Developing Operator Models for UAV Search Scheduling

L. F. Bertuccelli,[*] N. W. M. Beckers,[†] and M. L. Cummings[‡]

With the increased use of Unmanned Aerial Vehicles (UAVs), it is envisioned that UAV operators will become high level mission supervisors, responsible for information management and task planning. In the context of search missions, operators supervising a large number of UAVs can become overwhelmed with the sheer amount of information collected by the UAVs, making it difficult to optimize the information collection or direct their attention to the relevant data. Novel decision-support methods that account for realistic operator performance will therefore be required to aid the operators. This paper considers a decision support formulation for sequential search tasks, and discusses a non-preemptive scheduling formulation for a single operator performing a search mission in a time-constrained environment. The formulation is then generalized to include operator performance obtained from previous human-in-the-loop experiments, and presents one of the principal contributions of the paper. The sensitivity of the proposed model is analyzed in the presence of uncertainty to the operator model and search times, and a comparison is made between the expected performance difference between this scheduling system and a greedy scheduling strategy representative of operator planning. The paper concludes with the design of a human-in-the-loop experiment for a scheduling, replanning task for a simulated UAV mission.

I.   Introduction

Use of Unmanned Aerial Vehicles (UAVs) has extended beyond the military ISR (Information, Surveillance, and Reconnaissance) roles in Iraq and Afghanistan, and has been demonstrated in recent humanitarian crises such as in the earthquake tragedy in Haiti. In the days and weeks following the earthquake, UAVs provided highly detailed imagery that was quickly broadcast to ground-based first responders [1, 2]. As UAVs increasingly become more sophisticated with higher levels of autonomy, it is anticipated that missions such as these will become more commonplace. Furthermore, it is expected that UAV operators will transition to a higher-level supervisory role, including the management of multiple UAVs under a single operator [3]. As such, the operators will become responsible for a large number of high-level planning tasks, such as information management of the mission, possibly in coordination with analysts observing the relayed data in close to real time [4]. Demonstration of recent DoD interest in such complex information-gathering missions, including wide-area surveillance, has been demonstrated in the upcoming deployment of the Gorgon Stare program for UAVs and Constant Hawk/Angel Fire programs in manned vehicles [5]. These programs have been developed to allow end users (ranging from ground controllers all the way to high-level decision makers) to obtain multiple viewpoints of the same data stream.

The increase in UAV sensing capabilities has highlighted the issue of information bottlenecking, in which operators have access to large databases of imagery and video to analyze but may have difficulty fully processing the data in a timely manner [4]. Answering questions such as "what imagery needs to be looked

[*]Luca F. Bertuccelli is a Postdoctoral Associate at the Massachusetts Institute of Technology, Dept of Aeronautics and Astronautics `lucab@mit.edu`

[†]Nicolaas W. M. Beckers is a Master Student at Delft University of Technology, Faculty of Aerospace Engineering `N.W.M.Beckers@Student.TUDelft.NL`

[‡]Mary L. Cummings is an Associate Professor at the Massachusetts Institute of Technology, Dept of Aeronautics and Astronautics, `missyc@mit.edu`
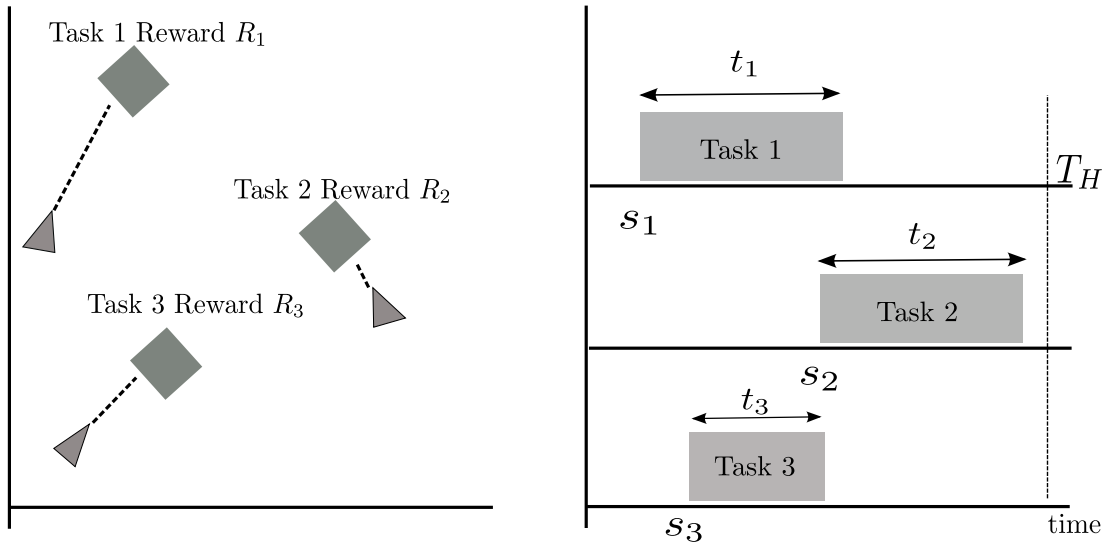
Figure 1. The ability to search for multiple tasks in an environment (left) can be aided by generating a schedule representation (right). In this simple representation, 3 tasks are available for search in a total of T_H mission time. UAV arrival times dictate that the tasks will be available at different times. An optimal schedule will output the task ordering to maximize an objective under the appropriate time constraints

at and when?" and "how much attention should be allocated to an imagery set?" may become a nontrivial task, particularly in the highly dynamic and stochastic nature of the environment. Operator constraints, such as attention and workload, will also play significant factors.

For example, consider a time constrained UAV search problem shown in Figure 1, where there are 3 heterogeneous tasks (with different rewards) to be investigated by 3 UAVs under the supervision of a single operator. Each task requires a different search time t_j, may have different start times s_j, and the operator only has a finite mission time T_H to identify as many tasks as possible. Due to the single operator limitation, tasks must be viewed sequentially (e.g., Task 1, then Task 2, etc.). Even for this simple example, however, it may not be immediately clear how to decide which task should be viewed and when. Searching Task 1 first may seem beneficial since it is the first task available, but also takes the longest. Likewise, searching Task 3 first could be beneficial since it takes the shortest amount of time, but requires waiting for a longer period of time, and may result in a lower accrued reward, since fewer tasks may be processed.

Furthermore, depending on the operator's workload, or the available time to make the decision, it may be unreasonable to expect an operator to make a good decision without external support from a properly tailored decision support system. An additional complexity in the design of decision support systems for operators in human supervisory control, is that the underlying algorithmic assumptions often are an approximation of the true system, and quantifying optimality of the decision support system may be a daunting task. For many HSC applications in fact, the success of a decision support can only be measured by the performance improvement that the decision support can provide to an operator.

