Proceedings of the Human Factors and Ergonomics Society Annual Meeting
2023, Vol. 67(1) 2128–2136

DOI: 10.1177/21695067231192549
journals.sagepub.com/home/pro
Sage

Development of a Real-Time Trust/ Distrust Metric Using Interactive Hybrid Cognitive Task Analysis

Shwen Zhou[1], Xioyun Yin[1], Matthew J. Scalia[1], Ruihao Zhang[1], Jamie C. Gorman[1], and Nathan J. McNeese[2]

Abstract

While there is increased interest in how trust spreads in Human Autonomy Teams (HATs), most trust measurements are subjective and do not examine real-time changes in trust. To develop a trust metric that consists of objective variables influenced by trust/distrust manipulations, we conducted an Interactive hybrid Cognitive Task Analysis (IhCTA) for a Remotely Piloted Aerial System (RPAS) HAT. The IhCTA adapted parts of the hybrid Cognitive Task Analysis (hCTA) framework. In this paper, we present the four steps of the IhCTA approach, including 1) generating a scenario task overview, 2) generating teammate-specific event flow diagrams, 3) identifying interactions and interdependencies impacted by trust/distrust manipulations, and 4) processing RPAS variables based on the IhCTA to create a metric. We demonstrate the application of the metric through a case study that examines how the influence of specific interactions on team state changes before and after the spread of distrust.

Keywords

Cognitive task analysis, Team performance/teamwork, Trust

Introduction

As artificial intelligence technologies advance, how trust spreads in Human-Autonomy Teams (HATs) becomes crucial to measure (Chen, 2018). A HAT constitutes humans and at least one autonomous teammate, defined as a type of technology that works with humans as an equal to perform essential tasks, make decisions, execute actions on its own, and communicate information with other teammates to complete tasks (McNeese et al., 2018). Trust between humans is defined by Mayer et al. (1995) as the "willingness of a party to be vulnerable to the action of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (p. 712). Trust between humans and autonomy is defined by Lee and See (2004) as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (p. 54). In a HAT, trust must not only be captured by the relationships between humans and autonomy but also between human teammates and autonomous teammates (Schelble et al., 2022).

There are few objective metrics for representing and measuring team dynamics (Kozlowski & Chao, 2018), which include trust in team dynamics. Current measurement tools of trust may only be sufficient to track gains and losses of trust between humans and trust in autonomous teammates resulting from human-human and human-autonomy interactions retroactively, primarily through subjective assessments (Schaefer et al., 2016). There are some behavioral and physiological evaluations of trust (Schaefer et al., 2016); however, none of these measures can be considered objective, real-time measures of system-level trust. We aimed to adopt a hybrid cognitive task analysis (CTA) framework to identify variables in HATs' interactions for measuring dynamic changes in trust and distrust in HATs (i.e., trust/distrust dynamics).

CTA can be used to elicit an understanding of how people successfully complete tasks and what may prevent them from completing their tasks (Crandall et al., 2006). However, traditional CTA methods require subject matter experts and previous implementations to generate assumptions for the current system and therefore, cannot be applied to futuristic systems (Scott et al., 2005). A hybrid CTA framework extends traditional CTA by analyzing a system

[1]Human Systems Engineering, Arizona State University, Mesa, AZ, USA
[2]Department of Human Centered Computing, Clemson University, Clemson, SC, USA

Corresponding Author:
Shiwen Zhou, Human Systems Engineering, Arizona State University, Interdisciplinary Science and Technology Building III 7418 E. Innovation Way South, Mesa, AZ 85212, USA.
Email:shiwen.zhou@asu.edu

from a high-level task goal or a scenario description. This framework can provide information requirements and recommendations for futuristic systems (Nehme et al., 2006). This paper adopts the hybrid CTA approach as we anticipated developing a trust/distrust metric to assess not only how trust and distrust spread in a single HAT but also futuristic multi-HAT constellations. In this paper, we present a hybrid CTA conducted on a HAT system to develop an objective, real-time measure of trust/distrust dynamics. A case study then describes a measure generated from the hybrid CTA. We propose that this combined approach can be used in other HAT settings in which measuring trust/distrust dynamics is key to system effectiveness.

Interactive Hybrid CTA

The hybrid CTA approach was initially developed for designing futuristic interfaces (Nehme et al., 2006). Our Interactive hybrid CTA (IhCTA) approach focuses on team interaction and adopts the first step of the original hybrid CTA by generating a scenario task overview. The second step combines components of an event flow diagram and a decision ladder to create teammate-specific event flow diagrams. The third step of the IhCTA combines the teammate-specific event flow diagrams to identify teammate interactions that overlap in time (i.e., "intersections") that should be impacted by the spread of trust and/or distrust. Finally, the last step uses the identified intersections to combine system variables to create a real-time trust/distrust metric. Each of these steps is described in detail in the context of an experimental task in which human and/or Wizard of Oz (WoZ) autonomy can spread trust and/or distrust to human participants through communication and behavior.

