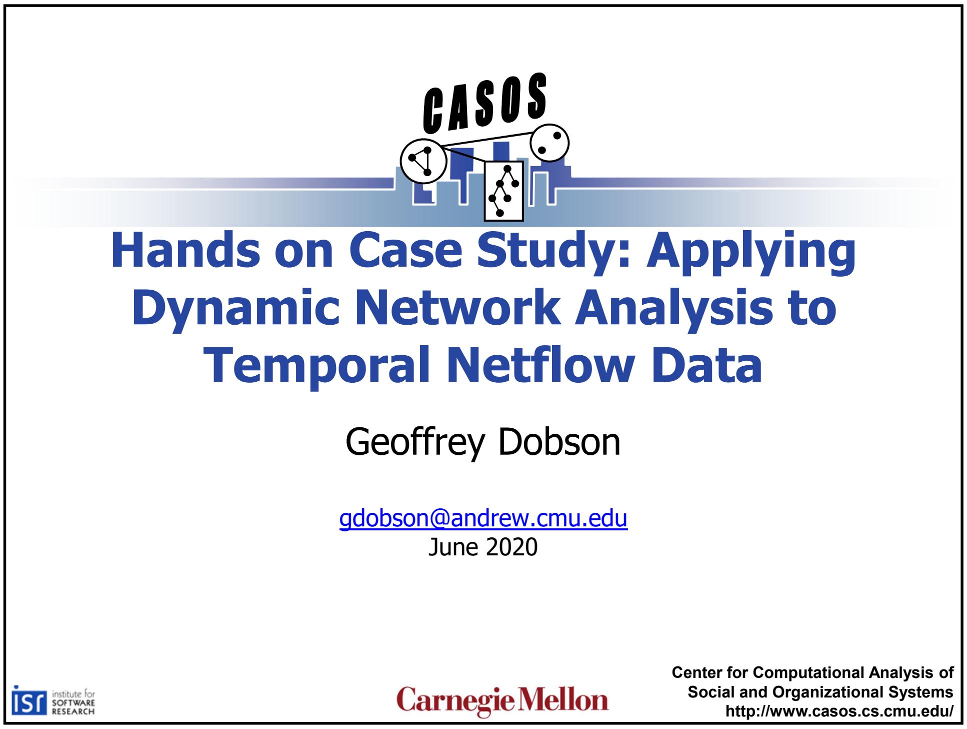


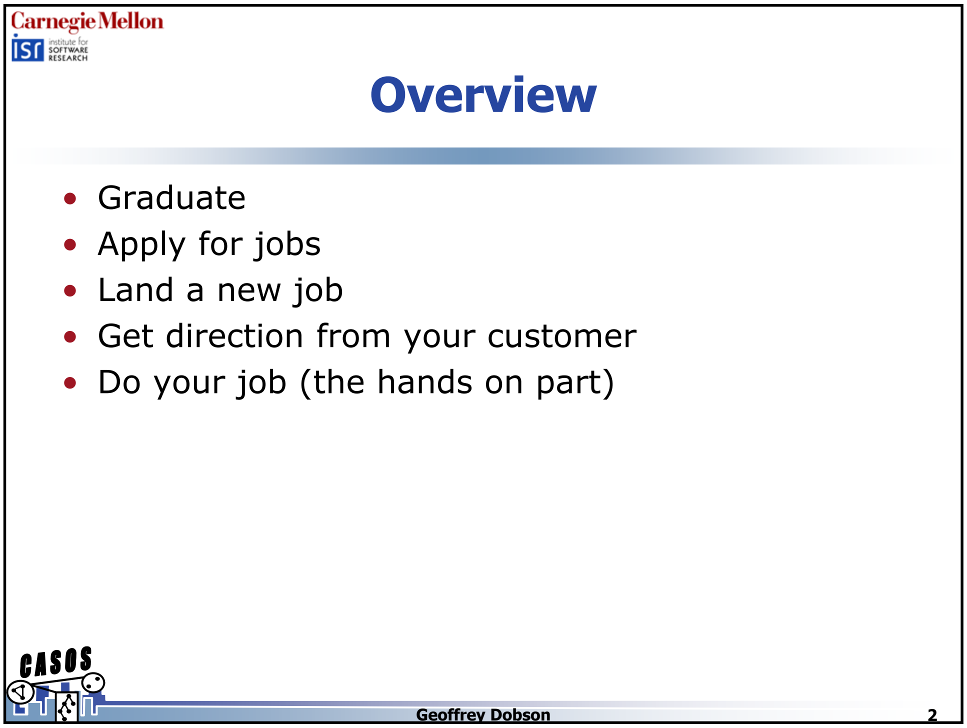
Hands on Case Study: Applying Dynamic Network Analysis to Temporal Netflow Data

Geoffrey Dobson

gdobson@andrew.cmu.edu

June 2020

Center for Computational Analysis of
Social and Organizational Systems
http://www.casos.cs.cmu.edu/

