Human-in-the-loop vehicle routing policies for dynamic environments

Ketan Savla Tom Temple Emilio Frazzoli

***Abstract*— In this paper we design coordination policies for a routing problem requiring human-assisted classification of targets through analysis of information gathered on-site by autonomous vehicles. More precisely, we consider the following problem: Targets are generated according to a spatio-temporal Poisson process, uniformly in a region of interest. It is desired to classify targets as *friends* or *foes.* In order to enable human operators to classify a target, one of the vehicles needs to travel to the target's location and gather sufficient information. In other words, the autonomous vehicles provide access to on-site information, and the human operator provide the judgment capabilities necessary to process such information. The objective of our analysis is to design joint motion coordination and operator scheduling policies that minimize the expected time needed to classify a target after its appearance. In addition, we analyze how the achievable system performance depends on the number of autonomous vehicles and of human operators. We present novel coordination policies between the vehicles and operators and compare the performance of these policies with respect to asymptotic performance bounds.**

I. INTRODUCTION

One of the prototypical missions involving Uninhabited Aerial Vehicles (UAVs), e.g., in environmental monitoring, security, or military setting, is wide-area surveillance. In such a mission, low-altitude UAVs must provide coverage of a region and investigate events of interest as they manifest themselves. In particular, we are interested in cases in which close-range information is required on targets detected by high-altitude aircraft, spacecraft, or ground spotters, and the UAVs must proceed to the locations to provide a target-specific service, possibly enabling further action under direct operator supervision. Possible services can include tasks like gathering on-site information (such as video or still images), target classification, localization, etc.

The vehicle routing formulation that we adopt in this paper was originally proposed in the name of the Dynamic Traveling Repairperson Problem (DTRP) in [1]. Variations of problems falling in that class have been studied by numerous researchers recently, e.g., see [2], [3], [4]. However, all of the prior work assumes that UAVs are perfectly autonomous and do not require any supervision from a human operator at any time during the mission. Even though one can foresee UAVs to have completely automated guidance and navigation modules in the near future, the role of human operators will be indispensable when the servicing of targets involves on-site decision making that require high level of cognitive capabilities provided only by a human. In particular, the role of a human operator becomes critical in decision making processes where the penalty for taking a wrong decision is substantial. For instance, incorrectly identifying an unknown object in a surveillance mission can have dire consequences.

In this paper, we consider a binary decision task, loosely inspired by the US Air Force COUNTER program [5], where it is desired to classify these targets as, e.g., *friends* or *foes.* In order to enable human operators to classify a target, one of the vehicles needs to travel to the target location and gather sufficient information. In other words, the autonomous vehicles provide access to on-site information, and the human operator provide the judgment capabilities necessary to process such information. The objective of our analysis is to design joint motion coordination and operator scheduling policies that minimize the expected time needed to classify a target after its appearance; in addition, we want to analyze how the achievable system performance depends on the number of autonomous vehicles and of human operators.

The inspiration for the problem setup of this paper can be traced to the field of *human supervisory control*, e.g., see [6], [7], where the idea is to allocate mundane tasks to the automation system and leave complex tasks for human supervisors. The scenario that we consider in this paper is also related to the one considered in [8]. The architecture discussed in this paper for the classification task is also reminiscent of a simple queuing network [9] with two components in series, first of which is a spatial queue with vehicle as servers and the second is a *conventional* queue with human operator as servers. The human decision-making process suitable for the binary task considered in this paper has been studied in the form of a two-alternative forced-choice task, e.g., see [10], [11], where the task for a human to choose between two alternatives, under time pressure and with uncertain information has been studied.

The contributions of this paper are threefold. First, we formulate a novel vehicle routing problem requiring human decision making and provide a framework for studying such problems. Second, we propose novel joint coordination and task allocation strategies between vehicles and human operators for the dynamic target classification task and third, we provide an analysis for the dependence of the classification time on the number of vehicles and human operators for meaningful scenarios.

Due to space limitations, the proofs are either briefly sketched or completely omitted and will be presented in a future extended version.

