THE INTERNATIONAL
C2 JOURNAL

VOLUME 2, NUMBER 2, 2008

SPECIAL ISSUE
Representing Human Decision Making in Constructive Simulations for Analysis

GUEST EDITOR
James Moffat
Dstl

The Impact of Heterogeneity on Operator Performance in Future Unmanned Vehicle Systems

Carl Nehme
Brian Mekdeci
Jacob W. Crandall
M. L. Cummings

CCRP

THE INTERNATIONAL C2 JOURNAL

About the Journal

The International C2 Journal was created in 2006 at the urging of an international group of command and control professionals including individuals from academia, industry, government, and the military. The Command and Control Research Program (CCRP, of the U.S. Office of the Assistant Secretary of Defense for Networks and Information Integration, or OASD-NII) responded to this need by bringing together interested professionals to shape the purpose and guide the execution of such a journal. Today, the Journal is overseen by an Editorial Board comprising representatives from many nations.

For more information

Visit us online at: www.dodccrp.org
Contact our staff at: publications@dodccrp.org

The Impact of Heterogeneity on Operator Performance in Future Unmanned Vehicle Systems

Carl Nehme, Brian Mekdeci, Jacob W. Crandall, and M. L. Cummings (Massachusetts Institute of Technology, USA)

Abstract

Recent studies have shown that with appropriate operator decision support and with sufficient automation, inverting the multiple operators to single-unmanned vehicle control paradigm is possible. These studies, however, have generally focused on homogeneous teams of vehicles, and have not completely addressed either the manifestation of heterogeneity in vehicle teams, or the effects of heterogeneity on operator capacity. An important implication of heterogeneity in unmanned vehicle teams is an increase in the diversity of possible team configurations available for each operator, as well as an increase in the diversity of possible attention allocation schemes that can be utilized by operators. To this end, this paper introduces a discrete event simulation (DES) model as a means to model a single operator supervising multiple heterogeneous unmanned vehicles. The DES model can be used to understand the impact of varying both vehicle team design variables (such as team composition) and operator design variables (including attention allocation strategies). The model also highlights the sub-components of operator attention allocation schemes that can impact overall performance when supervising heterogeneous unmanned vehicle teams. Results from an experimental case study are then used to validate the model, and make predictions about operator performance for various heterogeneous team configurations.

Introduction

Increasing use of automation in unmanned vehicle (UV) systems will shift the human operator's responsibility from manually controlling vehicles to commanding vehicles at the supervisory control level. This shift will be critical in order to realize the single operator-multiple UV control paradigm outlined in the Office of the Secretary of Defense Roadmap for unmanned aerial systems (UASs) (Office of the Secretary of Defense 2005), as well as teams of UVs from multiple domains as suggested by the Committee on Autonomous Vehicles in Support of Naval Operations (Naval Studies Board 2005). At the supervisory control level, implementation details of higher-level tasking initiated by the human operator is delegated to the automation onboard these vehicles (Sheridan 1992). The reduced workload afforded by supervisory control has several implications for unmanned system operations. One such ramification is an increase in operator idle time, which can be used as a force multiplier that allows operators to supervise multiple vehicles simultaneously, hence inverting the current many-to-one ratio of operators to vehicles. Inverting the operator to vehicle ratio can also be used to reduce manning in situations where the number of vehicles needed to accomplish missions exceeds that of available operators, which is currently a significant problem in the Predator community.

An increasing body of literature has addressed one or more aspects of human-UV interaction within the supervisory control of multiple UVs. For example, role allocation between the human operator and the vehicles has been addressed in the research of alternate autonomy architectures such as adjustable autonomy (Goodrich et al. 2001; Miller and Parasuraman 2007; Parasuraman, Barnes, and Cosenzo 2007). Similarly, research has investigated role allocation between the different vehicle team members by addressing task allocation issues including level of decentralization in task assignment (Alighanbari and How 2005) and embedding health management into vehicle mission tasking (Valenti et al. 2007).

This paper focuses on modeling the combined human-UV system in order to identify vehicle, operator and system limitations. Some limitations such as the capacity of single operators to supervise multiple UVs have been addressed when considering homogeneous UV teams (Ruff, Narayanan, and Draper 2002; Cummings et al. 2007; Olsen and Wood 2004). However, as UV system mission goals become increasingly demanding, the composition of UV teams is likely to involve vehicles of varying capabilities. For example, the military has proposed future operational concepts such as Network Centric Warfare (Alberts, Garstka, and Stein 1999) and the Future Combat System (FCS) (Feickert 2005) that require interoperability among UVs of varying attributes. In addition to heterogeneity across vehicle types, even a single UV can have multiple payloads which will ultimately lead to heterogeneity in operator tasks.

These multiple dimensions of heterogeneity introduce a number of problems in applying previous models of homogeneous UVs to the heterogeneous case. These previous models fail to account for the different demands that heterogeneous vehicles/tasks could require of the operator. This could cause these models to under or overestimate the efficiency of the overall interaction. Moreover, the heterogeneity in vehicles/tasks is likely to place stronger emphasis on the importance of operator cognitive processes such as the ability to maintain situational awareness, also not included in the extant homogenous models. In addition, since the vehicles and associated tasks are disparate in the heterogeneous case, there is a larger diversity in possible attention allocation schemes than in the homogeneous case. The method by which operators allocate their attention to the heterogeneous vehicles/tasks is likely to affect system performance. Therefore, capturing the various operator management strategies and their effect on system performance is another important variable that must be considered. Not accounting for the variability that heterogeneous teams introduce could result in a misunderstanding of the impact of heterogeneity and could render any design recommendations from extant models inaccurate.

This paper addresses these problems by introducing a discrete event simulation model that incorporates design variable inputs that include team structure and composition as well as attention allocation strategies that define the operator's interaction with the UV team. An experimental case study is then used to validate the ability of the model to replicate human-in-the-loop data extracted from the experiments. Finally, the paper concludes by predicting the effects of alternate operator attention allocation strategies in the supervision of a heterogeneous team.