RPAS IhCTA

The IhCTA was performed on the Cognitive Engineering Research on Team Tasks-Remotely Piloted Aircraft System-Synthetic Task Environment (CERTT-RPAS-STE; Cooke & Shope, 2005). The CERTT-RPAS-STE comprises three task-role stations connected via chat communications for conducting three-person team missions. The three roles are navigator (DEMPC), pilot (AVO), and payload operator (PLO). The mission goal is to take photographs of color-coded strategic target waypoints while avoiding color-coded hazard waypoints. In the current scenario, either the AVO or DEMPC can be played by WoZ autonomous agents. The IhCTA included six event types with six event processes and corresponding action tables. IhCTA legends are described in Table 1. These legends are used in steps 2, 3, and 4 of the IhCTA

Step 1: Scenario Task Overview

The purpose of generating the scenario task overview is to determine a definition of the mission statement in terms of different phases and different tasks in each phase. The phase goals and task sub-goals are defined to create a foundation for the subsequent task analysis steps.

For each target, there are three phases that the team needs to go through. First, an information phase (I) begins when DEMPC plans a route and ends when DEMPC provides information about a waypoint to other teammates. The phase goal for the I phase is for all team members to have the waypoint information that is sufficient for them to proceed with the task. The task sub-goal is for DEMPC to share target information and restrictions with other teammates successfully.

A negotiation phase (N) is next, which begins when AVO sets and provides a target's airspeed and altitude to PLO. PLO may need to negotiate airspeed and/or altitude adjustments with AVO according to the PLO's camera setting requirements. The N phase ends when AVO finalizes the airspeed, altitude, and bearing. The phase goal for the N phase is for the RPA to have stable airspeed, altitude, and bearing before entering the effective radius of the target. The task sub-goal is for AVO to set airspeed and altitude to fit the PLO's photograph requirements.

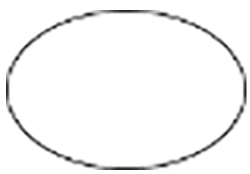
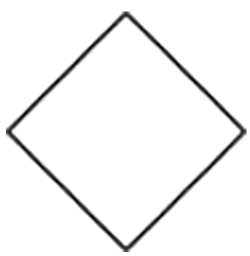
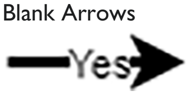
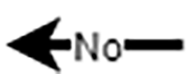
Lastly, there is a feedback phase (F), which begins when the RPA enters the effective radius of a target and ends when PLO sends feedback to both DEMPC and AVO that a good photo of that target has been taken. The goal for the F phase is for all team members to know that there is a good photo and that the team can proceed to the next waypoint. The task sub-goal is for PLO to take a good photo and share feedback as soon as possible. The mission consists of these three phases looping for every waypoint until the team reaches the last waypoint or until mission time (40 min) runs out. These loops are detailed in the next section.

Step 2: Teammate-Specific Event Flow Diagrams

After the task scenario is outlined, event flow diagrams are constructed to examine how events are not only temporally related within each teammate's actions but also interconnected across teammates. Teammate-specific event flow diagrams were created to represent each team member's tasks during each mission. Figure 1 illustrates part of an event flow diagram. Complete diagrams can be found at http://bit.ly/3YpYJTY.

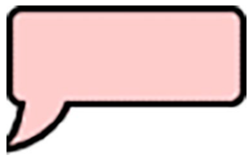
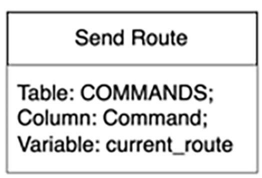
Navigator (DEMPC). After the start of the mission, DEMPC enters the I phase. In this phase, the diagram depicts DEMPC creating and sending the initial route list to AVO. The I phase ends when DEMPC sends upcoming waypoint information to AVO and/or sends upcoming target information to PLO. At the N phase, if the upcoming waypoint is an entry or exit waypoint, DEMPC will wait for the RPA to enter the effective radius of the waypoint and then update the route. This loop sends DEMPC back to the I phase. If the upcoming waypoint is a target, DEMPC will wait for the RPA to fly into the effective radius of the target. Once the effective radius for the target has been reached, the F phase begins.

Table 1. IhCTA legends.

