Adaptive Attention Allocation in Human-Robot Systems

Vaibhav Srivastava Amit Surana Francesco Bullo

***Abstract*— We propose an optimization framework to study two fundamental attention control aspects in human-robot systems: Where and how much attention should the operator allocate? In other words, which information source should be observed by the operator, and how much time duration should be allocated to the information feed in order to optimize the overall performance of the human-robot system? The proposed framework incorporates (i) operator performance constraints, such as error rates and service times based utilization history, (ii) sensor constraints, such as processing/travel time, and (iii) task constraints, such as prioritization. We use a receding horizon approach to solve the resulting dynamic program, leading to efficient policies for operator time duration allocation and sensor selection. We demonstrate our methodology in a distributed surveillance problem.**

I. Introduction

While emergence of mobile and fixed sensor networks operating at different modalities, mobility, and coverage, has enabled access to an unprecedented amount of data, human operators are unable to extract actionable information in a timely manner. As a consequence, there has been an increasing need of sensor management automation, and has spawned extensive research, see for example special issues [9] and [4] dedicated to topics of design of cooperative control and coordination strategies for teams of mobile sensors. While the use of automation will continue to increase, human operators will remain indispensable, since they can often bring knowledge and experience to bear in the complex problems such as surveillance and intelligence gathering, where a wrong decision can have lethal consequences. Given the complex interaction that can arise between the operator and the automated sensors, it is necessary to develop decision support systems that exploit not only operator strengths, but also account for their decision making inefficiencies, such as error rates and loss of situational awareness.

Several studies have focused on using operator performance models to optimize interaction with automation. Queuing model with operator as server has emerged as a popular paradigm [11]. A task release controller was designed in [10]. The authors consider operator's service time modeled by the Yerkes-Dodson curve and determine the maximum stabilizing arrival rate and a task release policy such that the situational awareness of the operator is at a desired level. Cooperative strategies between UAVs and humans have been studied in [12], [11]. The authors develop policies for joint target assignment for the operator and the UAVs such that the average time between appearance and classification of a target is minimized. Optimal time duration allocation policies for the human operator have been developed in [16], [14]. The authors consider queues with decision making tasks. Each task has a deadline which is incorporated as a soft constraint. The human decision making is modeled by a sigmoid function and the optimal allocation schemes are determined through the solution of an MDP. Further, the optimal task release rate is also determined. Bertuccelli et al [2] study the queues with re-look tasks. They consider a queue in which the arriving difficult tasks are placed in an orbit queue for later viewing while simple tasks are immediately processed. The authors show that in visual search tasks with re-queuing the accuracy of the detection improves but the number of targets searched decreases with re-queuing probability. The problem of determining the optimal sequence in which tasks should be processed by the operator has been studied in [1]. Here, the authors also point to the issue of high sensitivity of the optimal solution to operator model uncertainty. Certain decision making models for the human operator have been utilized in [18], [17], [15] to determine optimal sensor selection policies for human-robot systems.

In this paper we propose a novel framework to study two fundamental questions: "Which information source should be observed by the operator, and how much time duration should be allocated to the information feed in order to optimize the overall performance of the human-robot system?" Addressing these problems together in context of human-supervised distributed surveillance is one of the novel contribution of this paper. Additionally, we incorporate operator performance models related to both error rates and situational awareness in the optimization framework for determining optimal time-duration allocation. These two aspects of operator performance in decision making queues have only been studied separately in previous works listed above. We consider several variations of decision making queues including (i) time constrained static queues, (ii) dynamic queues with latency penalty and (iii) queues with prioritization and mandatory tasks. Using this optimal solution of time-duration allocation problem, we next close the loop on sensor selection by treating the operator input as a "sensor observation." Specifically, we run spatial quickest detection algorithm on

The funding provided by United Technology Research Center for this work is greatly appreciated. This work was supported in part by AFOSR MURI award FA9550-07-1-0528.

Vaibhav Srivastava and Francesco Bullo are with the Center for Control, Dynamical Systems, and Computation, University of California, Santa Barbara, Santa Barbara, CA 93106, USA, {vaibhav, bullo}@engineering.ucsb.edu

Amit Surana is with the United Technologies Research Center, East Hartford, CT 06108, USA, SuranaA@utrc.utc.com

operator inputs and derive an expression for the detection delay taking into account time varying operator performance in addition to sensor constraints such as travel time. We use this expression for optimizing sensor selection policy in such a way that the operator is presented with information which is most relevant for detecting anomalies. This setup leads to quickest detection in distributed surveillance applications.

The rest of the paper is organized as follows. In Section II we describe the problem setup and the assumptions. In Section III we survey relevant operator models and incorporate them in an optimization framework for determining optimal time-duration allocation in decision making queues. We illustrate certainty equivalent and receding horizon solution approaches through several examples. Section IV covers the coupling of attention allocation and sensor selection problems for quickest detection. Here we also demonstrate our overall methodology in a distributed visual surveillance problem. Finally, we conclude with future directions in Section V.

II. Problem Setup

We consider a distributed surveillance problem of surveying n disjoint regions $\mathcal{R} = \{1, \dots, n\}$. An autonomous agent, at each iteration, uses a sensor selection policy to determine a region $k \in \mathcal{R}$ it should next travel to, travels to that region and collects information, e.g., takes a picture, and sends it to a support system. The support system sends the collected information to a human operator. The support system decides on how much time human operator should spend on each information feed. The information waits in the queue till the human operator processes it. The human operator incorporates the first come first serve policy to process the incoming information and decides on the presence/absence of any anomaly. The decisions made by human operator may be erroneous. In order to make accurate decisions, the support system runs quickest detection algorithm on the decisions made by the human operator and decides on the presence or absence of any anomaly in any region. Based on the decisions made by the human operator, the support system adapts the region selection policy such that the region with high probability of being anomalous is sampled with high probability. The problem setup is shown in Fig. 1. We

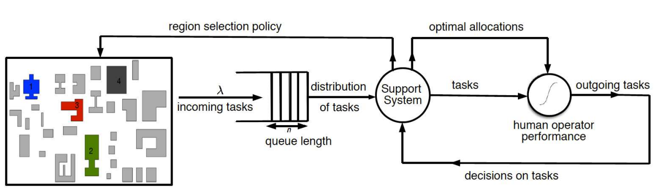


Fig. 1. Problem Setup

address this problem in two stages. First, in Section III we determine optimal time duration allocation policies for the human operator, that is, how much time the support system should ask the human operator to spend on each information feed? Second, in Section IV we determine simultaneous time duration allocation and sensor selection policies, where the human operator optimally processes the information feeds and her decisions are utilized to determine the region selection policy that ensures quickest detection of anomalies.

