Proceedings of the Modern Artificial Intelligence and Cognitive Science Conference. April 2015.

Measuring Trust in Automation: A New Approach

Joseph Nuamah, Seeung Oh, and Younho Seong

Department of Industrial and Systems Engineering

North Carolina A & T State University

1601 East Market Street, 406 McNair Hall

Greensboro, NC 27411

jknuamah@aggies.ncat.edu

Abstract

The trust that a decision maker has in an automated decision aid (ADA) may have serious consequences depending on the task he or she is performing. While many approaches exist to subjectively measure trust, there exists no known objective approach for measuring trust. Brain Computer Interface (BCI) techniques can produce remarkable images of brain structure and activity. Using Electroencephalography (EEG), the cognitive activities in the brain of decision makers can be objectively measured. We believe that brain images will be different for intuitive vs. analytic cognition and by correspondence intuitive trust vs. analytic trust.

Introduction

The correspondence between an ADA's capabilities and the decision maker's trust in that automation is critical for optimum performance (Bisantz et al. 1999). Trust in ADAs has both cognitive and affect components. Whereas the cognitive-based components are based on the decision maker's intellectual perceptions of the system's characteristics, the affect-based components are based on the decision maker's emotional responses to the system. Where the decision maker has insufficient knowledge upon which to base a cognitive decision, the overall affect-based component becomes the primary determinant of trusting behavior (Madsen and Gregor 2000). The Cognitive Continuum Theory (CCT) may be useful in evaluating trust. In this theory, human cognition dynamically varies on a continuum ranging from analytic to intuitive depending on the properties of the task (Hammond et al. 1987). Likewise, trust may be defined using a similar continuum which is termed the analytic-intuitive trust continuum. Trust is categorized and placed along a scale of trust continuum, varying from intuitive to analytic. Analytic trust refers to an aspect of trust that may be empirically established through interaction among humans and automated systems. Sample metrics include reliability and robustness. Intuitive trust, however, is based on human judgment and intuition and must be derived subjectively. Under these conditions, the Cognitive Continuum Theory may be useful in evaluating trust. In this theory, human cognition dynamically varies on a continuum ranging from analytic to intuitive depending on the properties of the task. Thus, trust may be defined using a similar continuum which is termed the trust continuum.

Trust Continuum

The understanding of the analytic-intuitive trust continuum could be examined using human perception of trust in automation. This perception may occur at levels on a continuous scale from low to high. For simplicity of measurement, a discrete binary metric of low trust and high trust is used in the proposed study. Thus, the interaction of the two factors on the trust continuum and on the two levels of trust perception yields a two dimensional metric of trust continuum and trust level. This leads to the development of four quadrants on a trust grid demonstrated in Figure 1 below.

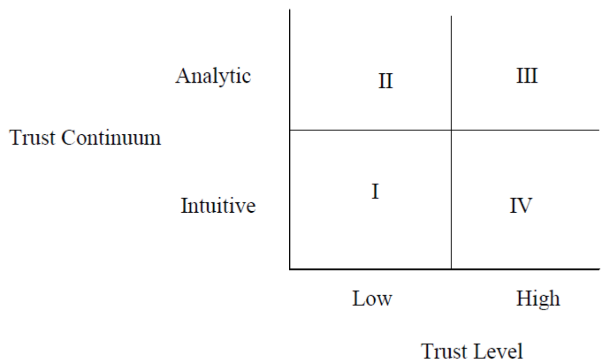

Figure 1. Trust Grid.

Quadrant 1 is the intuitively derived trust and low trust level. In this quadrant, trust is based on psychosocial factors such as bias and faith, and the interaction of these factors with poor automation leads to perception of low levels of trust. Quadrant 2 is the analytically derived trust and low trust level. Here, the human operator combines trust factors to derive a single composite score of automation performance that allows his/her evaluation of low trust in automation. This result may be attributed to many physical factors such as automation reliability, human-induced errors, robustness, and consistency among human and automation decision making. Quadrant 3 is the analytically derived trust and high level trust level. This quadrant captures the situation where human trust has been developed through consistent interaction with automation that demonstrated "good" behaviors. Quadrant 4 is the intuitively derived trust and high trust level. In this situation, humans overuse and/or over-rate automation through a high level perception of trust assignment. Intuitively, this result may be a result of positive bias and faith in automation. This quadrant captures the psychosocial aspects of human interaction with automation.

Proposed Study

In the proposed study, both ends of the trust continuum will be examined concurrently in order to understand how each of the quadrants in Figure 1 affects interpretation human trust in interactions with automated systems. The objective is to conduct simulated experiments and to develop empirical methods for calibrating trust metrics among human-automation interactions. This is designed to provide a proof-of-concept trust calibration using a two-interaction between two levels of trust (low and high) and two levels of trust continuum (intuitive and analytical). Furthermore, we intend to objectively measure trust by using BCI techniques. Our belief is that the brain images will be different for intuitive vs. analytic cognition.

References

Hammond, K. R.; Hamm, R. M.; Grassia, J.; and Pearson, T. 1987. Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *Systems, Man and Cybernetics, IEEE Transactions on*, *17*(5), 753-770.

Madsen, M.; and Gregor, S. 2000. Measuring human-computer trust. In *Proceedings of Eleventh Australasian Conference on Information Systems* (pp. 6-8).

Bisantz, A. M.; Finger, R.; Seong, Y.; and Llinas, J. 1999. Human performance and data fusion based decision aids. In *Proceedings of the FUSION* (Vol. 99, pp. 918-925).

Acknowledgement

This material is partially based on research sponsored by Air Force Research Laboratory and OSD under agreement number FA8750-15-2-0116. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory and OSD or the U.S. Government."