5-7-2019

Identifying a Possible Function for Artificial Agent Adaptation in Variable Task Rate Environments

David J. Canzonetta

Michael E. Miller

IDENTIFYING A POSSIBLE FUNCTION FOR ARTIFICIAL AGENT ADAPTATION IN VARIABLE TASK RATE ENVIRONMENTS

David J. Canzonetta and Michael E. Miller
Air Force Institute of Technology
Wright-Patterson, AFB, OH

The current research sought to identify a method to calculate agent response time (ART) as a function of inter-arrival time (IAT), which balances human-agent team performance, human engagement, and human workload. A human-in-the-loop experiment evaluated human-agent team performance, as measured by team score, human engagement, as measured by the number of manually performed tasks, and workload, as measured through a subjective questionnaire, as a function of IAT and ART combination. Results demonstrated that task IAT strongly correlated with performance, engagement, and workload, while ART strongly related to engagement. Optimization was applied to the resulting data to determine ARTs which maximized performance while sustaining desirable levels of human engagement and workload. The optimization produced an ART function for application in future research to judge the effectiveness of adapting ART to boost human-agent team performance.

Humans and artificial agents can be teamed together to complete intricate and vital tasks. Successful task completion relies on the balance of human engagement and workload within these teams. For example, an unengaged human operator experiencing underload can face decreased alertness (Parasuraman, 2008). Dynamic function allocation is a common adaptive automation method for maintaining proper workload balance (Schneider, Bragg, Henderson, & Miller, 2018). However, this type of function allocation can force the human to maintain awareness of their present tasks within the current allocation, effectively increasing mental workload (Kaber, Riley, Tan, & Endsley, 2001).

Previous research conveyed that agent responsiveness within a human-agent team can affect human engagement (Goodman, Miller, Rusnock, & Bindewald, 2017). This discovery suggests that a well-timed agent response could provide an alternative approach to achieving the proper balance between human engagement and human workload in systems employing adaptive automation. For situations where environmentally-imposed inter-arrival time (IAT) heavily influences operator workload, calculation of optimal agent response time (ART) as a function of IAT becomes a possible method for task load sharing. The current study varied IAT and ART, measuring their effects on human-agent team performance, human engagement, and human workload. The data collected from this study produced a function for desired ART as a function of IAT to support future research.

Method

Participants

The experiment involved 14 participants (9 male and 5 female). Two participants were left-handed. Mean participant age was 25.4 and ranged from 20 to 31. One participant had previous Space Navigator experience. All but one participant exhibited normal color vision using the Ishihara Color Deficiency Charts (Ishihara, 2012). The participant with apparently irregular color vision obtained the third highest recorded score, indicating their ability to successfully identify the items in the game. Therefore, the analysis included their data. Participants self-reported spending an average of 48.7 hours per week using a computer or similar machine.

Apparatus and Environment

The experiment used a touch-screen tablet application titled "Space Navigator." Space Navigator closely resembles commercially-available air-traffic-control games. In this game, a human and agent work together as peers to achieve the highest score possible. The object of Space Navigator is to navigate red, blue, yellow, or green ships that spawn onto the screen to planets of their corresponding color, while obtaining randomly-appearing bonuses during their routes. The human-agent team receives 100 points upon successful navigation of ships to their corresponding planet. Ships are removed from the screen when they arrive at their appropriate planet. Additionally, the human-agent team receives 50 points for navigating ship paths through bonuses that appear on the screen. A bonus appears at a random on-screen location once every 10 seconds and remains on-screen until collected by a ship. The team loses 200 points when two ships collide. The human can physically draw a ship path with their finger, but if the human does not draw a path within a specified time window, the artificial agent presents a straight-line path from the ship to its appropriate planet. However, this agent path does not account for any bonuses or the paths of any other ships on screen. The human can draw or redraw a route at any time. The agent cannot overwrite a human-drawn route. Participants played all games on a Microsoft Surface Pro 4 in a quiet and secluded location.