III. Optimal Time Duration Allocation

A. Operator Models

Speed-accuracy trade-off in human decision making:

Consider the scenario where, based on the collected evidence, the human has to decide on one of the two alternatives H_0 and H_1. The evolution of the probability of correct decision has been studied in cognitive psychology literature [8], [3].

Pew's model: The probability of deciding on hypothesis H_1, given that hypothesis H_1 is true, at a given time $t \in \mathbb{R}_{\geq 0}$ is given by

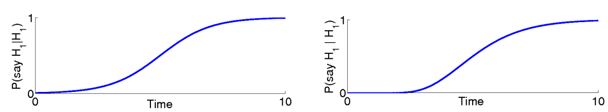
$$\mathbb{P}(\text{say } H_1 | H_1, t) = \frac{p_0}{1 + e^{-(at-b)}},$$

where $p_0 \in [0, 1]$, $a, b \in \mathbb{R}$ are some parameters which depend on the human operator [8].

Drift diffusion model: Conditioned on the hypothesis H_1, the evolution of the evidence for decision is modeled as a drift-diffusion process [3]. Given a drift rate $\beta > 0$, and diffusion rate σ, with a decision threshold η, the conditional probability of the correct decision is

$$\mathbb{P}(\text{say } H_1 | H_1, t) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \int_{\eta}^{\infty} e^{\frac{-(\Lambda - \beta t)^2}{2\sigma^2 t}} d\Lambda,$$

where $\Lambda \equiv \mathcal{N}(\beta t, \sigma^2 t)$ is the evidence at time t.



(a) Pew's model (b) Drift diffusion model

Fig. 2. The sigmoid evolution of the probabilities of correct decision

Yerkes-Dodson Law and Situational Awareness:

The Yerkes-Dodson law [19] is an empirical relation between the operator utilization and performance of a human operator. It states that the performance of a human operator is unimodal, that is, an inverted-U function of the utilization. In other words, the performance of a human operator increases with increasing operator utilization only upto a point. Any increase in operator utilization beyond this point results in decreased performance of the operator.

The Yerkes-Dodson law has inspired situational awareness models [6]. These models suggest that the human operator is less situational aware at low utilization as well as at high utilization, and accordingly, takes more time to service a task. Therefore, the expected service time can be modeled by a convex function of the utilization ratio, that is, the ratio of operator utilization to its maximum utilization.

The utilization ratio x of a human operator is captured by the following differential equation [10]

$$\dot{x}(t) = \frac{b(t) - x(t)}{\delta},$$
$$\text{where} \quad b(t) = \begin{cases} 1, & \text{operator is busy at time } t \\ 0, & \text{otherwise,} \end{cases}$$

and δ is a constant that depends on operator's sensitivity. Therefore, if the initial utilization ratio of an operator is x_0 and she serves a task for time t and rests for time r after processing the task, her utilization ratio evolves to

$$x(t+r) = (1 - e^{-\frac{t}{\delta}} + x_0 e^{-\frac{t}{\delta}})e^{-\frac{r}{\delta}}. \tag{1}$$

B. Time Constrained Static Queue

Consider the scenario where the human operator has to process N decision making tasks (e.g., identification of anomalous regions) within a given time T. The objective of the operator is to maximize the expected number of correct decisions. The expected number of correct decisions is the sum of probabilities of correct decision for each task. Let the performance function of the operator on task $\ell \in \{1, \dots, N\}$ be the sigmoid function $f_\ell : \mathbb{R}_{\geq 0} \to [0, 1[$. Some of the tasks may be more important than others, and accordingly, a weight is assigned to each task. Let w_ℓ be the weight on task ℓ. Let x_ℓ be the utilization ratio of the operator before starting task ℓ, and $S_\ell : [0, 1] \to \mathbb{R}_{>0}, \ell \in \{1, \dots, N\}$ be the function of x_ℓ that captures expected service time of the operator on task ℓ. Note that S_ℓ is the expected service time under the natural response of the operator. The support system wishes to control the time the operator allocates to a task. The time allocation on task ℓ suggested by the support system to the operator should be compatible with S_ℓ. In order to capture this feature, we constrain the allocations suggested by the support system to be more than S_ℓ. The support system should also take into account the situational awareness effects and hence, optimally control the utilization ratio of the operator. To do so, support system recommends a rest time r_ℓ to the operator after processing task $\ell \in \{1, \dots, N\}$. Let the initial utilization of the human operator be x_1. The following optimization problem captures the objective of support system:

$$\begin{aligned} \underset{\boldsymbol{t},\boldsymbol{r},\boldsymbol{z}}{\text{maximize}} \quad & \sum_{\ell=1}^{N} z_\ell w_\ell f_\ell(t_\ell) \\ \text{subject to} \quad & \sum_{\ell=1}^{N} z_\ell (t_\ell + r_\ell) = T \\ & x_{\ell+1} = (1 - e^{-\frac{z_\ell t_\ell}{\delta}} + x_\ell e^{-\frac{z_\ell t_\ell}{\delta}})e^{-\frac{z_\ell r_\ell}{\delta}} \\ & z_\ell t_\ell \geq z_\ell S_\ell(x_\ell) \\ & x_\ell \in [x_{\min}, x_{\max}] \\ & z_\ell \in \{0, 1\}, t_\ell \in [0, T], r_\ell \in [0, T], \end{aligned} \tag{2}$$

for all $\ell \in \{1, \dots, N\}$, where z_ℓ is the variable that determines whether or not the task ℓ would be processed, $x_{\min}, x_{\max} \in [0, 1]$ are the desired bounds on the utilization ratio of the operator, and $\boldsymbol{t}, \boldsymbol{r}, \boldsymbol{z}$ are N-vectors of t_ℓ, r_ℓ, and z_ℓ, respectively.

