SoURCE CODE

SECURING OUR UNDERLYING RESOURCES IN CYBER ENVIRONMENTS

INTELLIGENCE VALUE

SoURCE CODE seeks to improve law enforcement and intelligence forensic capabilities to measure similarity and estimate demographics of cyber tools used in criminal enterprises.

The SoURCE CODE program will begin in 2024 and aims to disrupt criminal cyber capabilities by improving law enforcement and Intelligence Community (IC) forensic capabilities. The volume, methods, and complexity of cyberattacks on companies and infrastructure has grown significantly and will continue to evolve over time. These cyberattacks – whether on government systems or private companies – pose a serious threat to national security. At the same time, there is a shortage of cyber expertise to fill all necessary cyber-focused positions in the commercial world. This challenge holds true with cyber-forensic experts, who play an important role in attributing these attacks to assist with informing companies and governments on the threats they are facing.

Attribution of attacks generally requires significant evidence to be built up. While forensic experts do well at this task, the SoURCE CODE program is seeking to provide novel automated technologies to assist forensic experts in making determinations. The program will explore full feature spaces in binary code and source code files to measure the similarity between files and provide demographic information to forensic experts to the likely origins (country, groups, etc.). Similarity matching provides capabilities that can act as a force multiplier to the forensic experts within law enforcement and the IC. This capability would allow the automated matching of similar binaries from known samples, allowing the IC to speed up the attribution of malicious attacks to improve law enforcement and IC responses.

PRIME PERFORMERS

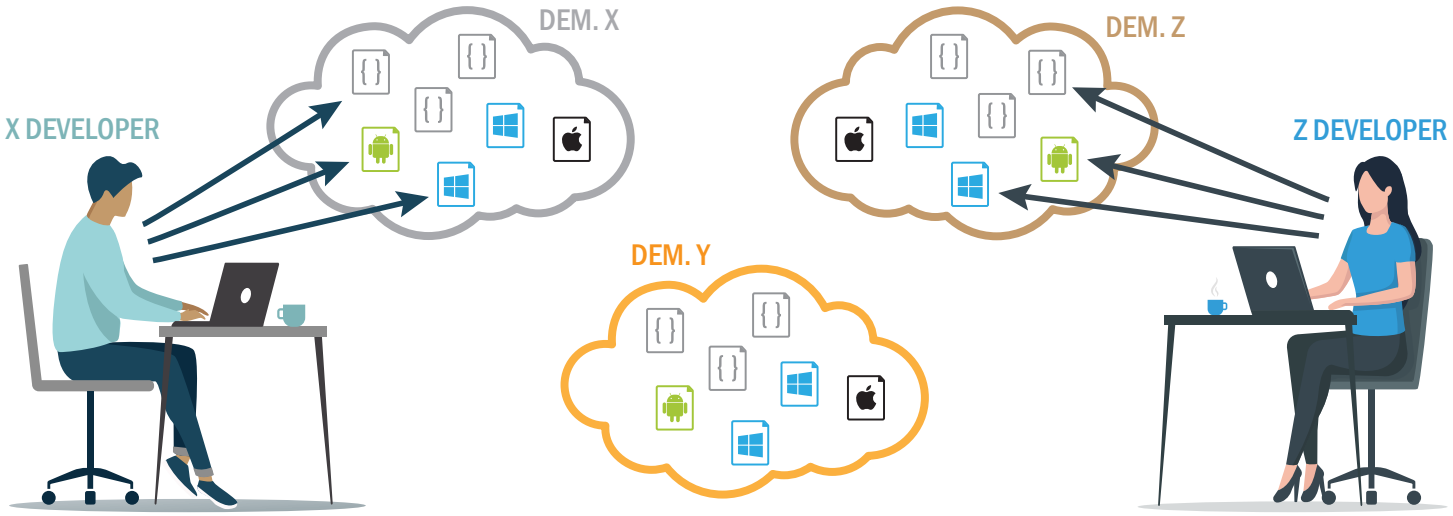
- TBD

TESTING AND EVALUATION PARTNERS

- Sandia National Laboratory
- Lawrence Livermore National Laboratory
- Software Engineering Institute

KEYWORDS

- Programming style
- Malware attribution
- Software forensics



The SoURCE CODE program seeks to identify and utilize features within code and binaries to sort and separate out Developers of demographic groups and measure the similarity of the tools being deployed by these groups So that Developer X and Z sort into their own, respective demographic clusters X and Z.

PROGRAM MANAGER
Kristopher Reese, Ph.D.
Phone: (301) 243-2086
kristopher.reese@iarpa.gov