Experimental Design and Procedure

The input variables to this study were agent response time (ART) and inter-arrival time (IAT). ART is the time an agent waits to draw a route for a new ship. IAT is the number of seconds between the times that two subsequent ships appear. Previous research narrowed and tested a range of IAT and ART values from 2s to 4s and 2.6s to 8.6s, respectively (Schneider et al., 2018). This research analyzed how the ratio of ART to IAT, referred to as the Adaptation Coefficient (AC), affects score, engagement, and workload (Schneider et al., 2018).

Decreasing IATs result in more ships appearing within a given time. This has the apparent and desired effect of increasing task load by requiring the human-agent team to provide more routes within a given time interval. Since these ships remain in the environment for a period of time to transit to their destination planet, the density of ships in the environment increases, increasing the probability of collisions, and reducing the number of possible collision free routes within the environment. This effect further increases task load as the human must draw or redraw longer and more complex routes.

Figure 1 displays the IAT and ART points used in this experiment, illustrated by points with markers "x" and "o", respectively. Past studies narrowed the sampling area to boundaries and points featured in Figure 1 by demonstrating team performance in the experiment environment remained similar for IAT values greater than 3.4s (Goodman et al., 2017; Schneider et al., 2018). The dashed lines that create the top and bottom boundaries represent AC of 2.0 and 0.5, respectively. These AC were chosen because they represent locations of manageable human workload in the Space Navigator environment, as discovered in previous research (Schneider et al., 2018), although human-agent team performance varied within this range. When IAT is significantly less than 2.6s, the human will struggle to keep up with new tasks, thereby experiencing overload. When IAT is significantly greater than 2.6s, the human will experience large breaks between new tasks, thereby experiencing underload. As ART decreases, the human typically draws routes slower than the agent, which could prevent the human from drawing and thereby decrease human engagement. Conversely, as ART increases, the human can draw routes faster than the agent, so one might assume that human engagement increases.

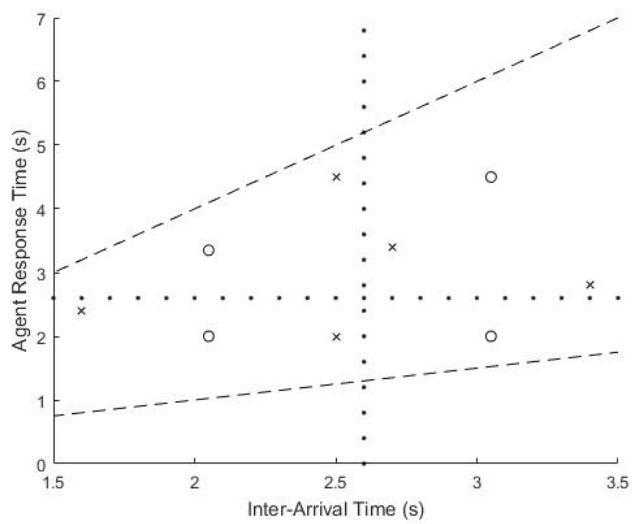

Figure 1. Depiction of Inter-Arrival Time (IAT) and Agent Response Time (ART) points sampled during the current experiment (shown as x's and o's). The vertical and horizontal dotted lines indicate the average human draw time of 2.6 s. The sloped dashed lines indicate a range of values useful for human-machine teaming based on previous research. Points marked with an "o" in Figure 1 represent the centroid of each region within the boundaries provided by the dashed and dotted lines. Points marked with an "x" were selected to be near the boundary extremes to provide insight into human performance near these transition regions.

For each experimental session, the research administrator provided a demonstration of Space Navigator to participants from a narrated script. The participants then played three, 2.5 minute practice rounds, each with an agent teammate, to become familiar with the Space Navigator environment. Practice rounds contained slower than average IAT and ART values to give participants time to understand the mechanics of the game and touchscreen response. Participants received no gameplay strategies during training.

The experimental session for each participant contained two blocks. Each block included nine, 1.75 minute trials with a workload questionnaire administered after each trial. Game time remained constant in all trials. Each block presented each input point described in Figure 1 to participants in a random order. A five-minute break separated the two blocks.