We intend to solve optimization problem (2) via dynamic programming. To do so, we introduce a new variable a_ℓ that measures aggregate allocation till task ℓ. Let $a_0 = 0$. The optimization problem (2) is equivalent to

$$\begin{aligned} \underset{\boldsymbol{t},\boldsymbol{r},\boldsymbol{z}}{\text{maximize}} \quad & \sum_{\ell=1}^{N} z_\ell w_\ell f_\ell(t_\ell) \\ \text{subject to} \quad & a_\ell = a_{\ell-1} + z_\ell(t_\ell + r_\ell) \\ & x_{\ell+1} = (1 - e^{-\frac{z_\ell t_\ell}{\delta}} + x_\ell e^{-\frac{z_\ell t_\ell}{\delta}})e^{-\frac{z_\ell r_\ell}{\delta}} \\ & z_\ell t_\ell \geq z_\ell S_\ell(x_\ell) \\ & x_\ell \in [x_{\min}, x_{\max}], a_\ell \in [0, T] \\ & z_\ell \in \{0, 1\}, t_\ell \in [0, T], r_\ell \in [0, T], \end{aligned} \tag{3}$$

for all $\ell \in \{1, \dots, N\}$.

The optimization problem (3) has the structure of a dynamic program. The processing of a task and the rest following it define the stage of the dynamic program. a_ℓ and x_ℓ are the states at stage ℓ and t_ℓ, r_ℓ and z_ℓ are the controls at stage ℓ. The stage reward is $z_\ell f_\ell(t_\ell)$. We now elucidate on the solution of the optimization problem (3) with an example.

Example 1 (Time Constrained Static Queue): Suppose the human operator has to serve $N = 10$ tasks within time $T = 25$. Let the performance of the operator on task ℓ be given by sigmoid function $f_\ell(t_\ell) = 1/(1 + \exp(-a_\ell t_\ell + b_\ell))$, where a_ℓ and b_ℓ are the ℓth entries in the N-vectors

$$\boldsymbol{a} = [1\ 2\ 1\ 3\ 2\ 4\ 1\ 5\ 3\ 6], \quad \text{and}$$
$$\boldsymbol{b} = [5\ 10\ 3\ 9\ 8\ 16\ 6\ 30\ 6\ 12].$$

Let the weights on tasks be $\boldsymbol{w} = [2\ 5\ 7\ 4\ 9\ 3\ 5\ 10\ 13\ 6]$, and the initial utilization ratio of operator be $x_1 = 0.7$. Let $S_\ell(x_\ell) = b_\ell(37x_\ell^2 - 37x_\ell + 15)/6a_\ell$. We pick $x_{\min} = 0.5$ and $x_{\max} = 0.9$. The optimal solution to the time constrained static queue for $\delta = 1$ is shown in Fig. 3.

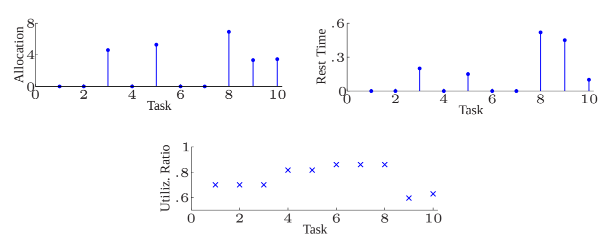


Fig. 3. Time constrained static queue. The top figures show the optimal allocations and the rest time. The bottom figure shows the optimal evolution of the operator utilization.

C. Dynamic Queue with Latency Penalty

Consider the scenario where the human operator has to serve a queue of decision making tasks. We assume that the tasks arrive according to a Poisson process with rate λ. Let Γ be a countable set, and assume that each task is sampled from a probability mass function $p : \Gamma \to [0,1]$ (accordingly, a probability density function if Γ is an arbitrary set). For a task $\gamma \in \Gamma$, let $f_\gamma : \mathbb{R}_{\geq 0} \to [0,1]$ be the performance function of the human operator and let w_γ be the reward allocated to the task, according to its priority. We assume that each task comes with a processing deadline and incorporate this deadline as a soft constraint. In particular, we let the task γ lose value at a rate c_γ while waiting in the queue. For example, if the task γ needs to be processed within time $T_\gamma^{\text{dead}} > t_\gamma^{\text{inf}}$ after its arrival, where t_γ^{inf} is the inflection point of the sigmoid function f_γ, then one may pick $c_\gamma = f'_\gamma(T_\gamma^{\text{dead}})$. The objective of the support system is to maximize its expected benefit, that is, the total expected reward obtained by processing the tasks minus the penalty incurred due to delay in processing the tasks.

In order to determine the optimal allocations that the support system should suggest to the human operator, we first determine the certainty equivalent solution, i.e., the solution considering the expected evolution of the system. Then, we modify this certainty-equivalent solution and determine optimal duration allocation policies for the dynamic queue. For the certainty-equivalent solution, define $\bar{c} = \mathbb{E}_p[c_\gamma]$, $\bar{w} = \mathbb{E}_p[w_\gamma]$, and $\bar{f} : \mathbb{R}_{\geq 0} \to [0,1]$ by $\bar{f}(t) = \mathbb{E}_p[w_\gamma f_\gamma(t)]/\bar{w}$. Let the initial queue length be n_1. In the spirit of the formulation in Subsection III-B, the certainty equivalent objective of the support system is:

$$\begin{aligned}
\underset{\boldsymbol{t},\boldsymbol{r},\boldsymbol{z}}{\text{maximize}} \quad & \lim_{L\to\infty} \frac{1}{L} \sum_{\ell=1}^{L} z_\ell \Big(\bar{w}\bar{f}(t_\ell) - \bar{c}(t_\ell + r_\ell)\bar{n}_\ell \\
& \qquad\qquad - \frac{\bar{c}\lambda(t_\ell + r_\ell)^2}{2}\Big) \\
\text{subject to} \quad & \bar{n}_{\ell+1} = \max\{1, \bar{n}_\ell - 1 + \lambda(t_\ell + r_\ell)\} \\
& x_{\ell+1} = (1 - e^{-\frac{z_\ell t_\ell}{\delta}} + x_\ell e^{-\frac{z_\ell t_\ell}{\delta}})e^{-\frac{z_\ell r_\ell}{\delta}} \\
& z_\ell t_\ell \geq z_\ell S_\ell(x_\ell) \\
& x_\ell \in [x_{\min}, x_{\max}] \\
& z_\ell \in \{0,1\}, t_\ell, r_\ell \in \mathbb{R}_{\geq 0}, \ \forall \ell \in \mathbb{N},
\end{aligned} \tag{4}$$

where $\bar{n}_\ell$ is the expected queue length just before processing task ℓ.