This paper formulates the sequential search task problem using a scheduling framework. We exploit previously gathered empirical data obtained from supervisory control experiments, and formulate the problem of allocating tasks to the operator as an optimal non-preemptive scheduling problem. We build on previous work of operator scheduling [6–9]. The work of Ref. [7, 8] proposed a scheduling protocol for the efficient allocation of operator workload, but only used a limited lookahead to reschedule tasks for the operator and treated the operator performance as deterministic. The work of Ref. [9] sought to emphasize the allocation of attention, but treated a different optimization problem and treated operator performance as deterministic. Numerous applications of operator scheduling outside the realm of UAV supervisory control include the telephony, manufacturing, and healthcare industries [10–13]. Ref. [10] formulates the optimization of operators using network flows, while Ref. [11] uses a mixed integer linear programming approach similar to that pursued in this paper. Yamada [13] follows a genetic algorithm approach to optimize the schedules of the

telephone operators, while Henderson [12] uses a heuristic approach. While much work has addressed the issue of uncertainty in scheduling tasks highlighting the ability for humans to plan under uncertainty (see Ref. [14, 15] for a review), little work has addressed the issues of how operators can be supported efficiently in UAV sequential search tasks while accounting for operator detection models, which is a key contribution of our work.

This paper is outlined as follows: Section II presents the scheduling formulation for sequential search tasks, while Section III presents a modification of this scheduling formulation that explicitly accounts for operator models obtained from previous supervisory control experiments. Sensitivity of this model is investigated empirically. Section IV then compares the performance of the proposed scheduling formulation with a greedy approximation to the operator predicted performance. An experiment description for validating these comparisons is presented in Section V, and we conclude with future work in Section VI.

II.   Scheduling UAV Search Tasks

We assume that there are N heterogeneous search tasks known to the operator, where each task $\mathcal{T}_j$ is defined by the following 3-tuple $\mathcal{T}_j = \{R_j, t_j, s_j\}, \quad \forall j = 1, \ldots, N$. In this description, R_j is the reward for succesfully completing task j, t_j is the total time required to complete or service the task, and s_j indicates the time at which the task becomes available to the operator. The operator scheduling problem is defined in this paper as the goal of maximizing the total accumulated reward obtained from a sequence of search tasks in a finite mission time T_H. A schedule, $\mathcal{S}$, is an ordered list of tasks, consisting of $\mathcal{S} = \{k_1, k_2, \ldots, k_{N^*}\}$ where k_i is the index of the task in the ith location in the schedule, and $N^* \leq N$. The reward of the task is specified clearly to the operator; hence, rational decision theory (which states that rational subjects will seek to maximize their reward) can be applied [16].

The start time s_j for each vehicle is used to model the fact that different tasks may become available at different times, either due to vehicle constraints or because image feeds are unavailable due to communication dropouts. UAVs that are not being used by the operator to search can loiter around their intended targets. In this paper we assume that each UAV is assigned exactly one task, as an abstract representation of real-world surveillance systems such as Gorgon Stare [5].

In the special case when all UAVs are loitering on the targets, and the tasks are available for search ($s_j = 0, \ \forall j$), maximizing the accumulated reward can be formulated as a knapsack problem [17]

$$\textbf{KP:} \quad \left\{ \max_{x_j \in \{0,1\}} \sum_{j=1}^{N} R_j x_j \ \bigg| \ \sum_j t_j x_j \leq T_H, \ \sum_j x_j \leq N \right\} \tag{1}$$

where $x_j = \{0,1\}$ is a binary decision variable that is equal to 1 if the task is chosen by the operator. Since the tasks are all available at the beginning of the allocation problem, the restrictions are that up to a maximum of N tasks can be serviced in the total time T_H (given by the two constraints). Note that switch times from task $\mathcal{T}_j$ to the next task $\mathcal{T}_k$ are an additional realistic issue and can be a significant contributor to the total time that an operator spends searching a task, switching times/costs can be included by encoding them in the total service times t_j.

II.A.   Scheduling formulation

In the more general case when the UAVs arrive at tasks at different times, the arrival times are non-zero ($s_j \neq 0, \ \forall j$), and the problem requires the inclusion of additional specifications. With the assumption of a single operator, the problem can be formulated as a modification of the well-studied single machine scheduling problem with arbitrary release dates [18]. The distinguishing feature from the classical machine scheduling problem is that for our problem formulation, the operator seeks to maximize the accumulated reward of each job, such as identifying a high value target, while classical machine scheduling problems optimize the total delay, total number of tardy jobs, and flowtimes [18].

Table 1. Schedule parameters of target rewards with associated start and search duration times

Target #	1	2	3	4	5
Target reward R_j	0.4	5.1	6.9	3.1	7.4
Expected reward $\mathbf{E}[R_j]$	0.2	2.4	2.4	0.7	0.8
Avg search time $\bar{t}_j$	52	64	78	96	118
Start time s_j	28	25	22	26	24
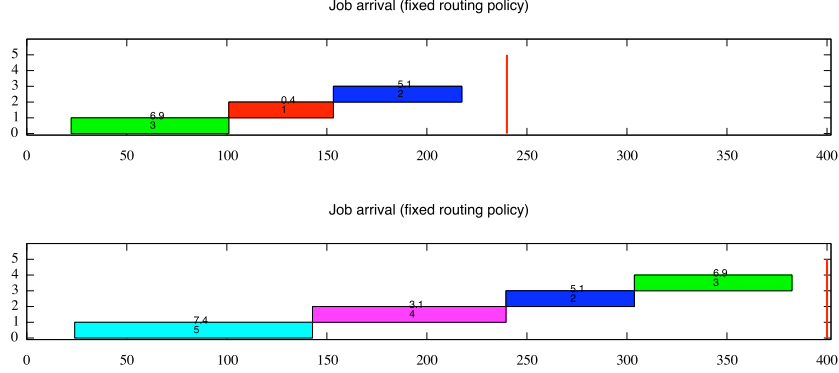
Figure 2. Sample schedules with increasing time horizon T_H: $T_H = 245$ (in red line, top); $T_H = 400$ (in red line, bottom). Task index is shown with task reward shown above the task (for example, in top figure, Task 3 in green has a reward of 6.9). In the top figure only task 3, 1, and 2 can be accomplished that maximizes reward in the allotted time, while tasks 5, 4, 2, and 3 can be completed with 400 time units. Note that task 1 is now not performed since the additional time allows the operator to perform tasks 5 and 4

With the goal of maximizing the accumulated reward, the scheduling formulation can be formulated as follows:

$$
\mathbf{SP:} \begin{cases}
& \max_{x_k^j \in \{0,1\}} \sum_j \sum_k R_j x_k^j \\
\text{subject to:} & \sum_k x_k^j \le 1 & \forall j \\
& \sum_j x_k^j \le 1 & \forall k \\
& \sum_j (s_j + t_j) x_k^j \le C_k, & \forall k \\
& C_{k-1} + \sum_j t_j x_k^j \le C_k, & \forall k \\
& C_k \le T_H, & \forall k
\end{cases}
$$

In problem **SP** the notation x_k^j denotes the position k in which task j appears. Note that with this formulation introduces an additional non-negative variable, C_k which denotes the completion time of the kth task (which cannot exceed the total available time T_H).[a] Two additional constraints are required for this variable: first, the kth task cannot be completed prior to the sum of its scheduled time $\sum_j s_j x_k^j$ and the search time $\sum_j t_j x_k^j$. Second, the next task k cannot be completed before the previous task is completed, and the completion time of the previous task, C_{k-1}. Note that this paper does not assume preemption [18], meaning that a task cannot be interrupted in order to initiate another task.[b] An example of this scheduling algorithm is shown in Figure 2 where a total of 5 tasks can be allocated in different mission times T_H. (The numerical data is found in Table 1, and shows the reward, search time, and start time for each task. The expected

---

[a]Note that the schedule and/or the completion times need not be unique, since if there is sufficient time to perform the tasks, it may be possible to either rearrange the task list, as well as delay the start of the task. Delaying the task may be beneficial in a setting when the tasks are arriving randomly.

