk) or selectees (i.e., those who have not been cleared, based on available information known about them [5]). Each assessed threat value is defined as the probability that a passenger carries a threat item, and the *p.d.f.* for passengers' threat values is denoted by f, with X_j indicating the threat value of passenger j. The capacity of the selectee class (i.e., the number of available screening devices associated with the selectee class) is c, and n denotes the capacity of the non-selectee class. Define the security level to be the conditional probability of detecting a passenger with a threat item given that they are classified as selectees or non-selectees, and let L_{S} and L_{NS} be the security levels associated with the selectee and non-selectee classes. Moreover, let $\gamma_j = 1$ and $\gamma_j = 0$ denote the jth passenger assignment as a selectee and a non-selectee, respectively. The *total security* for this setting is defined as

$$\sum_{j=1}^{n} X_j \left[L_{\mathrm{S}}\gamma_j + L_{\mathrm{NS}}(1-\gamma_j)\right],$$

where the objective is to find a policy for assigning passengers to classes as they check in so as to minimize the probability of the total security failing to achieve the target τ.

In the airport security problem the decision-maker needs to make sure that the total reward obtained is at least as great as a specified value with high probability. In other words, it is critical to obtain a stable level of security at all times. Note that although this security level might not be necessarily the highest possible, a critical goal is to maintain a reasonable security level at all times. Section 3 studies the n-stage TSSAP with discrete task values, describes it as a MDP, and presents optimality equations so as to find a policy that minimizes the threshold probability.

3. MODEL DESCRIPTION

Consider the original SSAP introduced by Derman et al. [4] where n workers are available to perform n i.i.d. sequentially arriving tasks. A random variable X_j denotes the value of the jth task that arrives during time period j, with a fixed value (or success rate) p_i associated with worker i. If the ith worker is assigned to the jth task with observed value x_j, the worker becomes unavailable for future assignments, and the expected reward due to this assignment is given by $p_i x_j$. Throughout this paper, it is assumed that the number of tasks equals the number of workers. To relax this assumption, let m denote the number of tasks, while n is the number of workers. If $m > n$, then we add $m - n$ phantom workers with success rates of 0, while if $m < n$, the $n - m$ workers with the smallest values are dropped so that only those m workers with the highest success rates can be chosen. With such a modification, the number of tasks equals the number of workers, and hence, the problem is simplified to the n-task, n-worker model. Unlike the existing SSAP literature, the problem studied here is under an objective function other than maximizing the expected total reward; specifically, the goal is to find a policy that minimizes the probability (or risk) of the total reward failing to achieve a target value, after assigning all the sequentially-arriving tasks to available workers.

Consider the n-stage TSSAP and assume that the i.i.d. sequentially arriving tasks take on values in the set $\mathcal{S} \subseteq [0, +\infty)$; in addition, a vector $P = (p_1, p_2, \ldots, p_n)$ is given with p_i denoting the success rate (or value) of the ith worker, where worker values are considered to be strictly positive. Given P and for $k = 1, 2, \ldots, n$, let

$$\mathcal{W}_k^P := \left\{ (q_1, q_2, \ldots, q_n) \middle| q_i \in \{0, p_i\} \text{ for } i = 1, 2, \ldots, n \text{ such that } \sum_{j=1}^{n} I_{\{q_j \neq 0\}} = n - k + 1 \right\},$$

denote the set of all possible vectors of worker values at time period k before the assignment of the kth task. Any element of the set $\mathcal{W}_k^P$ has $k - 1$ zero entries, which correspond to the workers that are no longer available, since they have been assigned to previous tasks over the first $k - 1$ time periods. Note that by definition, $\mathcal{W}_1^P = \{P\}$.

Let $\tilde{s}_k$, a_k, and r_k denote the state of the system, the action taken by the decision-maker, and the reward obtained at time period k, respectively. Then, the state of the system at stage k is defined by

$$\tilde{s}_k = (x_k, P^{(k)}) \in \tilde{\mathcal{S}}_k := \mathcal{S} \times \mathcal{W}_k^P,$$

where x_k and $P^{(k)}$ indicate the observed value of the kth task and the vector of success rates at time period k upon the arrival of the kth task, respectively. The state space of the system is thus defined by $\tilde{\mathcal{S}} := \cup_{k=1}^{n} \tilde{\mathcal{S}}_k$. In this section, assume that the state space of task values $\mathcal{S}$ is countable (i.e., $\{X_j\}$ are discrete random variables); the n-stage TSSAP where the task values are continuous random variables with an uncountable state space $\mathcal{S} \subseteq [0, +\infty)$ is studied in Section 5. The objective of the TSSAP signifies the decision-maker's need to consider the target level along with the original state of the system at each decision instance, and hence, the state space of the MDP must be enlarged so as to incorporate the target level at each time period. To this end, define $\tilde{\mathcal{D}} := \cup_{k=1}^{n} \tilde{\mathcal{D}}_k$ to be the updated state space of the MDP (referred to as the state space of the decision-maker) where $\tilde{\mathcal{D}}_k := \tilde{\mathcal{S}}_k \times \mathbb{R}$. Note that $\tilde{\mathcal{S}}$ is the state space of the system and should not be confused with $\tilde{\mathcal{D}}$. At time period k, the action space A at each state is given by the set of workers available for assignment at that state, and hence, $A(x_k, P^{(k)}) = \{ i | P_i^{(k)} \neq 0, 1 \leq i \leq n \}$ where $P_i^{(k)}$ denotes the ith element of $P^{(k)}$, and $\mathcal{A} := \cup_{s \in \tilde{\mathcal{S}}} A(s)$ is the overall action space. It is obvious from this definition that $A(x_k, P^{(k)})$ is independent of x_k; therefore, the action space at a given state $(x_k, P^{(k)})$ can also be denoted by $A(P^{(k)})$. If at time period k and upon the arrival of the kth task, x_k, action $a_k = i \in A(P^{(k)})$ is chosen, then the target level is decreased by the realized reward amount, $r_k = p_i x_k$. Given that the value of the task at time period $k + 1$ is

x, the conditional transition probability corresponding to this state change is defined by

$$f(x) = P\left\{\tilde{d}_{k+1} = ((x, P^{(k)} - p_i e_i), t - p_i x_k) \middle| \tilde{d}_k = ((x_k, P^{(k)}), t), a_k = i\right\},$$

for $\tilde{d}_k \in \tilde{\mathcal{D}}_k$, $\tilde{d}_{k+1} \in \tilde{\mathcal{D}}_{k+1}$, and $i \in A(P^{(k)})$, and $k = 1, 2, \ldots, n-1$, where f is the underlying probability mass function (*pmf*) of task values with support $\mathcal{S}$.

Let H_k denote the set of all admissible histories up to time period k. Given H_k, a *decision rule* ϕ_k at time period k is a conditional probability measure on the action space $\mathcal{A}$ such that

$$\phi_k(A(P^{(k)}) \mid h_k) = 1,$$

for all $h_k \in H_k$ and $k = 1, 2, \ldots, n$. A decision rule ϕ_k, which is applied at time period k upon the arrival of the kth task, is *deterministic* if it is a mapping from H_k onto $\mathcal{A}$ (i.e., $\phi_k(h_k) \in A(P^{(k)})$ for any $h_k \in H_k$). Consider ϕ_k, an arbitrary deterministic decision rule at time period k. ϕ_k is called a *continuous* decision rule in the target value over the interval $[0, \tau] \subset \mathbb{R}$ if for each $((x, P^{(k)}), t) \in \tilde{\mathcal{S}}_k \times [0, \tau]$, there exists $\epsilon > 0$ such that $\phi_k((x, P^{(k)}), s) = \phi_k((x, P^{(k)}), t)$, for all $s \in (t - \epsilon, t + \epsilon)$. Moreover, a sequence $\phi = (\phi_n, \phi_{n-1}, \ldots, \phi_1)$ of decision rules is called a *policy* for a n-stage TSSAP. If ϕ_k only depends on the current state at time k for all $k = 1, 2, \ldots, n$, then the policy ϕ is a *Markov* policy. In addition, ϕ is called a *deterministic* policy if ϕ_k is deterministic for all $k = 1, 2, \ldots, n$. Let Φ, Φ_D, Φ_{M}, and Φ_{DM} denote the sets of all policies, all deterministic policies, all Markov policies, and all deterministic Markov policies, respectively. Finally, define

$$\phi\,|_l := (\phi_n, \phi_{n-1}, \ldots, \phi_{n-l+1}),$$

for any $1 \leq l \leq n$.

Fix an arbitrary policy $\phi \in \Phi$, and define the target-dependent risk measure over the last k time periods under $\phi\,|_k$ as

$$V_k^\phi((x, P^{(n-k+1)}), t) := P_\phi\left\{R_k \leq t \middle| \tilde{d}_{n-k+1} = ((x, P^{(n-k+1)}), t)\right\},$$

for all $\tilde{d}_{n-k+1} \in \tilde{\mathcal{D}}_{n-k+1}$, where $\tilde{d}_{n-k+1}$ denotes the state of the decision-maker at time period $n - k + 1$ (before the assignment of the $(n - k + 1)$th task) and the superscript ϕ in R_k^ϕ is dropped to simplify the notation. Therefore, the optimal value function over n time periods is given by

$$\begin{aligned} V_n^{\phi^*}((x, P), t) &= \inf_{\phi \in \Phi} P_\phi\left\{R_n \leq t \middle| \tilde{d}_1 = ((x, P), t)\right\} \\ &= \inf_{\phi \in \Phi} V_n^\phi((x, P), t), \end{aligned}$$

for all $\tilde{d}_1 \in \tilde{\mathcal{D}}_1$. Observe that for each $\phi \in \Phi$, $V_n^\phi((x, P), t) = 0$ if $t < 0$ since task values are assumed to be non-negative, and hence, $V_n^{\phi^*}((x, P), t) = 0$ for all $t < 0$.

