

FloCon 2019

Using Data to Defend

Graph Measures for Network Traffic Analysis

Timothy Shimeall, Ph.D.
Joshua Fallon, Ph.D.

(with thanks to Adam Tse, M.S.)

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Copyright 2018 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.

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 use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

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

DM18-1364

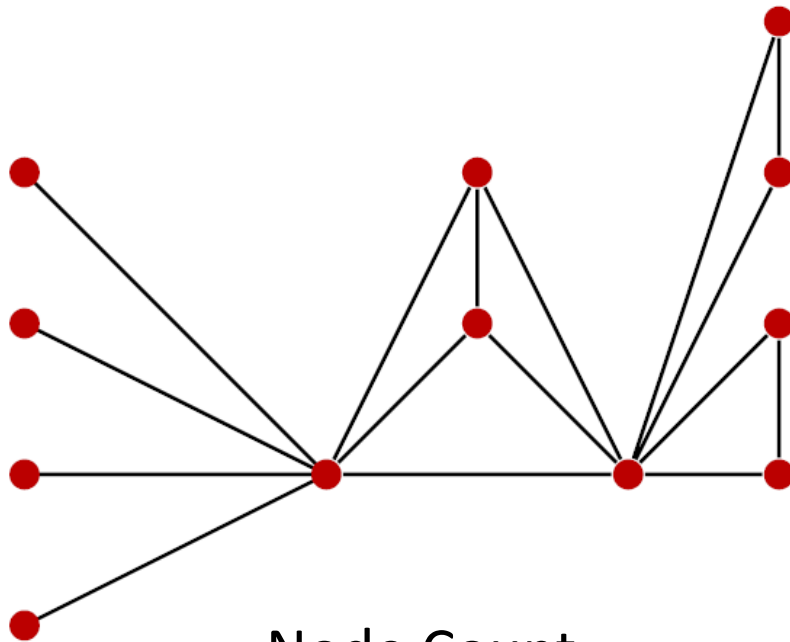
Agenda

Graph Measures

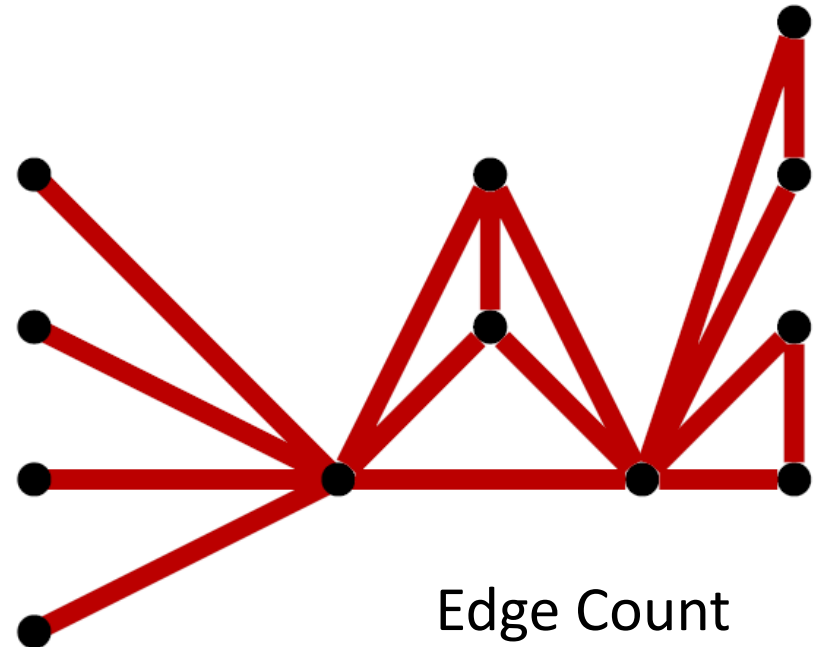
Assessing Network Traffic

Example

Node and Edge Count

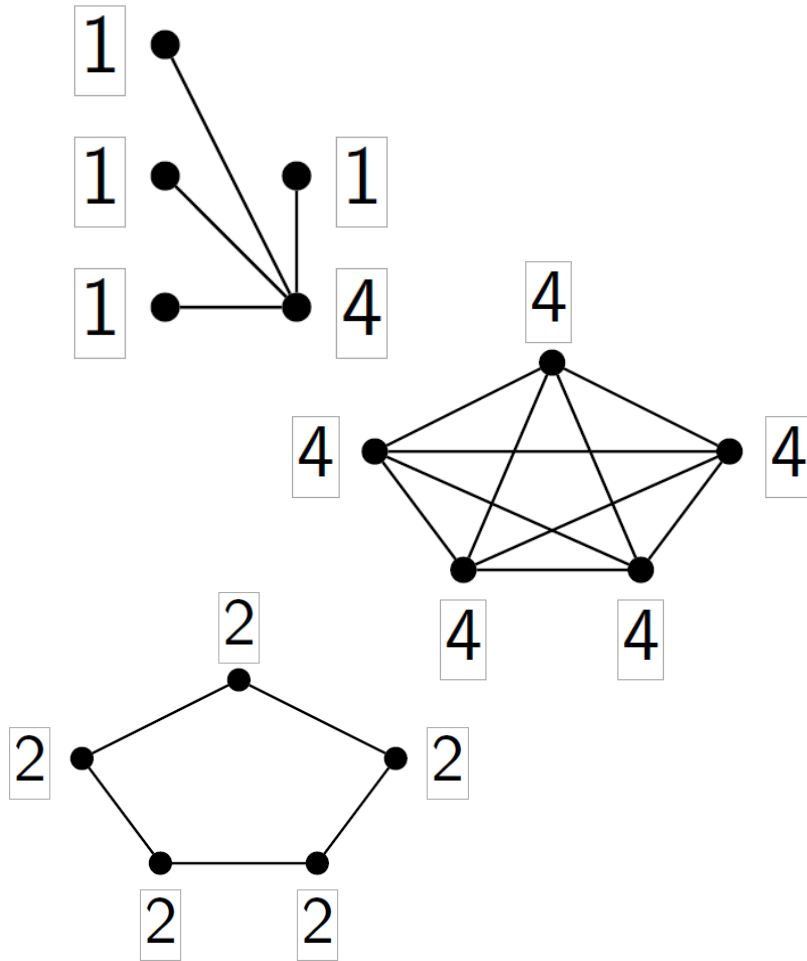


Node Count



Edge Count

Components

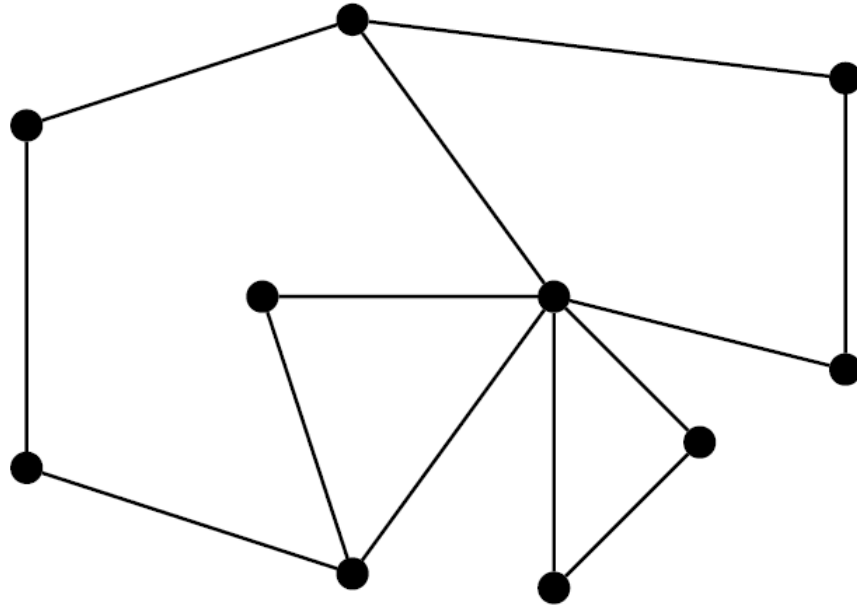


3 Components

Average Degree ~ 2.53

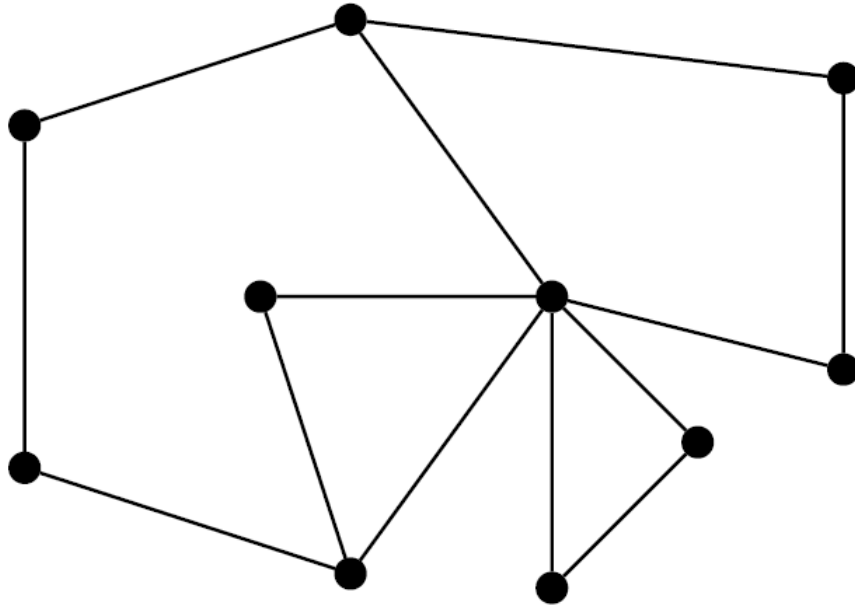
Link Density ~ 0.18

Ego Network



Fragmentation: Describes how much of the network is in highly-connected communities