II. PRELIMINARIES

A. Notations

Let $\mathcal{Q} \subset \mathbb{R}^2$ be a convex, compact domain on the plane, with non-empty interior; we will refer to $\mathcal{Q}$ as the *environment*. Let $\mathcal{A}$ be the area of $\mathcal{Q}$. A Poisson process generates

Ketan Savla, Tom Temple and Emilio Frazzoli are with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology, {ksavla,ttemple,frazzoli}@mit.edu

targets over time, which are associated to points in $\mathcal{Q}$, with finite time intensity $\lambda > 0$. Furthermore, the points are sampled from an absolutely continuous spatial distribution described by the density function $\varphi : \mathcal{Q} \to \mathbb{R}_+$. The spatial density function φ is normalized such that $\int_{\mathcal{Q}} \varphi(q)\, dq = 1$. In this paper we concentrate on the case in which the targets are generated uniformly in $\mathcal{Q}$, i.e., $\varphi(q) = 1/\mathcal{A}$.

This process can be thought of as a collection of functions $\{\mathcal{P} : \overline{\mathbb{R}}_+ \to 2^{\mathcal{Q}}$ such that, for any $t > 0$, $\mathcal{P}(t)$ is a random collection of points in $\mathcal{Q}$, representing the targets generated in the time interval $[0, t)$, and such that

- the total number of targets generated in two disjoint time-space regions are *independent* random variables;
- the *expected* number of targets generated in a measurable region $S \subseteq \mathcal{Q}$ during a time interval of length Δt is given by:

$$\mathrm{E}[\mathrm{card}\left((\mathcal{P}(t+\Delta t) - \mathcal{P}(t)) \cap \mathcal{S}\right)] = \lambda \Delta t \cdot \frac{\mathrm{Area}(\mathcal{S})}{\mathcal{A}}.$$

The dynamically arriving targets are assumed to be homogeneous. These dynamically arriving targets are to be classified as *friends* or *foes*. The classification task is to be carried out by a team composed of m autonomous vehicles and n remotely located human operators. The team is clearly heterogeneous, and each class of agent provides a unique capability: the vehicles provide access to on-site information, and the human operators provide the judgment capabilities necessary to process such information. In order to collect information about a target, one of the m vehicles needs to travel to the target location; the amount of information gathered is proportional to the amount of on-site time spent by the vehicle at the target location. The vehicles move with speed v, are identical and have unlimited capacity.

B. Human Decision Making

We now describe a state-space model for human decision making in the context of the classification task of this paper. We associate a real valued *classification state* $x_i(t)$ with the classification status of target i at time t, and let the evolution of the classification state in presence of information be described by a stochastic differential equation which is written in *Langevin form* as follows

$$\dot{x}_i(t) = f(x_i(t), \mathcal{T}_i, \mathcal{C}) \quad x_i(0) = 0 \quad \text{for every target } i,$$

where $\mathcal{C}$ denotes the set of cognitive parameters associated with the decision making process for the human operators. We assume that $\dot{x}_i(t) = 0$ in absence of any external information, i.e., the operator does not engage in any cognitive task related to the classification of a target in absence of any information about it. The observable[1], $y_i(t) \in \{0, 1\}$ is the actual classification status of target, where $y_i(t) = 1$ implies that the classification of target i has been completed by time t and $y_i(t) = 0$ implies otherwise. The relation between $y_i(t)$ and $x_i(t)$ is described as follows: an operator sets $y_i(t) = 1$ when $x_i(t)$ crosses a threshold, e.g., it can classify the target i as *friend* when x_i crosses 1 for the first time or classify it as *foe* when x_i crosses -1 for the first time. Even though the classification state $x_i(t)$ is not observable by the system, we assume that it can be shared perfectly among the operators. This is possible, for example, by *tagging* a target with information representative of its classification state which can be *deciphered* by any operator to extract the classification state. This assumption allows the possibility of an operator finishing or continuing classification of a target that was left incomplete (for instance, due to lack of sufficient information at that time) without any loss of information from the past.

[1]The classification status of the target is observable in the sense that an operator can notify the system as soon as it thinks that it has classified a target, for example by pressing appropriate button on its console, etc.