Background

The supervisory control of UVs requires a human operator handling intermittent events as they arrive, dividing his/her attention according to some allocation scheme. Due to the time-critical, event-driven nature of human supervisory control, discrete event simulation (DES), which models a system as it evolves over time by representation of events (Law and Kelton 2000), can be used to model supervisory control systems. Examples of this include the use of queuing-based DES models to describe the method by which humans allocate their attention when presented with intermittent sensory inputs such as when pilots have to make sense of alerts and cues from cockpit instrumentation (Carbonell, 1966), or when air traffic controllers have to manage aircraft in their sector (Schmidt, 1978). While many different modeling technologies, including agent-based models and Petri Nets, could potentially be used to capture the human-UV interactions, this research has focused on DES due to its ability to capture the temporal aspects of human-UV interactions. These temporal aspects of a system, which include wait times (Cummings and Mitchell 2008), interaction times, and neglect times (Crandall et al. 2005; Olsen and Goodrich 2003), determine the limitations of the system.

Previous research that examined the capacity of operators supervising multiple homogeneous robots by Olsen and Goodrich (2003) introduced several temporal-based metrics to describe how operators interact with UVs. Neglect Time (NT) was defined as the expected amount of time that a robot (which is representative of any UV) can be ignored before its performance drops below some acceptable threshold. Interaction Time (IT) was defined as the average time it takes for a human to interact with the robot to ensure that it is still working toward mission accomplishment. In the single robot example (Figure 1a), the operator interacts with the robot for length of time IT and then ignores it for length of time NT, and then repeats this process after time NT by interacting with the robot once again. In the multiple robot case, an operator would interact with one robot at a time while neglecting all other robots (Figure 1b).

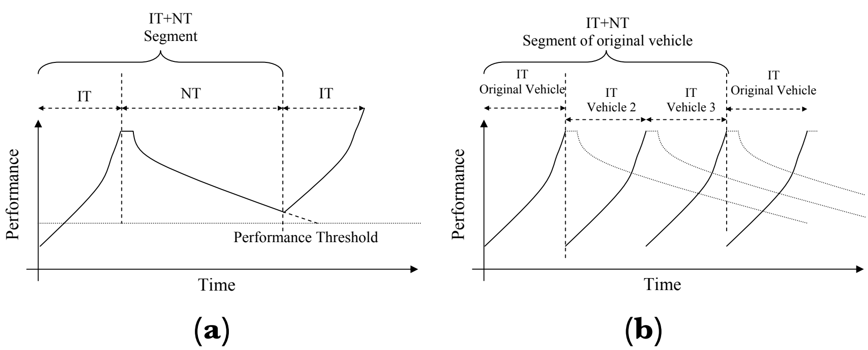


Figure 1. The relationship of NT and IT for (a) a single vehicle, and (b) multiple vehicles.

One drawback to this earlier work is the lack of accounting for human interaction delays and decision making inefficiencies. An additional critical variable needed when modeling human control of multiple vehicles is the concept of Wait Times (WT). Although it is possible for human beings to multi-task, humans act as serial processors in that they can only solve a single complex task at a time (Welford 1952; Broadbent 1958). While operators can rapidly switch between cognitive tasks, any sequence of tasks requiring complex cognition will form a queue and consequently, wait times will build

(Cummings and Mitchell 2008; Cummings et al. 2007). Wait times can occur when 1) a vehicle is neglected while the operator is busy interacting with another vehicle, or 2) when an operator requires re-orientation time while switching between vehicles, or 3) when a vehicle is neglected due to lack of operator situation awareness.

In this paper, a queuing-based model that builds on the concepts of NT and IT and captures the different wait times is proposed. This queuing model forms the basis of the discrete event simulation described in subsequent sections, which is used to model a single operator controlling multiple heterogeneous vehicles.

Discrete Event Simulation Model

By capitalizing on the event-driven nature of human-supervisory control, a discrete event simulation (DES) model based on queuing theory was developed to examine the impact of changing vehicle team structure, as well as operator attention allocation strategies on overall system performance. The previously introduced concepts of neglect time and interaction time are utilized in this model to capture vehicle autonomy and event service times respectively.

Overview

The operator model in Figure 2 was constructed under the assumption that the operator is acting in a supervisory control mode and that the different vehicles in the team are highly autonomous. Therefore, the vehicles generally only require operator interaction for tasks that require human judgment and reasoning. The operator can interact with a UV when either a) an event occurs that requires human judgment and reasoning, something the automation is incapable of handling, or b) the automation is not acting as expected and the operator believes that interaction can increase performance. For example, in the case of a UAV that is assigned a laser designation task, the operator could re-plan the vehicle path generated by automation in order to better meet a time-on-target restriction. The operator's judgment is also critical in deciding whether a specific target is the one that should be designated.

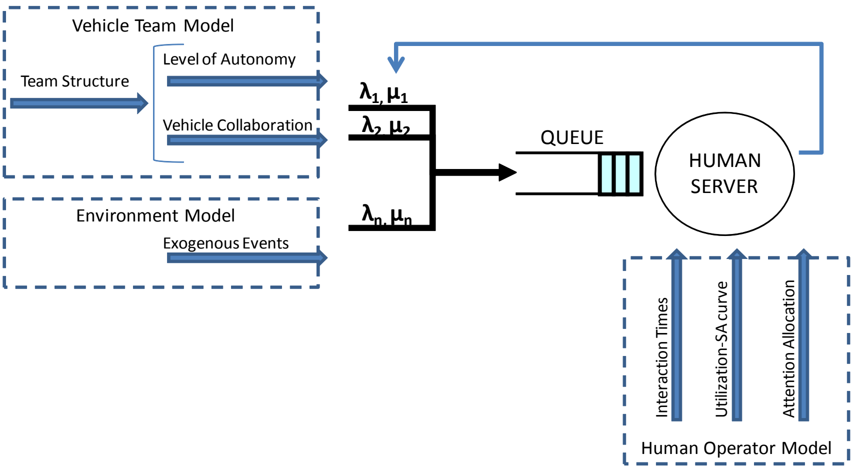


Figure 2. A high level representation of the discrete event simulation model including vehicle team and human operator input variables.

The model inputs in Figure 2 are composed of variables related to the vehicle team (team structure, level of autonomy and vehicle collaboration), the human operator (interaction times, operator attention allocation strategies, and the operator utilization-SA characteristic curve), and a model of environment unpredictability. These are discussed below in further detail.

