



 **SEI WEBINAR SERIES** | Keeping you informed of the latest solutions

# Copyright 2017 Carnegie Mellon University

All Rights Reserved.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

Internal use:\* Permission to reproduce this material and to prepare derivative works from this material for internal use is granted, provided the copyright and "No Warranty" statements are included with all reproductions and derivative works.

External use:\* This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other external and/or commercial use. Requests for permission should be directed to the Software Engineering Institute at [permission@sei.cmu.edu](mailto:permission@sei.cmu.edu).

\* These restrictions do not apply to U.S. government entities.

CERT® is registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM17-0708

# Carnegie Mellon University

This video and all related information and materials (“materials”) are owned by Carnegie Mellon University. These materials are provided on an “as-is” “as available” basis without any warranties and solely for your personal viewing and use.

You agree that Carnegie Mellon is not liable with respect to any materials received by you as a result of viewing the video, or using referenced websites, and/or for any consequences or the use by you of such materials.

By viewing, downloading, and/or using this video and related materials, you agree that you have read and agree to our terms of use ([www.sei.cmu.edu/legal/](http://www.sei.cmu.edu/legal/)).

Distribution Statement A: Approved for Public Release; Distribution is Unlimited

© 2017 Carnegie Mellon University.

# Four Valuable Data Sources for Network Security Analytics

Timothy Shimeall, Ph.D.

Software Engineering Institute  
Carnegie Mellon University  
Pittsburgh, PA 15213

# Overview

Four data sources

Analytic process

Analytic examples

Discussion



Four Valuable Data Sources for Network Security Analytics  
**Data and Process Description**



# Polling Question 1

What information sources do your organization use to inform network security?

- a. Mostly intrusion detection/prevention alerts
- b. Mostly network packet monitoring
- c. Mostly network flow collection (or traffic traces)
- d. Mostly third-party reports (Threat intelligence)
- e. Mostly vulnerability scanning
- f. Mostly host-based logs
- g. A balanced mix of sources

# Domain Resolution Data

Domain Name System (DNS) Records

Passive DNS

Polled active DNS

Host name or domain

Record result (success, no such domain, server fail)

Request type (A, AAAA, PTR, ...)

Date/time

Collector location



# Network Device Inventory/Configuration Data

Security Content Automation Protocol (SCAP) – Periodic report

## Common Vulnerability Enumeration

- Identity
- Severity score (CVSS)
- Systems affected per network

## Common Checklist Enumeration

- Identity
- Checklist items
- Systems affected per network

## Common Platform Enumeration

- Identity
- Systems affected per network

<http://nvd.nist.gov/>

# Network Flow Data

Identifying information:

- Source address
- Source port
- Destination address
- Destination port
- Transport protocol
- Sensor

Aggregate information:

- Bytes
- Packets
- Communication flags
- Start time
- End time

# Network Intrusion Detection/Prevention Alerts

Source

Destination

Alert identity

Confirming information

Sensor location

Alert time

# Process

Explore

Model

Test

Analyze

Refine

# Explore

Needs analysis – is there a prior analytic that addresses this?

Research analytic

- vendor documentation
- published papers
- data feeds

Identify unique attributes

- ports
- protocols
- associations
- behaviors

# Model

Lessons learned from prior analytics

Build model

- identified behavior
- similar behavior

Program model

- Shell
- Python
- other

# Test

Execute programmed model

- Monitor progress
- Debug

Save test results

- 'raw' files
- 'set' files
- 'bag' files
- Other formats



# Analyze

Review test results.

Reduce false positives.

Reduce false negatives.

Identify improvements.

# Refine

Apply improvements

Update programs

Repeat

Mature the process

- Templates
- Regression testing
- Code reuse / Analytics libraries

# Polling Question 2

What is a difficult step for your organization in developing security analytics?

- a. Getting dependable data
- b. Handling large data volumes
- c. Turning data into behavior observations
- d. Prioritizing significance on behavior observations
- e. Matching behaviors with threats
- f. Automating the process with the tools available
- g. Communicating efficiently with management

# Four Valuable Data Sources for Network Security Analytics

## Analytic Examples



# Example: Co-located Generated Domains

Explore: Scan DNS queries for computer-generated domains (several algorithms), couple with network flow data looking for propagation attempts, and intrusion detection alerts for compromise attempts.

Model: Identify timeframe from DNS queries. Identify sources from IDS alerts or from scanning / service probe detection through network flow.

Test: Apply at several scattered points throughout day (early workday, morning peak, noon, afternoon peak, end-of-workday, late evening). Watch for recurring sites and ongoing activity

Analyze: Correlate against third party reporting; remove contracted or internal security scanning

Refine: Revise model to improve throughput and to whitelist sources

# Example: Assessing Patch Efficiency

Explore: Patch efficiency – mitigations are applied for significant (serious and exploitable) vulnerabilities prior to exploitation. Couple detected responses to scanning against reported vulnerability patching and detected changes in behavior.

Model: Use inventory data to identify decreasing vulnerability, then query network flow data for responses to identified scans on previously-vulnerable services. Apply more in-depth flow analysis to profile and contrast service behavior before and after patching.

Test: Apply for very common services (web, email, DNS) and less common (database, file transfer) services.

Analyze: Compare changes in behavior for patched services against those for non-patched services.

Refine: Revise model to distinguish significant vs. coincidental changes.

Campbell, G. "MEASURES and METRICS In CORPORATE SECURITY." *Security Executive Council Publication Series*. January 2008.

# Example: Quantifying Vulnerability Exposure

Explore: Vulnerability exposure – probable loss associated with vulnerabilities in a given network service. Identify services with increasing vulnerabilities and pivot to associate with critical missions, characteristic behaviors, and threat activity.

Model: Profile reported vulnerabilities with new reports or increasing counts of affected systems. Identify intrusion detection reports for associated services in appropriate timeframe. Pivot against network flow data for overall traffic levels in these services, changes in behavior profile, and service timelines.

Test: Apply for very common services (web, email, DNS) and less common (database, file transfer) services.

Analyze: Compare changes in behavior for more vulnerable services against those for all services or less-vulnerable services.

Refine: Revise model to distinguish significant vs. coincidental changes.





Four Valuable Data Sources for Network Security Analytics  
**Understanding and Improving Security**

# Understanding and Improving

## Understanding:

- Data overload
- Observer bias
- Incomplete observation

## Improving:

- Response to change
- Associating threat and risk
- Focusing on what can be improved

# Contact Information

Tim Shimeall, Ph.D.

[Netsa-contact@cert.org](mailto:Netsa-contact@cert.org)

Software Engineering Institute  
4500 Fifth Ave  
Pittsburgh PA 15213

[www.flocon.org](http://www.flocon.org)