[b] Preliminary investigation in the role of preemption has been addressed in a previous paper investigating use of relooks to improve the likelihood of detection [19]. Use of preemption for this work is investigated in an ongoing human-in-the-loop experiment.
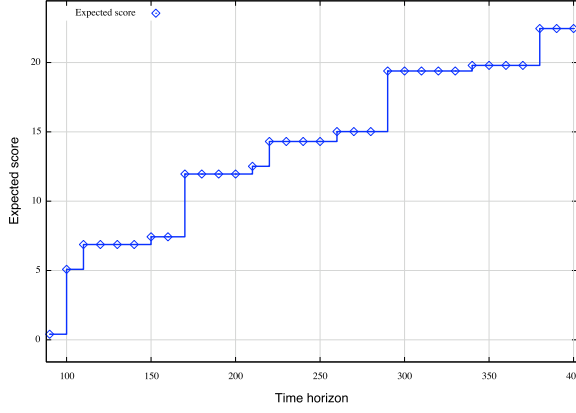
(a) Accumulated score, no operator model
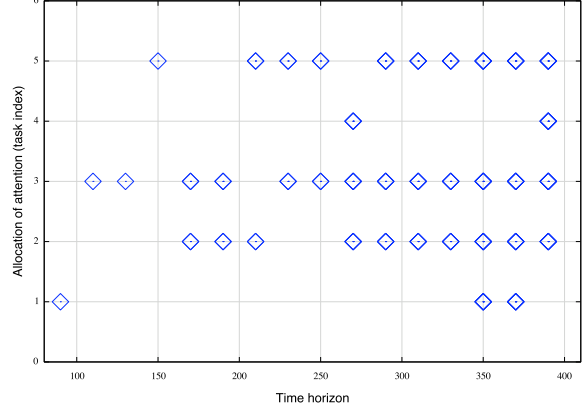
(b) Optimal assignment: each diamond corresponds to a task

Figure 3. Accumulated score with increased time (left) and optimal assignment (right) while solving SP

reward will be used in the next section.) In the first case (Figure 2 (top)), the mission time is set at $T_H = 245$ time units, and the operator maximizes reward by scheduling tasks 3, 1, and 2 for a total reward of 12.4. As the mission time is increased to $T_H = 400$ time units (Figure 2 (bottom)), the total reward is increased to 22.5; interestingly, in order to maximize the reward, task 1 is switched with tasks 4 and 5.

An example of the increase in mission time T_H is shown in Figure 3 with a new set of 5 tasks. Figure 3(a) shows the increased accumulated reward by increasing mission time, and shows unique regions of different sensitivity to increase in T_H. For example, an increase from $T_H = 160$ to $T_H = 170$ can increase the accumulated reward from 7.5 to 12, while a mission time increase from $T_H = 150$ to $T_H = 160$ does not generate any increase in accumulated reward. Figure 3(b) shows the task allocation as a function of the time T_H, where the diamonds correspond to which tasks were selected at a specific T_H. For example, for $T_H = 90$, only task 1 was selected, while for $T_H = 210$, tasks 2 and 5 were selected. Note that as expected, and as seen in Figure 2, the tasks scheduling can remove previously desirable tasks: for example, when increasing T_H from $T_H = 370$ to $T_H = 390$, task 1 is replaced with task 4.

II.B. Sensitivity to uncertainty in search times

In reality, the start times s_j and search times t_j may not be known with certainty since these search times vary with task difficulty, operator workload, and operator skill. Previous work has shown that search times can be modeled with a corresponding lognormal distribution [19]. A more detailed discussion on the sensitivity of the scheduling algorithms to the knowledge of search times in the next sections, but one method for handling the uncertainty in the start (respectively, search) times is to substitute the expected values, $\bar{s}_j$ (respectively, $\bar{t}_j$), and find the optimal schedule as follows:

$$
\mathbf{SP_{CE}}: \begin{cases} & \max_{x_k^j \in \{0,1\}} \sum_j \sum_k R_j x_k^j \\ \text{subject to:} & \sum_k x_k^j \leq 1 & \forall j \\ & \sum_j x_k^j \leq 1 & \forall k \\ & \sum_j (\bar{s}_j + \bar{t}_j) x_k^j \leq C_k, & \forall k \\ & C_{k-1} + \sum_j \bar{t}_j x_k^j \leq C_k, & \forall k \\ & C_k \leq T_H, & \forall k \end{cases}
$$

III. Scheduling with operator model

While the deterministic formulations of **SP** and **SP**$_{\mathbf{CE}}$ are appealing due to their simplicity, they lacks a key fundamental feature that does not make it amenable to implementation in a human supervisory control setting. For example, neither **SP** nor **SP**$_{\mathbf{CE}}$ account for the fact that accumulated reward and search times are random variables in real life. The reward of a task, R_j, relies on the stochastic (search) outcome of the human operator, and is therefore a random variable described by a specific probability distribution function. The search times t_j are also random variables. This section outlines appropriate modifications for the optimization problem to account for these realistic operator descriptions.

III.A. Operator modeling

One method for modeling the operator information gathering process is with use of 2-alternative choice models (2-AC) [20–25], a mathematical description of human accumulation of evidence. 2-AC models abstract the information accumulation process as a stochastic diffusion process

$$dx = \alpha dt + \sigma dW(t) \tag{2}$$

where $x(t)$ represents the integrated evidence in favor of one of the two alternatives,[c] $A = \{0,1\}$; σ denotes the intensity of the Wiener process $W(t)$; and α is the drift rate. It can be shown [26] that key characteristics of the diffusion model of Equation (2) can be summarized by two random variables: the probability of choosing one alternative, and mean response time

$$P_A \;=\; \frac{1}{1+\exp(2\eta\theta)} + \frac{1-\exp(-2x_0\theta)}{\exp(2\eta\theta)-\exp(-s\eta\theta)} \tag{3}$$

$$\bar{T} \;=\; \theta\tanh(\eta\theta) + \frac{2\eta\left(1-\exp(-2\eta\theta)\right)}{\exp(2\eta\theta)-\exp(-s\eta\theta)} - x_0 \tag{4}$$

where $\eta = (\alpha/\sigma)^2$ acts like a signal-to-noise ratio; $\theta = x_{th}/\alpha$ is related to the internal threshold x_{th} of the human decision maker; and P_A denotes the probability of choosing alternative $A = 1$. A key benefit of using detection probability and mean search times is that they can be determined experimentally.

Unfortunately, 2-AC models do not include important features in supervisory control applications, such as vigilance, boredom, or high workload. For example, a key feature in the 2-AC model is that the detection probability increases with the time spent learning about the underlying task. However, it is known that operator accuracy can degrade with time [27]. Additionally, 2-AC models may be overly simplistic in situations where operators make more complex decision tasks; i.e., in cluttered image search tasks, the operator is not simply tasked with declaring the absence or presence of the signal (the target). Rather, the operator must also prescribe the location of the target accurately, requiring the more complex Free-Response paradigm [28–30]. Experimental data obtained from previous experiments in supervisory control [31] has validated operator detection models that predict that operator accuracy can decrease with time.

