Community Detection and Associativity to Assess Network Threats

Akshay Peshave, Tim Oates

Dept. of Computer Science and Electrical Engineering
University of Maryland, Baltimore County

IBM RESEARCH
COGNITIVE
COLLOQUIUM

September 2017

New York, USA

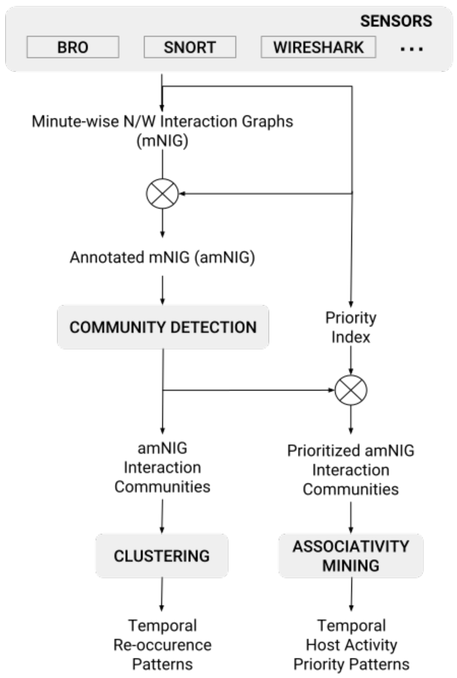
Introduction

A large collection of software and hardware sensors exist for monitoring network traffic at different granularity and alerting when suspicious traffic is encountered. The sensors utilize large and diverse rule-sets to detect malicious network traffic patterns. The data generated by these sensors can be utilized to provide a holistic assessment and reason about network threat patterns. We propose an analytic pipeline which applies graph theoretic and machine learning methods to achieve this.

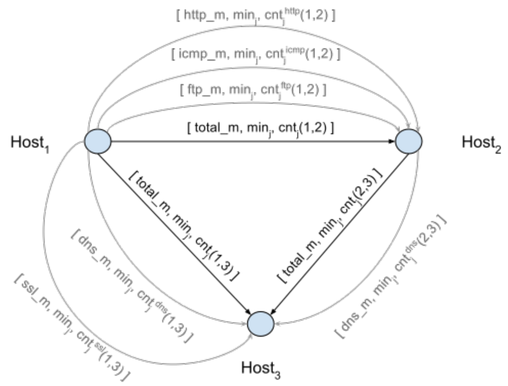
Dataset

The MACCDC2012 [2] dataset is used in this work. The annotated network traffic data for this dataset is available and acquired from SecRepo in the form of Bro and Snort logs.

Analytics Pipeline

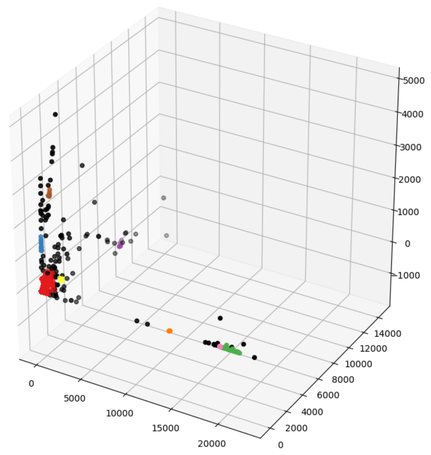


Minute-wise Network Interaction Graph (mNIG)



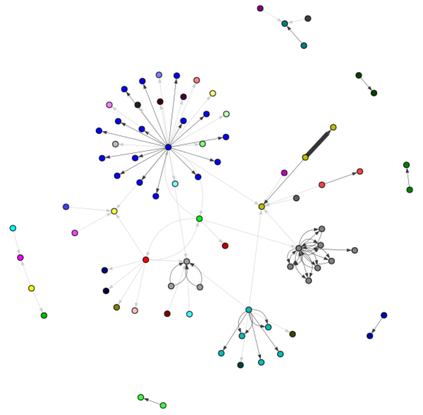
Nodes: hosts active on the network in minute 'j'.
Edge-sets: One per traffic category and one representing net interaction flow between pair of hosts.
Granularity: One NIG per minute of network traffic.

Community Clusters By Snort Class Vectors



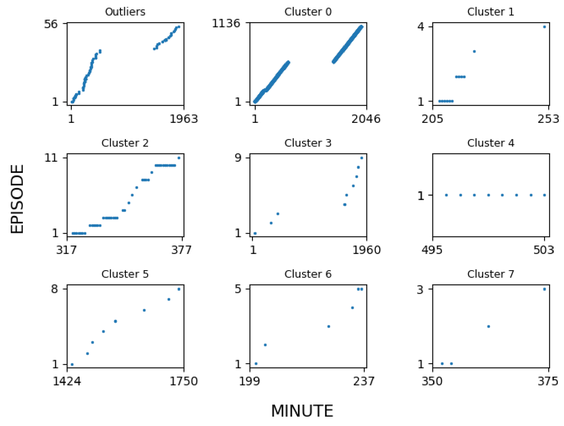
Communities clustered using their malicious class vectors provide insight into common, co-occurring malicious traffic classes within them. Bulk of the communities exhibit a common malicious class pattern in the MACCDC dataset.

Community Detection for Dense Traffic Patterns



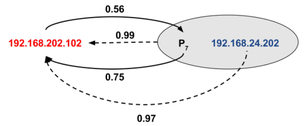
We partition each of the mNIGs into high flow density partitions using walktrap community detection. Each community is also characterized by a malicious traffic vector.

Community Cluster Episodes



Density-based clustering of temporal occurrence of community cluster types help identify episodes of malicious traffic patterns. The most common cluster occurs frequently while other malicious patterns are less frequent but episodic during short time windows.

Associativity Analysis



Snort alert classes are mapped to an ordinal threat scale (priority index in pipeline) using the concepts of severity and progression of an attack discussed in [1]. High confidence associativity between hosts and threat index can be computed for any time window.

Conclusion

The proposed analytics pipeline allows a holistic assessment of network traffic patterns at custom temporal granularity. Further, temporal coccurrence of host interactions and associativity can help discover possible collusion and attack campaign signatures. This automated workflow is extendable and customizable by adding new computation blocks and an interactive, human-in-the-loop experience.

References

[1] P. Giura and W. Wang.
A context-based detection framework for advanced persistent threats.
In *2012 International Conference on Cyber Security*, pages 69–74, Dec 2012.

[2] U.S. National CyberWatch.
Mid-Atlantic Collegiate Cyber Defense Competition (MACCDC), 2012.
Data retrieved from SecRepo, `http://secrepo.com`.

Contact Information

- Akshay Peshave: peshave1@umbc.edu
- Tim Oates: oates@umbc.edu