The amount of information required by the human operators for classification of a target is a random variable, which we denote by s. Let $\overline{s}$ and σ_s^2 be the mean and the variance of this random variable. We assume that these quantity are independent of the type of the target, i.e., it being a friend or a foe, and that the SS has knowledge about the distribution of s, e.g., from data collected from previous missions, etc.

Example: Drift Diffusion Models for Decision Making

We now give an example for a human decision making model which fits our requirements as stated in the earlier section. In the following, we briefly describe the Drift Diffusion Models (DDM) for modeling decision making process of humans when faced with choosing between two alternatives, under time pressure and uncertain information. In the following, we have adopted the description of DDM from [11] to fit into our framework.

In DDM, one accumulates the difference between the amounts of evidence supporting the two hypotheses, which in our case are classifying the target as a friend or a foe. Let $x_i(\tau)$ denote the accumulated value of this difference on the basis of τ amount of external information and assume that $x_i = 0$ represents equality in the amounts of integrated evidence. In the pure DDM, one starts with unbiased evidence, i.e., $x_i(0) = 0$, and accumulate according to

$$dx_i = A d\tau + c dW, \quad x_i(0) = 0,$$

where dx_i denotes the change in dx_i as a result of additional $d\tau$ amount of information, which is comprised of two parts: the constant drift $Ad\tau$ represents the average increase in evidence supporting the correct choice per unit amount of information. The second term, cdW, represents white noise, which is Gaussian distributed with mean 0 and variance $c^2 d\tau$. Hence, x_i grows at rate A on average, but solutions also diffuse due to accumulation of noise. In the free-response paradigm, the decision is made when x_i reaches one of the two fixed thresholds, each corresponding to a hypothesis. If both the alternatives are equally probable, the thresholds are symmetric. The value of τ at the *first cross-over point*, i.e, the point at which the x_i crosses one of the two thresholds for the first time represents the total amount of information required for target classification.

C. Support System Architecture

We propose an automation Support System (SS) to facilitate the cooperation between vehicles and the human operators. The overall architecture of the SS is depicted in Figure 1, where we show the interaction between the

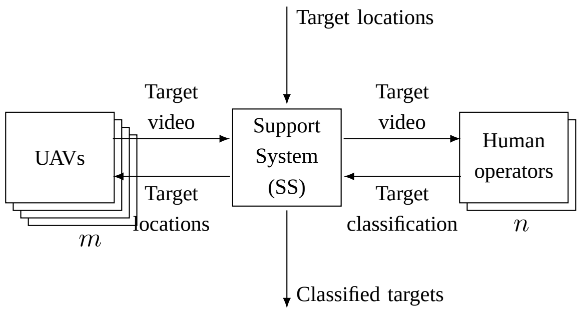

Fig. 1. Support System Architecture

SS, vehicles and the operators for the scenario where the target information is in the form of video. The detailed interaction between the vehicles, human operators and the SS is described in the following.

The target locations are assumed to be transmitted to the SS instantaneously with their generation. Before the start of the mission, the SS assigns loitering stations to the vehicles, i.e., locations where the vehicles wait when they do not have any targets alloted to them. During the course of the mission, the SS allocates target locations in batches to each vehicle. In addition, it also specifies the *mode of information collection* for each of those target locations. There are two modes of information collection: *online* and *offline*. Roughly speaking, the online mode involves collecting information under the direct supervision of an operator and the offline mode implies autonomous information collection by the vehicle which would be later used by the operator. While specifying offline mode for a particular target, the SS also specifies the amount of time the vehicle should spend collecting information at that location. Note that the SS could, on recommendation from the operator, also add further specifications in the offline mode, e.g., collecting video from a particular view angle, etc. However, we assume that, for a given amount of time over which sensor data are collected, the *information content* of the data is independent of other factors.

Every vehicle deals with the batches in the order that they are received. For each batch of target locations assigned to it, the vehicle determines the shortest tour through them starting with its current location. We assume that each vehicle navigates autonomously along straight line paths between two locations. Hence, the shortest tour for the vehicle over a collection of target locations would be the Euclidean Traveling Salesperson tour over those target locations. For every target, once the vehicle has reached its location, it does the following:

(i) if the mode of information collection for that target is online, the vehicle issues an *attention request* to the SS and waits until it gets attention of a human operator. Once it gets attention from a human operator, it collects information under the supervision of the operator until asked by the SS to stop,
(ii) if the mode of information collection for that target is offline, the vehicle collects information for the specified amount and concurrently transmits that information to the SS.