The model for the detection probability $P(t_j)$ has been found from previous empirical results obtained in UAV single operator search tasks. Figure 4(a) shows the best estimate of the probability of detection (solid line) as a function of search time for a visual search task run from previous experimental data in multi-UAV simulated missions [31]. In this case, the estimate of detection probability $P(t)$ was modeled using a logistic regression of the form

$$P(t_j) = \frac{1}{1+\exp(\beta^T \mathbf{t})} \tag{5}$$

---

[c]For example, one alternative could be "target present", while the other alternative could be "target absent". The initial condition $x_0 = 0$ applies if the initial state is unbiased.

(a) Detection probability vs. time

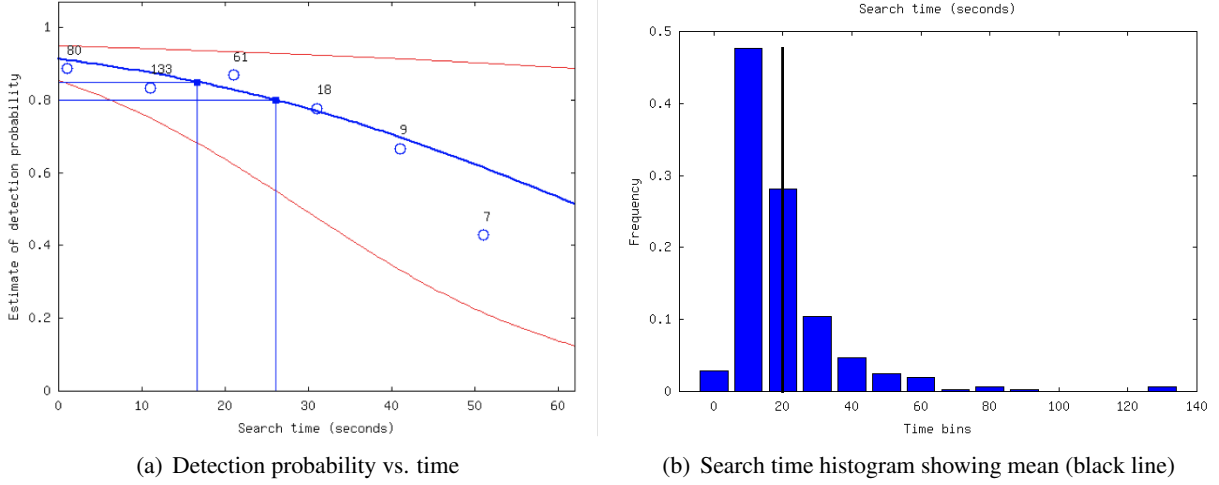(b) Search time histogram showing mean (black line)

Figure 4. Operator detection models obtained from previous experimental results: detection probability (left) and search times (right)

with $\mathbf{t} = [1, t_j]$; t_j is the total search time; and $\beta = [\beta_0, \beta_1]$ is the vector of parameters of the logistic regression. The regression on the parameter β was found be be described by $\beta = [-2.3, -.037]$, and the 2-σ range of the regression coefficients was given by $\beta_0 \in [-2.78, -1.82]$ and $\beta_1 \in [-.05, -.025]$. Operator search times are shown in Figure 4(b), and are well approximated with a lognormal distribution with estimated mean and variance parameters $\{\hat{\mu}, \sigma^2\}$. Note that when the outcome of the search is stochastic (due to random search times and detection probability), a reasonable formulation is that an operator would seek to maximize an expected accumulated reward $\max \mathbf{E}_p[R_{T_H}]$ where the total accumulated reward up to time T is given by R_T, and we explore this further in the next section.

III.B. Operator detection model

Using the expected value of the reward, a scheduling formulation for the sequential search task problem with operator models can be formulated as follows

$$\textbf{SPOM} \begin{cases} & \max_{x_k^j \in \{0,1\}} \mathbf{E}\left[\sum_j \sum_k R_j x_k^j\right] \\ \text{subject to:} & \sum_k x_k^j \leq 1 & \forall j \\ & \sum_j x_k^j \leq 1 & \forall k \\ & \sum_j (s_j + t_j) x_k^j \leq C_k, & \forall k \\ & C_{k-1} + \sum_j t_j x_k^j \leq C_k, & \forall k \\ & C_k \leq T_H, & \forall k \end{cases}$$

where $\mathbf{E}\left[\sum_j \sum_k R_k x_k^j\right] = \sum_j \sum_k \mathbf{E}[R_k] x_k^j = \sum_j \sum_k P(t_j) R_k x_k^j$. Here, $P(t_j)$ denotes the probability of detection of the search task. Note that the **SPOM** formulation is a minor modification to the **SP** formulation.

An example of the increase in mission time T_H with the operator model and schedule using **SPOM** is shown in Figure 5 (using the same set of tasks as those used in Figure 3). Figure 5(a) shows the increased accumulated reward by increasing mission time, and shows significant differences in the original sensitivity regions of problem **SP** (e.g., $T_H = 170$ is still sensitive, while $T_H > 300$ does not affect expected accumulated reward much). Figure 5(b) shows the task allocation as a function of the time T_H. The diamonds correspond to which tasks were selected at a specific T_H. Note that the tasks selected can vary considerably from the optimization in **SP** (which does not consider the operator model). For $T_H = 250$, optimization **SP** returns that
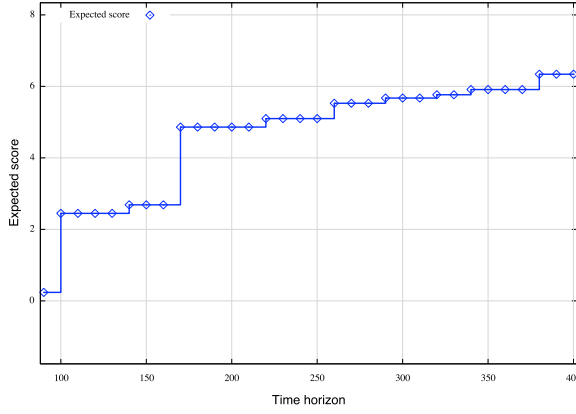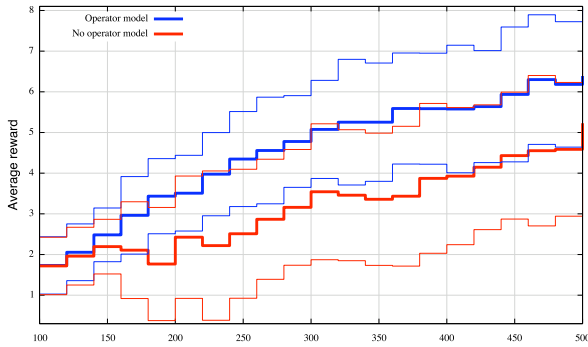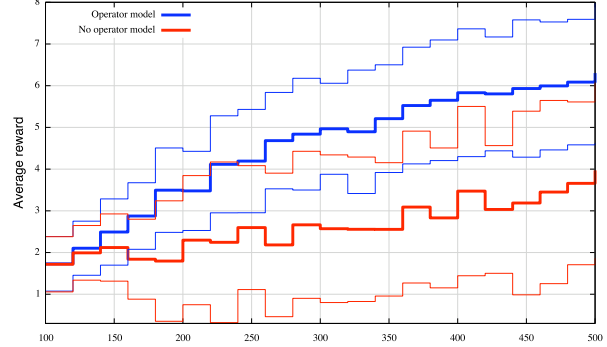
(a) Accumulated score, with operator model
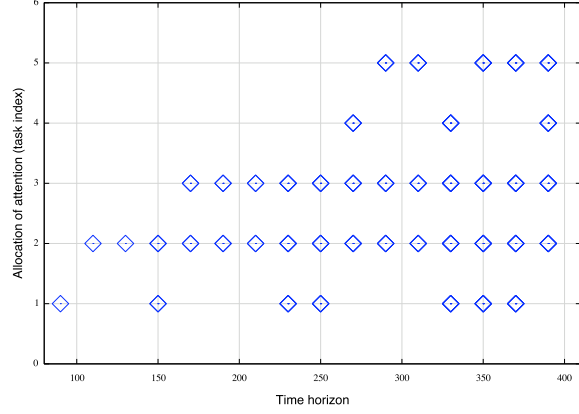
(b) Optimal assignment, with operator model