Data Analysis

Each experimental round contained the same game duration but employed different IAT. Thus, a different number of ships appeared in each experimental round. Therefore, it was inappropriate to compare the number of routes drawn and the total score across each experimental round as changes in IAT influenced these variables. To account for this difference, performance was measured as the percentage of the maximum possible score obtained in a game. Furthermore, engagement was calculated through two measures: human draws (HD) per ship and HD per second. When experiencing small IATs, the user may struggle to draw a route for every ship, even if the user desires to draw a manual route per ship. However, this does not mean the user is less engaged in the task than rounds where the user is physically capable of drawing a route for every ship. Therefore, it was desirable to use HD per second to measure overall engagement of a human at each IAT and ART point. However, HD per ship still proved useful for defining thresholds (i.e. we can say the human must at least engage with one in every five ships). Workload was measured using a subjective questionnaire containing three questions from NASA-TLX on a 0-20 scale. These questions were selected as previous studies found a correlation between the workload categories of temporal demand, effort, and performance with changes in IAT (Schneider et al., 2018). Workload values were standardized using min-max normalization within each participant to allow comparison across all participants. Total workload for a single Space Navigator round was calculated as the sum of the normalized workload values for each of the three workload questions.

Relationships between our independent and dependent variables were investigated using multiple regression analysis. This analysis contained two steps. First, multiple regression analysis on output variables was conducted to the third order. Second, insignificant effects were removed one at a time until only significant effects remained. Regression analysis was applied for each output variable across all participants. If large participant variability caused no significance for IAT and ART across all participants, regression analysis was conducted on the mean output values for each input IAT and ART combination.

Results and Discussion

Table 1 displays correlations of IAT and ART with human-agent team performance, human engagement, and workload. Results indicated IAT strongly correlates with score ($r(8) = 0.9229, p = 0.0004$), engagement ($r(8) = -0.7969, p = 0.0642$), and workload ($r(8) = -0.9578, p < 0.0001$). Results also indicated that ART strongly correlates with engagement ($r(8) = 0.8481, p = 0.0039$). From Table 1, it becomes evident that as IAT increases, the percent of maximum possible score increases, human draws per ship increases, and workload increases. Additionally, Table 1 illustrates that as ART increases, participant engagement with the system increases. These results are consistent with data obtained in preceding research (Schneider et al., 2018).

Table 1.
Correlations between variables. Values in bold represent significant correlation at α = 0.05. Italicized data points represent significant correlation at α = 0.10.

	Avg. % Max Score	Avg. HD per Ship	Avg. HD per Sec	Avg. Std Workload
IAT (IV)	**0.9229**	*0.6385*	**-0.7969**	**-0.9578**
ART (IV)	0.0018	**0.8481**	0.3955	0.0006

Multiple regression analysis on the data across all participant trials indicated that there was a collective significant effect between IAT and ART on percentage of max score, F(5, 246) = 25.4565, $p < 0.0001$, $R^2 = 0.3410$. Further examination of the predictors indicated that IAT (t = 6.14, $p < 0.0001$, β = 0.1404), IAT to the second degree (t = -3.42, p = 0.0007 , β = -0.1288), ART (t = -3.15, p = 0.0018, β = -0.1263), ART to the second degree (t = -2.96, p = 0.0034, β = -0.0812), and ART to the third degree (t = 2.84, p = 0.0049 , β = 0.0757) were significant predictors in this model.

Multiple regression on data across all participant trials indicated there was a collective significant effect between IAT and ART on human engagement represented as human draws per ship, F(2, 249) = 16.1716, $p < 0.0001$, $R^2 = 0.1150$. Further examination of the predictors indicated that IAT (t = 2.78, p = 0.0058, β =0.0890) and ART (t = 4.29, $p < 0.0001$, β = 0.0746) were significant predictors in this model.

