C3E Challenge Problems

Dan Wolf and Don Goff, Co-PIs

17 September 2024

Value to C3E Workshops

- Provide an analytic activity supporting the overall C3E
- Address a gap in near-term cybersecurity research
- Provide expert innovative research to an actual government “need”
- Move researchers to publish technical papers relevant to a C3E technical problem
- Offer another outcome to C3E
- Offer opportunity to share concepts with a broader community

History of C3E Cybersecurity Challenge Problems

- Identity Discovery Challenge (2012)
- APT Infection Discovery Using DNS Data (2013)
- Metadata-based Malicious Cyber Discovery (2014)
- Novel Approaches to Avoid Misattribution (2015)
- Modeling Consequences of Ransomware (2016-17)
- Adversarial Machine Learning (ML), connections with Explainable Artificial Intelligence (XAI), and Decision Support Vulnerabilities (2018)
- Cognitive Securing and Human-Machine Teaming (2019)
- Economics of Security and Continuous Assurance (2020-2021)
- Supply Chain Software Static Analysis Coverage, AI, and Victimology (2021-2022)

Challenge Problems

- **Devolve from the work at the annual C3E Workshop**
- **Are vetted with the C3E committee and Workshop organizers for value and relevance**
- **Are posted about two months later on the CPS-VO web page**

NSF Support

- **Since 2016, the National Science Foundation has provided support to perform follow-up research on the Challenge Problem topics**
- **Researchers receive a modest honorarium to pursue deeper studies**
- **Results are presented at the following Workshop**
- **Posters, papers, and video presentations are also submitted**

2023-24 Challenge Problems

- **Cyberpsychology Aspects of Foreign Malign Influence**
- **Generative AI and Large Language Models**
- **AI/ML and the Human Element**
- **Resilience, Architecture, and Autonomy**

2023-2024 Challenge Problems

Cyberpsychology Aspects of Foreign Malign Influence

- What gaps are there in this approach?
- What are the long-term implications of this approach?
- What tools might be developed to actively scan social media to identify and assess false information?
- How can AI be used to counter such information campaigns?
- How can attribution be determined for the source of such campaigns?

2023-2024 Challenge Problems

Generative AI and Large Language Models

- Trustworthiness
- Validation related to meeting user intentions
- Concerns related to automation bias, the ELIZA effect, and fluency
- Prompt engineering in eliciting high-utility output from large language models
- Identifying and characterizing the bounds of coherent generation
- Human-Computer Interaction (HCM) approaches that address what the appropriate interface is to help a user make sense of LLM generations

2022-2023 Carry Overs

- **AI/ML and the Human Element**
  - Closing the Understanding Gap Between Simulations and Real-World Situations using Realistic Human Models
  - Cyber Analyst Assistant AI
  - Trust in AI
- **Resilience, Architecture, and Autonomy**
  - Active Agents for Resiliency and Autonomy
  - Resilient Architectures
  - Trust Factor in Resilient and Autonomous Systems

Research Projects 2023-2024

"From Generative Pre-trained Models to Verifiable Protocols for Security and Privacy Preservation“ Ufuk Topcu

“Language Models and Protocol Models Arise From the Same Logic of Attention” Dusko Pavlovic

“Cyber Ecosystem Observability and Resilience” Ryan Hilger

“Cyberpsychology Aspects of Foreign Malign Influence” Mia Bloom and Sophia Moskalenko

“Generative AI and LLMs: Security and Robustness” George Kesidis and David J. Miller

Research Projects 2023-2024 (cont'd)

- "Comp-HuSim: Complex Human Simulations" Trenton W. Ford and Michael G. Yankoski
- "Syntax-Guided Synthesis (SyGuS) with LLM and Predicate Sub-Typing" Stéphane Graham-Lengrand
- "Rational Resilience: Incentivizing Rational Systems" Spencer Oriot
- "LLMs for Robotics and Autonomy: Safety Perspective" Sayan Mitra

2025: Sociotechnical Aspects of Human-AI Systems

Sociotechnical Aspects of Human-AI Systems:

Expertise from multidisciplinary perspectives such as social psychology, behavioral economics and decision science, philosophy, linguistics, and cognitive, computer and data sciences.

This Challenge Problem will ask technologists and social/behavioral researchers to identify areas of future research.

2025: Generative Technologies and Verification

Systems Engineering and the use of generative technologies to assist with development of a specification:

System engineering involves a delicate interplay among specification, implementation, and verification. Novel generative AI technologies have emerged that can assist with implementation and verification.

This Challenge Problem will focus on the use of generative technologies to assist with development of a specification.

Schedule

- **About October 15, 2024:**
  - **We'll notify ~150 potential researchers of funding and solicit proposals on the two track outcomes from C3E 2024**
  - **Form a review panel to evaluate proposals**
  - **Offer honorariums based on reviews of those proposals**
- **Researchers will:**
  - **Provide posters and video presentations at C3E 2025**
  - **Final paper due at end of CY25**

Questions?