Figure 5. Accumulated score with increased mission time (left) and optimal assignment (right) including the operator model


(a) Average reward for 7 tasks

(b) Average reward for 9 tasks

Figure 6. Comparison of average expected reward using SP and SPOM(1-σ deviation shown in thinner lines). The average reward for SP was calculated by evaluating the expected reward with the schedule obtained by solving SP. (Note that accumulated rewards increase mononotically when using additional MC simulations.)

the operator should look at tasks 3 and 5, for a total reward of 14.3 (see Figure 3(a)), while the optimization **SPOM** would allocate tasks 1, 2, and 3, for a total expected reward of 5.19, considerably lower than that predicted by **SP**. Tasks 1-3 are selected since their expected reward is much higher than that of tasks 3 and 5 combined. Note that the reward obtained by **SP** significantly overestimates the expected reward obtained by **SPOM**.

III.C.   Comparison between SP and SPOM

To validate the proposed model, the expected reward of the two problems **SP** and **SPOM** were compared in 100 Monte Carlo simulations. The task schedule for **SP**, $\mathscr{S}_{\mathbf{SP}}$, was optimized based on the actual rewards R_j and the expected reward of this task list was evaluated as $\sum_{j \in \mathscr{S}_{\mathbf{SP}}} E[R_j] = \sum_{j \in \mathscr{S}_{\mathbf{SP}}} R_j P(t_j)$, where t_j was the search time for task j. The schedule for **SPOM**, $\mathscr{S}_{\mathbf{SPOM}}$, was also evaluated, and the reward for this task list was evaluated as $\sum_{j \in \mathbf{SPOM}} E[R_j]$. The results of a 7- and 9- task example are shown in Figures 6(a) and 6(b), where the average expected reward for problem **SPOM** (in blue) is compared to the average expected reward for problem **SP** (in red). The average expected reward was averaged across the 100 Monte Carlo simulations, and 1-σ standard deviation lines are shown in the thinner lines. As anticipated, the true expected performance of problem **SP** is much lower than **SPOM** (which accounts for the operator model).
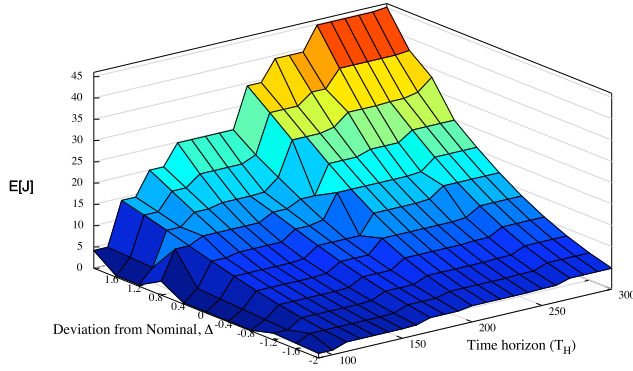
(a) Sensitivity analysis of expected score for 9 tasks
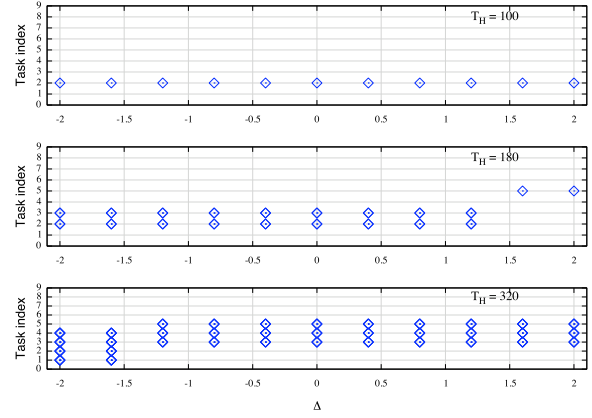
(b) Task selection sensitivity

Figure 7. Sensitivity to operator model $P(t_j)$ for different time T_H

III.D. Sensitivity to operator detection model

One of the core assumptions of the scheduling problem with the operator model **SPOM** is the accuracy of the operator detection model $P(t_j)$. Since the detection probability model was estimated from previous experimental results [31], it is fundamental to understand how sensitive the schedule and the schedule performance are to the operator detection model $P(t_j)$. Unfortunately, sensitivity analysis of scheduling problems is challenging task [32, 33], and scheduling under uncertainty is still an ongoing research topic.

For this paper, the preliminary investigation will be empirical, and we consider the sensitivity of the detection probability as follows

$$P(t_j) = \frac{1}{1 + \exp(\tilde{\beta}^T \mathbf{t})} \tag{6}$$

with $\tilde{\beta} = \beta \pm \Delta \sigma_\beta$; β is the nominal parameter set of the logistic regression; and σ_β is the standard deviation. The tuning parameter Δ introduces the amount of variability in the detection model. This parameter was then varied in the interval $\Delta \in [-2, 2]$ to cover the expected variability of the logistic regression model.

Some preliminary results are shown in Figure 7(a) and Figure 7(b). Figure 7(a) shows the sensitivity of a 9-task schedule, where the expected accumulated reward is parametrized by the time T_H and the deviation from the nominal model Δ. The schedule was re-evaluated for each value of T_H and Δ, and can be seen that for the choices of $\{R_j, t_j, s_j\}$, the schedule performance can in fact be quite sensitive to the detection probability model, particularly (and as expected), for longer mission times, when more tasks are available to be placed in the optimal schedule. While the sensitivity of the optimal cost is important, it is not clear whether the change in performance arises from the change in the reward evaluated for each task, or a change in the task selection (which task is chosen in which order). Therefore it is more crucial to confirm the sensitivity of the optimal task choice.

Figure 7(b) shows the effect of the change in the probability model (shown by Δ) and the change in the task assignment for three difference times T_H ($T_H \in \{100, 180, 320\}$). Note that for the problem parameters in the shorter mission time, the task choice is not sensitive to the operator model. For a time of $T_H = 180$ units, there is no change in the optimal schedule: it is optimal to choose task 2. As a longer mission time is chosen, it is optimal to perform task 2 followed by task 3 under the nominal model ($\Delta = 0$). If the model is underestimated ($\Delta < 0$), the optimal schedule does not change, but if the model is overestimated, the optimal choice is task 5. For even longer times, the nominal schedule ($\Delta = 0$) is to select tasks 5, 4, and 3 (in that order). If the model is underestimated, however, the optimal schedule is to perform tasks 2, 1, 4, and 3.