Vehicle Team Inputs

The level of vehicle autonomy is captured through the previous discussed concept of neglect time (NT). Since NTs represent the time a vehicle can operate without human intervention, they effectively represent degrees of autonomy. Discrete events in this system represent both endogenous and exogenous situations that the operator must address. Exogenous events are events that create the need for operator interaction which result from unexpected external environmental conditions, such as an emergent threat area which require re-planning vehicle trajectories.

Endogenous events are events created internally within the UV system, such as when an operator elects to re-plan an existing UV path in order to reach a goal in a shorter time. Endogenous events can be either vehicle-generated or operator induced in which case the interaction may not be required by the system but operator-induced with the intention of improving performance. The team structure variable, which represents the number and type of vehicles included in the system being modeled, is captured in the number of event streams that arrive to the operator queue as well as the arrival processes associated with each stream. Lastly, the model captures the effect of vehicle collaboration by taking into account the effect of servicing a particular event belonging to one vehicle on the arrival process of another event belonging to another vehicle.

Human Operator Inputs

The operator model in Figure 2 is based on the single server queue with multiple input streams. The operator can attend to only one complex event at a time, and this is captured by the single-server architecture such that any events that arrive while the operator is busy will wait in a queue. The length of time it takes the operator to deal with an event, interaction time, is captured through a probability distribution of event service times (such that the probability distribution captures the variability of performance between different operators and variability in the performance of a single operator). Interaction times occur for a single vehicle task, so in order to model the effect of an operator controlling multiple vehicles, the model should consider how and when operators elect to attend to the vehicles, also known as attention allocation (Crandall and Cummings

2007a). When supervising multiple UVs, the operator attention allocation strategy will dictate the method by which the operator will supervise the different vehicles. Our model captures two attention allocation strategies that can impact the effectiveness of human-UV interaction; a) the operator management strategy and b) the operator switching strategy.

Operator Management Strategy. The first strategy is the operator management strategy and affects the amount of operator re-planning. Since this model supports endogenous events that are both vehicle-generated and operator-induced, the rate at which operator-induced events arrive to the system depends on the operator's desire to interact with the vehicles beyond unavoidable vehicle-generated events. The management strategy can vary per vehicle, and can be thought of as the scheme by which the operator distributes his/her attention across the different vehicles. One type of management strategy is a macro-management strategy where the operator services the vehicle only when necessary and otherwise allows the vehicle's automation to undertake tasks. On the other hand, a micro-management strategy is one where the operator constantly interferes with the vehicle's automation. Other management strategies can exist between these two extremes.

Operator Switching Strategy. The second component of human-attention allocation is the order by which the different vehicles are serviced. When multiple vehicles require operator attention simultaneously, the operator must select the next vehicle to be serviced. Whereas this selection process is relatively simple in the homogeneous case, it is much more involved in the heterogeneous case. In the heterogeneous case, the difference in vehicle capabilities and their assigned tasks allows for more diverse selection strategies. For example, an operator that is supervising two UAVs with heterogeneous tasks can service the vehicles on a first come, first serve basis (FIFO) or allocate attention to the UAVs based on the priority of their tasks (preemptive priority queuing). The order by which the vehicles are serviced affects the total time that vehicles spend in the system, including the time they spend waiting for service as well as their processing time (Mau and Dolan 2006; Sheridan and Tulga 1978). In addition to having an effect on wait times, when a human operator switches between two different tasks, this is accompanied by a mental model switch that comes at a time (or switch) cost (Goodrich, Quigley, and Cosenzo 2005; Squire, Trafton, and Parasuraman 2006). Thus switching between different combinations of heterogeneous vehicles can lead to different switch costs. In order to model the switching strategy of the operator, the type of queue can be varied to represent different strategies. Examples of switching strategies that can be modeled include the first-in-first-out (FIFO) queuing scheme as well as the highest attribute first (HAF) strategy (Pinedo 2002). The HAF strategy is similar to a preemptive priority scheme in that high priority events are serviced first except that there is no pre-emption. Therefore if an event is generated with a priority higher than any of the events already in the system, it will be moved to the front of the queue but will not preempt a lower priority vehicle that is already being serviced.

Situational awareness (SA) is defined as the combination of perception of elements in the environment, the comprehension of their meaning, and the projection of their status in the future (Endsley 1995). The effect of low SA is to create additional vehicle wait times due to loss of situational awareness (WTSA), which increase the time it takes the operator to notice the needs of the system (Cummings et al. 2007). In order to capture SA, this model builds on an assumption that SA is related to operator utilization (Endsley 1993). When operators are under high levels of utilization, it is assumed that they are too busy to accumulate the information that is required to build SA. At the same time, when operators are under-utilized, it is presumed that due to a low level of arousal and complacency, they could overlook information from the environment, which would also lead to low SA.

While this model accounts for variation in human performance through its stochastic representation of operator performance, we recognize that other variables are likely to influence individual operator performance such as training and fatigue. Since the purpose of this initial modeling attempt is to determine the impact of system variables (i.e., number and type of vehicle), we leave investigation of the impact of individual operator attributes as the subject of future work.

Model Architecture

Since an endogenous event associated with a specific vehicle generally requires attention before an event of the same type can be generated by the same vehicle (or human-triggered), the arrival process is one of correlated arrivals. For example, if a vehicle A requires the operator to analyze a captured image, he or she must finish servicing that event before vehicle A can next generate another "analyze image" event. In order to model this phenomenon, the model uses a closed queuing network paradigm such that each endogenous event type in the system (where each endogenous event type is associated with a specific vehicle) has a population of one (Lazowska et al. 1984). In this manner, once an event of a specific endogenous type for a specific vehicle in the team is generated, no other endogenous event of the same type and belonging to the same vehicle can be generated. Thus inter-arrival times for a stream are the times between the completion of service for an event and the arrival of the next event.

In order to capture interaction effects between two or more event types, the servicing of one event type can be modeled to have an effect on the arrival process of another. For example, a UV might be modeled through two event types; a) the need for operator interaction whenever the operator is required to review imagery captured by the UV, and b) the need for operator interaction once the image reviewing process is complete and the vehicle requires a new assignment. In this case, event type (b) is generated only after event type (a) is serviced by the operator. Similarly, the model can also represent vehicle collaboration by accounting for the influence of one vehicle's event servicing on a second vehicle's event arrival. This same modeling idea can be extended to model collaboration between three or more vehicles.