It can be seen that the optimization problem (4) has a stage structure and hence, can be solved using dynamic programming. We solve this dynamic program using a receding horizon strategy, i.e., at each iteration, we solve the optimization problem (4) for a finite horizon length; apply the control to the present stage; move to next stage, and recompute the finite horizon solution with new state as initial condition. In the following example, we show some features of the certainty equivalent solution.

Example 2: We consider the same set of sigmoid functions and weights as in Example 1. We normalize the weights by their mean. We pick $\bar{c} = 0.01$, $\lambda = 0.7$ and sample tasks from a uniform distribution. The optimal allocations for the receding horizon policy where the optimization problem (4) with horizon length 10 is solved at each stage is shown in Fig. 4. The optimal benefit per unit task, i.e., the total reward minus the cost incurred due to penalty, and the optimal benefit rate is also shown. Note that the optimal benefit per unit task is almost constant till a critical value of arrival rate, and then starts decreasing. Similarly, the optimal benefit rate increases till a critical value of the arrival rate and then saturates. At this critical value of the arrival rate, one expects an arrival as soon as the current task is finished. This is the arrival rate at which the system should be designed.

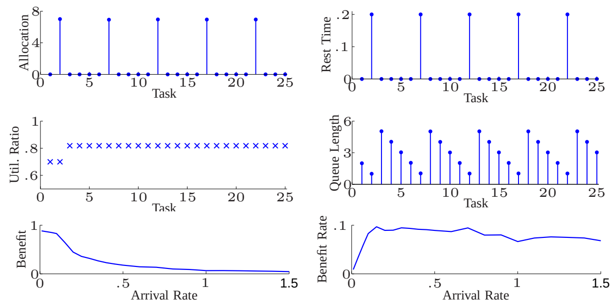


Fig. 4. Certainty Equivalent Solution. The top figures show the optimal allocation to each task and rest time after processing it, respectively. The middle figures show the optimal evolution of the expected queue length and the utilization ratio, respectively. The bottom figures show the benefit per unit task and benefit per unit time as a function of arrival rate, respectively.

Modified Receding Horizon Policy:

In the certainty-equivalent formulation, each task is equivalent and equal to the average task. However the information about the nature of the current task is available and should be incorporated to determine the optimal policies. We incorporate this information in the following way: Let $V^* : \mathbb{N} \times [0,1] \to \mathbb{R}$ be the value function associated with the optimization problem (4), and let $\hat{n}_{\ell+1}, \hat{x}_{\ell+1}$ be the expected values of the queue length and utilization ratio for a duration allocation $z_\ell t_\ell$ followed by rest time $z_\ell r_\ell$ at stage ℓ. Define $J_\ell : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \{0,1\} \to \mathbb{R}$ by

$$\begin{aligned}
J_\ell(t_\ell, r_\ell, z_\ell) = z_\ell(w_{\gamma_\ell} f_{\gamma_\ell}(t_\ell) - c_{\gamma_\ell} t_\ell - \bar{c}(n_\ell - 1)(t_\ell + r_\ell) \\
- \frac{\bar{c}\lambda(t_\ell + r_\ell)^2}{2}) + V^*(\hat{n}_{\ell+1}, \hat{x}_{\ell+1}),
\end{aligned}$$

where f_{γ_ℓ} is the sigmoid function and c_{γ_ℓ} is the penalty rate associated with task ℓ. Note that the function also incorporates the exact state n_ℓ at stage ℓ. The modified receding horizon policy at stage ℓ determines optimal allocations

$$\{t_\ell^*, r_\ell^*, z_\ell^*\} \in \text{argmax } J_\ell(t_\ell, r_\ell, z_\ell). \tag{5}$$

Example 3: For the same set of data as in Example 2, we determine the optimal policies for a sample evolution of the queue. The simulation results are shown in Fig. 5.

Handling Mandatory Tasks: Consider the case when some tasks arrive with a token of mandatory processing. Once a mandatory task arrives, we set the action space for the mandatory task to $z = 1$, where z is the action variable that determines whether the task is processed. Due to the

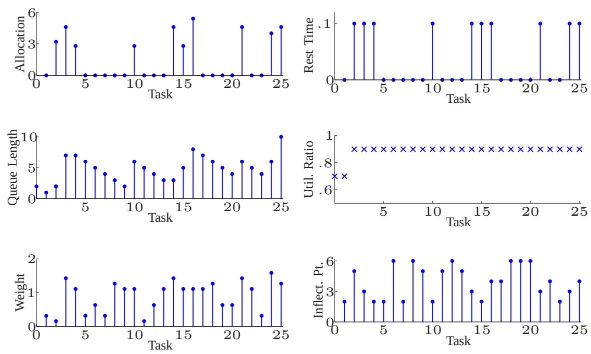
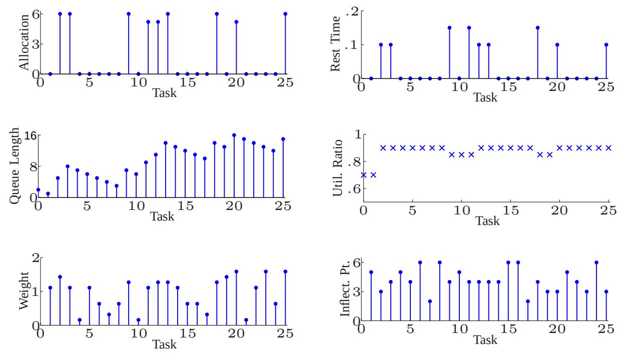
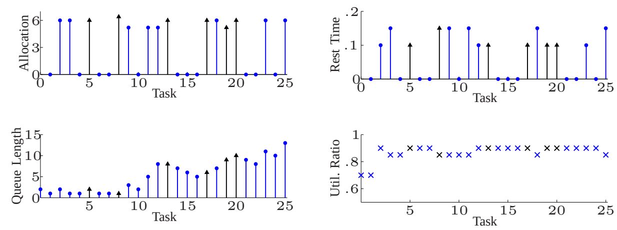