(a) Under-estimation of expected score

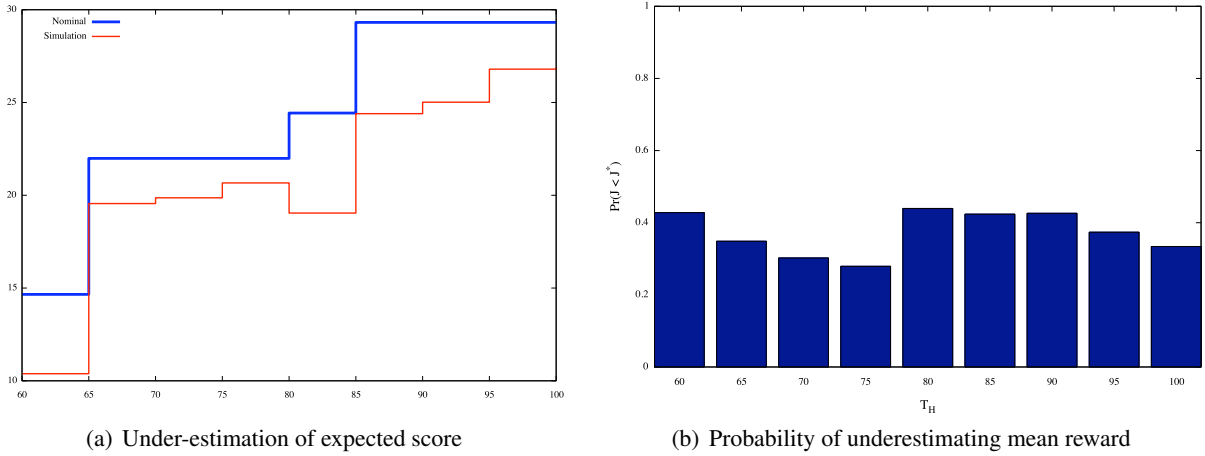(b) Probability of underestimating mean reward

Figure 8. Sensitivity of optimal reward to precise knowledge of search time

Future work will be devote to determining whether it is possible to determine the sensitivity of the optimal schedule using on-line methods. An additional item of concern is the uncertainty in the search times.

III.E.   Sensitivity to uncertain search times

Since the optimization of **SPOM** is a discrete decision problem, it can also be sensitive to the knowledge of the search times. As for the case of the detection model, a thorough sensitivity analysis can be much more complex, particular because the temporal effect appears in the constraints of the optimization. One way of handling the fact that the search times are uncertain, is to substitute the mean value, $\bar{t}$, and use this value as the *true* value of the optimization. Thus, one can solve the following problem

$$
\textbf{SPOM}_{\textbf{CE}}: \quad
\left\{
\begin{array}{l}
\max \mathbf{E}_{p,t}\left[\sum_j \sum_k R_k x_k^j\right] \\
\text{subject to:} \quad \sum_k x_k^j \leq 1 \quad \forall j \\
\phantom{\text{subject to:}} \quad \sum_j x_k^j \leq 1 \quad \forall k \\
\phantom{\text{subject to:}} \quad \sum_j (s_j + \bar{t}_j) x_k^j \leq C_i \\
\phantom{\text{subject to:}} \quad C_{i-1} + \sum_j \bar{t}_j x_k^j \leq C_i \\
\phantom{\text{subject to:}} \quad C_i \leq T
\end{array}
\right\}
$$

The **SPOM**$_{\textbf{CE}}$ approach can lead to a loss of performance when implemented on the actual scheduling system, since the optimized schedule may not be feasible to implement on the real system, since the search time durations may vary significantly from the assumed mean. For example, the schedule calculated by **SPOM**$_{\textbf{CE}}$ that optimistically assumes the nominal search times, can actually lead to fewer of the tasks being found in the allotted time T_H. An example of this is shown in a 100-Monte Carlo simulation result of a 5-task schedule with increasing time horizon T_H (Figure 8(a) shows an example of Monte Carlo simulations under a particular scenario). The optimal schedule calculated with **SPOM**$_{\textbf{CE}}$ is shown in blue, while the expected value of the schedule (evaluated with realizations of the search time), is shown in red; the accumulated reward decreases by almost 10% simply by using the realized search times. The Monte Carlo results using 100 different search times are shown in Figure 8(b), which plots the probability of underestimating the accumulated reward (calculated as $\Pr(J < J^*)$, where J^* is the optimal reward of **SP**$_{\textbf{CE}}$). Underestimation can occur as frequently as 40%. Sensitivity analysis, as for the case of the detection model, is an ongoing research focus.
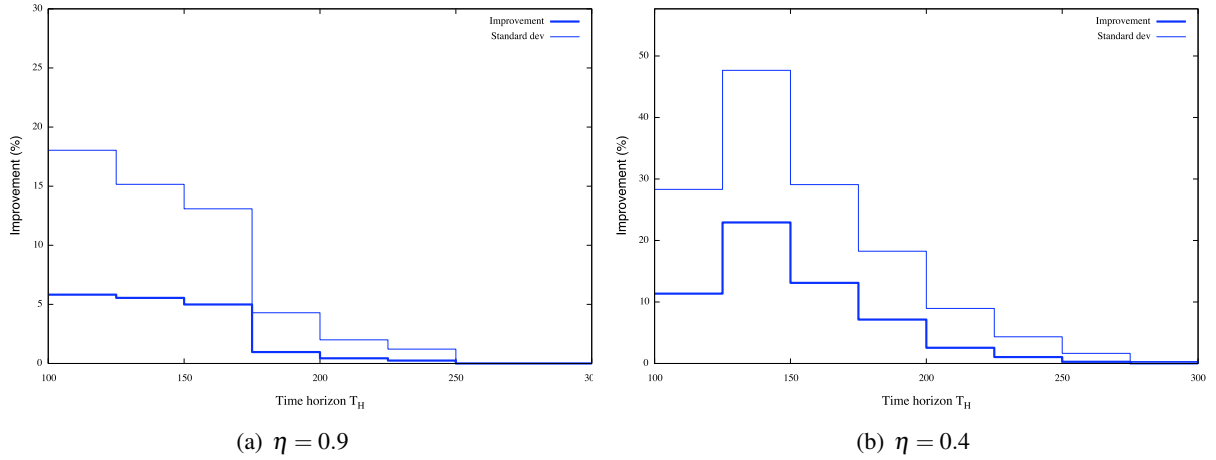
(a) $\eta = 0.9$ (b) $\eta = 0.4$

Figure 9. **Receding horizon greedy comparison showing improvement of the scheduling algorithm over the greedy algorithm (in percentage): mean improvement $\bar{J}$ and standard deviation σ_J are calculated from simulation data**

IV. Modeling the Human Operator: A Greedy Approach

Thus far, the modeling approach has been primarily focused on previously observed statistical features, such as detection probability and decision times. However, a rich set of qualitative and quantitative behavioral information has also been inferred in previous scheduling experiments that can be encoded in these simulations. This section considers the operator high-level decision making from a behavioral perspective from previous literature, and, encoding the previous statistical models, makes performance predictions from the output of a decision support algorithm to what could be expected from a typical operator.