Unlike endogenous events, exogenous events stem from sources external to the vehicle (weather, target movements, etc.) and are generally generated in an independent manner. For example, many emergent threats can arise simultaneously, each requiring operator intervention. Therefore the arrival process in the case of exogenous events is generally one of independent arrivals. For both endogenous and exogenous events, the arrival process can be described by a probabilistic distribution over a random variable (X'_i) which is a function of two main components; a) the random variable X_i, and b) operator loss of situational awareness (Equation 1).

$$X'_i = X_i + \chi * X_i \qquad (1)$$

The first term in Equation 1, X_i, is a random variable that describes the time between one service and the next arrival in the case of endogenous events and the time between arrivals in the case of exogenous events. The generation of a task does not necessarily imply that the operator notices the generated task as the rate excludes any effects due to loss of SA.

The second term in Equation 1, $\chi * X_i$, represents a penalty due to operator loss of situational awareness (SA), with χ taking a value of zero when the operator has complete SA and higher χ indicating degraded SA. The χ variable in Equation 1 is related to operator utilization through a parabolic function that is concave upwards (Figure 3). This implies that at both high and low operator utilization, χ increases according to a quadratic law and therefore increases X' correspondingly. The parabolic relationship is inspired by the Yerkes Dodson Law (Yerkes and Dodson 1908), which relates operator utilization to performance. The χ variable is multiplied by X in order to capture the effect on X' due to loss of SA, which is a function of the rate at which the vehicle generates tasks that require operator intervention. Vehicles that produce tasks infrequently are serviced less often, and are therefore more likely to be overlooked than vehicles that are serviced more frequently.

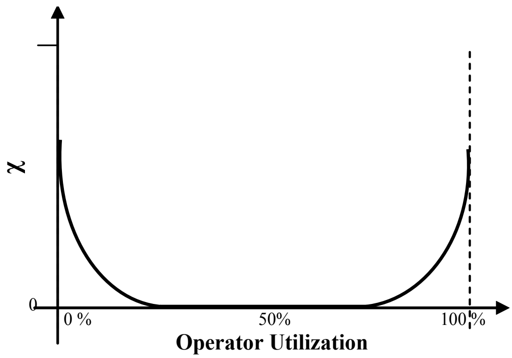


Figure 3. χ curve.

Also associated with each input stream is a service rate which is based on the length of time it takes the operator to interact with a particular event. The service process can be described by a probabilistic distribution over a random variable (Y'_i) which is a function of two main components; a) the random variable Y_i, and b) wait times due to interaction (Equation 2).

$$Y'_i = Y_i + WTI \qquad (2)$$

The first term in Equation 2, Y_i, is the random variable that describes the length of time for which the operator must interact with event type *i*. The probability distribution describing Y_i is mainly a function of interface and decision support quality. The second term in Equation 2, WTI, is a function of the wait times due to context switching that arise when servicing a specific vehicle. The switch cost is not limited to switching between cognitively complex tasks, but exists even when humans switch between cognitively simple ones (Rogers and Monsell 1995). For example, Goodrich et al. (2005) demonstrated that the existence of context switching costs in multi-vehicle control is unavoidable, and that the amount of time required to switch between vehicles can be substantial. The effect of switching times creates additional interaction wait times (WTI) which increases Y', due to the operator taking longer to interact with the vehicle.

Experimental Case Study

To evaluate the ability of the model to accurately replicate the performance characteristics of human-UV teams, outputs from the model were compared to results from an experimental study conducted to investigate operator performance issues in the control of multiple simulated homogeneous UVs conducting a simulated search-and-rescue mission. A homogeneous UV simulator was used because the test bed has been successfully used in a number of previous human-in-the-loop experiments in order to validate discrete event simulation predictions (Crandall and Cummings 2007a; Crandall and Cummings 2007b; Pina et al. 2008). The experimental study and model parameters are discussed below.

Experimental Apparatus

Three aspects of the experimental test bed used in the user study are described in this subsection, which include *mission, interface,* and *UV behavior*.

Mission. The human-UV team (which consisted of the participant and multiple simulated UVs) was assigned the task of removing as many objects as possible from a maze in an 8-minute time period. The objects were randomly spread through the maze, which was initially unknown. However, as each UV moved about the maze, it created a map which it shared with the participant and the other UVs in the team. The team could only see the positions of six of the objects initially. In each minute of the session, the locations of two additional objects were shown. Thus, there were 22 possible objects to collect during a session.

An object was removed from the maze (i.e., collected) using a three-step process. First, a UV moved to the location of the object in the maze (i.e., target designation, mission planning, path planning, and UV monitoring). Second, the UV "picked up" the object (i.e., sensor analysis and scanning). In the real world, performing such an action might require the human operator to assist in identifying the object with video or laser data. To simulate this task, we asked users to identify a city on a map of the mainland United States using *Google Earth*-style software. Third, the UV carried the object out of the maze via one of two exits.

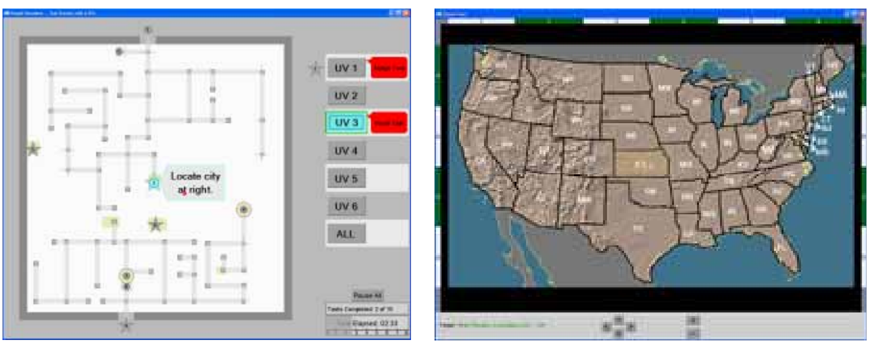


Figure 4. Two-screen interface by which an operator directed the UVs.

Interface. The human-UV interface was the two-screen display shown in Figure 4. On the left screen, the map of the maze was displayed, along with the positions of the UVs and (known) objects in the maze. The right screen was used to locate the cities. A participant could only control one UV at a time. When a user desired to control a certain UV, he/she clicked a button on the interface corresponding to that UV (labeled UV1, UV2, etc.). Once the participant selected the UV, he/she could direct the UV by designating a goal location and modifying the UV's intended path to that goal. Designating a goal for the UV was done by dragging the goal icon corresponding to the UV in question to the desired location. Once the UV received a goal command, it generated and displayed the path it intended to follow. The participant was allowed to modify this path using the mouse.