	Legend	Description
Event Flow Diagram	 Starting Points	Indicate the start of the task allocated to a specific teammate.
	 End Points	Represent the end of the allocated task.
	 Action Points	Behavioral actions taken by each teammate.
	 Knowledge States	Capture the teammate's information-processing activities.
	 Decision Points	Indicate when a teammate needed to make a knowledge-based decision.
	 Blank Arrows	These unlabeled arrows show the event flow temporally.
	 "Yes" Arrows	"Yes" labeled arrows are temporally constrained flows that originate from decision points. They incorporate temporal information and a knowledge-based decision.
	 "No" Arrows	"No" labeled arrows are temporally constrained flows that originate from decision points. They incorporate temporal information and a knowledge-based decision.

(continued)

Table 1. (continued)

	Legend	Description
	 Phases	Indicate the phases (i.e., I, N, F).
Intersection (Interdependency)	 INF action Indicators	Indicate the individual task that should be done within the information, negotiation phase, and feedback phase.
	 Communication Points	Indicate part of the event flow diagram where trust and distrust manipulations are spread between teammates via the communication between them.
	 Interaction Points	Indicate parts of the event flow diagram where trust and distrust may spread between teammates due to a specific teammate action involving the system in use.
Database Variables	 Actions Tables	For assigning Access Database Tables/Variables to Actions; these are used to identify variables in the system for input into the real-time trust metric.

At the F phase, DEMPC waits to receive a feedback message from PLO stating that a good photo of the target has been taken. Once that message is received, DEMPC determines if this is the last target for the mission. If it is the last target for the mission, DEMPC waits for the RPA to enter the effective radius of the last exit waypoint, and the mission ends. If the target is not the last target for the mission, DEMPC must update the route. Thus, the event flow diagram depicts DEMPC re-entering the I phase via a feedback loop. If, however, a feedback message is not received by DEMPC and the RPA has left the effective radius of the target, the DEMPC event flow diagram loops back to the N phase. As AVO flies the RPA back into the target's effective radius, the DEMPC event flow enters the F phase again. If the feedback message for the last target is received, the DEMPC event flow diagram ends along with the mission once the last exit waypoint is reached.

Pilot (AVO). When the mission starts, the AVO requests a route immediately and asks DEMPC for the first waypoint's information. The AVO event flow diagram marks this as the beginning of the I phase. Once DEMPC provides the waypoint's information and restrictions, AVO can set the waypoint the RPA goes to. AVO then asks DEMPC about the next waypoint's information and ensures that the CERTT-RPAS-STE system interfaces display the current waypoint it is heading to and the queued waypoint for all team members. AVO also sets airspeed, altitude, and bearing according to the information that DEMPC provides. If the RPA is heading to a nontarget, the I phase for AVO will loop back to the beginning of the I phase once the RPA enters the effective radius of a waypoint. If the RPA is heading to a target, the I phase for AVO will transition to the N phase when AVO receives information and restrictions for a target from DEMPC.

The AVO event flow diagram depicts the start of the N phase when the RPA is en route to a target waypoint. AVO sends the planned airspeed and altitude information to PLO and negotiates with PLO to determine if airspeed and/or altitude need adjustment. AVO ensures that all adjustments can be made before the RPA enters the effective radius of the target so that PLO has a stable environment to take a photo. The F phase shown in the AVO event flow diagram starts when the RPA enters the target's effective radius. AVO waits for PLO's feedback and might ask whether PLO has a good photo if PLO does not send feedback after a while. Under unexpected circumstances, the N phase may overlap with the F phase if AVO asks whether PLO has a good photo and PLO requests airspeed or altitude adjustment after that. In some extreme cases, when PLO cannot take a photo inside the effective radius due to a mismatch between the RPA's current altitude and the camera setting, the status will transition from the F phase to the N phase and back to the F phase once the RPA re-enters the target's effective radius. The F phase ends when AVO receives photo-taken feedback from PLO, and the flow will loop back to the I phase if the next waypoint is a target. If not, then the I phase restarts once the RPA again heads to a target waypoint.

Payload Operator (PLO). Once the mission starts, PLO can either use the "Entry Waypoint-Target Waypoints-Exit Waypoint" logic (i.e., after an exit waypoint is an entry waypoint followed by a target) to plan ahead to achieve maximum efficiency or wait to gather this information later during the mission. If the RPA is heading to a nontarget and PLO does not plan ahead, PLO will not enter any phase.