Fig. 5. The optimal policies for the sample evolution of a decision making queue with heterogeneous tasks. The top figures show the optimal duration allocation to each task and the rest time after processing each task, respectively. The middle figures show the expected evolution of the queue, and the evolution of utilization ratio dynamics, respectively. The bottom figures, respectively, show the weight assigned to each task, and the difficulty level of each task, captured through the inflection point of associated sigmoid function.



(a) No Mandatory Task



(b) Few Mandatory Tasks

Fig. 6. The optimal policies for an expected evolution of the decision making queue with heterogeneous tasks. (a) No task is mandatory. The optimal policies for decision making queue with heterogeneous tasks. The top figures show the optimal duration allocation to each task and the rest time after processing each task, respectively. The middle figures show the expected evolution of the queue, and the evolution of utilization ratio dynamics, respectively. The bottom figures, respectively, show the weight assigned to each task, and the difficulty level of each task, captured through the inflection point of associated sigmoid function. (b) The tasks with black triangular head are mandatory. The top figures show the optimal duration allocation to each task and the rest time after processing each task, respectively. The bottom figures show the expected evolution of the queue, and the evolution of utilization ratio dynamics, respectively.

change in the action space, the state at next stage may differ significantly from the state originally predicted by the optimal policy. The receding horizon solution offers a natural way of handling this issue by computing the optimal policy from the new state.

Example 4: For the same set of data as in Example 2, we consider the case when each sampled task is mandatory with probability 0.2. A comparison of the case with mandatory tasks with the standard case is shown in Fig. 6.

IV. Adaptive Attention Allocation

A. Spatial Quickest Detection

Consider an autonomous agent that surveys a set of regions $\mathcal{R} = \{1, \dots, n\}$. Let D be the Euclidean distance matrix between the regions. At each iteration, the autonomous agent visits the region $k \in \mathcal{R}$ with probability q_k, collects evidence, and sends it to a fusion center. Let the probability distribution function of the evidence be p_k^1 and p_k^0, respectively, conditioned on the presence of an anomaly at region k, and otherwise. Also, let the expected time the autonomous agent takes to collect, process, and transmit information at region k be T_k. The fusion center runs parallel CUSUM tests to detect any anomaly in any of the regions. For stationary region selection probabilities, that the worst case expected detection delay [15] at region k is

$$\mathbb{E}_{p_k^1}[T_\delta^k] = \frac{|e^{-\eta} + \eta - 1|}{q_k \mathcal{D}(p_k^1, p_k^0)} (\boldsymbol{q} \cdot \boldsymbol{T} + \boldsymbol{q} \cdot D\boldsymbol{q}),$$

where η is the CUSUM threshold, $\boldsymbol{q}$ and $\boldsymbol{T}$ are the n-vectors of region selection probabilities and expected processing times, respectively. The adaptive spatial quickest detection Algorithm 1 utilizes the current observations to adapt the region selection policy (see [15] for details.)

B. Spatial Quickest Detection with Human Input

We wish to run the spatial quickest detection algorithm on the decision made by the human operator. We assume that, conditioned on presence of an anomaly, the probability of correct decision at region k is determined by sigmoid function $f_k^1 : \mathbb{R}_{\geq 0} \to [0, 1]$ of the time allocated to the feed from region k. Similarly, let $f_k^0 : \mathbb{R}_{\geq 0} \to [0, 1]$ determines the probability of correct decision, conditioned on no anomaly being present at region k. Thus, conditioned on the presence of an anomaly, the decision of the human operator after allocating time t to a task from region k is sampled from a Bernoulli's distribution with probability of correct decision $f_k^1(t)$, and similar statement holds for the case when no anomaly is present. We now state an important property of the Kullback-Leibler divergence between two Bernoulli distributions.

Lemma 5 (Monotonicity of Kullback-Leibler Divergence): Consider two Bernoulli distributions with probability of success p_1 and p_2, respectively. The Kullback-Leibler divergence between the two distributions, conditioned on the first distribution, increases with increasing p_1 and decreases with increasing p_2, provided $p_1 > p_2$.

Algorithm 1 Adaptive Spatial Quickest Detection

1: **Given:** $\mathcal{R}$, D, $\boldsymbol{T}$, p_k^0, p_k^1, for each $k \in \mathcal{R}$
2: **Given:** prior probability of anomaly π_k^1, for each $k \in \mathcal{R}$
3: set $\Lambda_0^j = 0$, for all $j \in \mathcal{R}$, and $\tau = 0$
4: set $w_i = e^{\Lambda_\tau^i} / (\sum_{j \in \mathcal{R}} e^{\Lambda_\tau^j})$, for each $i \in \mathcal{R}$
5: obtain solution $\boldsymbol{q}^* = \operatorname{argmin}\{\sum_{k\in\mathcal{R}} w_k \mathbb{E}_{p_k^1}[T_\delta^k]\}$
6: at time $\tau \in \mathbb{N}$, select a random region $k \in \mathcal{R}$ according to the probability distribution $\boldsymbol{q}^*$
7: collect sample y_τ from region k
8: update the statistic at each region

$$\Lambda_\tau^j = \begin{cases} \left(\Lambda_{\tau-1}^k + \log \frac{p_k^1(y_\tau)}{p_k^0(y_\tau)}\right)^+, & \text{if } j = k, \\ \Lambda_{\tau-1}^j, & \text{if } j \in \mathcal{R} \setminus \{k\}. \end{cases}$$