UV Behavior. The UVs' map of the maze took the form of an undirected graph. Each edge of the graph was an ordered pair (u, v) representing a connection between vertices u and v in the graph. Associated with each edge was a weight indicating the cost for a UV to move along that edge. Since the maze was not fully known, a UV had to choose between (a) moving along the shortest path of the known maze to its user-specified goal and (b) exploring the unknown portions of the maze in hopes of finding a shorter path. To make this decision, a UV assumed that an unmapped edge from a known vertex v led directly to the goal position with a cost equal to the Manhattan distance from v to the UV's goal, plus some cost of exploration (CE). Each UV used Dijkstra's algorithm on the resulting graph to determine the path it intended to follow.

Using this approach, the constant CE determines the degree to which the UVs explore the unknown maze. Higher values of CE result in less exploration. We used a small value of CE for a UV that was searching for an object, and a higher value for a UV that was carrying an object. Since users sometimes felt that the resulting behavior was undesirable, they were allowed to modify a UV's path if they desired.

Two different versions of UV autonomy were employed in the user study. In the first condition, called the *no-decision support* (NDS) condition, each UV's goal destination was determined completely by the human operator. Once the UV arrived at its user-defined goal destination, it did not move again until it received a new command from the user.

In the second condition, called the *full-decision support* (FDS) condition, each UV automatically selected a new goal when it was left idle. Specifically, a management-by-exception level of automation was used in which a UV left idle at its goal destination, but not on an object in the maze, waited 15 seconds for the user to intervene. If the user did not intervene, the UV automatically moved to the nearest unassigned object (if the UV was searching for an object) or the nearest exit (if the UV was already carrying an object). Additionally, if the user did not intervene, UVs automatically chose to exit the maze via the (estimated) nearest exit in the final 45 seconds of a session. The FDS condition also had one other additional decision support tool to assist the user in locating cities on the map (to "pick up" objects). This decision support tool decreased the search time for a city on the map by about 5 seconds on average.

Participants and Experimental Procedure

The experimental design was a 2x4 factor study. The decision support condition (NDS or FDS) was a between-subjects factor. UV team size was a within-subjects factor; each participant performed the search-and-rescue mission for team sizes of two, four, six, and eight UVs. The order in which the participants used each team size was counter-balanced throughout the study. Each participant was first randomly assigned to a decision support condition (NDS or FDS), and then was trained on all aspects of the system. They then completed three comprehensive practice sessions. Following these practice sessions, each participant performed four test sessions (each with a different team size). Participants were paid \$10 per hour; the highest scorer also received a \$100 gift certificate. Thirty-two participants between the ages of 18 and 45 (mean 24.4 years old) participated in the study, 16 in each condition.

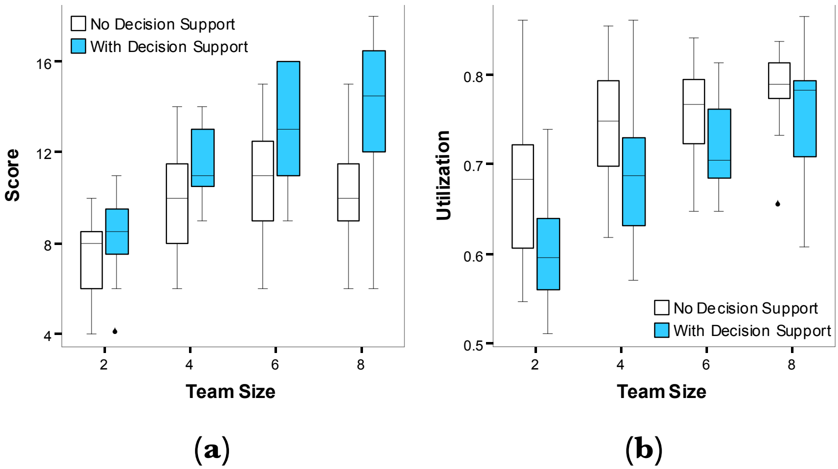


Figure 5. Results from and case study for two, four, six, and eight vehicles with decision support and without decision support for (a) score, and (b) utilization.

Results & Discussion

Human-in-the-Loop Experimental Results

The results from the case study are shown in Figure 5. A repeated measures ANOVA showed that team size had a significant main effect on the score ($F(3, 90) = 41.874$, $p < 0.001$). Pair-wise comparisons showed a significant difference in score between all team sizes ($p = 0.04$ between the four and eight UV team sizes, $p = 0.01$ between the four and six UV team sizes, and $p < 0.0001$ for the rest), except between the six and eight UV team sizes, which was not significant. Analysis of decision support type showed a significant main effect on the score variable ($F(1, 30) = 9.84$, $p = 0.004$).

The repeated measures ANOVA for utilization showed that team size had a significant main effect ($F(3, 84) = 27.97$, $p < 0.001$). Pair-wise comparisons showed a significant difference in utilization between all team sizes ($p = 0.005$ between the four and eight team sizes, and $p < 0.0001$ for the rest), except between the four and six UV team sizes ($p = 0.43$) and between the six and eight UV team sizes ($p = 0.12$). The decision support type also showed a significant main effect for utilization ($F(1, 28) = 8.45$, $p = 0.007$).

DES Results

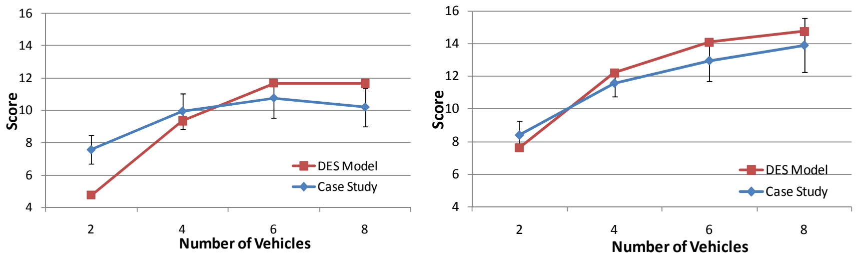
In order to compare the DES model results to the human-in-the-loop experimental results, vehicle-generated and operator-induced distributions in the case study were identified. Vehicle-generated events include both locating a city on the map and goal assignment in the case of the NDS condition, but only locating a city on the map in the FDS condition (since the UV could function without goal-assignments from the user in this condition). Thus, five data sets were measured from the experimental data: (1) arrival rate of vehicle-generated events after the UV was serviced for a vehicle-generated event, (2) arrival rate of vehicle-generated events after the UV was serviced for an operator-induced event, (3) service times of vehicle-generated events, (4) arrival rate of operator-induced events, and (5) service times of operator-induced events.