Inquiring DEMPC about the target waypoint's effective radius marks the I phase's beginning in the PLO event flow diagram. PLO also relies on the information provided by the CERTT-RPAS-STE system interface to determine the camera settings and desired altitude for each target. The N phase begins when PLO receives airspeed and altitude information for the target waypoint from AVO. PLO then negotiates with AVO to take a good photo based on the camera setting and the RPA status. PLO can also choose to request information about the effective radius of the target waypoint from DEMPC, therefore, overlaps between the I phase and the N phase can occur. Once PLO receives the effective radius from DEMPC and ensures the RPA's current status aligns with the camera settings and is stable, the event flow diagram moves to the F phase from the I phase and/or the N phase.

During the F phase, PLO can still negotiate with AVO to finalize the RPA's airspeed and altitude, creating an overlap between the N and F phases. When the RPA is inside the effective radius of the target, PLO can take a photo and decide if the photo should be accepted. Once a photo is accepted, PLO must report this to the other teammates, marking the end of the F phase. After sending the photo confirmation, the PLO event flow diagram either loops back to the I phase to continue the mission or ends after receiving DEMPC's notification about the mission's completion. PLO can also request AVO to fly back to the target if the RPA passes the target waypoint and is out of the effective radius, which creates a back-and-forth transition between the F phase and the N phase. Throughout the phases, PLO must also monitor and fix several alarms when they go off.

Step 3: Interaction and Interdependency

Interaction between teammates—which can include talking to one another, sharing tasks, etc.—is crucial for team performance. Specific jobs require more than one person to complete, which creates interdependency within the team and intersections of activities across diagrams.

Each target's INF loop includes communication among the three teammates. In a typical scenario, AVO and PLO need DEMPC to send the target's information. They usually ask DEMPC for that information (e.g., Figure 2). In some exceptional circumstances, AVO and PLO will alert DEMPC to send the data using a route request button, eliminating the need to send request messages via chat.

AVO and PLO engage the most during the N phase, wherein PLO depends on AVO to proceed (e.g., Figure 3). Depending on the trust/distrust manipulation, the AVO may use this opportunity to spread trust or distrust via behavior or communication. DEMPC may send information for the next waypoint while the AVO and PLO are in the N phase. This overlap will be described later. During the F phase, once a good photo has been taken, the entire team will be informed by PLO (e.g., Figure 4).

Interdependency occurs not only within a phase but between phases. For example, if a photo is not successfully taken, AVO and PLO will need to modify the altitude and airspeed once more and return to the target. Another example is when PLO takes a photograph, AVO can request the information for the next waypoint from DEMPC, corresponding to the interdependency between the I and N phases.

1. Step 4: Mapping Data Sources

The final step is to create action tables based on the interactions and interdependencies identified in step 3. These tables include the variables used for input into a real-time trust/distrust metric. The data are from simulated real-time sensors for communication events, user interface control inputs, RPA status, and system warning and alarm states stored in a Microsoft Access database. The complete diagram with action tables is available at http://bit.ly/3YpYJTY. Next, a case study demonstrates the final step of pulling data based on the IhCTA-developed action tables to analyze changes in patterns of team influence before and after spreading distrust. The distrust manipulation was constructed based on Lewicki et al. (1998)'s definition of distrust, which encompasses "the fear of, a propensity to attribute sinister