Multiple regression analysis to on data across all ART and IAT combination averages indicated there was a significant effect between IAT and ART on workload, F(4, 4) = 130.1843, p = 0.0002, $R^2 = 0.9924$. Further examination of the predictors indicated that IAT (t = -18.71, $p <$ 0.0001, β = -0.2345), ART (t = 4.94, p = 0.0078, β = 0.0377), ART to the second degree (t = -3.39, p = 0.0275, β = -0.0265), and the interaction of ART and IAT (t = 3.08, p = 0.0370, β = 0.0535) were significant predictors in this model

Derivation of Near-Optimal Agent Response Function

To determine the optimal ART, the regression equations derived in the previous section were applied within an optimization problem. The optimization problem was solved for the ART at each IAT value between zero and four seconds on a 0.001s interval. This optimization sought to maximize the percentage of maximum score subject to the constraints that the participant would draw at least one route for every five ships and would have a mean standardized workload between the mean, plus or minus one standard deviation of the workload from this experiment (between 0.423 and 0.561).

The optimization determined that when IAT is less than approximately 1.5s, the optimal ART is 0s. In this range, IAT is much lower than the average human response time, so the human will likely struggle to match the pace at which new tasks appear. Therefore, the human will likely require shorter ART. Once IAT is greater than 1.5s, the ART increases as IAT increases, permitting the human to take on a more involved role since they can better keep up with the slower rate at which tasks appear. As IAT approaches the average human response time, it disrupts the linear function. This permits a constant ART for IAT near the average human response time. ART then continues to increase once IAT is greater than the average human response time. Violation of the constraints specified in the function occurred at IAT greater than 3s. For this reason, ART at IAT greater than 3s was extrapolated from the function starting at IAT of 2.7s. As IAT increases from 2.7s, the human has more time to complete present tasks until the next ship arrives. Therefore, human need for agent assistance remains low at IAT levels greater than 2.7s. Optimization produces a piecewise linear function for the calculation of the optimal ART based on IAT. Equation 1 provides this piecewise linear function.

$$\begin{aligned}
&For\ IAT < 1.485, ART = 0 \\
&For\ 1.485 \leq IAT < 2.206, ART = 3.5327 * IAT - 5.2461 \\
&For\ 2.206 \leq IAT < 2.735, ART = 2.5471 \\
&For\ IAT \geq 2.735, ART = 5.2807 * IAT - 11.8955
\end{aligned} \tag{1}$$

Conclusion

Results from this study indicate that IAT is strongly correlated with human-agent team performance, human engagement, and workload. Furthermore, ART is correlated with human engagement. This study produced a method for computing ART as a function of IAT. The ART function was obtained by gathering data at logical IAT and ART points and calculating which ART produced the maximum percentage of possible team score while following workload and engagement constraints. The proposed ART function will be applied in subsequent research to determine if ART calculated from IAT can effectively balance workload and engagement while maintaining equal or better performance than a constant ART agent.

Disclaimer and Acknowledgement

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Air Force, Department of Defense, nor the U.S. Government. The authors gratefully acknowledge the financial support of the Air Force Office of Scientific Research, Computational Cognition and Machine Intelligence Program.

References

Goodman, T. J., Miller, M. E., Rusnock, C. F., & Bindewald, J. M. (2017). Effects of agent timing on the human-agent team. *Cognitive Systems Research*, *46*, 40–51. https://doi.org/10.1016/j.cogsys.2017.02.007

Ishihara, S. (2012). Ishihara's design charts for colour deficiency of unlettered persons. Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0,36&q=ishihara%27s+design+charts+for+colour+deficiency

Kaber, D. B., Riley, J. M., Tan, K.-W., & Endsley, M. R. (2001). On the Design of Adaptive Automation for Complex Systems. *International Journal of Cognitive Ergonomics*, *5*(1), 37–57. https://doi.org/10.1207/S15327566IJCE0501_3

Parasuraman, R. (2008). *Supporting Battle Management Command and Control: Designing Innovative Interfaces and Selecting Skilled Operators*. Fairfax, VA. Retrieved from http://www.dtic.mil/docs/citations/ADA480645

Schneider, M. F., Bragg, I. L., Henderson, J. P., & Miller, M. E. (2018). Human Engagement with Event Rate Driven Adaptation of Automated Agents. In *2018 IISE Annual Conference*. Orlando, FL.