For the n-stage TSSAP described here, the state space of the system $\tilde{\mathcal{S}}$ is countable; in addition, the action space $\mathcal{A}$ is finite. Therefore, Theorem 1 in Wu and Lin [17] implies that an optimal policy exists and is in fact deterministic Markovian (i.e., $\phi^* \in \Phi_{\mathrm{DM}}$); moreover, the following recursive optimality equations are used to derive the optimal policy and the

https://doi.org/10.1017/S0269964813000028 Published online by Cambridge University Press

minimum risk of failing to achieve the target value in the n-stage TSSAP:

$$V_1^{\phi^*}((x, P^{(n)}), t) = I_{\{t \geq p_i x\}}, \tag{3.1}$$

for all $((x, P^{(n)}), t) \in \tilde{\mathcal{D}}_n$ where $A(P^{(n)}) = \{i\}$, and

$$V_k^{\phi^*}((x, P^{(n-k+1)}), t) = \min_{a \in A(P^{(n-k+1)})} E_a V_{k-1}^{\phi^*}(x, P^{(n-k+1)}, t), \tag{3.2}$$

for all $k = 2, 3, \ldots, n$, where

$$E_a V_{k-1}^{\phi}(x, P^{(n-k+1)}, t) := \sum_{y \in \mathcal{S}} V_{k-1}^{\phi}(y, P^{(n-k+1)} - p_a e_a, t - p_a x) f(y),$$

for $(x, P^{(n-k+1)}, t) \in \tilde{\mathcal{S}}_{n-k+1} \times [0, \tau]$, $a \in A(P^{(n-k+1)})$, and $\phi \in \Phi$.

Section 4 discusses the exact method to solve the optimality equations given by (3.1)–(3.2); furthermore, the algorithm proposed by Boda et al. [3] to approximate the optimal value function and the optimal policy is studied, and useful properties of this algorithm are presented.

4. THE APPROXIMATE ALGORITHM

In Section 3, the n-stage TSSAP is formulated in a form similar to classical dynamic programming problems, and hence, it can be solved by the associated backward recursion algorithm. For notational simplicity, let $((x, P^{(n-k+1)}), t)$ be denoted as $(x, P^{(n-k+1)}, t)$ henceforth, for any $((x, P^{(n-k+1)}), t) \in \tilde{\mathcal{S}}_{n-k+1} \times [0, \tau]$. Also, for any $k = 1, 2, \ldots, n$, $P^{(n-k+1)} \in \mathcal{W}_{n-k+1}^P$, and $t \in [0, \tau]$, define the function $V_k^{\phi}(., P^{(n-k+1)}, t) : \mathcal{S} \to [0, 1]$ to be equal to the target-dependent risk-measure V_k^{ϕ} over the last k time periods under $\phi\,|_k$ where $P^{(n-k+1)}$ and t are fixed. Let $V_k^{\phi}(x, ., t) : \mathcal{W}_{n-k+1}^P \to [0, 1]$ and $V_k^{\phi}(x, P^{(n-k+1)}, .) : [0, \tau] \to [0, 1]$ be defined in a similar fashion. The algorithm proposed by Wu and Lin [17] can be modified to compute optimal value functions and optimal policies for the n-stage TSSAP provided that $\mathcal{S}$ is finite. For any $1 \leq k \leq n$, each given $x \in \mathcal{S}$, and $P^{(n-k+1)} \in \mathcal{W}_{n-k+1}^P$, it follows from this recursive algorithm that $V_k^{\phi^*}(x, P^{(n-k+1)}, t)$ is a step distribution function of t with finite jump points. Let $\{\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_z\}$ be the set of possible rewards that can be obtained during a single arbitrary time period. In addition, let $J_k = \{u_1, u_2, \ldots, u_{j_k}\}$ be the set of all jump points obtained when solving for $V_k^{\phi^*}$ in (3.2). Arrange the $\{u_l + r_i\}$ in ascending order for all $l = 1, 2, \ldots, j_k$ and $i = 1, 2, \ldots, z$ so as to obtain ordered values $v_1 < v_2 < \cdots < v_M$. It is shown by Wu and Lin [17] that all the jump points of $V_{k+1}^{\phi^*}$ belong to the set $\{v_1, v_2, \ldots, v_M\}$, and hence, for any given state of the system $(y, P^{(n-k)})$, one only needs to evaluate $V_{k+1}^{\phi^*}(y, P^{(n-k)}, .)$ at the points in $\{v_1, v_2, \ldots, v_M\}$. Furthermore, $V_{k+1}^{\phi^*}(y, P^{(n-k)}, t) = 0$ for $t < v_1$ and $V_{k+1}^{\phi^*}(y, P^{(n-k)}, t) = 1$ for $t \geq v_M$.

The algorithm proposed by Wu and Lin [17] quickly becomes computationally inefficient since a growing number of jump points must be considered as one moves backward through each successive stage. In fact, the number of points to consider and the computation time to perform the algorithm grow exponentially as the state space and the action space expand. Therefore, a straightforward computational substitute for solving (3.1)–(3.2) is used in which computations are done on a suitable fixed grid of target values (see Boda et al. [3]). To this end, one can focus on the interval $[0, \tau]$ where τ is the largest target value that is needed to be considered. Note that taking the lower bound of this interval to be zero is well-justified, since task values are assumed to be non-negative. The interval $[0,\tau]$ is then divided into m subintervals using the grid $B=\{t_0,t_1,\ldots,t_m\}$, where $t_0=0$, $t_m=\tau$, and $t_i<t_{i+1}$ for $i=0,1,\ldots,m-1$. For $k=1,2,\ldots,n$ and each $(x,P^{(n-k+1)})\in\tilde{\mathcal{S}}_{n-k+1}$, the target-dependent risk measure $V_k^{\phi^*}(x,P^{(n-k+1)},.)$ on $[0,\tau]$ is approximated by a set of values $\{(t_0,V_k^{\phi^*}(x,P^{(n-k+1)},t_0)),(t_1,V_k^{\phi^*}(x,P^{(n-k+1)},t_1)),\ldots,(t_m,V_k^{\phi^*}(x,P^{(n-k+1)},t_m))\}$. Note that regardless of the iteration index and the time period, the grid set B is kept fixed. This approximate algorithm is referred to as the Grid Method (GM) for solving the n-stage TSSAP.

Let ϕ_m and $V_k^{\phi_m}$ denote the approximate policy and the approximate risk measure over the last k time periods obtained from the GM using the grid set $B=\{t_0,t_1,\ldots,t_m\}$, respectively. Analogous to (3.1)–(3.2), $V_k^{\phi_m}$ is evaluated at the t_i's recursively from $V_{k-1}^{\phi_m}$, and the interpolation of $V_k^{\phi_m}$ between the grid points is performed using any desired approximation. The GM approximation equations are given by

$$V_1^{\phi_m}(x,P^{(n)},t):=V_1^{\phi^*}(x,P^{(n)},t), \tag{4.1}$$

for all $(x,P^{(n)},t)\in\tilde{\mathcal{S}}_n\times[0,\tau]$, and

$$V_k^{\phi_m}(x,P^{(n-k+1)},t_i):=\min_{a\in A(P^{(n-k+1)})}E_aV_{k-1}^{\phi_m}(x,P^{(n-k+1)},t_i), \tag{4.2}$$

for all $(x,P^{(n-k+1)})\in\tilde{\mathcal{S}}_{n-k+1}$, $t_i\in B$, and $k=2,3,\ldots,n$, with the interpolation

$$V_k^{\phi_m}(x,P^{(n-k+1)},t):=V_k^{\phi_m}(x,P^{(n-k+1)},t_{i-1}), \tag{4.3}$$

if $t\in[t_{i-1},t_i)$ for some $1\le i\le m$. Observe that by (4.1)–(4.3), $V_k^{\phi_m}(x,P^{(n-k+1)},.)$ is a step function, with its jump points belonging to the grid set B.

Lemma 4.1: *The GM defined by (4.1)–(4.3) with the grid set* $B=\{t_0,t_1,t_2,\ldots,t_m\}$ *provides a lower bound for the optimal target-dependent risk measure* $V_n^{\phi^*}$. *Therefore,*

$$V_n^{\phi_m}(x,P,t)\le V_n^{\phi^*}(x,P,t), \tag{4.4}$$

for all $(x,P)\in\tilde{\mathcal{S}}_1$ *and* $t\in[0,\tau]$. *Moreover,* $V_n^{\phi_m}(x,P,.)$ *is a non-decreasing step function on* $[0,\tau]$ *for each* $(x,P)\in\tilde{\mathcal{S}}_1$.