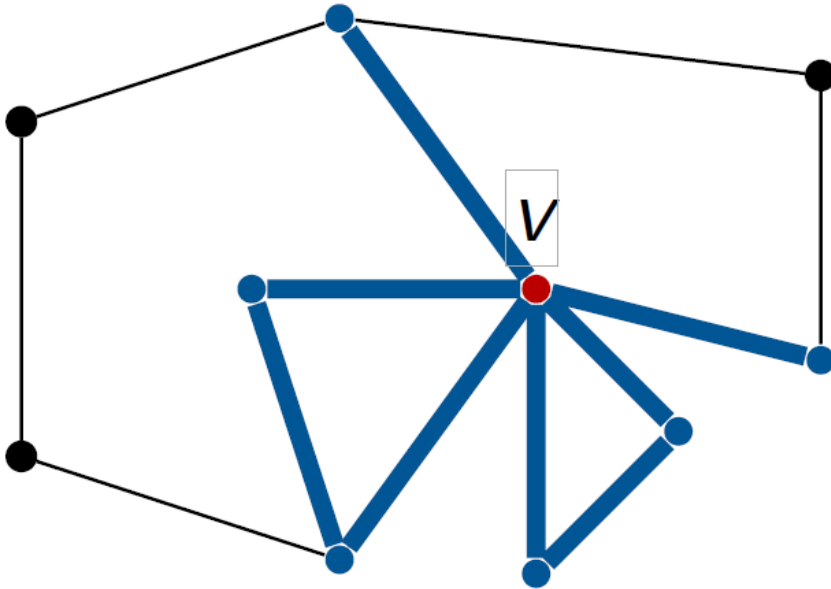
Ego Network



Fragmentation: Describes how much of the network is in highly-connected communities

Local Clustering Coefficient: Link density among a node's neighbors

Ego Network



Fragmentation: Describes how much of the network is in highly-connected communities

Local Clustering Coefficient: Link density among a node's neighbors

Ego Network: The local network defined by a node and its neighbors

Assessing Network Traffic

Bin protocols

Subdivide by autonomic and human-directed traffic (and other)

Take measures

Follow up ego networks for hosts of interest

Advantages:

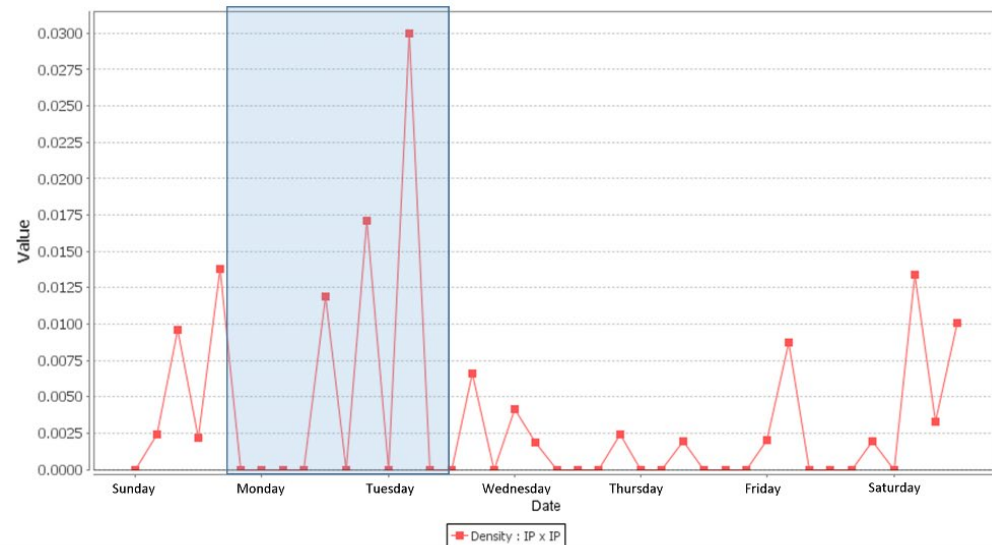
- Speed of analysis: faster operations restrict focus of slower ones
- Objectivity: very data-driven analysis
- Repeatability: over time, thresholds and break points will be known

Issues:

- Which are the best measures?
- How much is organization dependent?
- How much is an artifact of sensors, placement, or data format?

