

Haystack is an Agent-based System for Simulating an Enterprise's Cyber Social Interactions. A hybrid approach that combines real world data with simulation, it creates large streams of realistic multi-degree social interactions.

Cyber Security Testbed and Dataset Generation

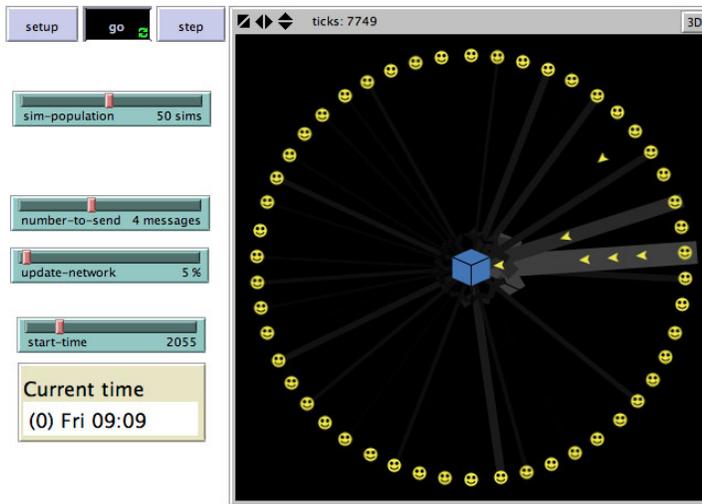
Objective

CyberNET is a research infrastructure that enables cybersecurity scientists and engineers to rigorously develop, experiment, and evaluate hypotheses, tools, and algorithms.

Approach

We are taking a three-pronged approach to solving the identified problems. First, the CyberNET research infrastructure is providing “sandboxes” for researchers to develop, experiment, and evaluate their algorithms and tools. Second, in

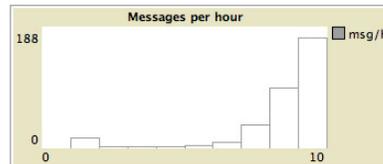
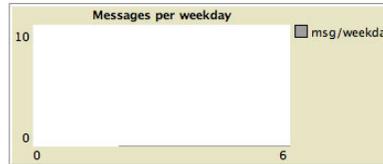
modeling and simulating, we are standing up the reference architecture—a virtual enterprise that represents a real world software development organization. To further improve realism, HAYSTACK is creating background “noise” by emulating real human and machine communications patterns on the network. To make the noise as realistic as possible, we are emulating the behavioral characteristics of human users in a real enterprise setting. Third, we are stewarding several large datasets that were collected during an enterprise's operation and during cyber training exercises.



```

464839.742055: Sent from 38 to 11, 55983 bytes
464875.322652: Received from 38 to 11, 55983 b
464829.389656: Sent from 41 to 24, 153051 byte
464872.287855: Received from 41 to 24, 153051
464840.112207: Sent from 34 to 24, 112093 byte
464842.656595: Received from 34 to 24, 112093
464883.643389: Sent from 12 to 39, 404448 byte
464929.812907: Received from 12 to 39, 404448

```



Haystack Simulating Cyber Social Interactions

Achievements

- Stood up the CyberNET research infrastructure under an OpenStack cloud computing platform
- Instantiated the network piece of the reference architecture using the CyberNET research infrastructure
- Analyzed PNNL email monitoring data, extracting communications behaviors and interactions
- Developed a prototype demonstrating HAYSTACK cyber-social interactions
- Stewarding three datasets; the first is from the operation of the PNNL, and the other two were collected during training exercises.

Impact

CyberNET research infrastructure and datasets will accelerate scientists and engineers and reduce costs, time, and redundancies across the ARC Initiative. Enhanced modeling and simulation, supported by real world datasets, will increase realism in models, leading to more relevant research. The HAYSTACK cyber-social interaction models and framework will improve existing cyber-user behavior frameworks. These efforts will improve the state of science throughout ARC.

Future Work

For FY 2014, HAYSTACK focused on email-based anthropogenic system activity models. Next year, we will be focusing on modeling anthropogenic web traffic. We are proposing to integrate the HAYSTACK framework with an existing cyber-user behavior modeling mechanism, such as the Cyber Range Toolkit. The overall goal is to create dynamic, fully synthetic datasets that can be shared among all ARC projects.

ABOUT

The Asymmetric Resilient Cybersecurity Initiative

Researchers at PNNL are delivering the theory, processes, methodologies, and algorithms that will enable a resilient cyber infrastructure with an asymmetric advantage to thwart adversaries who seek to infiltrate and damage our national security through digital means. This exploratory science in Laboratory Directed Research and Development effort is made possible by the Pacific Northwest National Laboratory through funding provided by the U.S. Department of Energy.

For more information on the science you see here, please contact:

Tom Carroll

Pacific Northwest National Laboratory
P.O. Box 999, MSIN: K3-12
Richland, WA 99352
(509) 371-6731
thomas.carroll@pnnl.gov



Pacific Northwest
NATIONAL LABORATORY

U.S. DEPARTMENT OF
ENERGY

Proudly Operated by **Battelle** Since 1965