The data sets collected were then used to generate random distributions that were used by the model. In most cases, the distribution that best fit the data was the lognormal distribution. This was expected for both service times and arrival rates since they are skewed to the left, corresponding to the cases where an abnormally long amount of time passes between events or while the operator is servicing a vehicle.

The complete model of the human-UV team also requires a performance model. In the user study, the team scored points when an object was removed from the maze. In the NDS condition, this required two vehicle-generated events to occur (goal-assignment and locating a city). Thus, the DES Model awarded a point for the servicing of every two vehicle-generated events. In the FDS condition, only one vehicle-generated event (locating a city) needed to be performed. Thus, in this condition, the DES Model awarded a point for every serviced vehicle-generated event.

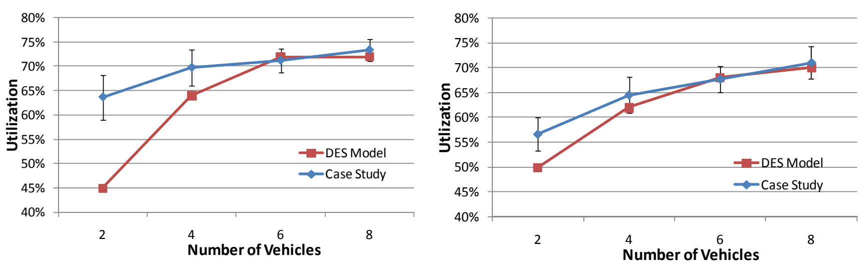
Using the distributions of arrival rates and service times generated from the data in the user study, 10,000 trials were conducted with the DES to replicate system performance and operator utilization for each condition of the study, in order to compare the results with the human-in-the-loop experiment. Operator utilization is included in the analysis since users who are more than 70% busy typically demonstrate degraded performance (Schmidt 1978; Rouse 1983; Cummings and Guerlain 2007).

The observed system performance and operator utilization from the user study are compared with the model's estimates in Figures 6 and 7, respectively. For the FDS condition, the model's results for system performance are all within the 95% confidence intervals. Likewise, the FDS utilization results are all within the 95% confidence intervals except in the 2 UV case. In this case, the model underestimates operator utilization by approximately one standard deviation. In the NDS condition, the model's results of system performance are within the 95% confidence intervals for the 4 and 6 UV conditions, but were low (1.6 standard deviations away from the mean) in the 2 UV condition and slightly high in the 8 UV condition (0.6 standard deviations from the mean). Additionally, results of operator utilization are within the 95% confidence intervals for 6 and 8 UV teams, but not the 2 and 4 UV teams, where results are off by 2 and 0.8 standard deviations respectively.



(a) (b)

Figure 6. Comparison of mean performance scores from DES model and case study for 2, 4, 6, and 8 vehicles for (a) the no decision support case, and (b) the full decision support case.



(a) (b)

Figure 7. Comparison of mean utilizations from DES model and case study for 2, 4, 6, and 8 vehicles for (a) the no decision support case, and (b) the full decision support case.

Two observations about the accuracy of the model can be made from these results. First, model results are more accurate for larger teams than small teams. This trend appears to be caused, at least to some degree, by overly high penalties associated with low utilization in the SA model (Equation 1 and Figure 3). As a result, the human is modeled as not servicing UVs as often as necessary in the 2 UV condition. This leads to the under estimation of both operator utilization and system performance in this condition.

The second trend observed in Figures 6 and 7 is that the model's estimates are better in the FDS condition than in the NDS condition. Again, it appears that this trend is due to difficulties in modeling human behavior, as human behavior is more difficult to model than is automation. Thus, systems that rely more on human behavior (i.e., the NDS condition) are more difficult to accurately model than systems that rely more on automated behavior (i.e., the FDS condition).

Despite variations in the accuracy of the model's results, they capture the general trends in system performance and operator utilization as the size of the team and the level of decision support change. This is important since it means that the model gives adequate descriptions of the behavior of different system architectures in a cost effective manner. Given this positive result, the model is now used to predict system performance and operator utilization in the same mission for heterogeneous teams.

The Extension to Heterogeneous Teams

The teams considered in the user study consisted of UVs with homogeneous capabilities. In this subsection, simulated heterogeneous UV teams of two, four, six, and eight UVs with NDS and FDS capabilities are considered. Such teams reflect situations in which operators must simultaneously supervise both legacy UVs with less autonomy (such as NDS) and newer, more automated UVs (such as FDS). The performance of each of these teams using various operator strategies is compared to the performance of the homogeneous teams. Different switching strategies and management strategies are considered.

Switching Strategies. Recall that switching strategies refer to the order in which the operator attends to UVs that need to be serviced. This requires that UVs and tasks be given specific priorities, and that the operator services the UV with the highest priority. For the heterogeneous model predictions, a first-in-first-out (FIFO) switching scheme as well as two different priority schemes are considered. In the first priority scheme (referred to as *NDS-preferred*), vehicle-generated events by UVs with NDS capabilities are given the highest priority, followed by vehicle-generated events by UVs with FDS capabilities, followed by operator-induced events. For the second priority scheme (referred to as *FDS-preferred*), the highest priority is given to vehicle-generated events from UVs with FDS capabilities, followed by vehicle-generated events by UVs with NDS capabilities, followed by operator-induced events.

