

The workshop will focus on the application of artificial intelligence to problems in cyber security. This year's AICS emphasis will be on human-machine teaming within the context of cyber security problems and will specifically explore collaboration between human operators and AI technologies. The workshop will address applicable areas of AI, such as machine learning, game theory, natural language processing, knowledge representation, automated and assistive reasoning and human machine interactions. Further, cyber security application areas- with a particular emphasis on the characterization and deployment of human-machine teaming- will be the focus. Additional areas can be discussed with similar challenges and solution spaces (e.g. genomic big data, astronomy, and cyberbiosecurity).

As cyber security has rapidly matured, data collection has become easier to instrument, implement, and collect. This has led to a massive increase of the amount of data that must be analyzed to achieve situational awareness- the scale of which is beyond human capabilities. Additionally, with the concurrent advancements in machine learning capabilities, there are algorithms and tools with the impressive ability to automatically analyze and classify massive amounts of data in complex scenarios, but deploying them in specific domains can be challenging. Together, this has created an environment of increased reliance on AI-based systems for humans to interact with the scale of cyber security problems.

Because humans must interact with at least parts of these AI systems, many challenges and arise. Principally among them are: 1) Determining optimal techniques to improve AI performance given targeted, limited human input, 2) understanding the extent to which the interaction between humans and AI introduces an attack surface for adversarial techniques to influence the performance of both the human and computer systems, 3) establishing and quantifying trust between humans and AI systems, 4) providing explainable AI where humans are required to do 'last mile' synthesis of information provided from a black box algorithm, and 5) defining the scope in which an AI system can operate autonomously in distinct cyber security domains while maintaining safety. A successful framework for the interaction between humans and AI is extremely important as machine learning based AI capabilities become incorporated into everyday life. Human-computer interactions will continue to increase. If they are not accurate, robust, trustworthy, explainable, and safe the systems will be prone to failure even if the underlying algorithms and/or people are individually effective.

For this workshop we consider general challenges 1-5 in the domain of cyber security as a focus application area. Cyber security is difficult to perform because of its high reliance on subject matter expertise to recognize anomalies in cyber data. Because AI systems are not yet well suited for this context-generating tasks for cyber, there is a human-in-the-loop requirement for most cyber security applications. Cyber security thus provides a unique case study in exploring the relationship between AI systems and humans because each rely on input and parse output from the other.

This year we are asking the AI for cyber security community to submit solutions to a challenge problem. The challenge problem (<http://aics.site/AICS2020/challenge.html>) is focused on a representative cyber security task that generally requires human interaction.

Understanding and addressing challenges associated with systems that involve human-machine teaming requires collaboration between several different research and development communities including: artificial intelligence, cyber-security, game theory, machine learning, human factors, as well as the formal reasoning communities. This workshop is structured to encourage a lively exchange of ideas between researchers in these communities from the academic, public, and commercial sectors.

Topics of interest include, but are not limited to:

- Human-machine teaming and human computer interactions
- Adversarial machine learning
- Cyberbiosecurity
- Machine learning approaches to make cyber systems secure and resilient
- Formal reasoning, with focus on human element, in cyber systems
- Game Theoretic reasoning in cyber security
- Robust AI metrics
- Multi-agent interaction/agent-based modeling in cyber systems
- Modeling and simulation of cyber systems and system components
- Decision making under uncertainty in cyber systems
- Automation of data labeling and ML techniques that learn to learn
- Quantitative human behavior models with application to cyber security
- Operational and commercial applications of AI

Challenge Problem

For information on this year's AICS challenge problem:

<http://aics.site/AICS2020/challenge.html>

Workshop Format

Invited speakers, presentations, panel and group discussions

Submission Requirements

One of two submissions is solicited:

- Full-length papers (up to 8 pages in AAAI format)
- Challenge problem papers (up to 8 pages in AAAI format)

Submissions are not anonymized. Please submit PDF via AICS Workshop website.

The deadline to submit papers is November 15, 2019.

Workshop URL

<http://aics.site/AICS2020>

Publication

Accepted papers will be published in the AICS Workshop

proceedings on arXiv after the event.

The AAI-20 Workshop on
Artificial Intelligence for Cyber Security (AICS)
February 7 or 8, 2020

New York, New York, USA

Part of the Association for the Advancement of Artificial Intelligence 2020 Conference

<https://aaai.org/Conferences/AAAI-20/>

Workshop Co-Chairs

- Dennis M. Ross, MIT Lincoln Laboratory, MA, USA
- Diane P. Staheli, MIT Lincoln Laboratory, MA, USA
- David R. Martinez, MIT Lincoln Laboratory, MA, USA
- William W. Streilein, MIT Lincoln Laboratory, MA, USA
- Arunesh Sinha, Singapore Management University, Singapore
- Milind Tambe, Harvard University, MA, USA

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

This material is based upon work supported under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Air Force.