% detect change if the threshold is crossed
9: **if** $\Lambda_\tau^k > \eta$, **then** declare anomaly detected at region k and set $\Lambda_\tau^k = 0$
10: continue to step 4:

Proof: The Kullback-Leibler divergence between the two distributions is

$$\mathcal{D}(p_1, p_2) = p_1 \log \frac{p_1}{p_2} + (1-p_1) \log \frac{1-p_1}{1-p_2}.$$

It can be verified that

$$\frac{\partial \mathcal{D}}{\partial p_1} = \log \frac{p_1(1-p_2)}{p_2(1-p_1)}, \text{ and}$$
$$\frac{\partial \mathcal{D}}{\partial p_2} = \frac{p_2 - p_1}{p_2(1-p_2)},$$

and if $p_1 > p_2$ the first derivative is positive, while the second derivative is negative. ■

We assume that if a task is processed, then the optimal policy allocates a time that is more than the expected time suggested by the Yerkes-Dodson law. We further assume that the minimum expected service time for a task from region k suggested by the Yerkes-Dodson law is $S_k^{\min}$. We also assume that if the a task is not processed, then the operator's decision is sampled from a Bernoulli's distribution with probability of success 0.5. Notice that under this assumption, the probability of success increases above and below 0.5 for any positive allocation to the task in presence and absence of an anomaly, respectively. Therefore, it follows from Lemma 5 that the Kullback-Leibler divergence between the two Bernoulli distributions will increase with the time allocated to the task. Also, let $\mathcal{D}_k^{\min}$ be the Kullback-Leibler divergence between two Bernoulli distributions with probability of success $f_k^1(S_k^{\min})$ and $f_k^0(S_k^{\min})$, respectively.

For the CUSUM with input from human operator, the expected processing time on an evidence also includes the time duration human allocates to it. It can be shown that the time duration allocated to an evidence from region k is upper bounded by $t_k^{\max} = \max\{t \in \mathbb{R}_{\geq 0} \mid {f_k^1}'(t) = c_k\}$. Thus, the expected processing time at region k is upper bounded by $T_k^{\text{upper}} = T_k + t_k^{\max}$.

Theorem 6 (Expected Sample Size): For the stationary region selection policy $\boldsymbol{q}$, and the CUSUM algorithm with input from human operator, conditioned on the presence of an anomaly at region k, the following statements hold:

i) the worst case expected sample size N_k^d for anomaly detection at region k satisfies
$$\mathbb{E}[N_k^d] \leq \frac{|e^{-\eta} + \eta - 1|}{q_k \mathcal{D}_k^{\min}};$$

ii) the worst case expected detection delay T_k^d at region k satisfies
$$\mathbb{E}[T_k^d] \leq (\boldsymbol{q} \cdot \boldsymbol{T}^{\text{upper}} + \boldsymbol{q} \cdot D\boldsymbol{q})\mathbb{E}[N_k^d],$$
where $\boldsymbol{T}^{\text{upper}} = [T_1^{\text{upper}}, \dots, T_n^{\text{upper}}]$.

Proof: In the spirit of the standard proof for the CUSUM algorithm [13], we consider the associated SPRT and determine the upper bounds on the expected sample size for it. Consider the CUSUM algorithm for region k, and let μ_k be the stopping time for the associated SPRT. Let the optimal duration allocation for these samples be $\{t_i\}_{i\in\{1,\dots,\mu_k\}}$. Let $\{\upsilon_i^k\}_{i\in\{1,\dots,\mu_k\}}$ be the log-likelihood ratios for decisions made by the operator and let $\{\xi_i^k\}_{i\in\{1,\dots,\mu_k\}}$ be the associated Kullback-Leibler divergences, conditioned on the presence of an anomaly at region k. Note that if the task at iteration i is not from region k, then $\upsilon_i^k = 0$. Consider the sequences $\{S_i^k\}_{i\in\{0,\dots,\mu_k\}}$, defined by $S_i^k = S_{i-1}^k + \upsilon_i^k - q_k \xi_i^k$, and $S_0^k = 0$. Also, let $\{\mathcal{F}_i^k\}_{i\in\{1,\dots,\mu_k\}}$ be the filtration defined by the sigma algebra of decisions and regions visited till iteration i. It is easy to verify that the sequence $\{S_i^k\}_{i\in\{0,\dots,\mu_k\}}$ is a martingale with respect to filtration $\{\mathcal{F}_i^k\}_{i\in\{1,\dots,\mu_k\}}$. It follows from the optional stopping theorem [5] that

$$E[S_{\mu_k}^k | \mathcal{F}_{\mu_k}] = E[S_0^\ell] = 0 \implies \sum_{i=1}^{\mu_k} \upsilon_i^k = q_k \mathbb{E}[\sum_{i=1}^{\mu_k} \xi_i^k].$$

It follows from Lemma 5 that if $\xi_i^k > 0$, then it is lower bounded by $\mathcal{D}_k^{\min}$. Therefore, it follows that

$$\sum_{i=1}^{\mu_k} \upsilon_i^k \geq q_k \mathcal{D}_k^{\min} \mathbb{E}[\mu_k] \implies \mathbb{E}[\mu_k] \leq \frac{1}{q_k \mathcal{D}_k^{\min}} \mathbb{E}[\sum_{i=1}^{\mu_k} \upsilon_i^k].$$

The remainder of the proof is similar to the standard proof for the CUSUM algorithm [13] that involves Wald's approximation, followed by relating CUSUM with the associated SPRT.