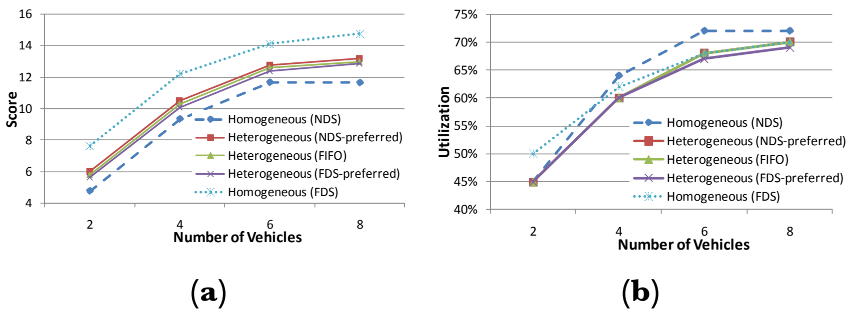


Figure 8. Comparison of (a) mean performance scores, and (b) mean utilizations when using alternate switching strategies for teams with two, four, six, and eight vehicles.

In Figure 8, the performance score and utilization of each heterogeneous team for each of the three switching strategies (generated using the DES) are compared to the results associated with the two homogeneous teams from the previous section. Due to the congruity of the vehicles, it can be assumed that operators exhibited a FIFO switching strategy in the two homogeneous cases. Figure 8a shows that all three heterogeneous team results have nearly identical performance scores that fall in between the performances of the NDS and FDS homogeneous teams. For teams of six and eight UVs, operator utilization is saturated in the heterogeneous case for all three switching strategies, just like in the homogeneous FDS case (Figure 8b). However, performance for the homogeneous FDS case exceeds the performance in the heterogeneous cases. This is expected since the heterogeneous operators have a higher task load due to the inclusion of NDS vehicles in the teams, which require more regular operator attention. In the homogeneous FDS case, operators take advantage of the reduced task load by doing extra re-planning, which results in improved performance.

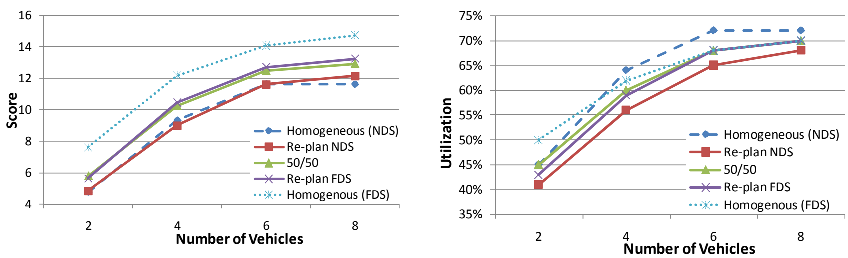
Thus our model predicts that vehicle team size and the level of decision support will have a more profound impact on system performance for this particular search-and-rescue mission than does the operator's choice of switching strategy. For this simulated mission, the system was relatively robust in terms of operator strategy, which should be a design goal in human supervisory control systems. Because military operators receive significant on-the-job training and experience high turnover rates, a well-designed system that can tolerate wide variability in strategies is crucial for future multiple-UV systems.

Management Strategies. Recall that an operator's management strategy refers to how likely he/she is to re-program a UV at any point in time. In the heterogeneous case, it can be important to use different management strategies for different UV systems. For example, a highly autonomous UV may need to be re-planned at a different rate than a less autonomous UV.

For the heterogeneous model predictions, three management strategies are analyzed. In the first strategy (referred to as *50/50*), the operator applies an equal management strategy to UVs with NDS and FDS capabilities. In the second strategy (called *Re-plan NDS*), the operator chooses to re-plan UVs with NDS capabilities, but not UVs with FDS capabilities. In the third strategy (called *Re-plan FDS*), the operator re-plans UVs with FDS capabilities, but not UVs with NDS capabilities.

The predicted system performance and operator utilization of these teams with the three operator strategies are compared to NDS and FDS homogeneous teams in Figure 9. Figure 9a shows that as with the previous results, the homogeneous team operators with full decision support achieved the highest performance due to the lower task and mission complexity. In addition, those heterogeneous operators that focused on replanning the NDS vehicles performed at the same degraded level as those operators with a homogeneous team with no decision support. Interestingly, these same operators experienced the lowest utilization, so they had spare cognitive resources but performed at a degraded level because they did not efficiently allocate their attention. These results illustrate the need to provide decision support, regardless of the team composition, as well as directed decision support that makes it clear to the operator what tasks are the most critical.

Another interesting result is that when operators re-plan UVs with FDS capabilities (i.e., in 50/50 and Re-plan FDS), the performance of the team improves. These results imply that when transitioning from legacy non-highly autonomous vehicles to more autonomous vehicles, operators supervising mixed teams can achieve higher levels of performance, but that they need to understand when and how to override the automation when replanning.



(a) (b)

Figure 9. Comparison of (a) mean performance scores, and (b) mean utilizations when using alternate management strategies for teams with 2, 4, 6, and 8 vehicles.

Conclusions

In this paper, a discrete-event simulation model was developed to investigate the effect of alternate operator strategies and team compositions in the supervisory control of multiple heterogeneous UVs. The model was used to replicate the system performance and operator utilization of multiple simulated homogeneous UV teams. Comparisons of these results with those observed in the human-in-the-loop experiments show that the model adequately captures these system dynamics for these homogeneous cases.

Because this model allows for alternative configurations both in vehicle assignments and in operator strategies, it is useful for modeling other system configurations. Along these lines, given the search-and-rescue setting of our simulation, we demonstrated that vehicle team size has a large impact on system performance, as compared to an operator's switching strategy. These results are paralleled in the air traffic control domain where number of aircraft has been shown to be the primary source of complexity for an air traffic controller (Kopardekar 2003). Moreover, our results showed that the presence of some automated decision support improves operator performance, but this is tempered by the need for operators to understand when and how to override the automation.

While these discrete event simulation predictions are intriguing, they are not yet validated with an actual system, primarily because single operator control of multiple UAVs in a search and rescue task is still not an operational reality, much less the control of heterogeneous UVs (including those from the ground and underwater domains). Work is underway through funding from the Office of Naval Research to build such a system in order to test these predictions as well as control algorithms and improved human interaction interfaces.

Acknowledgements

The research was supported by Charles River Analytics, the Office of Naval Research (ONR), and MIT Lincoln Laboratory.

References

Alberts, David S., John J. Garstka, and Frederick P. Stein. 1999. *Network Centric Warfare: Developing and Leveraging Information Superiority*. CCRP.

Alighanbari, Mehdi, and Jonathan P. How. 2005. Cooperative Task Assignment of Unmanned Aerial Vehicles in Adversarial Environments. Paper read at IEEE American Control Conference.