Example

TCP Link Density – 7 days

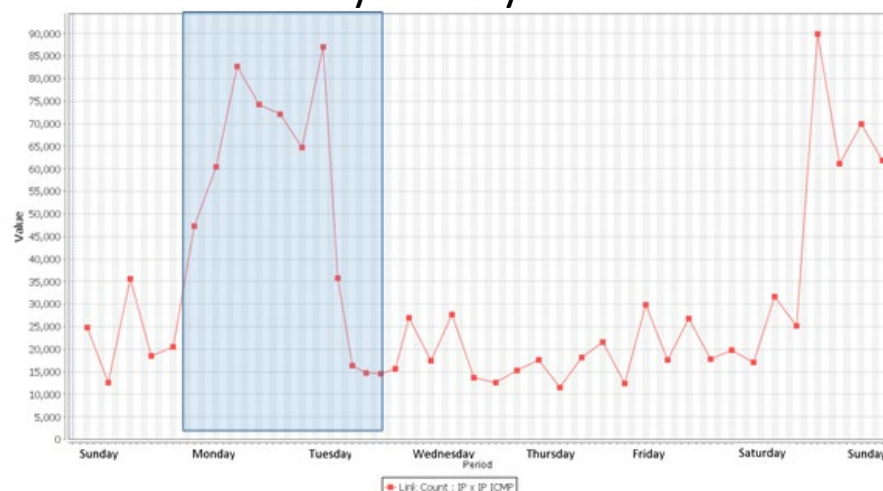


Data from a large-scale enterprise network

Flash crowd event over two days

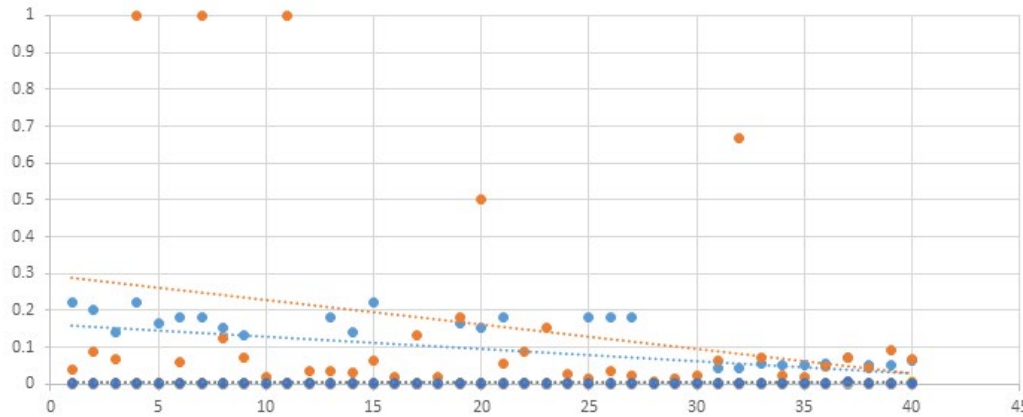
Network impacted, but delivered service

ICMP Link Density – 7 days



Example Ego Network Data

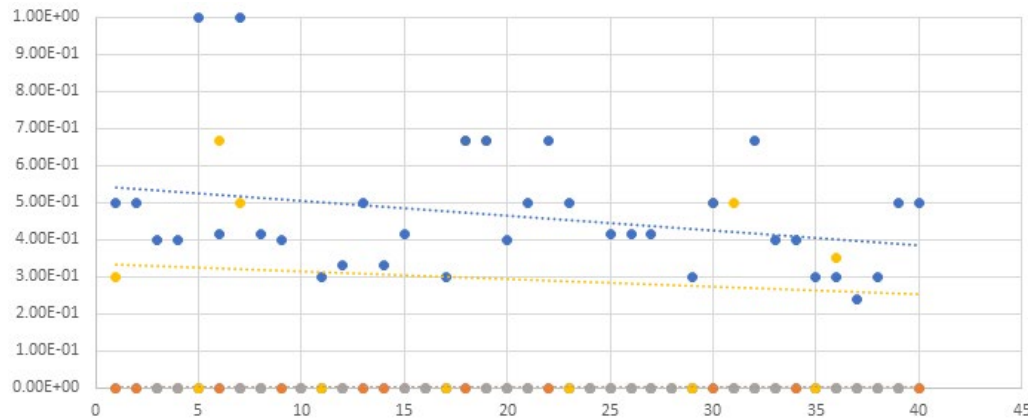
TCP Non-abuse node Link Density



Subdivided most changed
TCP nodes based on watch
list

Generated ego networks

Time series of link density
color coded by node
identity



TCP Abuse node Link Density

Summary

Graph models allow repeatable assessment of network traffic

Community and relational perspective – patterns in relationships on nodes and links, rather than record-specific issues

Binning data matters: aggregation interval, types of traffic, groups of nodes

Ongoing work to understand how this can be generalized