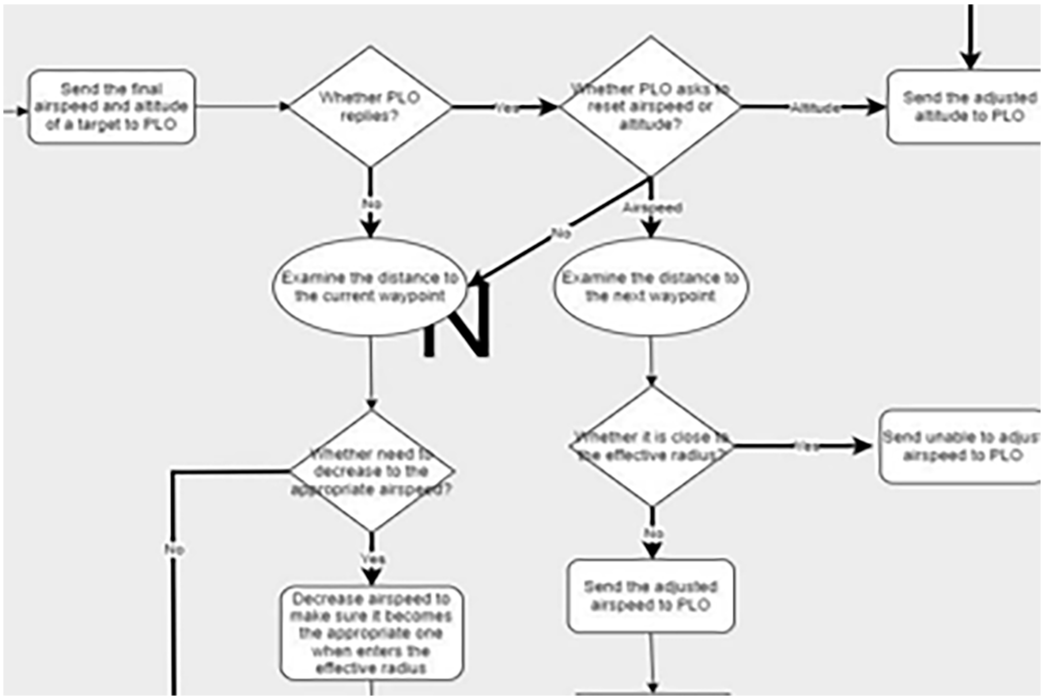


Figure 1. Partial illustration of a team-specific event flow diagram. Full diagrams can be found at http://bit.ly/3YpYJTY.

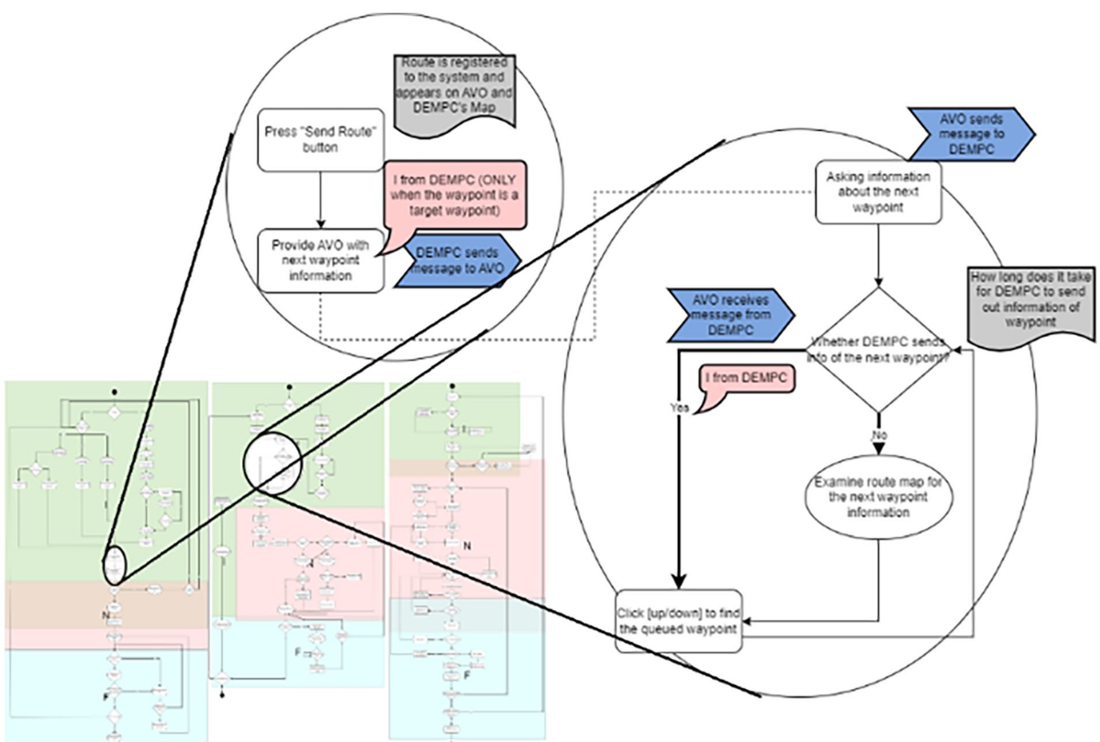


Figure 2. Example of interdependence (diagram intersection) between DEMPC and AVO during the information phase. Full diagrams can be found at http://bit.ly/3YpYJTY.