In an attempt to model the anticipated operator strategy, we take inspiration from previous results in sequential choice (Bandit-like) problems [34, 35] and "least-commitment" strategies in manufacturing [15] which have shown that humans will likely only exhibit a finite-horizon lookahead. We assume that operators will greedily select the search tasks based on value or availability [9], and subjects are assumed to maximize the reward while ensuring the task can be completed. The subjects then pick the most rewarding and feasible task first, remove it from the available tasks, and repeat. Note that operator greedy strategies have been seen in previous work [9, 36, 37].

The greedy operator approach model was tested against the scheduling problem **SPOM** in $N_{MC} = 200$ Monte Carlo (MC) simulations. For each MC simulation, search times t_j and start time s_j were generated randomly from first generating random instances of the mean and variance of the log-normal distribution

$$\mu = 2.92 + 0.15 * [1:N]. * \mathtt{rand}, \quad \sigma^2 = 0.6 + 0.05 * [1:N]. * \mathtt{rand} \tag{7}$$

The search times were calculated as $t_j = \mu(j) + \eta\sigma(j)$ where η was a tuning parameter that reflected much longer search times. The optimal (**SPOM**) and greedy schedules were then calculated for different times $T_H \in \{100, 125, \ldots, 300\}$. The expected reward for the greedy schedule was calculated based on the regression model of Equation (5). The expected accumulated reward for the greedy and scheduled (J_g and J_s respectively) was calculated for each instance of the search times and each T_H, and the average and standard deviation of the accumulated reward ratio of each schedule were calculated. These results are shown in Figure 9 for two different values of η. The case for $\eta = 0.9$ (Figure 9(a)) shows that as the mission time increases, the difference between the greedy and scheduled algorithm decreases as expected. For moderate times, however, this problem is fairly constrained, and so the difference between the scheduled assignment and the greedy allocation is on the order of 5%. When there is a much longer time to schedule the tasks, as in the case of $\eta = 0.4$ (Figure 9(b)), the benefit of the scheduling algorithm is predicted to improve performance on the order of 20%.

V.   Experimental Design

An ongoing experiment is being conducted that tests the scheduling performance of operators in a group of real-time planning tasks. The experiment takes inspiration from the obtained simulation results, and has two goals in mind. First, we seek to validate the previous literature and determine whether or not human subjects are efficient at scheduling search tasks as they are in scheduling job-shop problems. The primary differences are that search tasks are primarily perceptual, and exhibit a very high variability between human operators. Secondly, this experiment is interested in determining whether or not the addition of supplementary information is beneficial in the allocation of sequential search tasks. This additional information takes the form of a 'preview panel', in which the subjects see a restricted version of the search task: presumably, by inferring the kind of environment and the goal of the search task, the subjects should be able to obtain more information on the difficulty of the overall task, and use this information to generate more efficient search schedules.

Due to the tuning requirements (and sensitivity) of some of the problem parameters, this experiment will first investigate whether or not problem $\textbf{SP}_{\textbf{CE}}$ can provide good decision support to the operator. Implementation of a decision support that uses **SPOM** will be future work.

V.A.   Scheduling Modes

The operator will have a number of tasks available, which will appear stochastically. At any given time an operator has to optimize his accumulated reward by scheduling and conducting as many search tasks as possible. Moreover, operators are evaluated on their ability to schedule vehicle search tasks introduced in Section II in three different scheduling modes:

1) **Human Only (HO) scheduling**: the operator is solely responsible for the scheduling of the search tasks

2) **Algorithm Aided (AA) scheduling**: the scheduling algorithm solves problem **SP** and provides the optimized sequence of tasks to the operator, who will then choose whether to implement this schedule or not. The operator can still reschedule tasks by interacting with the optimization algorithm

3) **Algorithm Only (AO) scheduling**: the scheduling algorithm solves problem **SP**, and presents the tasks to the operator in the optimized order. The operator is not able to change the optimized schedule; the operator has to accept it; however, the operator is able to skip an individual task when engaged

V.B.   HOSS Simulation Environment

The experiment is performed in the HOSS environment (Helping Operators Schedule Search tasks), a simulation environment tailored to investigate human-in-the-loop search task scheduling scenarios. A typical HOSS interface is shown in Figure 10. A single operator has an overview of a list of tasks which are, or will be, available in the near future in Pane B (the task overview pane). The tasks are defined by a unique ID number j, a reward R_j (Low, Medium, or High), the task start time s_j, and the expected search time $\bar{t}_j$.

The objective of the operator is to maximize the accumulated reward by constructing appropriate schedules in the **HO** mode, by modifying the schedules in the **AA** mode, and by accepting the schedules in the **AO** mode. In the **HO** and **AA** modes, the operator can drag tasks from the task overview pane to the active task list (Pane D). The operator has the ability to reschedule a task. The scheduled task are also graphically represented in a timeline (Pane F).

Once scheduled in the active task list, depending on the start time, the task can be engaged, by clicking the 'Engage' button, initiating a search task in the payload search pane (Pane A). A search task consists of a identifying a target in an image loaded in the payload search pane. The operator can actively pan and zoom
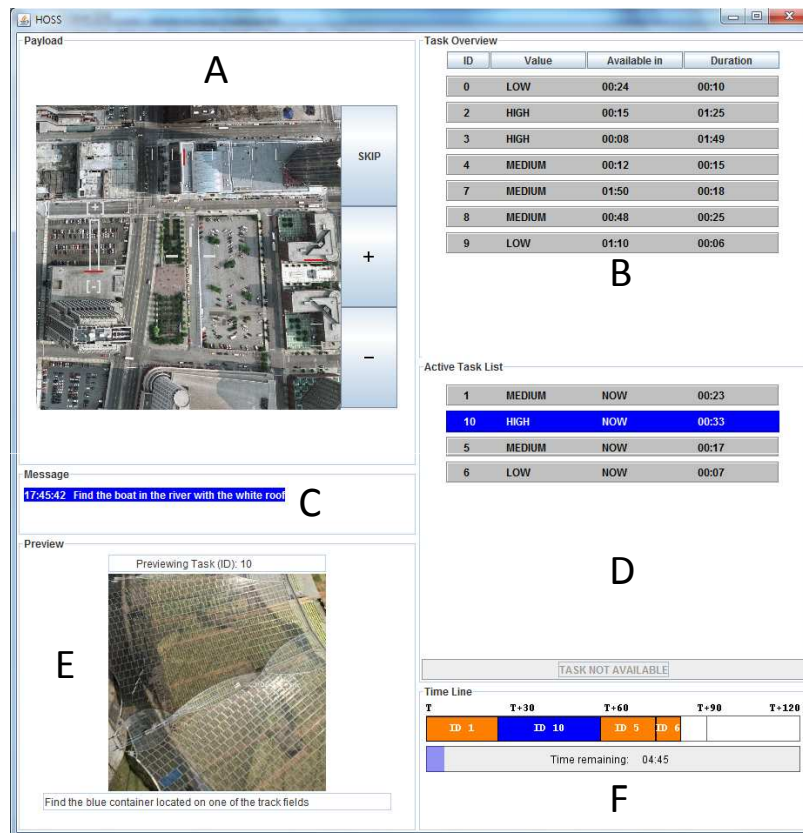
Figure 10. HOSS interface. The payload search pane is denoted by A, B is the task overview pane, C is the message box, the active task list pane is given by D, E is the preview pane and finally, E is the timeline pane

in the payload search pane in order to identify and submit the target. The target description is given in the message pane (Pane C).