The SS schedules these attention requests and offline information from the vehicles to the human operators. In addition, the SS also keeps track of the *notes* from earlier classification rounds. While allocating a human operator to a particular target, it also supplies that operator with the notes (if any) of that target.

For every task assigned by the SS, a human operator deciphers the notes for that target (if any) and does the following:

(i) If the task is attending an attention request from a vehicle, the operator processes the information online, i.e., as it is being collected by the vehicle until it makes a classification, at which point it notifies the SS about the classification,
(ii) If the task involves offline processing of information, the operator notifies the SS if he/she has classified the target after or at any time while processing the information. We assume that the amount of time required to process information by an operator is same as the amount of time required to collect it.

The classification of a target is finished as soon as an operator notifies the SS about it. If an operator has not classified a target after processing a given segment of offline information, that target is retained as unclassified and it needs additional information. As mentioned before, the UAVs navigate autonomously and do no require assistance from the SS or human operators for the same. In summary, a typical control policy for the SS consists of:

(i) assigning loitering locations for the vehicles, forming batches of unclassified targets and allocating them to the vehicles while specifying the mode of information collection at those targets, and
(ii) scheduling and allocating incoming attention requests and offline information from the vehicles to the operators

In the course of this paper, we would be interested in designing control policies for the SS that yield a desired quality of service.

D. Problem Formulation

Let the set $\mathcal{D}(t) \subset \mathcal{Q}$ represent the *demand*, i.e., the targets whose classification is outstanding at time t, and we define $n(t) = \mathrm{card}(\mathcal{D}(t))$. Let π be a typical control policy of the SS. The objective is the design of a policy that allows the timely classification of the targets. A policy π for the SS is said to be *stabilizing* if, under its effect, the expected number of outstanding unclassified targets, $\overline{n}_\pi$, does not diverge over time, i.e., if

$$\overline{n}_\pi = \lim_{t\to\infty} \mathrm{E}\left[n(t) : \mathrm{SS\ \ executes\ } \pi\right] < \infty.$$

Intuitively, a policy is stabilizing if it facilitates cooperation between the team of vehicles and the human operators in such a way that the rate of classification of targets is, on average, at least as fast as the rate at which new targets are generated. Let T_j be the time elapsed between the generation of the j-th target, and the time it is classified. If the system is stable, then the following balance equation (also known as Little's formula [12]) holds: $\overline{n}_\pi = \lambda \overline{T}_\pi$, where $\overline{T}_\pi := \lim_{j\to\infty} \mathrm{E}[T_j]$ is the system time under policy π, i.e., the expected time a target must wait before being classified, given that the SS implements the policy π. Note that the system time $\overline{T}_\pi$ can be thought of as a measure of the quality of service collectively provided by the SS.

At this point we can finally state our problem: we wish to design a policy for the SS that is (i) stabilizing, and (ii) yields a quality of service (i.e., system time) achieving, or approximating, the theoretical optimal performance given by

$$\overline{T}_{\rm opt} = \inf_{\pi \text{ stabilizing}} \overline{T}_\pi.$$

In the following, we are interested in designing computationally efficient control policies that are within a constant factor of the optimal, i.e., policies π such that $\overline{T}_\pi \leq \kappa \overline{T}_{\rm opt}$ for some constant κ.

III. Lower bounds

In this section, we provide various lower bounds on the average classification time. These results will then be used as a standard to compare performances of the algorithms for asymptotic working conditions. To define these asymptotic working conditions for the vehicles and the operators we define the vehicle load factor $\rho_v := \lambda\overline{s}/m$ to be the fraction of time spent by the vehicles collecting on-site information about the targets. Similarly, we define $\rho_h := \lambda\overline{s}/n$ to be the fraction of time spent by the human operators processing the information provided by the vehicles. Using these definitions, one can immediately see that that a necessary condition for stability is that $\rho_v \leq 1$ *and* $\rho_h \leq 1$. We would say that human operators are working under light load conditions if ρ_h is close to zero and that they are working under heavy load conditions when ρ_h is close to one. We define light and heavy load conditions for vehicles in a similar fashion.