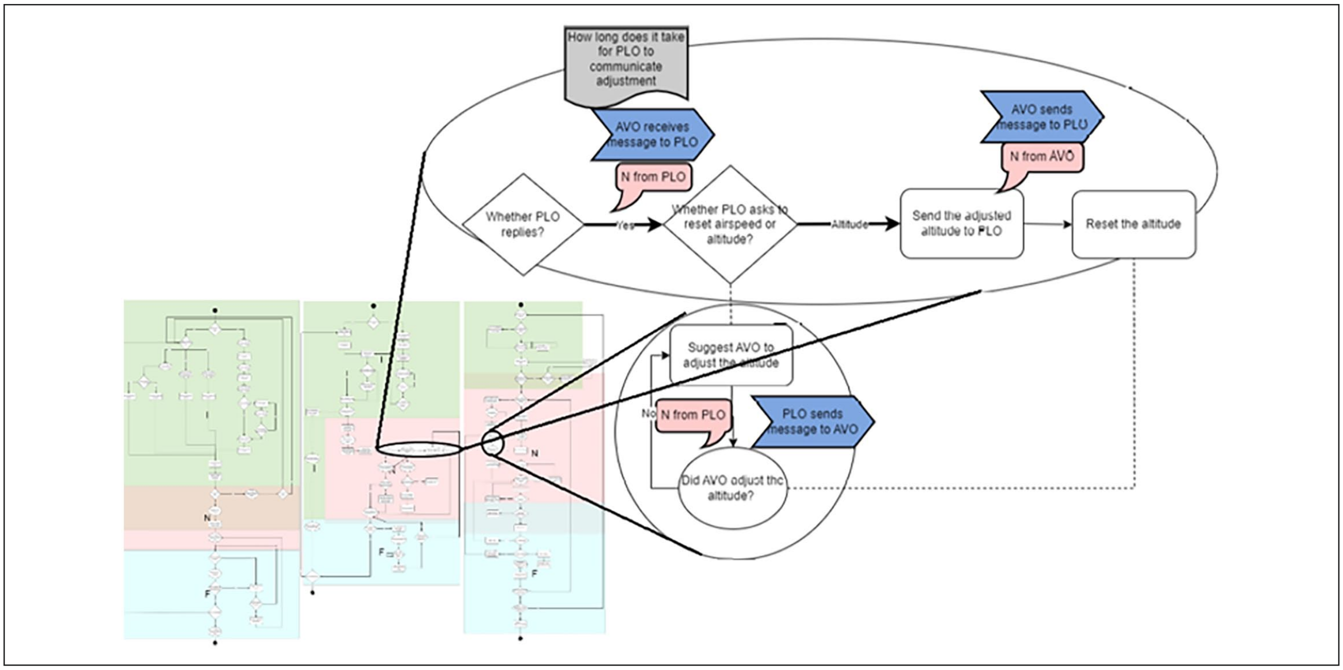


Figure 3. Example of interdependence (diagram intersection) between AVO and PLO during the negotiation phase. Full diagrams can be found at http://bit.ly/3YpYJTY.