Proof: The proof is by induction on n starting with $n=2$ tasks as the base case. Observe that for $i=1,2,\ldots,m$ and $(x,P^{(n-1)})\in\tilde{\mathcal{S}}_{n-1}$,

$$\begin{aligned}V_2^{\phi_m}(x,P^{(n-1)},t_i)&=\min_{a\in A(P)}E_aV_1^{\phi_m}(x,P^{(n-1)},t_i)\\&=\min_{a\in A(P)}E_aV_1^{\phi^*}(x,P^{(n-1)},t_i)=V_2^{\phi^*}(x,P^{(n-1)},t_i),\end{aligned} \tag{4.5}$$

where the second equality follows from (4.1). In other words, $V_2^{\phi_m}$ and $V_2^{\phi^*}$ coincide at all the grid points in B, and hence, (4.4) holds true if $t\in B$. Now, assume that $t\in(t_i,t_{i+1})$ for some $0\le i\le m-1$, and note that

$$V_2^{\phi_m}(x,P^{(n-1)},t)=V_2^{\phi_m}(x,P^{(n-1)},t_i)=V_2^{\phi^*}(x,P^{(n-1)},t_i)\le V_2^{\phi^*}(x,P^{(n-1)},t),$$

where the first and second equalities follow, respectively, from (4.3) and (4.5), and the inequality is obtained since $V_2^{\phi^*}(x,P^{(n-1)},t)$ is a distribution function (and hence, a non-decreasing function) of t on $[0,\tau]$ (see Wu and Lin [17]). It is also inferred that $V_2^{\phi_m}(x, P^{(n-1)}, .)$ is a non-decreasing step function due to (4.5) and the fact that $V_2^{\phi^*}(x, P^{(n-1)}, .)$ is non-decreasing on $[0,\tau]$. For the induction step, assume that (4.4) holds true for $n-1$ where $n \geq 3$ and that $V_{n-1}^{\phi_m}$ is a non-decreasing step function on $[0,\tau]$. Now fix an arbitrary $t \in [0,\tau]$, and without loss of generality assume that $t \in [t_i, t_{i+1})$ for some $0 \leq i \leq m-1$. To prove the lemma for n, observe that

$$V_n^{\phi_m}(x,P,t) = V_n^{\phi_m}(x,P,t_i) = \min_{a \in A(P)} E_a V_{n-1}^{\phi_m}(x,P,t_i),$$

and note that since (4.4) holds for $n-1$, it follows that $E_a V_{n-1}^{\phi_m}(x,P,t_i) \leq E_a V_{n-1}^{\phi^*}(x,P,t_i)$ for any $a \in A(P)$. Therefore, $V_n^{\phi_m}(x,P,t_i) \leq V_n^{\phi^*}(x,P,t_i) \leq V_n^{\phi^*}(x,P,t)$, and hence, (4.4) holds true for n. That $V_n^{\phi_m}(x,P,.)$ is a non-decreasing function on $[0,\tau]$, follows from the induction assumption and Proposition 5 in [14]. ■

Lemma 4.2 studies the behavior of the GM defined by (4.1)–(4.3) as the size of the grid set B increases.

Lemma 4.2: *Consider two sets of breakpoints $B_1 = \{t_0, t_1, \ldots, t_{m_1}\}$ and $B_2 = \{v_0, v_1, \ldots, v_{m_2}\}$ where $m_1 < m_2$ and $B_1 \subset B_2$ which implies that B_2 provides a finer grid on $[0,\tau]$. Let ϕ_{m_1} and ϕ_{m_2} denote the approximate policies obtained from the GM with grid sets B_1 and B_2, respectively. Then,*

$$V_n^{\phi_{m_1}}(x,P,t) \leq V_n^{\phi_{m_2}}(x,P,t), \tag{4.6}$$

for all $(x,P,t) \in \tilde{\mathcal{S}}_1 \times [0,\tau]$.

Proof: The proof is by induction on n, with the base of induction starting from $n=2$. It suffices to show that (4.6) holds for all the elements of B_2 since $V_n^{\phi_{m_1}}(x,P,.)$ and $V_n^{\phi_{m_2}}(x,P,.)$ are step functions whose jump points are elements of B_1 and B_2, respectively. Fix an arbitrary subinterval $[t_i, t_{i+1})$ from B_1 for some $0 \leq i \leq m_1 - 1$, and re-label the breakpoints of B_2 (if any) that lie within this subinterval as $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_{d_i}\}$. Note that

$$V_2^{\phi_{m_2}}(x, P^{(n-1)}, \bar{v}_l) = V_2^{\phi^*}(x, P^{(n-1)}, \bar{v}_l) \geq V_2^{\phi_{m_1}}(x, P^{(n-1)}, \bar{v}_l),$$

for all $l = 1, 2, \ldots, d_i$ where the inequality follows from Lemma 4.1, and hence, the base of induction is proven. Now, assume that (4.6) holds for $n-1$ where $n \geq 3$, and observe that

$$\begin{aligned} V_n^{\phi_{m_2}}(x,P,\bar{v}_l) &\geq \min_{a \in A(P)} E_a V_{n-1}^{\phi_{m_1}}(x,P,\bar{v}_l) \geq \min_{a \in A(P)} E_a V_{n-1}^{\phi_{m_1}}(x,P,t_i) \\ &= V_n^{\phi_{m_1}}(x,P,t_i) = V_n^{\phi_{m_1}}(x,P,\bar{v}_l), \end{aligned}$$

for all $l = 1, 2, \ldots, d_i$ where the first and second inequalities follow, respectively, from the induction assumption and Lemma 4.1, and the last equality follows from (4.3). ■

Lemma 4.1 together with Lemma 4.2 indicate that the GM defined by Eqs. (4.1)–(4.3) with the grid set $B = \{t_0, t_1, \ldots, t_m\}$ provides a lower-bound step-function approximation for the optimal target-dependent risk measure $V_n^{\phi^*}(x,P,.)$ on $[0,\tau]$ for any initial state of the system $(x,P) \in \tilde{\mathcal{S}}_1$; moreover, one obtains better approximations to $V_n^{\phi^*}$ as the size of the grid set B increases. Recall from Wu and Lin [17] that $V_n^{\phi^*}(x,P,.)$ is a step distribution function on $[0,\tau]$ for a given $(x,P) \in \tilde{\mathcal{S}}_1$. Now, fix an arbitrary initial state $(x,P) \in \tilde{\mathcal{S}}_1$, and let $J = \{t_1^*, t_2^*, \ldots, t_d^*\}$ and h_j denote the set of all jump points and the jth jump size of

$V_n^{\phi^*}(x, P, .)$, respectively. Consider an element of J, t_j^*, that lies within the open subinterval formed by two consecutive elements of B; equivalently, $t_i < t_j^* < t_{i+1}$ for some $1 \leq j \leq d$ and $0 \leq i \leq m-1$. Arbitrarily fix $t \in [t_j^*, t_{i+1})$, and observe that

$$V_n^{\phi^*}(x, P, t) \geq V_n^{\phi^*}(x, P, t_j^*), \tag{4.7}$$

since $t \geq t_j^*$; moreover,

$$V_n^{\phi^*}(x, P, t_j^*) - V_n^{\phi^*}(x, P, t_i) \geq h_j, \tag{4.8}$$

since $t_i < t_j^*$ where t_j^* is the jth jump point of $V_n^{\phi^*}$. Now, note that

$$V_n^{\phi_m}(x, P, t) = V_n^{\phi_m}(x, P, t_i) \leq V_n^{\phi^*}(x, P, t_i), \tag{4.9}$$

which implies that

$$\begin{aligned} V_n^{\phi^*}(x, P, t) - V_n^{\phi_m}(x, P, t) &\geq V_n^{\phi^*}(x, P, t_j^*) - V_n^{\phi_m}(x, P, t) \\ &\geq V_n^{\phi^*}(x, P, t_j^*) - V_n^{\phi^*}(x, P, t_i) \geq h_j, \end{aligned} \tag{4.10}$$

where the inequalities follow from (4.7), (4.9), and (4.8), respectively . Although it is shown with numerical examples by Boda et al. [3] that the GM (applied to a similar problem in a different context) approximates the optimal value function extremely well, while reducing the computation time dramatically, (4.10) implies that the gap between $V_n^{\phi^*}(x, P, t)$ and $V_n^{\phi_m}(x, P, t)$ has a lower bound equal to the jth jump size of $V_n^{\phi^*}(x, P, .)$ which is independent of m. Equivalently, no matter how much m increases, there is a difference between the approximation provided by the GM and the optimal risk measure for all $t \in [t_j^*, t_{i+1})$ such that $t_i < t_j^* < t_{i+1}$ for some $1 \leq j \leq d$ and $0 \leq i \leq m-1$. Even by increasing m, such that the smallest grid point that is greater than t_j^* gets closer and closer to t_j^*, this difference cannot be reduced to a value lower than h_j, and hence, the only possible way of eliminating this gap is to pick t_{i+1} such that it coincides with t_j^*. However, choosing the grid set B so that it contains all the jump points of $V_n^{\phi^*}$ is counter-productive since as mentioned before, the exact method to obtain the optimal policy (as defined by Wu and Lin [17]) quickly becomes computationally inefficient since more and more jump points must be considered as one moves backward through each successive stage. Therefore, the set B must have a small number of grid points compared to the number of jump points of $V_n^{\phi^*}$ so that the foremost goal of the GM, which is to provide a good approximation in a reasonable amount of time, is not undermined. In other words, having such gaps between the values of $V_n^{\phi^*}$ and $V_n^{\phi_m}$ is inevitable.