Broadbent, D. E. 1958. *Perception and Communication*. Oxford: Pergamon.

Crandall, J. W., and M. L. Cummings. 2007a. Identifying predictive metrics for supervisory control of multiple robots. *IEEE Transactions on Robotics – Special Issue on Human-Robot Interaction* 23 (5):942-951.

Crandall, Jacob W., and M. L. Cummings. 2007b. Attention Allocation Efficiency in Human-UV Teams. In AIAA *Infotech@Aerospace Conference*. Rohnert Park, CA.

Crandall, Jacob W., Michael A. Goodrich, Dan R. Olsen, and Curtis W. Nielsen. 2005. Validating Human-Robot Interaction Schemes in Multitasking Environments. *Systems, Man, and Cybernetics* 35 (4):438-449.

Cummings, M. L., and S. Guerlain. 2007. Developing Operator Capacity Estimates for Supervisory Control of Autonomous Vehicles. *Human Factors* 49 (1):1-15.

Cummings, M. L., and P. J. Mitchell. 2008. Predicting Controller Capacity in Supervisory Control of Multiple UAVs. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 38 (2):451-460.

Cummings, M. L., Carl E. Nehme, Jacob Crandall, and P. J. Mitchell. 2007. Predicting Operator Capacity for Supervisory Control of Multiple UAVs. In *Innovations in Intelligent UAVs: Theory and Applications*, edited by L. Jain.

Endsley, M. R. 1995. Toward a theory of situation awareness in dynamic systems. *Human Factors* 37 (1):32-64.

Endsley, Mica R. 1993. Situation Awareness and Workload: Flip Sides of the Same Coin. Paper read at 7th International Symposium on Aviation Psychology, at Columbus, Ohio.

Feickert, Andrew. 2005. The Army's Future Combat System (FCS): Background Issues for Congress: Congressional Research Service.

Goodrich, Michael A., Dan R. Olsen, Jacob W. Crandall, and Thomas J. Palmer. 2001. Experiments in Adjustable Autonomy. Paper read at IJCA101 Workshop on Autonomy, Delegation, and Control: Interacting with Autonomous Agents, at Seattle.

Goodrich, Michael A., Morgan Quigley, and Keryl Cosenzo. 2005. Task Switching and Multi-Robot Teams. Paper read at 3rd International Workshop on Multi-Robot Systems, at Washington D.C.

Kopardekar, P. 2003. Measurement and Prediction of Dynamic Density. Paper read at ATM 2003 R&D Seminar, at Budapest.

Law, A. M., and W. D. Kelton. 2000. *Simulation modeling and analysis*. 3rd ed. Boston: McGraw-Hill International Series.

Lazowska, E. D., J. Zahorjan, G. S. Graham, and K. C. Sevcik. 1984. *Quantitative System Performance: Computer System Analysis Using Queueing Network Models*. Upper Saddle River, NJ: Prentice Hall, Inc.

Mau, Sandra, and John M. Dolan. 2006. Scheduling to Minimize Downtime in Human-Multirobot Supervisory Control. In *International Workshop on Planning and Scheduling for Space*. Baltimore, Maryland.

Miller, Christopher A., and Raja Parasuraman. 2007. Designing for Flexible Interaction Between Humans and Automation: Delegation Interfaces for Supervisory Control. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49 (1):57-75.

Naval Studies Board. 2005. Autonomous Vehicles in Support of Naval Operations. Washington DC: National Research Council.

Office of the Secretary of Defense. 2005. Unmanned Aircraft Systems (UAS) Roadmap, 2005-2030. Washington DC DoD.

Olsen, D. R., and M. A. Goodrich. 2003. Metrics for Evaluating Human-Robot Interactions. In *Performance Metrics for Intelligent Systems*. Gaithersburg, MD.

Olsen, D. R., and S. B. Wood. 2004. Fan-out: Measuring Human Control of Multiple Robots. In SIGCHI *conference on Human factors in Computing Systems*. Vienna, Austria.

Parasuraman, Raja, Michael Barnes, and Keryl Cosenzo. 2007. Adaptive Automation for Human-Robot Teaming in Future Command and Control Systems. *The International C2 Journal* 1 (2).

Pina, P. E., M. L. Cummings, Jacob W. Crandall, and Mauro Della Penna. 2008. Identifying Generalizable Metric Classes to Evaluate Human-Robot Teams. Paper read at Metrics for Human-Robot Interaction Workshop at the 3rd Annual Conference on Human-Robot Interaction, at Amsterdam, Netherlands.

Pinedo, M. 2002. *Scheduling: Theory, Algorithms, and Systems*. 2nd ed. Englewood Cliffs, NJ: Prentice Hall.

Rogers, R. D., and S. Monsell. 1995. Costs of a Predictable Switch between Simple Cognitive Tasks. *Journal of Experimental Psychology: General* 124 (2):207-231.

Rouse, W. B. 1983. *Systems Engineering Models of Human-Machine Interaction:* New York: North Holland.

Ruff, H. A., S. Narayanan, and M. H. Draper. 2002. Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles. *Presence* 11 (4):335-351.

Schmidt, D. K. 1978. A Queuing Analysis of the Air Traffic Controller's Workload. IEEE *Transactions on Systems*, Man, and Cybernetics 8 (6):492-498.

Sheridan, Thomas B. 1992. *Telerobotics, Automation, and Human Supervisory Control* Cambridge, MA: MIT Press.

Sheridan, Thomas B., and M. Kamil Tulga. 1978. A Model for Dynamic Allocation of Human Attention Among Multiple Tasks. Paper read at 14th Ann. Conf. on Manual Control, at NASA. Ames Res. Center.

Squire, Peter, Greg Trafton, and Raja Parasuraman. 2006. Human Control of Multiple Unmanned Vehicles: Effects of Interface Type on Execution and Task Switching Times. Paper read at HRI, at Salt Lake City, Utah.

Valenti, M., B. Bethke, J. How, D. Pucci de Farias, and J. Vian. 2007. Embedding Health Management into Mission Tasking for UAV Teams. Paper read at IEEE American Controls Conference, at New York, NY.

Welford, A. T. 1952. The psychological refractory period and the timing of high-speed performance – a review and a theory. *British Journal of Psychology* 43:2-19.

Yerkes, R. M., and J. D. Dodson. 1908. The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology* 18:459-482.