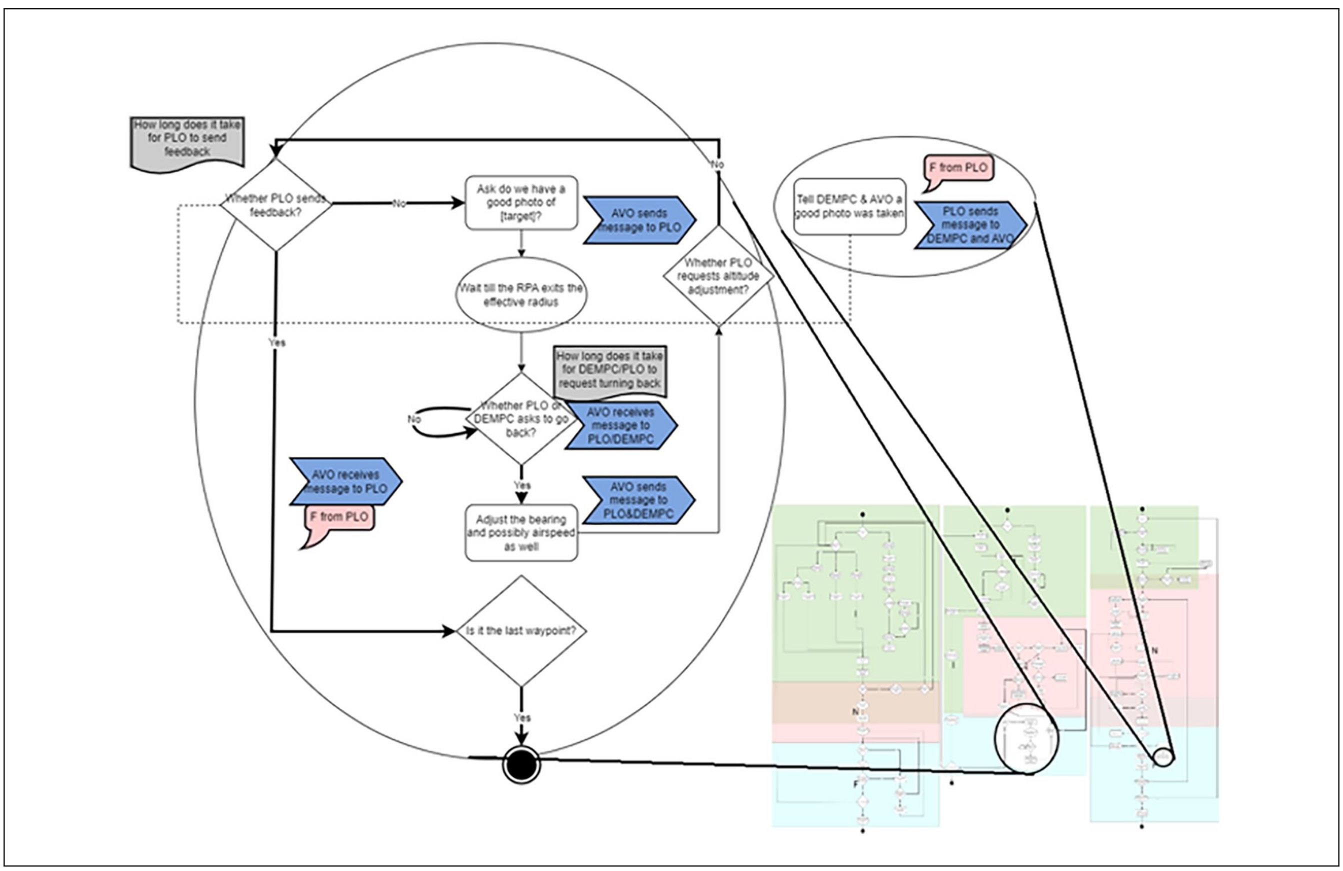


Figure 4. Example of interdependence (diagram intersection) between AVO and PLO during the feedback phase. Full diagrams can be found at http://bit.ly/3YpYJTY.

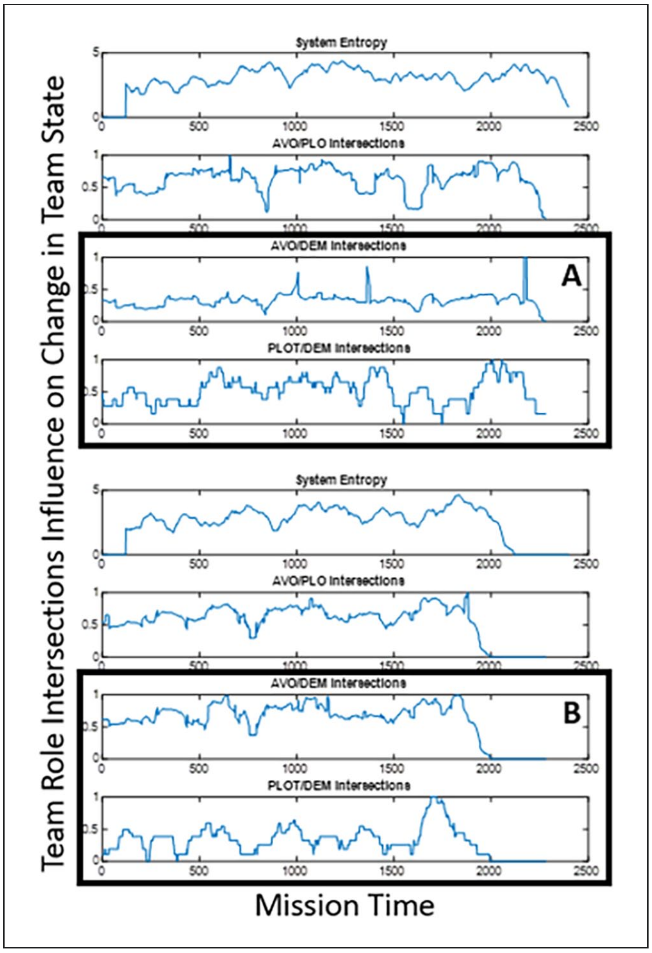


Figure 5. (A) unique team member intersections influence on overall system state prior to the spread of distrust; (B) the same intersections influence on system state after the spread of distrust.

intentions to, and a desire to buffer oneself from the effects of another's conduct" (p. 439).

Case Study

For any mission, unique intersections of team member states (e.g., AVO is turning right while DEMPC is sending the route; PLO taking a photo while AVO is ascending) were identified across all phases of the INF loop. These intersections comprise over a hundred unique combinations of system states that quantify changes in team member interdependency following the spread of trust and/or distrust. In this section, we present two example uses of a real-time trust/distrust metric that captures changes in how role intersections influence change in team state following the spread of distrust.