The second statement follows from Wald's identity (see [15] for details.) ■

Remark 7: The upper bounds obtained in Theorem 6 assume that each task is processed by the human operator. As exemplified in Section III, the optimal policy may not process each task, this may not be a valid assumption. In such cases, a task dropping factor may be introduced into the optimization problem, e.g., if β_k is the fraction of tasks from the region k that are processed, then the effective region selection probability becomes $\beta_k q_k$. Since, we are interested in determining the maximum arrival rate at which no task is dropped, we will assume that $\beta_k = 1$, for each $k \in \mathcal{R}$ in the following discussion. □

To determine the optimal stationary policy, we state the following optimization problem:

$$\underset{\boldsymbol{q} \in \Delta_{n-1}}{\text{minimize}} \quad \sum_{k \in \mathcal{R}} \frac{w_k |e^{-\eta} + \eta - 1|}{q_k \mathcal{D}_k^{\min}} (\boldsymbol{q} \cdot \boldsymbol{T}^{\text{upper}} + \boldsymbol{q} \cdot D\boldsymbol{q}), \quad (6)$$

where w_k is the weight associated with region k and Δ_{n-1} is the $n-1$ dimensional probability simplex.

C. Simultaneous quickest detection and duration allocation

We propose a simultaneous quickest detection and duration allocation procedure in Algorithm 2. The algorithm, at each iteration, determines the optimal allocation to the current task in the queue using the receding horizon strategy discussed in Section III. If the current task is from region k, the performance function of that task is determined by weighed sum of the performance functions in presence/ absence of an anomaly. The support system runs parallel CUSUM tests on the decisions made by the human operator. The CUSUM statistics are used to adapt the weights for each region and the optimal region selection policy is determined by solving optimization problem (6). Note that the processing and collection of the evidence are asynchronous, and the region selection policy remains the same while an evidence is processed. Once, the region selection policy is adapted, the autonomous agent first finishes its current assignment and then implements the adapted policy at the next iteration.

Algorithm 2 Adaptive Attention Allocation

1: **Given:** $\mathcal{R}$, D, $\boldsymbol{T}$, η
2: **Given:** performance functions f_k^0, f_k^1, for each $k \in \mathcal{R}$
3: **Given:** prior probability of anomaly π_k^1, for each $k \in \mathcal{R}$
4: set $\Lambda_0^k = 0$, for all $k \in \mathcal{R}$, and $\tau = 0$
5: set $w_k = e^{\Lambda_\tau^k} / (\sum_{j \in \mathcal{R}} e^{\Lambda_\tau^j})$, for each $k \in \mathcal{R}$; $\tau = \tau + 1$
6: set $f_k = w_k f_k^1 + (1 - w_k) f_k^0$
7: obtain solution $\boldsymbol{q}^*$ of the optimization problem (6)
8: solve optimization problem (5) with $p = \boldsymbol{q}^*$ and performance functions f_k to obtain allocation t_τ^*
9: collect evidence from regions sampled according to the probability distribution $\boldsymbol{q}^*$
10: collect operator decision d_τ on the current task from region k_τ
11: update the CUSUM statistic at each region k_τ

$$\Lambda_\tau^j = \begin{cases} \left(\Lambda_{\tau-1}^j + \log \frac{\mathbb{P}(d_\tau | \text{anomaly }, t_\tau^*, k_\tau)}{\mathbb{P}(d_\tau | \text{no anomaly }, t_\tau^*, k_\tau)}\right)^+, & \text{if } j = k_\tau, \\ \Lambda_{\tau-1}^j, & \text{if } j \in \mathcal{R} \setminus \{k_\tau\}. \end{cases}$$

% detect change if the threshold is crossed
12: **if** $\Lambda_\tau^k > \eta$, **then** declare anomaly detected at region k and set $\Lambda_\tau^k = 0$
13: continue to step `5:`

Example 8: We consider 4 regions surveyed by an autonomous agent. The probability of the correct decision at region k evolves as sigmoid functions $f_k^0 = 1/(1+e^{-a_k^0 t + b_k^0})$ and $f_k^1 = 1/(1 + e^{-a_k^1 t + b_k^1})$, respectively, in absence and presence of any anomaly. We picked the array of sigmoid parameters as $a^0 = [1 \ 2 \ 1 \ 1]$, $b^0 = [2 \ 5 \ 1.5 \ 3]$, $a^1 = [1 \ 2 \ 1 \ 1]$, and $b^1 = [5 \ 8 \ 3 \ 5]$. The Euclidean distance matrix between the regions is

$$D = \begin{bmatrix} 0 & 25 & 9 & 17 \\ 25 & 0 & 18 & 28 \\ 9 & 18 & 0 & 14 \\ 17 & 28 & 14 & 0 \end{bmatrix}.$$

Let processing time at each region be unity and the arrival rate of the tasks be 0.3. Let the Yerkes-Dodson curve and the latency penalty be the same as in Example 3. The anomalies appear at regions 3, 1, 2, and 4, when collecting task number 2, 10, 15, and 20, respectively. The simulation results are shown in Fig. 7.

Region 3 is closest to all other regions while region 2 is farthest from each of the other regions. For detection of an anomaly, the evidence obtained from region 3 are easiest to process, followed by region 2, region 1, and region 4. The difficulty of the task is captured by the inflection point of the associated sigmoid function. The geometry of the region, difficulty of the task, and probability of anomaly in that region jointly determine the region selection probability, e.g., when the CUSUM statistics at each region is zero, there are two opposing effects on selection probability of region 3. The convenient location of region 3 makes its selection probability high, while the easy nature of the task decreases its selection probability. The ease of task decreases the selection probability, since less samples are required to make decision on easy tasks. One may notice that as the CUSUM statistic increases at a region, its selection probability increases accordingly.

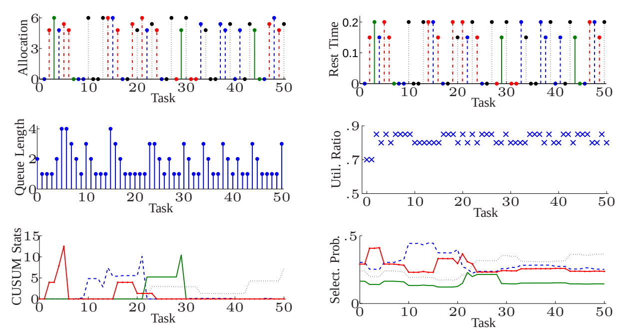


Fig. 7. An autonomous agent surveys 4 regions depicted by dashed blue, solid green, solid red with dots and dotted black lines, respectively. The top figures show the duration allocation to each task and rest time following it, respectively. The color scheme represents the region from which the task has come. The middle figures show the evolution of the expected queue length and the utilization ratio, respectively. The bottom figures show evolution of the CUSUM statistics and region selection probabilities, respectively. Note that the regions with higher value of CUSUM statistics are chosen with higher probability. The region selection probability is also a function of location of the region, and the difficulty of task coming from that region. Once an anomaly is detected the associated CUSUM statistic is set to zero.