Before stating our first lower bound, we briefly review a problem from geometric optimization. Given a set $\mathcal{Q} \subset \mathbb{R}^2$ and a set of points $p = \{p_1, p_2, \ldots, p_m\} \in \mathcal{Q}^m$, the expected distance between a random point q, sampled from a uniform distribution over $\mathcal{Q}$, and the closest point in p is given by

$$\begin{aligned}\mathcal{H}_m(p,\mathcal{Q}) :=\ & \mathrm{E}\left[\min_{i\in\{1,\ldots,m\}} \|p_i - q\|\right] \\ &= \sum_{i=1}^{m} \int_{\mathcal{V}_i(p)} \|p_i - q\| dq,\end{aligned}$$

where $\mathcal{V}(p)$ is the Voronoi partition of the set $\mathcal{Q}$ discussed earlier. The function $\mathcal{H}_m$ is known in the locational optimization literature as the *continuous multi-median function*; see, for example, [13] and references therein.

The m-median of the set $\mathcal{Q}$ is the global minimizer

$$p^*_m(\mathcal{Q}) = \operatorname*{argmin}_{p\in\mathcal{Q}^m} \mathcal{H}_m(p,\mathcal{Q}).$$

We let $\mathcal{H}^*_m(\mathcal{Q}) = \mathcal{H}_m(p^*_m(\mathcal{Q}),\mathcal{Q})$ be the global minimum of $\mathcal{H}_m$. We will not pursue the issue of computation of the m-median and of the corresponding $\mathcal{H}^*_m(\mathcal{Q})$, but will assume that these values are available.

We now state our first lower bound on the average classification time.

Theorem 3.1: The classification time $\overline{T}_{\rm opt}$ satisfies the following lower bound.

$$\overline{T}_{\rm opt} \geq \frac{\mathcal{H}^*_m(\mathcal{Q})}{v} + \overline{s}.$$

The lower bound in Theorem 3.1 is obtained by assuming that, at the time of generation of every target: (i) there are no other outstanding targets (ii) the vehicles know the exact time of generation of the target but do not know its location; they position themselves to minimize the expected distance from the target location to the nearest vehicle, and (iii) the human operators are idle. Theorem 3.1 is useful to compare performances of proposed policies when the vehicles and the human operators are performing under light load conditions. Next, we state a lower bound which is inspired by a lower bound on the system time for the multi-vehicle DTRP in [1].

Theorem 3.2: There exists a constant $\bar{\gamma} \approx 0.07$ such that the classification time satisfies the following lower bound.

$$\overline{T}_{\rm opt} \geq \bar{\gamma}\frac{\lambda\mathcal{A}}{m^2 v^2 (1-\rho_v)^2} - \frac{\overline{s}(1-2\rho_v)}{2\rho_v}.$$

The lower bound in Theorem 3.2 is obtained by assuming that a human operator is always available to process information. We then use lower bound for the system time for the DTRP from [1]. Theorem 3.2 is useful to compare performances of proposed policies when the human operators are performing under light load conditions and the vehicles are working under high load conditions. Finally, we state a lower bound on the classification time for the case when the human operators are working under heavy load conditions.

Theorem 3.3: The classification time satisfies the following lower bound under heavy load conditions.

$$\overline{T}_{\rm opt} \geq \frac{\lambda\sigma_s^2}{2n^2(1-\rho_h)} + \overline{s}, \text{ as } \rho_h \to 1.$$

This lower bound is obtained by assuming that, at the time of generation of every target: (i) there is at least one vehicle which is idle and is present at the target location, (ii) moreover, this vehicle knows the exact amount of information required by a human operator to classify that target, and (iii) the vehicle acquires that information instantaneously. These assumptions are then used results along with results for the system time for a G1/G/n queue [9] to arrive at the lower bound.