It is obvious that this gap results from jump points that exist in the graph of $V_n^{\phi^*}(x, P, .)$ over $[0, \tau]$ (or equivalently, the jump-point discontinuities of $V_n^{\phi^*}(x, P, .)$). One might wonder whether the performance of the GM would be affected positively if these jump points had not existed. Section 5 shifts the focus to a n-stage TSSAP with continuous task values, presents sufficient conditions for the existence of a deterministic Markov optimal policy, and provides optimality equations to obtain the optimal policy. Moreover, the behavior of the GM is studied where it is shown that, under certain mild conditions, the optimal target-dependent risk measure has no jump points.

5. TSSAP WITH CONTINUOUS TASK VALUES

The task values have been assumed to be discrete random variables. This assumption is relaxed in this section, where an n-stage TSSAP with continuous task values is considered.

This results in the state space of the system to be extended to an uncountable set, since the task values can vary in an interval as opposed to a countable set of real numbers (as presumed by the existing literature in this area). Suppose that task values are continuous random variables following a Riemann integrable *p.d.f.* f with support $\mathcal{S} \subseteq [0, +\infty)$. Given a vector of worker values P of size n, the objective is to find a Markov policy $\phi^* \in \Phi_{\mathrm{M}}$ that minimizes the probability of the total reward failing to achieve a target value.

In order to proceed, some notation is introduced. Let $\tilde{\Phi} \subseteq \Phi$ be the set of all policies ϕ such that $V_l^{\phi}(x, P, .)$ is non-decreasing on $[0, \tau]$ for arbitrarily fixed $(x, P) \in \tilde{\mathcal{S}}_1$ and $1 \leq l \leq n$. Under a policy $\phi \in \tilde{\Phi}$ and when the initial state of the system is fixed, the probability of failing to achieve a target value increases as the target value grows. Likewise, let $\tilde{\Phi}_{\mathrm{M}}$ and $\tilde{\Phi}_{\mathrm{DM}}$, respectively, denote the set of all Markov and deterministic Markov policies that lie in $\tilde{\Phi}$; equivalently, $\tilde{\Phi}_{\mathrm{M}} := \tilde{\Phi} \cap \Phi_{\mathrm{M}}$ and $\tilde{\Phi}_{\mathrm{DM}} := \tilde{\Phi} \cap \Phi_{\mathrm{DM}}$. Also, define Δ to be the set of all Markovian decision rules that are Riemann integrable on $\mathcal{S}$, and let the E_a notation, which was introduced in Section 4, be modified to suit the continuity assumption in this section, as follows: For a given $(x, P^{(n-k)}, t) \in \tilde{\mathcal{S}}_{n-k} \times [0, \tau]$, $a \in A(P^{(n-k)})$, and an arbitrary decision rule $\delta \in \Delta$, define the operators E_a and E_δ as

$$
\begin{aligned}
E_a V_k^{\phi}(x, P^{(n-k)}, t) &:= \int_{\mathcal{S}} V_k^{\phi}(u, P^{(n-k)} - p_a e_a, t - p_a x) f(u)\, du, \\
E_\delta V_k^{\phi}(x, P^{(n-k)}, t) &:= \sum_{i \in A(P^{(n-k)})} \delta(i \mid x, P^{(n-k)}, t) E_i V_k^{\phi}(x, P^{(n-k)}, t),
\end{aligned}
\tag{5.1}
$$

for any $\phi \in \tilde{\Phi}_{\mathrm{M}}$ and $k = 1, 2, \ldots, n-1$. Note that $\delta(i \mid x, P^{(n-k)}, t)$ is the probability that worker i is chosen under decision rule δ given that the current state is $(x, P^{(n-k)}, t)$. Lemma 5.1 ensures that the operators E_a and E_δ are well-defined.

Lemma 5.1: *Assume that task values are continuous random variables following a bounded Riemann integrable p.d.f. f with interval $\mathcal{S}$ its support. Consider an arbitrary policy $\phi = (\phi_n, \phi_{n-1}, \ldots, \phi_1) \in \tilde{\Phi}_{\mathrm{M}}$ with $\phi_l \in \Delta$ for all $l = 1, 2, \ldots, n$. The following results hold for any $((x, P^{(n-k+1)}), t) \in \tilde{\mathcal{S}}_{n-k+1} \times [0, \tau]$ and $2 \leq k \leq n$:*

$$
\begin{aligned}
&(i)\ E_{\phi_{n-k+1}} V_{k-1}^{\phi}(x, P^{(n-k+1)}, t) \text{ is well-defined.} \\
&(ii)\ V_k^{\phi}(x, P^{(n-k+1)}, t) = E_{\phi_{n-k+1}} V_{k-1}^{\phi}(x, P^{(n-k+1)}, t).
\end{aligned}
\tag{5.2}
$$

Proof: Proof of the first statement in (5.2) is by induction on k, and the second equation in (5.2) follows from the first statement. Observe that $V_1^{\phi}(x, P^{(n)}, t) = P\{p_i x \leq t\} = I_{\{p_i x \leq t\}}$, for any $\phi \in \tilde{\Phi}_{\mathrm{M}}$ where $A(P^{(n)}) = \{i\}$, and hence, $E_{\phi_{n-1}} V_1^{\phi}(x, P^{(n-1)}, t)$ is well defined. The second equation in (5.2) follows directly from conditioning arguments and the Markov property.

For the induction step, assume that $E_{\phi_{n-k+1}} V_{k-1}^{\phi}(x, P^{(n-k+1)}, t)$ is well-defined. Let $\delta := \phi_{n-k+1}$ and $\pi := \phi|_k$ for notational simplicity, and note that $\pi = (\phi_n, \phi_{n-1}, \ldots, \phi_{n-k+2}, \delta)$. Fix an arbitrary $(x, P^{(n-k+1)}, t) \in \tilde{\mathcal{S}}_{n-k+1} \times [0, \tau]$, let $\bar{P} := P^{(n-k+1)}$, and observe that

$$
\begin{aligned}
&E_\delta V_{k-1}^{\phi}(x, \bar{P}, t) \\
&\quad = \sum_{a \in A(\bar{P})} \delta(a \mid x, \bar{P}, t) \int_{\mathcal{S}} V_{k-1}^{\phi}(u, \bar{P} - p_a e_a, t - p_a x) f(u)\, du
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{a \in A(\bar{P})} \delta(a \mid x, \bar{P}, t) \int_{\mathcal{S}} P_\phi\{R_{k-1} \le t - p_a x \mid \tilde{d}_{n-k+2} \\
&= (u, \bar{P} - p_a e_a, t - p_a x)\} f(u)\, du \\
&= \sum_{a \in A(\bar{P})} \delta(a \mid x, \bar{P}, t) P_\pi \left\{ R_k \le t \mid \tilde{d}_{n-k+1} = (x, \bar{P}, t), \delta(x, \bar{P}, t) = a \right\} \\
&= V_k^\pi(x, \bar{P}, t), \qquad (\mathbf{5.3})
\end{aligned}
$$

where the third equality follows from the Markov property. This completes the proof of the second statement.

According to (5.1) and (5.3), verifying whether $E_{\phi_{n-k}} V_k^\phi(x, P^{(n-k)}, t)$ is well-defined comes down to proving that the following integral is well-defined for all $a \in A(P^{(n-k)})$:

$$
\int_{\mathcal{S}} E_\delta V_{k-1}^\phi(u, P^{(n-k)} - p_a e_a, t - p_a x) f(u)\, du. \qquad (\mathbf{5.4})
$$

Recall that $E_\delta V_{k-1}^\phi(.) = \sum_i \delta(i \mid .) E_i V_{k-1}^\phi(.)$, and hence, the problem simplifies to showing that

$$
\delta(i \mid ., P', t') E_i V_{k-1}^\phi(., P', t') f(.) \qquad (\mathbf{5.5})
$$

is Riemann integrable on $\mathcal{S}$ for all i, where $P' := P^{(n-k)} - p_a e_a$ and $t' = t - p_a x$. Fix i and consider the following two cases:

- $\mathcal{S}$ is a bounded interval: Note that $E_i V_{k-1}^\phi(u, P', t') = \int_{\mathcal{S}} V_{k-1}^\phi(z, P' - p_i e_i, t' - p_i u) f(z)\, dz$ is well-defined (by the induction assumption) and is a monotone non-increasing function of u, since $\phi \in \tilde{\Phi}_{\mathrm{M}}$ by assumption. Any monotone function on a bounded interval in $\mathbb{R}$ can have at most countably many discontinuity points and is Riemann integrable. Recall that $\delta \in \Delta$, and f is a bounded Riemann integrable function, so (5.5) is Riemann integrable on $\mathcal{S}$.
- $\mathcal{S}$ is an unbounded interval: Observe that $\mathcal{S}$ can be represented as $\mathcal{S} = \cup_{j=1}^{+\infty} I_j$, where I_j's are bounded disjoint intervals. To show that (5.5) is Riemann integrable on $\mathcal{S}$, first show that (5.5) is Riemann integrable on I_j, for $j = 1, 2, \ldots$. To see this, fix j and note that

$$
E_i V_{k-1}^\phi(u, P', t') = \int_{\mathcal{S}} V_{k-1}^\phi(z, P' - p_i e_i, t' - p_i u) f(z)\, dz,
$$

since $E_i V_{k-1}^\phi$ is well-defined, by the induction assumption. Define

$$
h_l(u) := \int_{I_l} V_{k-1}^\phi(z, P' - p_i e_i, t' - p_i u) f(z)\, dz,
$$

and observe that for any $l \ge 1$, $h_l(u)$ is a monotone function of u and has at most countably many discontinuity points on I_j, which implies that $E_i V_{k-1}^\phi(u, P', t') = \sum_{l=1}^{+\infty} h_l(u)$ is Riemann integrable on I_j, and hence, (5.5) is Riemann integrable on I_j, for $j = 1, 2, \ldots$. Now, observe that

$$
0 \le \int_{I_j} \delta(i \mid u, P', t') E_i V_{k-1}^\phi(u, P', t') f(u)\, du \le \int_{I_j} f(u)\, du,
$$

for any $j \geq 1$, which implies that $0 \leq \sum_{j=1}^{+\infty} \int_{I_j} \delta(i \mid u, P', t') E_i V_{k-1}^{\phi}(u, P', t') f(u)$ $du \leq 1$, and hence, (5.5) is Riemann integrable on $\mathcal{S}$.

■

Before proceeding to Theorem 5.2, a notation should be introduced. Let $\phi = (\phi_n, \phi_{n-1}, \ldots, \phi_1) \in \tilde{\Phi}_{\mathrm{M}}$ with $\phi_l \in \Delta$ for any $1 \leq l \leq n$ be an arbitrary policy. Based on ϕ, define a deterministic decision rule $\delta_{\phi_l}^* : \tilde{\mathcal{S}}_l \times [0, \tau] \to \mathcal{A}$ such that $\delta_{\phi_l}^*(x, P^{(l)}, t) \in A(P^{(l)})$ and

$$\delta_{\phi_l}^*(x, P^{(l)}, t) := \underset{a \in A(P^{(l)})}{\arg\min}\ E_a V_{n-l}^{\phi}(x, P^{(l)}, t),$$

for all $l = 1, 2, \ldots, n-1$. For $l = n$, let $\delta_{\phi_n}^*$ be the deterministic policy which assigns the last arriving task X_n to the only remaining worker (this is in fact the only admissible policy at time period n). Theorem 5.2 provides sufficient conditions for the existence of a deterministic Markov optimal policy and specifies the optimality equations to obtain it.

THEOREM 5.2: *Consider the n-stage TSSAP where task values are continuous random variables following a bounded Riemann integrable p.d.f. f with an interval $\mathcal{S}$ as its support. The optimality equations*

$$V_1^{\phi^*}(x, P^{(n)}, t) = I_{\{p_i x \leq t\}}, \tag{5.6}$$

where $A(P^{(n)}) = \{i\}$, and

$$V_l^{\phi^*}(x, P^{(n-l+1)}, t) = \min_{a \in A(P^{(n-l+1)})} E_a V_{l-1}^{\phi^*}(x, P^{(n-l+1)}, t), \tag{5.7}$$

for $(x, P^{(n-l+1)}, t) \in \tilde{\mathcal{S}}_{n-l+1} \times [0, \tau]$ and $l = 2, 3, \ldots, n$ yield the optimal policy ϕ^* if

$$\delta_{\phi_{n-l+1}^*}^* \in \Delta, \tag{5.8}$$

for $l = 2, 3, \ldots, n$. Moreover, if (5.8) holds true, then $\phi^* \in \tilde{\Phi}_{\mathrm{DM}}$, and $\delta_{\phi_k^*}^*$ is the optimal decision rule at time period $k = 1, 2, \ldots, n$.

PROOF: The proof is by induction on k. At the final stage (i.e., when the value of the last task is observed), only one worker is remained with value p_i by assumption, and the final task must be matched with this worker independent of the policy applied during the previous stages. Therefore, $V_1^{\phi^*}(x, P^{(n)}, t) = I_{\{p_i x \leq t\}}$, and (5.6) is verified. For the induction step, assume that (5.7) holds for all l where $l \leq k$ with $\pi := \phi^*|_k \in \tilde{\Phi}_{\mathrm{M}}$ and that $\delta_{\phi_{n-k}^*}^* \in \Delta$. Define a policy $\phi_0 := (\pi, \sigma)$ where $\sigma := \delta_{\phi_{n-k}^*}^*$ is a deterministic decision rule. Note that

$$\begin{aligned} V_{k+1}^{\phi_0}(x, P^{(n-k)}, t) &= E_\sigma V_k^{\pi}(x, P^{(n-k)}, t) = E_\sigma V_k^{\phi^*}(x, P^{(n-k)}, t) \\ &= \min_{a \in A(P^{(n-k)})} E_a V_k^{\phi^*}(x, P^{(n-k)}, t), \end{aligned} \tag{5.9}$$

where the first, the second, and the last equalities follow from Lemma 5.1, the definition of π, and the definition of σ, respectively. Moreover, $\phi^*|_k \in \tilde{\Phi}_{\mathrm{M}}$ by the induction assumption, which implies that $E_a V_k^{\phi^*}(x, P^{(n-k)}, .)$ is non-decreasing for any $a \in A(P^{(n-k)})$. Therefore,

https://doi.org/10.1017/S0269964813000028 Published online by Cambridge University Press

it follows that $V_{k+1}^{\phi_0}$ is a non-decreasing function of t which implies that $\phi_0 \in \tilde{\Phi}_{\mathrm{M}}$. Observe that

$$\begin{aligned} V_{k+1}^{\phi^*}(x, P^{(n-k)}, t) &= \inf_{\phi \in \Phi_{\mathrm{M}}} V_{k+1}^{\phi}(x, P^{(n-k)}, t) \le V_{k+1}^{\phi_0}(x, P^{(n-k)}, t) \\ &= \min_{a \in A(P^{(n-k)})} E_a V_k^{\phi^*}(x, P^{(n-k)}, t). \end{aligned} \tag{5.10}$$

Now, consider an arbitrary policy $\phi = (\phi_n, \phi_{n-1}, \ldots, \phi_{n-k}) \in \Phi_{\mathrm{M}}$, let $\bar{P} := P^{(n-k)}$, and obtain the following:

$$\begin{aligned} &V_{k+1}^{\phi}(x, \bar{P}, t) \\ &\quad = \sum_{i \in A(\bar{P})} \phi_{n-k}(i \mid x, \bar{P}, t) P_{\phi}\left\{R_{k+1} \le t \middle| \tilde{d}_{n-k} = (x, \bar{P}, t), \phi_{n-k}(x, \bar{P}, t) = i\right\} \\ &\quad = \sum_{i \in A(\bar{P})} \phi_{n-k}(i \mid x, \bar{P}, t) P_{\phi}\left\{R_k \le t - p_i x \middle| \tilde{d}_{n-k} = (x, \bar{P}, t), \phi_{n-k}(x, \bar{P}, t) = i\right\} \\ &\quad = \sum_{i \in A(\bar{P})} \phi_{n-k}(i \mid x, \bar{P}, t) P_{\phi}\left\{R_k \le t - p_i x \middle| \tilde{d}_{n-k+1} = (X_{n-k+1}, \bar{P} - p_i e_i, t - p_i x)\right\} \\ &\quad \ge \sum_{i \in A(\bar{P})} \phi_{n-k}(i \mid x, \bar{P}, t) P_{\phi^*}\left\{R_k \le t - p_i x \middle| \tilde{d}_{n-k+1} = (X_{n-k+1}, \bar{P} - p_i e_i, t - p_i x)\right\} \\ &\quad = \sum_{i \in A(\bar{P})} \phi_{n-k}(i \mid x, \bar{P}, t) E_i V_k^{\phi^*}(x, \bar{P}, t) \\ &\quad \ge \min_{a \in A(\bar{P})} E_a V_k^{\phi^*}(x, \bar{P}, t), \end{aligned}$$

which implies that $V_{k+1}^{\phi^*}(x, P^{(n-k)}, t) \ge \min_{a \in A(P^{(n-k)})} E_a V_k^{\phi^*}(x, P^{(n-k)}, t)$. Combining this with (5.10) yields

$$V_{k+1}^{\phi^*}(x, P^{(n-k)}, t) = V_{k+1}^{\phi_0}(x, P^{(n-k)}, t) = \min_{a \in A(P^{(n-k)})} E_a V_k^{\phi^*}(x, P^{(n-k)}, t),$$

where $\phi^*|_{k+1} := \phi_0 = (\phi^*|_k, \delta^*_{\phi^*_{n-k}}) \in \tilde{\Phi}_{\mathrm{DM}}$. ■

Theorem 5.3 provides a sufficient condition for (5.8) to hold and presents useful properties of the optimal policy and the optimal value function under this condition.