Remark 9 (Poisson Approximation): In the simultaneous quickest detection and duration allocation formulation, we assumed that the arrival process of the decision making queue is Poisson. In general, this is not true and the arrival process is a general renewal process. For n regions, the arrival process can be thought of summation of n^2 renewal processes with mean arrival time $q_i q_j d_{ij} + T_i, i, j \in \{1, \ldots, n\}$. For large n, under certain conditions, this could be approximated by a Poisson process [7]. For the cases, where this approximation does not hold, the certainty-equivalent solution can be modified with an appropriate expected evolution of the queue. □

V. Conclusions and Future Directions

In this paper we have studied the problem of simultaneous information aggregation and processing in human machine systems. We incorporated human operator performance models related to error rates and situational awareness for determining the optimal duration allocation policies for the operators. We also demonstrated how the human input can be thought of as a "sensor measurement" and the decisions made by the operator can be used to adapt sensor selection policy. We proposed to run spatial quickest detection algorithm on operator inputs and derived an expression for detection delay taking into account time varying operator performance in addition to sensor constraints such as travel time. This expression was used for optimizing sensor selection policy such that the information from the region with maximum probability of being anomalous is chosen with high probability. We demonstrated our methodology in a distributed surveillance task where the selected information is send to the operator, the operator processes the information based on optimal attention allocation scheme, and the decisions of the operator are used to adapt sensor selection policy for quickest detection.

In future we plan to experimentally validate the framework proposed in this paper with human operator in loop supervising a camera network for identifying anomalous behavior. Such a study would be necessary for testing the applicability of operator performance models used in this paper. Additionally, there can be wide variability in operator performance model parameters. Developing robust optimization schemes which takes operator model uncertainty into account is another important direction of future work. In this paper, it was assumed that all the tasks which the support system selected, will be serviced by the operator. Another possibility is to give operators ability to requeue tasks which has been shown to improve overall performance [2]. Developing stable attention allocation schemes with requeuing poses additional challenges and needs to be further investigated.

VI. Acknowledgment

The authors thanks Luca Bertuccelli and Andrzej Banaszuk for their feedback and suggestions.

References

[1] L. F. Bertuccelli, N. W. M. Beckers, and M. L. Cummings. Developing operator models for UAV search scheduling. In *AIAA Conf. on Guidance, Navigation and Control*, Toronto, Canada, August 2010.

[2] L. F. Bertuccelli, N. Pellegrino, and M. L. Cummings. Choice modeling of relook tasks for UAV search missions. In *American Control Conference*, pages 2410–2415, Baltimore, MD, USA, June 2010.

[3] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen. The physics of optimal decision making: A formal analysis of performance in two-alternative forced choice tasks. *Psychological Review*, 113(4):700–765, 2006.

[4] F. Bullo, J. Cortés, and B. Piccoli. Special issue on control and optimization in cooperative networks. *SIAM Journal on Control and Optimization*, 48(1):vii–vii, 2009.

[5] R. Durrett. *Probability: theory and examples*. Cambridge University Press, 2010.

[6] C. Nehme, B. Mekdeci, J. W. Crandall, and M. L. Cummings. The impact of heterogeneity on operator performance in futuristic unmanned vehicle systems. *The International C2 Journal*, 2(2):1–30, 2008.

[7] G. F. Newell. Approximations for superposition arrival processes in queues. *Management Science*, pages 623–632, 1984.

[8] R. W. Pew. The speed-accuracy operating characteristic. *Acta Psychologica*, 30:16–26, 1969.

[9] C. A. Rabbath, C. Y. Su, and A. Tsourdos. Guest editorial introduction to the special issue on multivehicle systems cooperative control with application. *IEEE Transactions on Control Systems Technology*, 15(4):599–600, July 2007.

[10] K. Savla and E. Frazzoli. A dynamical queue approach to intelligent task management for human operators. *Proceedings of the IEEE*, 100(3):672–686, 2012.

[11] K. Savla, C. Nehme, T. Temple, and E. Frazzoli. On efficient cooperative strategies between UAVs and humans in a dynamic environment. In *AIAA Conf. on Guidance, Navigation and Control*, Honolulu, Hawaii, 2008.

[12] K. Savla, T. Temple, and E. Frazzoli. Human-in-the-loop vehicle routing policies for dynamic environments. In *IEEE Conf. on Decision and Control*, pages 1145–1150, Cancún, México, December 2008.

[13] D. Siegmund. *Sequential Analysis: Tests and Confidence Intervals*. Springer, 1985.

[14] V. Srivastava and F. Bullo. Hybrid combinatorial optimization: Sample problems and algorithms. In *IEEE Conf. on Decision and Control and European Control Conference*, pages 7212–7217, Orlando, FL, USA, December 2011.

[15] V. Srivastava and F. Bullo. Stochastic surveillance strategies for spatial quickest detection. In *IEEE Conf. on Decision and Control and European Control Conference*, pages 83–88, Orlando, FL, USA, December 2011.

[16] V. Srivastava, R. Carli, F. Bullo, and C. Langbort. Task release control for decision making queues. In *American Control Conference*, pages 1855–1860, San Francisco, CA, USA, June 2011.

[17] V. Srivastava, K. Plarre, and F. Bullo. Adaptive sensor selection in sequential hypothesis testing. In *IEEE Conf. on Decision and Control and European Control Conference*, pages 6284–6289, Orlando, FL, USA, December 2011.

[18] V. Srivastava, K. Plarre, and F. Bullo. Randomized sensor selection in sequential hypothesis testing. *IEEE Transactions on Signal Processing*, 59(5):2342–2354, 2011.

[19] R. M. Yerkes and J. D. Dodson. The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology*, 18(5):459–482, 1908.