Request for Information (RFI)
DARPA-SN-10-46
Suspected Malicious Insider Threat Elimination (SMITE)

"Trusted insiders … are targeting the US information infrastructure for exploitation, disruption, and potential destruction."

- National Counterintelligence Strategy of the United States of America (2008)

Information systems security personnel are drowning in ever expanding oceans of observational data from heterogeneous sources and sensors from which they must extract indicators of increasingly sophisticated malicious insider behavior. The Defense Advanced Research Projects Agency (DARPA) Information Processing Techniques Office (IPTO) is requesting information on areas of research related to the development of methods, tools, and techniques to reduce these enormous volumes of data to actionable information. Such technology must be flexible, scalable and highly interactive in order to cope with the dynamic nature of the insider threat. For the purposes of this RFI, we define *insider threat* as malevolent (or possibly inadvertent) actions by an already trusted person with access to sensitive information and information systems and sources.

The fundamental challenge is one of finding a poorly understood, subtle, or hidden signal (indicators of malicious behavior) buried in enormous amounts of noise (observational data of no immediate relevance) under the constraint that the measures of significance are themselves moving targets (based on dynamic context) that must be continually monitored and updated. The first step in meeting this challenge is to create a scalable, distributed infrastructure to securely collect, store, access, process, and correlate relevant data from heterogeneous sources over extended periods of time. The next step is to determine whether an individual or group of individuals is exhibiting anomalous behavior that is also malicious. However, this analysis is very heavily dependent on the context of the individual, groups of individuals and any data involved. Furthermore, context (e.g., location, time, roles and relations) is dynamic and so must be continually inferred, managed and applied automatically. Part of the challenge is detecting deceptive behavior. Deceptive behavior is characteristic of malicious intent which leads to the problem of assigning intent to observed behaviors.

Looking for clues that suggest an insider attack 1) can be anticipated, 2) is underway or 3) has already taken place could potentially be easier than recognizing explicit attacks. On the other hand, in both the real and virtual world, it is very difficult to do anything without leaving some evidence behind. Attempts to conceal or remove evidence generally create new evidence that, if detected, could be a strong indication of the perpetrator's intent. Security is often difficult because the defenses must be perfect, while the attacker needs to find only one flaw. An emphasis on forensics could reverse the burden by requiring the attacker and his tools to be perfect, while the defender needs only a few clues to recognize an intrusion is underway.

Forensic-like techniques can be used to find clues, gather and evaluate evidence and combine them deductively. Many attacks are combinations of directly observable and *inferred* events. Topics of interest to this RFI include, but are not limited to, techniques to (a) derive information about the relationship between deductions, the likely intent of inferred actions, and suggestions about what evidence *might* mean and (b) dynamically forecast context-dependent behaviors –

both malicious and non-malicious. Also of interest are on-line and off-line algorithms for feature extraction and detection in enormous graphs (as in *billions* of nodes) as well as hybrid engines where deduction and feature detection mutually inform one another.

DARPA is requesting white papers in the following three broad areas relating to malicious insider threat detection:

1. Data and Evaluation.
    a. Creation of data sets for development purposes.
    b. Planning, design, construction, execution and evaluation of verification and validation testing (evaluation methods and metrics) of developed technologies under realistic conditions of load and scale
2. Sensors and Algorithms that address the scale and complexity of current and projected insider threats
3. Novel Approaches and Methodologies, e.g.
    a. Social Behavioral Science (as opposed to signature based) methods
    b. Denial and Deception
    c. Red teaming including methods such as social engineering to more effectively understand and model the threat

DARPA encourages respondents to this RFI to submit ideas in one, two or all three of these areas and welcomes participation from teams composed of members from one or more, but not limited to, the following communities: (1) traditional insider threat, (2) deception detection, (3) pattern recognition, (4) automated reasoning, (5) analysis and algorithms for massive graphs and (6) computational psychology and sociology.

WORKSHOP

A DARPA-sponsored workshop is being planned for June 14-15, 2010 in Arlington, VA for the purpose of reviewing and discussing current and future research relevant to this RFI. Information discussed at this workshop may assist in the formulation of possible future areas of DARPA research with the objective of creating tools and techniques for the analysis and identification of malicious insider behavior.

Space for the workshop is limited and attendance will be by invitation only. Invitations will be based on white papers submitted, per the instructions below, no later than 1200 (ET), 26 May 2010. Participants will not be asked to make formal presentations. The workshop format will be group discussion. Invitations will be sent via email by 1500 (ET), 04 June 2010, and will provide further details on the workshop (times, location, etc.). All attendees will be encouraged to participate in general discussions and to make recommendations for future research in the area.

SUBMISSION FORMAT

Format specifications for white papers include 12 point font, single-spaced, single-sided, 8.5 by 11 inches paper, with 1-inch margins in either Microsoft Word or Adobe PDF format. Each white paper will consist of:

1. Cover Page (1 page)
    a. Title

   b. Organization
   c. Respondent’s technical and administrative points of contact (names, addresses, phones and fax numbers, and email addresses)
   d. Indication of willingness to attend the Workshop
2. Summary of technical ideas for 1 – 3 of the broad areas specified (1 page)
3. One chart summarizing ideas submitted. (1 page)
4. Team Bio: A brief summary of the team including ongoing or prior work (1 page)
5. Bibliography: Papers you think are particularly relevant (1 page)

Respondents are encouraged to be as succinct as possible while at the same time providing actionable insight.

SUBMISSION INSTRUCTIONS

Responses to this RFI must be submitted via email to DARPA-SN-10-46@darpa.mil between 1200 (ET), 18 May 2010 and 1200 (ET), 26 May 2010. Please include "SMITE RFI" in the subject line in all correspondence.

DISCLAIMER

This is an RFI issued solely for information gathering purposes; this RFI does not constitute a formal solicitation for proposals. In accordance with FAR 15.201(e), responses to this notice are not offers and cannot be accepted by the Government to form a binding contract. DARPA will not provide reimbursement for costs incurred in responding to this RFI or attending the workshop. Respondents are advised that DARPA is under no obligation to acknowledge receipt of the information received or provide feedback to respondents with respect to any information submitted under this RFI. Submission of a white paper is voluntary and is not required to propose to any subsequent solicitations on this topic, if any.

No classified information shall be included in the RFI response. White paper submissions containing proprietary data should have the cover page and each page containing proprietary data clearly marked as containing “proprietary” data. It is the respondent’s responsibility to clearly define to the Government what is considered proprietary data.

Submissions may be reviewed by: the Government (DARPA and partners); Federally Funded R&D Centers (such as MIT Lincoln Laboratory); and Scientific Engineering and Technical Assistance (SETA) contractors (such as Schafer Corporation, Science and Technology Associates, CACI International, and System Analysis, Inc, etc.).

POINT OF CONTACT

Dr. Rand Waltzman, Program Manager, DARPA/IPTO. All inquiries on this RFI must be submitted to DARPA-SN-10-46@darpa.mil. No telephone inquiries will be accepted.