THEOREM 5.3: *Consider the n-stage TSSAP where task values are continuous random variables following a bounded Riemann integrable p.d.f. f that has interval $\mathcal{S}$ as its support. Then:*

- *Eq. (5.8) is satisfied, and hence, the optimal policy $\phi^* \in \Phi_{\mathrm{DM}}$ is obtained by (5.6)–(5.7).*
- *The optimal decision rule at time period k is right-continuous in the value of the current task and in the current target level, for $k = 1, 2, \ldots, n-1$.*
- *$V_n^{\phi^*}(x, P, t)$ is continuous in x and in t and is a non-decreasing function of t. Moreover, it is a distribution function in t if $\mathcal{S}$ is bounded.*

PROOF: The proof is by induction on n starting with $n = 2$ tasks as the base case. Fix an arbitrary task value $x \in \mathcal{S}$ and a vector of worker values $P^{(n-1)}$ at time period $n = 2$, and assume without loss of generality that $A(P^{(n-1)}) = \{1, 2\}$. Note that if $E_i V_1^{\phi^*}$ is a continuous function of t for $i = 1, 2$ and $\delta^*_{\phi^*_{n-1}} \in \Delta$, then $V_2^{\phi^*}$ is continuous in t due to (5.7) and the fact that it is the minimum over a finite set of continuous functions. Observe that

$$E_2 V_1^{\phi^*}(x, P^{(n-1)}, t) = \int_{\mathcal{S} \cap [0,(t-p_2x/p_1)]} f(u)\, du = P\left\{X_2 \in \mathcal{S} \cap \left[0, \frac{t - p_2 x}{p_1}\right]\right\},$$

which implies that $E_2 V_1^{\phi^*}$ is continuous in t. Similarly, $E_1 V_1^{\phi^*}$ is shown to be continuous in t; therefore, the optimal decision rule at this time period (i.e., when X_{n-1} arrives) is right-continuous in t when $(x, P^{(n-1)})$ is kept fixed. Following the same argument as above results in $E_1 V_1^{\phi^*}$ and $E_2 V_1^{\phi^*}$ being continuous functions of x on $\mathcal{S}$. This implies that (5.8) is satisfied (i.e., $\delta^*_{\phi^*_{n-1}} \in \Delta$); thus, the optimality Eq. (5.7) is used to derive $\delta^*_{\phi^*_{n-1}}$ and $V_2^{\phi^*}$. It also follows that $V_2^{\phi^*}$ is continuous in x and in t and is a distribution function in t. For the induction step, assume that Theorem 5.3 holds true for a TSSAP with $n - 1$ tasks where $n \geq 3$. To prove the theorem for n, fix $(x, P) \in \tilde{\mathcal{S}}_1$ and $i \in A(P)$, and assume that the sequence $\{t_j\}$ converges to t_0 as $j \to +\infty$, where $t_0 \in [0, \tau]$ is arbitrarily fixed. To prove that $E_i V_{n-1}^{\phi^*}$ is continuous in t, consider the following two cases:

- $\mathcal{S}$ is a bounded interval: Observe that for any $u \in \mathcal{S}$,

$$\lim_{j \to +\infty} V_{n-1}^{\phi^*}(u, P - p_i e_i, t_j - p_i x) = V_{n-1}^{\phi^*}(u, P - p_i e_i, t_0 - p_i x), \tag{5.11}$$

  by the induction assumption that $V_{n-1}^{\phi^*}$ is continuous in its third argument. Let M be an upper bound of f on $\mathcal{S}$ and note that

$$0 \leq V_{n-1}^{\phi^*}(u, P - p_i e_i, t_j - p_i x) f(u) \leq M, \tag{5.12}$$

  for all $u \in \mathcal{S}$ and $j \geq 1$, since $V_{n-1}^{\phi^*}$ is a distribution function by the induction assumption. Hence, the bounded convergence theorem implies that $\lim_{j \to +\infty} E_i V_{n-1}^{\phi^*}(x, P, t_j) = E_i V_{n-1}^{\phi^*}(x, P, t_0)$ by (5.11), (5.12), and the fact that the right-hand side of (5.11) is integrable over $\mathcal{S}$ by the induction assumption. Therefore, $E_i V_{n-1}^{\phi^*}(x, P, t)$ is continuous in t.
- $\mathcal{S}$ is an unbounded interval: Observe that $\mathcal{S}$ can be represented as $\mathcal{S} = \cup_{j=1}^{+\infty} I_j$, where I_j's are bounded disjoint intervals. Note that $h_j(t) := \int_{I_j} V_{n-1}^{\phi^*}(u, P - p_i e_i, t - p_i x) f(u)\, du$ is continuous in t for all j, by the bounded convergence theorem. Also, $0 \leq h_j(t) \leq \int_{I_j} f(u)\, du$, for any $t \in [0, \tau]$ and $j = 1, 2, \ldots$, where $\sum_{j=1}^{+\infty} \int_{I_j} f(u)\, du = 1 < +\infty$. The Weierstrass M-test implies that $\sum_{j=1}^{+\infty} h_j(t)$ converges uniformly on $[0, \tau]$, and hence, $E_i V_{n-1}^{\phi^*}(x, P, t)$ is a continuous function of t.

Likewise, it is proven for each $i \in A(P)$ that $E_i V_{n-1}^{\phi^*}(x, P, t)$ is continuous in x. Continuity of $E_i V_{n-1}^{\phi^*}$ in x and in t for all $i \in A(P)$ implies that the optimal decision rule upon the arrival of the first task (i.e., $\delta^*_{\phi_1}$) is right-continuous in both x and t. Therefore, the optimality Eq. (5.7) is used to derive $V_n^{\phi^*}$, and hence, $V_n^{\phi^*}$ is continuous in both t and x since the action space $A(P)$ is finite. It also follows that $V_n^{\phi^*}$ is a distribution function in t by the induction assumption (see Proposition 5 in [14]) if $\mathcal{S}$ is bounded. ■

Remark: The results of Lemma 5.1, Theorem 5.2, and Theorem 5.3 can be generalized to the case where the support of f is an open subset of $[0,+\infty)$ or a countable union of disjoint intervals in $[0,+\infty)$. (Recall that every open set in $\mathbb{R}$ can be represented as a countable union of disjoint bounded open intervals.) The proof is similar to that presented for Lemma 5.1.

Lipschitz continuity of the optimal value function is proven in Theorem 5.4 for continuous task values with bounded integrable probability distribution functions.

THEOREM 5.4: *Consider the n-stage TSSAP where task values are continuous random variables following a bounded Riemann integrable p.d.f. f that has an interval $\mathcal{S}$ as its support. For any arbitrarily fixed pair $(x,P)\in\tilde{\mathcal{S}}_1$, $V_n^{\phi^*}(x,P,t)$ is Lipschitz continous in t:*

$$\left|V_n^{\phi^*}(x,P,t)-V_n^{\phi^*}(x,P,s)\right|\le C\mid t-s\mid, \tag{5.13}$$

for $s,t\in[0,\tau]$ where C is a positive constant, independent of x.

PROOF: The proof is by induction on n starting with $n=2$ tasks as the base case. Arbitrarily fix $(x,P^{(n-1)})\in\tilde{\mathcal{S}}_{n-1}$, and assume without loss of generality that $A(P^{(n-1)})=\{1,2\}$. Let $s,t\in[0,\tau]$ with $s\le t$ and $\tilde{P}:=P^{(n-1)}-p_2e_2$, and observe that

$$\begin{aligned}
&\left|E_2V_1^{\phi^*}(x,P^{(n-1)},t)-E_2V_1^{\phi^*}(x,P^{(n-1)},s)\right|\\
&\quad\le\int_{\mathcal{S}}\left|V_1^{\phi^*}(u,\tilde{P},t-p_2x)-V_1^{\phi^*}(u,\tilde{P},s-p_2x)\right|f(u)\,du\\
&\quad=\int_{\mathcal{S}\cap(\frac{s-p_2x}{p_1},\frac{t-p_2x}{p_1}]}f(u)\,du\\
&\quad\le\frac{M}{p_{\min}}(t-s),
\end{aligned}$$

where M is an upper bound of f on $\mathcal{S}$ and $p_{\min}:=\min_{i\in\{1,2,\ldots,n\}}p_i$. A similar argument can be made for $E_1V_1^{\phi^*}$, and hence, it follows that $|V_2^{\phi^*}(x,P^{(n-1)},t)-V_2^{\phi^*}(x,P^{(n-1)},s)|\le(M/p_{\min})|t-s|$, for all $(x,P^{(n-1)})\in\tilde{\mathcal{S}}_{n-1}$ and $s,t\in[0,\tau]$. For the induction step, assume that $V_{n-1}^{\phi^*}(x,P^{(2)},t)$ is Lipschitz continous on $[0,\tau]$ with the Lipschitz constant $(M/p_{\min})$ for all $(x,P^{(2)})\in\tilde{\mathcal{S}}_2$ and some $n\ge 3$. To show that (5.13) holds for n, fix $(x,P)\in\tilde{\mathcal{S}}_1$, and observe that

$$\begin{aligned}
&\left|E_iV_{n-1}^{\phi^*}(x,P,t)-E_iV_{n-1}^{\phi^*}(x,P,s)\right|\\
&\quad\le\int_{\mathcal{S}}\left|V_{n-1}^{\phi^*}(u,P-p_ie_i,t-p_ix)-V_{n-1}^{\phi^*}(u,P-p_ie_i,s-p_ix)\right|f(u)\,du\\
&\quad\le\frac{M}{p_{\min}}|t-s|,
\end{aligned}$$

for all $i\in A(P)$ and $s,t\in[0,\tau]$, which implies that $|V_n^{\phi^*}(x,P,t)-V_n^{\phi^*}(x,P,s)|\le(M/p_{\min})|t-s|$, due to the finiteness of the action space. ■

Corollary 5.5 presents an interesting property of the optimal value function.