As mentioned before, in order to see if an additional source of information in the form of a search task preview helps the operator scheduling, HOSS also provides the operator with a preview (Pane E). The preview is accompanied by a short task description. The preview feature adds another two experimental conditions to the three scheduling modes. The results of this experiment will be published in future work.

VI. Conclusion and Future Work

This paper has presented a novel scheduling formulation for the single operator, multi-UAV problem that is designed to maximize the operator's accumulated reward of the search tasks in a time-pressured environment. The formulation was generalized to include operator performance obtained from previous human-in-the-loop experiments. Comparisons were made between the expected performance difference between this scheduling system and a greedy scheduling strategy representative of operator planning, showing the potential for improvement of the proposed strategy.

Ongoing work is analyzing the experimental results of a human-in-the-loop experiment in which subjects were asked to optimize their scheduling of search tasks in a time-pressured environment. Comparisons between human-only and human-assisted conditions are being evaluated. Our ongoing work is also addressing the importance of uncertainty of search times in the scheduling recommendations made by the decision support algorithms.

VII. Acknowledgements

References

[1]L. Johnson, "Beale officials send Global Hawk to aid in Haiti earthquake relief efforts ." http://www.af.mil/news/story.asp?id=123185619.

[2]A. Moore, "Predator team deploys to support Haiti relief efforts." http://www.af.mil/news/story.asp?id=123187427.

[3]M. L. Cummings and S. Bruni, "Collaborative human-automation decision making," in Handbook of Automation (S. Y. Nof, ed.), pp. 437–447, Springer, 2009.

[4]C. Drew, "Military Is Awash in Data From Drones." http://www.nytimes.com/2010/01/11/business/11drone.html.

[5]M. Hoffman, "New Reaper sensors offer a bigger picture." http://www.airforcetimes.com/news/2009/02/airforce_WAAS_021609/.

[6]M. Tulga and T. Sheridan, "Dynamic decisions and work load in multitask supervisory control," IEEE Transactions on Systems, Man, and Cybernetics, vol. 10, no. 5, pp. 217–232, 1980.

[7]M. Cummings and P. Mitchell, "Managing Multiple UAVs through a Timeline Display," Proceedings of AIAA Info Tech, 2005.

[8]M. Cummings and P. Mitchell, "Automated scheduling decision support for supervisory control of multiple UAVs," AIAA Journal of Aerospace Computing, Information, and Communication, vol. 3, no. 6, pp. 294–308, 2006.

[9]J. Crandall and M. Cummings, "Identifying predictive metrics for supervisory control of multiple robots," IEEE Transactions on Robotics, vol. 23, no. 5, p. 942, 2007.

[10]M. Segal, "The operator-scheduling problem: A network-flow approach," Operations Research, vol. 22, no. 4, pp. 808–823, 1974.

[11]E. Keith, "Operator scheduling," IIE Transactions, vol. 11, no. 1, pp. 37–41, 1979.

[12]W. Henderson and W. Berry, "Heuristic methods for telephone operator shift scheduling: an experimental analysis," Management Science, vol. 22, no. 12, pp. 1372–1380, 1976.

[13]T. Yamada, K. Yoshimura, and R. Nakano, "Information operator scheduling by genetic algorithms," Lecture notes in computer science, pp. 50–57, 1999.

[14]H. Aytug, M. Lawley, K. McKay, S. Mohan, and R. Uzsoy, "Executing production schedules in the face of uncertainties: A review and some future directions," European Journal of Operational Research, vol. 161, no. 1, pp. 86–110, 2005.

[15]J. Cegarra, "A cognitive typology of scheduling situations: a contribution to laboratory and field studies," Theoretical Issues in Ergonomics Science, vol. 9, no. 3, pp. 201–222, 2008.

[16]J. Berk, E. Hughson, and K. Vandezande, "The price is right, but are the bids? An investigation of rational decision theory," The American Economic Review, vol. 86, no. 4, pp. 954–970, 1996.

[17]D. Bertsimas, Introduction to Linear Optimization. Athena Scientific, 1997.

[18]K. Baker and D. Trietsch, Principles of Sequencing and Scheduling. Wiley-Blackwell, 2009.

[19]L. Bertuccelli, N. Pellegrino, and M. Cummings, "Choice Modeling of Relooks Tasks in UAV Search Missions," IEEE American Control Conference, 2010.

[20]R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The Physics of Optimal Decision making: A Formal Analysis of Models of Performance in Two-Alternative Forced Choice Tasks," in Psychological Review, vol. 113, pp. 700–765, 2006.

[21]J. R. Busemeyer and A. Diederich, "Survey of Decision Field Theory," in Mathematical Social Sciences, vol. 43, pp. 345–370, 2002.

[22]J. R. Busemeyer and J. T. Townsend, "Fundamental Derivations From Decision Field Theory," in Mathematical Social Sciences, vol. 23, pp. 255–282, 2000.

[23]A. Diederich and J. R. Busemeyer, "Simple Matrix Methods for Analyzing Diffusion models of Choice Probability, Choice Response Time, and Simple Response Time," in Journal of Mathematical Pschology, vol. 47, pp. 304–322, 2003.

[24]A. Diederich, "Dynamic Stochastic Models for Decision Making under Time Constraints," in Journal of Mathematical Pschology, vol. 41, pp. 260–274, 1997.

[25]R. J. Herrnstein, "Rational Choice Theory: Necessary but Not Sufficient," in American Psychologist, vol. 45, pp. 356–357, 1990.

[26]C. Gardiner, Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences. Springer-Verlag, 1985.

[27] M. Mouloua and R. Parasuraman, Automation and Human Performance: Theory and Applications (Ch. 9). Lawrence Erlbaum Associates, 1996.

[28] J. Irvine, "Assessing target search performance: the free-response operator characteristic model," Optical Engineering, vol. 43, p. 2926, 2004.

[29] D. Chakraborty, "Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data," Medical Physics, vol. 16, p. 561, 1989.

[30] T. Hutchinson, "Free-response operator characteristic models for visual search," Physics in medicine and biology, vol. 52, pp. L1–L3, 2007.

[31] C. Nehme, Modeling Human Supervisory Control In Heterogeneous Unmanned Vehicle Systems. PhD thesis, MIT, 2008.

[32] B. Penz, C. Rapine, and D. Trystram, "Sensitivity analysis of scheduling algorithms," European Journal of Operational Research, vol. 134, no. 3, pp. 606 – 615, 2001.

[33] N. Hall and M. Posner, "Sensitivity analysis for scheduling problems," Journal of Scheduling, vol. 7, no. 1, pp. 49–83, 2004.

[34] N. Gans, G. Knox, and R. Croson, "Simple models of discrete choice and their performance in bandit experiments," Manufacturing & Service Operations Management, vol. 9, no. 4, p. 383, 2007.

[35] D. Acuna and P. Schrater, "Bayesian Modeling of Human Sequential Decision-Making on the Multi-Armed Bandit Problem," in Proceedings of the 30th Annual Conference of the Cognitive Science Society, 2008.

[36] L. Charny and T. Sheridan, "Satisficing decision-making in supervisory control," Unpublished paper. Cambridge, MA: MIT, Man-Machine Systems Laboratory, 1986.

[37] T. Sheridan, Telerobotics, automation, and human supervisory control. MIT Press, 1992.