Remark 3.4: We believe that the lower bound obtained in Theorem 3.3 is overly conservative since it is obtained by assuming that the target information does not incur any delay from the moment of the target generation to the moment it is broadcast to the SS and that the inter arrival time of the target information to SS has zero variance. We conjecture that one can obtain a tighter lower bound by proper consideration of these two factors in the classification time:

$$\overline{T}_{\rm opt} \geq \frac{\lambda(c/\lambda^2 + \sigma_s^2/n^2)}{2(1-\rho_h)} + \overline{s}, \text{ as } \rho_h \to 1,$$

for some $c < 1$.

After stating numerous lower bounds in this section, we now proceed to the next section where we design control policies for the SS and compare their performances to these lower bounds.

IV. Algorithms and upper bounds

The first policy, the Median Based Assignment Policy, that we propose assigns a *depot* for each vehicle and schedules the operators in such a way that every target is classified in the online mode. The policy is formally described as follows.

The Median Based (MB) Assignment Policy

The loitering locations for the vehicles are the m median locations for the region $\mathcal{Q}$. The SS forms batches of one target each and assigns it to the vehicle whose loitering location is closest to the target location. The mode of information collection is set to online for all the targets. The incoming attention requests from the vehicles are assigned to the human operator on a first come first serve basis.

Let $\overline{T}_{\text{MB}}$ be the system time obtained by implementing the median based assignment policy. We now analyze the performance of this policy first for the case when both the vehicles and the human operators are operating under light load conditions.

Theorem 4.1: The classification time obtained by implementation of the MB policy satisfies the following upper bound under light load conditions.

$$\overline{T}_{\text{MB}} = \frac{\mathcal{H}^*_m(\mathcal{Q})}{v} + \overline{s} \quad \text{as } \rho_v \to 0 \text{ and } \rho_h \to 0.$$

The proof of Theorem 4.1 is obtained by realizing that for a given m, n, $\rho_h \to 0$ implies $\lambda \to 0$. Under these conditions, there exists a finite time when all the vehicles are positioned at their depots and all the human operators are idle.

Remark 4.2: (i) Theorem 3.1 and Theorem 4.1 show that the MB policy gives the optimal performance when the vehicles as well as the human operators are operating under light load conditions.

(ii) Theorem 4.1 also implies that when the rate of generation of targets is so low that the vehicles and the vehicles and human operators are *idle* most of the time, adding or removing some of the human operators would not have any effect on the classification time as long as a human operator is always available when required. However, it is interesting to note that even when the vehicles are idle most of the time, the average classification time can be affected by changing the number of vehicles. This is because the term $\mathcal{H}^*_m(\mathcal{Q})$ in the expression of the average classification time in Theorem 4.1 is of the order $1/\sqrt{m}$.

In the last scenario, we observed that when the human operators are idle most of the time, the classification time is independent of the number of human operators.

We now propose a policy, the Median Team Based (MTB) assignment policy. This policy is an adaptation of the MB policy and is better suited for scenarios where the vehicles are under light load and the operators are under heavy load. For simplicity of exposition, we let m to be an integral multiple of n.

The Median Team Based (MTB) Assignment Policy

Form teams of vehicles with n vehicles in each team. All the vehicles in the same team share the same loitering location. The loitering locations for the teams are the m/n median locations of $\mathcal{Q}$. The SS forms batches of one target each and assigns it to the team whose loitering location is closest to the target location. Among the vehicles in the team, the target is assigned to an available vehicle that is closest to the target location. The mode of information collection is set to online for all the targets. The incoming attention requests from the vehicles are scheduled to match the order in which targets were generated and they are assigned to the human operators on a first come first serve basis.

Let $\overline{T}_{\text{MTB}}$ be the system time obtained by implementing the MTB policy. The following theorem states an upper bound on $\overline{T}_{\text{MTB}}$.

Theorem 4.3: When m is an integral multiple of n, the classification time given by the MTB policy is upper bounded as follows:

$$\overline{T}_{\text{MTB}} \leq \frac{\lambda(1/\lambda^2 + \sigma_s^2/n^2)}{2(1-\rho_h)} + 2\mathcal{H}^*_{m/n}(\mathcal{Q}) + \overline{s}.$$

Remark 4.4: (i) Although we assumed m to be an integral multiple of n, the policy can be implemented for any values of n and m and one could state a result similar to Theorem 4.3 at the expense of additional terms.