COROLLARY 5.5: *Consider the n-stage TSSAP where task values follow a bounded Riemann integrable p.d.f. f with an interval $\mathcal{S}$ as its support. For any sequence $\{t_j\}_{j=1}^{+\infty}$ converging*

https://doi.org/10.1017/S0269964813000028 Published online by Cambridge University Press

to an arbitrarily fixed $t \in [0,\tau]$, $V_n^{\phi^*}(x,P,t_j)$ *converges uniformly to* $V_n^{\phi^*}(x,P,t)$ *on* $\mathcal{S}$ *as* $j \to +\infty$*; that is,*

$$\sup_{x\in\mathcal{S}} \left| V_n^{\phi^*}(x,P,t_j) - V_n^{\phi^*}(x,P,t) \right| \to 0 \quad as \quad j \to +\infty.$$

PROOF: The proof follows from Theorem 5.4 and the Cauchy criterion for uniform convergence and is eliminated due to simplicity. ■

Before proceeding to Lemma 5.6, the GM presented in Section 4 for the discrete case is generalized to the continuous case using (4.1)–(4.3), where $E_i V_l^{\phi_m}$ is defined in (5.1). Lemma 5.6 proves that the operator E_i is well-defined for the GM.

LEMMA 5.6: *Consider the* n*-stage TSSAP with continuous task values following a bounded Riemann integrable p.d.f.* f *with an interval* $\mathcal{S}$ *as its support. Let* $B = \{t_0, t_1, \ldots, t_m\}$ *with* $t_0 = 0$ *and* $t_m = \tau$ *be a grid set for the GM. For* $l = 1, 2, \ldots, n$, $V_l^{\phi_m}(., P^{(n-l+1)}, t)$ *and* $V_l^{\phi_m}(x, P^{(n-l+1)}, .)$ *are non-increasing and non-decreasing functions, respectively. Moreover,* $E_i V_l^{\phi_m}$ *is well-defined for* $l = 1, 2, \ldots, n-1$.

PROOF: The proof of $V_l^{\phi_m}$ being a monotone function on $\mathcal{S}$ and on $[0,\tau]$ is by induction on l. That the operator $E_i V_l^{\phi_m}$ is well-defined follows from $V_l^{\phi_m}$ being a monotone function on $\mathcal{S}$. The proof is eliminated due to simplicity. ■

Note that the results of Lemma 4.1 and Lemma 4.2 are easily generalized to the GM defined for the continuous case. For an n-stage TSSAP with continuous task values, Proposition 5.7 studies the behavior of the GM as the grid on $[0,\tau]$ becomes finer.

PROPOSITION 5.7: *Consider the* n*-stage TSSAP with a given vector of workers* P *and continuous task values following a bounded Riemann integrable p.d.f.* f *that has an interval* $\mathcal{S}$ *as its support. For a grid set* $B = \{t_0^m, t_1^m, \ldots, t_m^m\}$ *where* $t_i^m := (\tau/m)i$, $V_n^{\phi_m}(x,P,t)$ *converges to* $V_n^{\phi^*}(x,P,t)$ *uniformly on* $\mathcal{S} \times [0,\tau]$ *with order one as* $m \to +\infty$*; that is,*

$$\sup_{x\in\mathcal{S}, t\in[0,\tau]} \left| V_n^{\phi^*}(x,P,t) - V_n^{\phi_m}(x,P,t) \right| = O\left(m^{-1}\right). \tag{5.14}$$

PROOF: Observe that by (4.1) and as in the discrete case, $V_2^{\phi^*}$ and $V_2^{\phi_m}$ coincide at the breakpoints; therefore, $V_2^{\phi^*}(x, P^{(n-1)}, t_i^m) = V_2^{\phi_m}(x, P^{(n-1)}, t_i^m)$, for $i = 0, 1, 2, \ldots, m$ and $(x, P^{(n-1)}) \in \tilde{\mathcal{S}}_{n-1}$. Fix an arbitrary $(x, P^{(n-1)}) \in \tilde{\mathcal{S}}_{n-1}$ and $t \in [0,\tau]$, and without loss of generality assume that t belongs to the $(k+1)$th interval defined by B on $[0,\tau]$ for some $0 \le k \le m-1$ (i.e., $t \in [t_k^m, t_{k+1}^m)$). Note that the GM provides a lower bound approximation for the optimal value function; therefore,

$$\begin{aligned} 0 &\le V_2^{\phi^*}(x, P^{(n-1)}, t) - V_2^{\phi_m}(x, P^{(n-1)}, t) \\ &\le V_2^{\phi^*}(x, P^{(n-1)}, t_{k+1}^m) - V_2^{\phi^*}(x, P^{(n-1)}, t_k^m) \le C\frac{\tau}{m}, \end{aligned}$$

where the second inequality follows from Lemma 5.6 (specifically, the fact that $V_2^{\phi^*}(x, P^{(n-1)}, .)$ is a non-decreasing function), and C is the Lipschitz constant defined in Theorem 5.4; hence, (5.14) is satisfied for a problem of size $n = 2$. For the induction step, assume that for some $n \geq 3$:

$$0 \leq V_{n-1}^{\phi^*}(u, P^{(2)}, t_{j+1}^m) - V_{n-1}^{\phi_m}(u, P^{(2)}, t_j^m) \leq (n-2)C\frac{\tau}{m},$$

for all $u \in \mathcal{S}$, $P^{(2)} \in \mathcal{W}_2^P$, and $j = 0, 1, \ldots, m-1$. To prove (5.14) for n, observe that

$$\begin{aligned} 0 &\leq V_n^{\phi^*}(x, P, t) - V_n^{\phi_m}(x, P, t) \\ &\leq V_n^{\phi^*}(x, P, t_{k+1}^m) - V_n^{\phi_m}(x, P, t_k^m) \\ &= \min_{i \in A(P)} E_i V_{n-1}^{\phi^*}(x, P, t_{k+1}^m) - \min_{i \in A(P)} E_i V_{n-1}^{\phi_m}(x, P, t_k^m). \end{aligned} \tag{5.15}$$

Since $t_{k+1}^m - t_k^m = (\tau/m)$ (or equivalently, $(t_{k+1}^m - p_i x) - (t_k^m - p_i x) = (\tau/m)$), it follows that there exists $k' < k$ such that $t_k^m - p_i x \in [t_{k'}^m, t_{k'+1}^m]$ and $t_{k+1}^m - p_i x \in [t_{k'+1}^m, t_{k'+2}^m)$. Therefore,

$$\begin{aligned} E_i V_{n-1}^{\phi^*}(x, P, t_{k+1}^m) &= \int_{\mathcal{S}} V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k+1}^m - p_i x) f(u)\, du \\ &\leq \int_{\mathcal{S}} V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k'+2}^m) f(u)\, du, \end{aligned}$$

and

$$\begin{aligned} E_i V_{n-1}^{\phi_m}(x, P, t_k^m) &= \int_{\mathcal{S}} V_{n-1}^{\phi_m}(u, P - p_i e_i, t_k^m - p_i x) f(u)\, du \\ &= \int_{\mathcal{S}} V_{n-1}^{\phi_m}(u, P - p_i e_i, t_{k'}^m) f(u)\, du, \end{aligned}$$

which leads to

$$\begin{aligned} & E_i V_{n-1}^{\phi^*}(x, P, t_{k+1}^m) - E_i V_{n-1}^{\phi_m}(x, P, t_k^m) \\ &\quad \leq \int_{\mathcal{S}} (V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k'+2}^m) - V_{n-1}^{\phi_m}(u, P - p_i e_i, t_{k'}^m)) f(u)\, du \\ &\quad = \int_{\mathcal{S}} (V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k'+2}^m) - V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k'+1}^m)) f(u)\, du \\ &\qquad + \int_{\mathcal{S}} (V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k'+1}^m) - V_{n-1}^{\phi_m}(u, P - p_i e_i, t_{k'}^m)) f(u)\, du. \end{aligned} \tag{5.16}$$

Now, recall from Theorem 5.4 that $V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k'+2}^m) - V_{n-1}^{\phi^*}(u, P - p_i e_i, t_{k'+1}^m) \leq C\frac{\tau}{m}$, for all $u \in \mathcal{S}$. Moreover, the second term on the right-hand side of Eq. (5.16) can be bounded above by the induction assumption. Therefore,

$$0 \leq E_i V_{n-1}^{\phi^*}(x, P, t_{k+1}^m) - E_i V_{n-1}^{\phi_m}(x, P, t_k^m) \leq (n-1)C\frac{\tau}{m}, \tag{5.17}$$

for all $i \in A(P)$, which implies that

$$0 \leq V_n^{\phi^*}(x, P, t) - V_n^{\phi_m}(x, P, t) \leq (n-1)C\frac{\tau}{m},$$

by (5.15) and from the finiteness of the action space. ■

Proposition 5.7 establishes the uniform convergence of the approximate value function obtained by the GM to the optimal value function as the grid on $[0, \tau]$ becomes finer. However, recall that for the n-stage TSSAP with discrete task values, there are intervals within $[0, \tau]$ in which a lower bound exists on the difference between the approximate and the optimal value function. As shown in Section 4, this lower bound is a constant and unaffected by increases in m (i.e., the number of grid points); equivalently, for any value of m which results in a reasonable computation time for the GM, there always exist intervals within $[0, \tau]$ with gaps (greater than a given constant) between the approximate and optimal value functions. Therefore, the GM has a better performance when applied to the n-stage TSSAP with countinuous task values. Section 6 provides numerical results to compare the performance of the SSAP and TSSAP.

