

Opportunity Title: Modeling insider threat with neural networks

Opportunity Reference Code: IC-18-11

Organization Office of the Director of National Intelligence (ODNI)

Reference Code IC-18-11

How to Apply **Create and release your Profile on Zintellect** – Postdoctoral applicants must create an account and complete a profile in the on-line application system. **Please note: your resume/CV may not exceed 2 pages.**

Complete your application – Enter the rest of the information required for the IC Postdoc Program Research Opportunity. The application itself contains detailed instructions for each one of these components: availability, citizenship, transcripts, dissertation abstract, publication and presentation plan, and information about your Research Advisor co-applicant.

Additional information about the IC Postdoctoral Research Fellowship Program is available on the program website located at:

<https://orau.org/icpostdoc/>.

If you have questions, send an email to ICPostdoc@orau.org. Please include the reference code for this opportunity in your email.

Application Deadline 3/12/2018 11:59:00 PM Eastern Time Zone

Description **Research Topic Description, including Problem Statement:**


An ancient problem for defense and intelligence efforts is insider threat: how do we know a colleague or collaborator is trustworthy and not pretending to be so as to gain some future advantage? In a separate domain, recent advances in machine learning have used artificial neural networks (ANNs) to create automated agents that can perform in complicated environments, such as Atari games. Can this new technology be used to address the older problem? The insider threat problem may be simulated by having multiple automated agents collaborate in a simulated environment (such as collaborative video games), but one of the agents has a different goal that runs counter to the others. Importantly, this goal may require collaborating with the other agents at least until a critical moment of defection (which may be explicit or implicit). Such simulations could allow for high-throughput examination and quantification of principles of insider threat.


- Questions could include:
 - When can a future defector be identified by their pre-defection behavior and when can they not? What are the critical factors that allow this detection, both of the collaborators' behavior and of the environment?
 - If the collaborative agent needs to be both performing in the environment and simultaneously watching for defectors, how does the computational burden of defector detection scale with the environment relative to the also-increasing burden of performing?


 **OAK RIDGE INSTITUTE**
FOR SCIENCE AND EDUCATION

ORISE GO

The ORISE GO mobile app helps you stay engaged, connected and informed during your ORISE experience – from application, to offer, through your appointment and even as an ORISE alum!

Visit ORISE GO 

GET IT ON
 Google Play

Download on the
 App Store

Opportunity Title: Modeling insider threat with neural networks

Opportunity Reference Code: IC-18-11

- It is presumably possible to test defectors by discretely putting them into simulated "gotchya" moments that would reveal their preferences. What are the dynamics when the defector is on the lookout for such tests?
- Instead of testing behavior, would having direct access to the ANN's code allow for identifying a defector? The "code" here would be the policy network and not any reward or loss function.
- For reinforcement learning, under what conditions is it possible to iteratively reward a potential defector into correct action such that there is little probability it will defect in the future? What are the guarantees or statistics of how that probability decreases with rewards?

Example Approaches:

- Approaches may not need to address all questions above, but they should consider using a "real-world" scenario of ANNs collaborating in an environment. It is possible to start with very simple ANNs and environments, then scale up to more complicated deep ANNs and environments. Environments could start with simple games like Pong (e.g. a collaborative Pong where one player is trying to maximize the total score while the other is trying to maximize their personal score). More complicated environments could include more complex Atari games or whatever the state-of-the-art performance will be for ANNs at the time of the research.
- The paradigm of Generative Adversarial Networks may particularly afford simulating actors in a cat-and-mouse game of defector detection and evasion. However, other techniques are welcome.

Qualifications Postdoc Eligibility

- U.S. citizens only
- Ph.D. in a relevant field must be completed before beginning the appointment and within five years of the application deadline
- Proposal must be associated with an accredited U.S. university, college, or U.S. government laboratory
- Eligible candidates may only receive one award from the IC Postdoctoral Research Fellowship Program.

Research Advisor Eligibility

- Must be an employee of an accredited U.S. university, college or U.S. government laboratory
- Are not required to be U.S. citizens

Eligibility • **Citizenship:** U.S. Citizen Only

Opportunity Title: Modeling insider threat with neural networks

Opportunity Reference Code: IC-18-11

- Requirements**
- **Degree:** Doctoral Degree.
 - **Discipline(s):**
 - **Chemistry and Materials Sciences** ([12](#))
 - **Communications and Graphics Design** ([6](#))
 - **Computer, Information, and Data Sciences** ([16](#))
 - **Earth and Geosciences** ([21](#))
 - **Engineering** ([27](#))
 - **Environmental and Marine Sciences** ([14](#))
 - **Life Health and Medical Sciences** ([45](#))
 - **Mathematics and Statistics** ([10](#))
 - **Other Non-Science & Engineering** ([5](#))
 - **Physics** ([16](#))
 - **Science & Engineering-related** ([1](#))
 - **Social and Behavioral Sciences** ([28](#))