(ii) Theorem 4.3 implies that, when $m/n \to \infty$ (i.e., when the number of vehicles per operator is large), the upper bound on the classification time is of the order $(1-\rho_h)^{-1}$.

(iii) When $m/n \to \infty$, i.e., $\rho_v/\rho_h \to 0$ and $\rho_h \to 1$, Theorems 3.3 and 4.3 imply that the MTB policy is within $n^2/(\lambda^2\sigma_s^2)$ factor of the optimal. However, the conjecture in Remark 3.4, once proven, will establish that the MTB policy is within $1/c$ factor of the optimal as $\rho_v/\rho_h \to 0$ and $\rho_h \to 1$.

We now propose an another policy which is inspired by a similar policy for the multi-vehicle DTRP in [1]. This policy is best suited for scenarios where both the vehicles and the human operators are working under heavy load conditions.

The Equipartition Based (EB) Assignment Policy

Select a location $\bar{q}$ in $\mathcal{Q}$ arbitrarily and let this be the loitering location for all the vehicles. For some fixed integer $l \geq 1$, divide $\mathcal{Q}$ into l sub-regions of equal area using radial cuts centered at $\bar{q}$ (i.e., form l wedges of area $\mathcal{A}/l$). As the targets are generated, form batches of k/l targets each, with targets in each batch coming from the same sub-region. Once a batch is formed, deposit them in a queue. The batches are assigned to the first available vehicle in a First Come First Serve (FCFS) basis. Optimize over k and l. The mode of information collection is set to online for all the targets. The incoming attention requests from the vehicles are assigned to the human operator on a first come first serve basis.

Let $\overline{T}_{\text{EB}}$ be the system time obtained by implementing the equipartition based assignment policy.

Theorem 4.5: There exists a constant $\bar{\beta} \approx 3.66$ such that the classification time given by the EB policy satisfies the following relations.

(i) For $\rho_h \leq \rho_v$,

$$\frac{\overline{T}_{\text{EB}}}{\overline{T}_{\text{opt}}} \leq \bar{\beta} \quad \text{as } \rho_v \to 1, \text{ and}$$

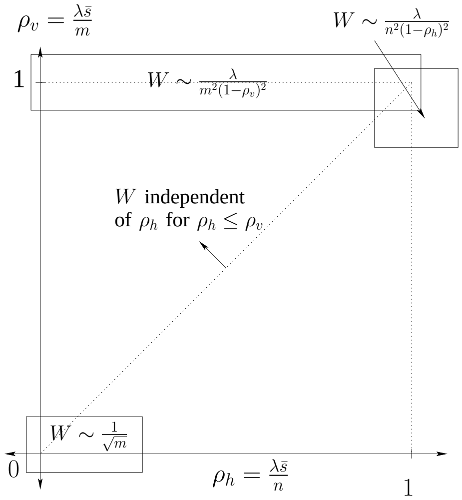

Fig. 2. A $\rho_h - \rho_v$ chart showing dependence of average wait time for a target, on the number of humans n and number of vehicles m for a given target generation rate λ and average information required for classification per target is $\bar{s}$.

(ii) For $\rho_v < \rho_h$,

$$\frac{\overline{T}_{\text{EB}}}{\overline{T}_{\text{opt}}} \leq \bar{\beta}\left(\frac{\rho_h}{\rho_v}\frac{1-\rho_v}{1-\rho_h}\right)^2 \quad \text{as } \rho_h \to 1.$$

Remark 4.6: Note that the EB policy utilizes at most m human operators (even when $n \geq m$). Part (i) of Theorem 4.5, along with Theorem 3.2, implies that the EB policy gives a $\bar{\beta}$ factor approximation to the optimal.

V. Conclusion

Our goal in this paper was to formulate and study a novel routing problem requiring human decision and design coordination policies for the vehicles and human human operators. We considered a problem where the autonomous vehicles provide access to on-site information, and the human operator provide the judgment capabilities necessary to process such information. We designed joint motion coordination and operator scheduling policies and proved that they performed within a constant factor of the optimal for relevant asymptotic cases. The summary of the results is depicted in Figure 2, where we have shown the dependence of the average wait time $W := \overline{T}_{\text{opt}} - \bar{s}$.