6. NUMERICAL RESULTS

This section compares the performance of the optimal policies obtained from the SSAP and TSSAP, using a numerical example. Consider a stochastic sequential assignment problem with $n = 10$ tasks arriving sequentially at each time period to be allocated to the available resources. Assume that the task values follow a Binomial distribution with parameters $(4, 0.3)$ and the vector of worker values is given by $P = (10, 50, 100, 150, 250, 400, 540, 600, 750, 950)$. We solve the TSSAP for each target value within the interval $[3,000, 13,000]$ with a step size of 50. For each fixed target value, a total of $s = 1,000$ TSSAPs are solved by simulating the arriving task values with the given Binomal distribution, and for every one of the $1,000$ problems simulated, a SSAP is solved as well. For a fixed target value τ, let r_{T}^{τ} and r_{S}^{τ} denote the number of times (out of a $1,000$

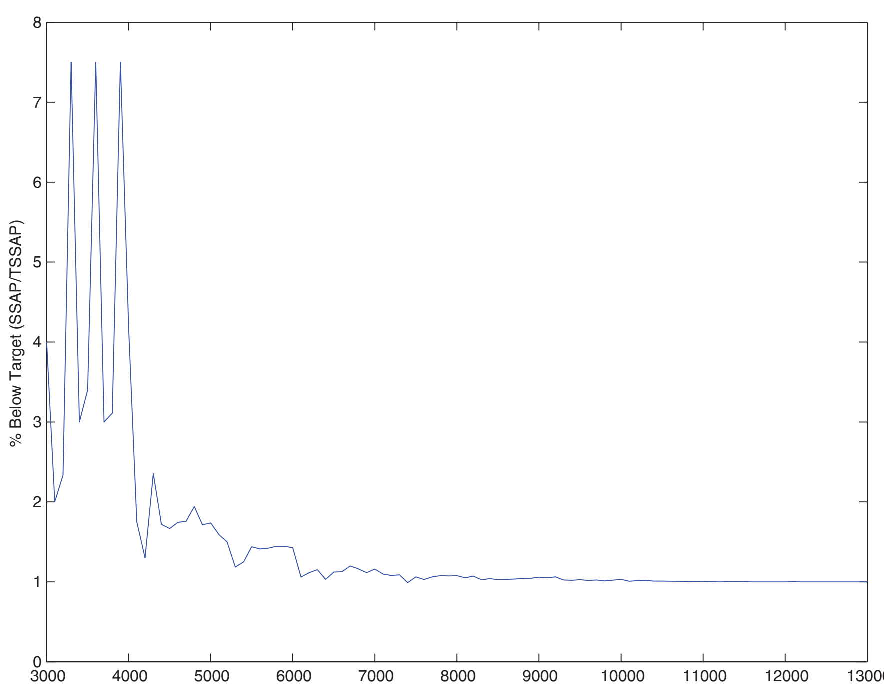

FIGURE 1. (Color online) Comparing the optimal policy of SSAP vs. TSSAP.

https://doi.org/10.1017/S0269964813000028 Published online by Cambridge University Press

simulations) that the TSSAP and the SSAP yield a total reward lower than τ. Figure 1 depicts the ratio $(r_{\mathrm{S}}^{\tau}/r_{\mathrm{T}}^{\tau})$ as a function of τ. As it can be seen from the figure, the optimal policy from the TSSAP performs significantly better than that of the SSAP for target values that are below or around the $E_f[X]\sum_{i=1}^{n} p_i$ (where $E_f[X]$ is the expected value of X_j and p_i is the success rate of the ith worker). As the target value increases, the ratio $(r_{\mathrm{S}}^{\tau}/r_{\mathrm{T}}^{\tau})$ decreases but stabilizes at one (which is intuitive).

7. CONCLUSION

This paper studies the SSAP under the threshold criterion, which attempts to minimize the probability of the total reward (obtained from the sequential assignment of tasks to available workers) failing to achieve a specified target value. The problem is modelled as an MDP for discrete task values and is then extended to the case where the state space of arriving tasks is uncountable (i.e., task values are considered to be continuous random variables). Sufficient conditions for the existence of a deterministic Markov optimal policy are derived along with fundamental properties of the optimal value function. An algorithm (referred to here as GM) is introduced to approximate the optimal value function and the optimal policy, since the problem becomes computationally inefficient and intractable as the number of arriving tasks increases. The behavior of GM is analyzed for the countable and the uncountable state space cases, and convergence of the approximate value function (obtained by GM) to the optimal value function is established.

It is assumed here that the underlying distribution function of task values is given beforehand, and further research is required to address the TSSAP in which task values follow a probability distribution with unknown parameters. In addition, a possible extension of the TSSAP is to the case where the total number of tasks is unknown until after the final arrival and follows a generic probability distribution. Other challenges include shifting one's attention from the i.i.d. sequence of tasks to a more general case with dependent task values and/or considering an infinite sequence of arriving tasks. Moreover, another research direction is extending the main results in this paper, which are obtained for the SSAP, to a more general MDP framework, where the action space is not necessarily finite.

Acknowledgments

This research has been supported in part by the Air Force Office of Scientific Research under Grant No. FA9550-10-1-0387, and the National Science Foundation under Grant No. $CMMI$-0900226. This material is based on work supported in part by (while serving at) the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Government, the Air Force Office of Scientific Research, or the National Science Foundation.

References

1. Albright, S.C. (1974). A markov-decision-chain approach to a stochastic assignment problem. *Operations Research* 22(1): 61–64.
2. Albright, S.C. (1974). Optimal sequential assignment with random arrival time. *Management Science* 21(1): 60–67.
3. Boda, K., Filar, J.A., Lin, Y., & Spanjers, L. (2004). Stochastic target hitting time and the problem of early retirement. *IEEE Transactions on Automatic Control* 49(3): 409–419.
4. Derman, C., Lieberman, G.J., & Ross, S.M. (1972). A sequential stochastic assignment problem. *Management Science* 18(7): 349–355.
5. McLay, L.A., Jacobson, S.H., & Nikolaev, A.G. (2009). A sequential stochastic passenger screening problem for aviation security. *IIE Transactions* 41(6): 575–591.

6. Nikolaev, A.G. & Jacobson, S.H. (2010). Stochastic sequential decision-making with a random number of jobs. *Operations Research* 58: 1023–1027.
7. Nikolaev, A.G., Jacobson, S.H., & McLay, L.A. (2007). A sequential stochastic seceurity system design problem for aviation security. *Transportation Science* 41(2): 182–194.
8. Ohtsubo, Y. (2003). Value iteration methods in risk minimizing stopping problems. *Journal of Computational and Applied Mathematics* 152: 427–439.
9. Ohtsubo, Y. (2004). Optimal threshold probability in undiscounted markov decision processes with a target set. *Applied Mathematics and Computation* 149: 519–532
10. Ohtsubo, Y. & Toyonaga K. (2002). Optimal policy for minimizing risk models in markov decision processes. *Journal of Mathematical Analysis and Applications* 271: 66–81.
11. Ohtsubo, Y. & Toyonaga, K. (2004). Equivalence classes for optimizing risk models in markov decision processes. *Mathematical Methods of Operations Research* 60: 239–250.
12. Righter, R.L. (1987). The stochastic sequential assignment problem with random deadlines. *Probability in the Engineering and Informational Sciences* 1: 189–202.
13. Sakaguchi, M. & Ohtsubo, Y. (2010). Optimal threshold probability and expectation in semi-markov decision processes. *Applied Mathematics and Computation* 216: 2947–2958.
14. Smith, J.E. & McCardle, K.F. (2002). Structural properties of stochastic dynamic programs. *Operations Research* 50(5): 796–809.
15. Su, X. & Zenios, S.A. (2005). Patient choice in kidney allocation: A sequential stochastic assignment model. *Operations Research* 53(3): 443–455.
16. White, D.J. (1993). Minimizing a threshold probability in discounted markov decision processes. *Journal of Mathematical Analysis and Applications* 173: 634–646.
17. Wu, C. & Lin, Y. (1999). Minimizing risk models in markov decision processes with policies depending on target values. *Journal of Mathematical Analysis and Applications* 231: 47–67.