AZSecure-data.org

Autumn Data Release

Data Infrastructure Buildings Blocks for Intelligence and Security Informatics (DIBBs for ISI)

Thank you to those who attended our DIBBS-ISI panel discussion at the IEEE-ISI meeting in Tucson this September, it was a pleasure meeting you. Mark your calendars for Women in Cybersecurity (WiCyS) in Tucson, March 31st - April 1st. Stay tuned for IEEE-ISI 2017 location and dates.

New Data

The AZSecure-data.org data portal has published new intelligence and security informatics-related data sets. Security researchers are encouraged to explore the collections and announce their availability to their students and other researchers.

The newly added data sets include the following:

- An in-progress collection of phishing websites joined our previous phishing collections in September. PhishMonger collects phishing websites as they are reported and before these ephemeral websites can disappear. The University of Virginia researchers David Dobolyi and Ahmed Abbasi are making this collection available as they compile it. A second batch was added to AZSecure-data.org in November. As of output #3649, the collection contains 171,360 websites collected between November 2015 and September 2016. The PhishMonger collection can be used for the development or improvement of anti-phishing tools, for usability studies and susceptibility experiments, as well as time-series and longitudinal analysis of phishing attacks and malware. New data will be added periodically.

- Ether Malware Analysis Dataset is our first collection specifically of malware and contains 25,118 unique malware instances collected between January and March 2008. They can be used to observe the behavior of malware in detecting and evading existing malware analyzer(s) or as input for training a new malware detection framework or malware analyzer. This dataset will be invaluable for research in Malware Analysis and detection.

- Intrusion Detection System research has relied on the KDD and UNM datasets, which are increasingly seen as outdated. ADFA IDS was created and released in 2013 as an IDS dataset representative of modern attack structure and methodology in the Linux and Windows

Hsinchun Chen, Ph.D.

Regents' Professor, Thomas R. Brown Chair of Management and Technology, and Director of the Artificial Intelligence Lab, University of Arizona.

PI and Director for Data Infrastructure Building Blocks for Intelligence and Security Informatics (DIBBs for ISI).

hchen@eller.arizona.edu

University of Arizona
Tucson, AZ 85721 USA
520.621.2748 tel
520.521.2433 fax

ai.arizona.edu

environments and is now hosted at AZSecure-data. This dataset contains around 5,000 files with intrusion detection patterns. This corpus is hand categorized into Training and Validation datasets which can be used out-of-the-box for training & testing new intrusion detection frameworks.

- Adding to our corpus of network traffic datasets is the Comprehensive Multi-Source Cyber Security Events Dataset collected over 58 consecutive days at Los Alamos National Laboratory. It contains data from authentication, process, DNS, network flows, and red team attacks on LANL's corporate internal network. This dataset will be useful for researchers in the dynamic network security domain. It can be used for threat/pattern analysis in networks for malicious software & intrusion detection.

Share Your Data

AZSecure-data continues to accept legacy and current datasets from the intelligence and security informatics domain. You can recommend or submit a dataset at any time by contacting ailab@eller.arizona.edu. **AZSecure-data.org is an excellent solution for funder-mandated data sharing.**

Tell Us Your Data Story!

Data stories help convey the importance of making sharable data easy to locate and reuse. Tell us your story about working with the datasets we host: which dataset(s) you used and how you used them, what you were investigating, and what you found out.

About the Project

With funding from the National Science Foundation, this Data Infrastructure Building Blocks for Intelligence and Security Informatics ("DIBBs for ISI") project is developing a research infrastructure for use by ISI scientists, students, and others studying a wide range of cyber- and security-related issues, whether in the computer, information, or social sciences. The project portal, AZSecure-data, provides researchers and students with free access to a wide variety of relevant data sets.This material is based upon work supported by the National Science Foundation Grant No. ACI-1443019.

Follow Us!

This newsletter is quarterly, but if you desire more frequent updates on ISI related topics, take a look at Intelligence and Security Informatics on LinkedIn, Facebook, and Twitter.

Interested in learning more?

Join our Intelligence and Security Informatics Group on LinkedIn, Facebook, and Twitter.

Do you have data to share?

We can help! Send an inquiry to ailab@email.arizona.edu.

This material is based upon work supported by the National Science Foundation under Grant No. ACI-1443019.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.