There are many outstanding issues within the scope of the problem formulated in this paper. It is worth pointing out that none of the policies in this paper utilized the offline mode of information collection. This is primarily because analyzing policies with offline information collection mode and re-look lead to open problems in queuing theory. However, the fact that the online mode based policies are optimal under certain conditions even under the possibility of offline mode also shows the futility of the offline mode in those conditions. We intend to investigate the performance of offline mode based policies for other problem conditions through means of extensive numerical experiments. Other open problems of interest include extending this framework to include routing under constraints on capacity, fuel, communication, sensing range and vehicle motion. The architecture proposed in this paper does not necessarily get the most out of human participation in the policy making. Specifically, we assume that the operators are unable to share the classification state $x_i(t)$ with the SS. In the future, we intend to consider different architectures which allow greater participation from the human operators at different levels depending upon the advantages that it brings along. This type of architecture would allow for much more sophisticated policies including, for instance, a scouting mission upon which a larger plan is based.

Another ongoing extension is to model human performance based upon his or her utilization history. This allows for a much more accurate model of performance and preliminary work has been reported in [14]. Finally, we intend to extend the framework introduced in this paper to other cooperative tasks involving human decision making.

Acknowledgments

This research was done with support from the Michigan/AFRL Collaborative Center on Control Sciences. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the supporting organizations.

References

[1] D. J. Bertsimas and G. J. van Ryzin, "Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles," *Operations Research*, vol. 41, no. 1, pp. 60–76, 1993.

[2] J. J. Enright, E. Frazzoli, K. Savla, and F. Bullo, "On multiple UAV routing with stochastic targets: performance bounds and algorithms," in *AIAA Conf. on Guidance, Navigation and Control*, Aug. 2005. Electronic Proceedings.

[3] K. Savla, E. Frazzoli, and F. Bullo, "Traveling Salesperson Problems for the Dubins vehicle," *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1378–1391, 2008.

[4] W. Li and C. Cassandras, "A cooperative receding horizon controller for multivehicle uncertain environments," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 242–257, 2006.

[5] D. Gross, S. Rasmussen, P. Chandler, and G. Feitshans, "Cooperative operations in UrbaN TERrain (COUNTER)," in *Defense Transformation and Network-Centric Systems*, vol. 6249 of *Proc. of SPIE*, pp. 62490G–1–11, SPIE, 2006.

[6] T. B. Sheridan, *Telerobotics, Automation and Human Supervisory Control*. Cambridge, MA: MIT Press, 1992.

[7] C. Wickens, S. Gordon-Becker, and Y. Liu, *An Introduction to Human Factors Engineering*. Prentice Hall, 1997.

[8] M. L. Cummings, S. Bruni, S. Mercier, and P. J. Mitchell, "Automation architecture for single operator, multiple uav command and control," *The International C2 Journal*, vol. 1, no. 2, pp. 1–24, 2007.

[9] L. Kleinrock, *Queueing Systems, Volume II: Computer Applications*. New York: Wiley, 1976.

[10] R. Ratcliff and P. L. Smith, "A comparison of sequential sampling models for two-choice reaction time," *Psychological Review*, vol. 111, pp. 333–367, 2004.

[11] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. Cohen, "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks," *Psychological Review*, vol. 113, no. 4, pp. 700–765, 2006.

[12] J. Little, "A proof of the queuing formula $l = \lambda w$," *Operations Research*, vol. 9, pp. 383–387, 1961.

[13] P. K. Agarwal and M. Sharir, "Efficient algorithms for geometric optimization," *ACM Computing Surveys*, vol. 30, no. 4, pp. 412–458, 1998.

[14] K. Savla, C. Nehme, T. Temple, and E. Frazzoli, "On efficient cooperative strategies between humans and uavs in a dynamic environment," in *AIAA Conference on Guidance, Navigation and Control*, Aug 2008. Electronic Proceedings.