Figure 5 shows real-time trust metrics obtained from the data identified in the IhCTA action tables. Specifically, team member intersections were obtained by coding the

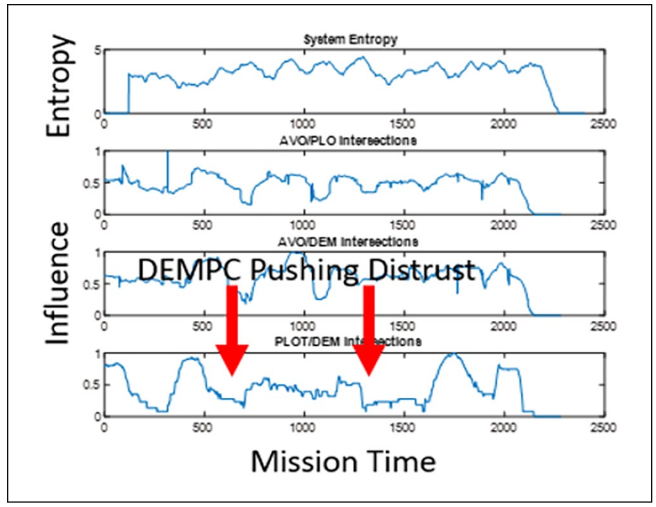


Figure 6. Within-mission changes in team member influence on team state in real-time.

event data in the IhCTA into mutually exclusive team member component states that can be added to obtain the intersection (at every second) of team members' actions. To measure how much each pair of team member intersections (e.g., AVO/PLO) influences overall system state, we computed the mutual information between the intersection and overall team state (i.e., the sum across all component states) at each second. Using this measure, we hypothesized that trust and distrust manipulations could be tracked in real-time as the altering of patterns of influence from different combinations of team member intersections on the system state, such that spreading distrust would result in less influence over time.

Figure 5A shows how the AVO/DEMPC intersections consistently influence system state, while PLO/DEMPC intersections have consistently moderate to large influence on the system state. The data in 5A were taken before the spread of distrust. Figure 5B shows a mission from the same team after the spread of distrust. Note the tradeoff between Figure 5A and 5B. After the spread of distrust, AVO/DEMPC intersections now have a more sizeable consistent influence on team state, while PLO/DEMPC influence on team state has diminished considerably.

Figure 6 illustrates another use of the metric. This example, taken from a different team, illustrates a within-mission drop in influence for PLO/DEMPC intersections following the spread of distrust from DEMPC to the human PLO.

Conclusion

The approach illustrated in this paper borrows from the hCTA framework to identify interaction-based variables for measuring the spread of trust or distrust. Using the action tables that IhCTA generates, metrics can be developed to measure real-time changes in influence across missions and within a mission due to spreading trust/distrust. This approach provides a foundation for developing objective real-time trust metrics.

Acknowledgments

This research was supported by AFOSR grant FA9550-21-1-0314. The authors are grateful to Yiwen Zhao, Anna Crofton, Kalia McManus, Edith Garner, Jessica Harley, Anya Polomis, Yawen Tan, Hruday Shah, and Vibha Mohan for their contributions to this research.

ORCID iD

Shiwen Zhou https://orcid.org/0000-0002-2851-1940

References

Chen, J. Y. (2018). Human-autonomy teaming in military settings. *Theoretical issues in ergonomics science*, *19*(3), 255-258.

Cooke, N. J., & Shope, S. M. (2004). Synthetic task environments for teams: CERTT's UAV-STE. In *Handbook of human factors and ergonomics methods* (pp. 476-483). CRC Press.

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Mit Press.

Kozlowski, S. W., & Chao, G. T. (2018). Unpacking team process dynamics and emergent phenomena: Challenges, conceptual advances, and innovative methods. *American Psychologist*, *73*(4), 576.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human factors*, *46*(1), 50-80.

Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *The Academy of Management Review*, *23*(3), 438–458. https://doi.org/10.2307/259288

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management review*, *20*(3), 709-734.

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human factors*, *60*(2), 262-273.

Nehme, C. E., Scott, S. D., Cummings, M. L., & Furusho, C. Y. (2006, October). Generating requirements for futuristic hetrogenous unmanned systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, No. 3, pp. 235-239). Sage CA: Los Angeles, CA: Sage Publications.

Schelble, B., Flathmann, C., Scalia, M., Zhou, S., Myers, C., McNeese, N., Gorman, J., & Freeman, G. (2022). Addressing the Spread of Trust and Distrust in Distributed Human-AI Teaming Constellations. *Computer-Human Interaction: Trust and Reliance in AI-Human Teams*.

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human factors*, *58*(3), 377-400.

Scott, R., Roth, E. M., Deutsch, S. E., Malchiodi, E., Kazmierczak, T. E., Eggleston, R. G., . . . Whitaker, R. D. (2005). Work-centered support systems: A human-centered approach to intelligent system design. *IEEE Intelligent Systems*, *20*(2), 73-8