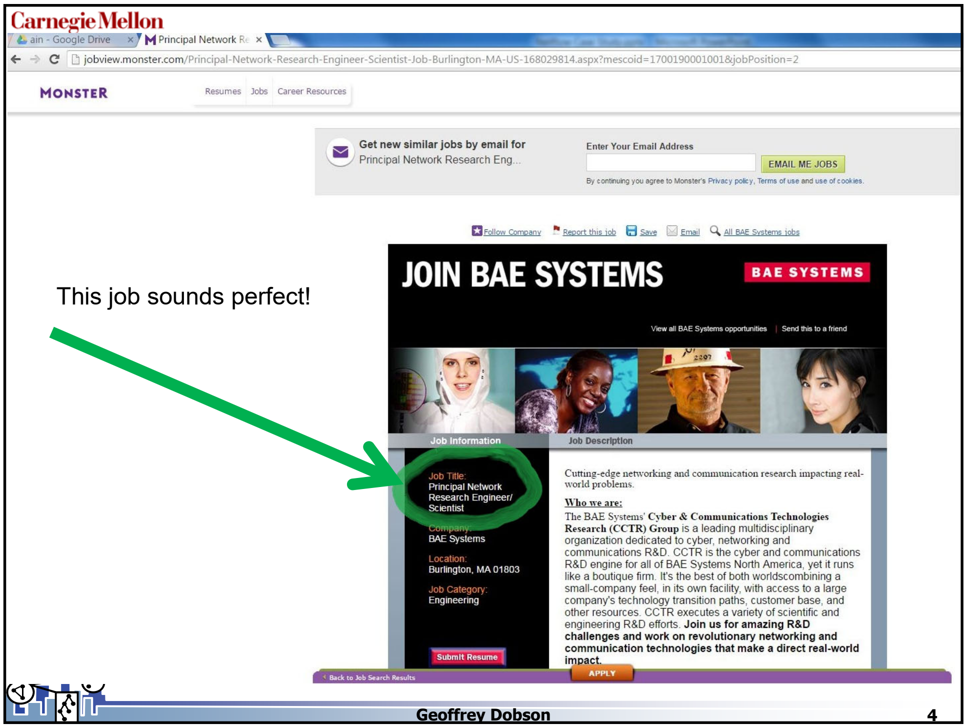


Overview

- Graduate
- Apply for jobs
- Land a new job
- Get direction from your customer
- Do your job (the hands on part)

Graduate

✓ Ph.D.



This job sounds perfect!

ain - Google Drive
Principal Network Re
jobview.monster.com/Principal-Network-Research-Engineer-Scientist-Job-Burlington-MA-US-168029814.aspx?mescoid=1700190001001&jobPosition=2
MONSTER
Resumes
Jobs
Career Resources

Get new similar jobs by email for
Principal Network Research Eng...

Enter Your Email Address

EMAIL ME JOBS

By continuing you agree to Monster's Privacy policy, Terms of use and use of cookies.

Follow Company
Report this job
Save
Email
All BAE Systems jobs

JOIN BAE SYSTEMS

BAE SYSTEMS
View all BAE Systems opportunities | Send this to a friend
2207

Job Information

Job Title:
Principal Network Research Engineer/ Scientist

Company:
BAE Systems

Location:
Burlington, MA 01803

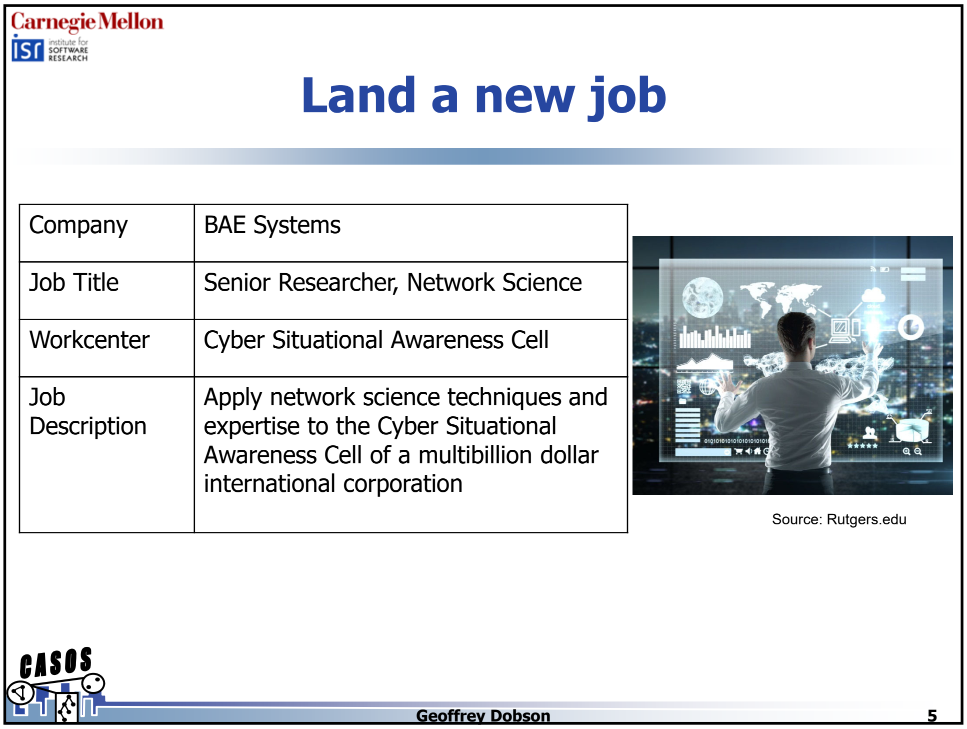
Job Category:
Engineering

Job Description

Cutting-edge networking and communication research impacting real-world problems.

Who we are:

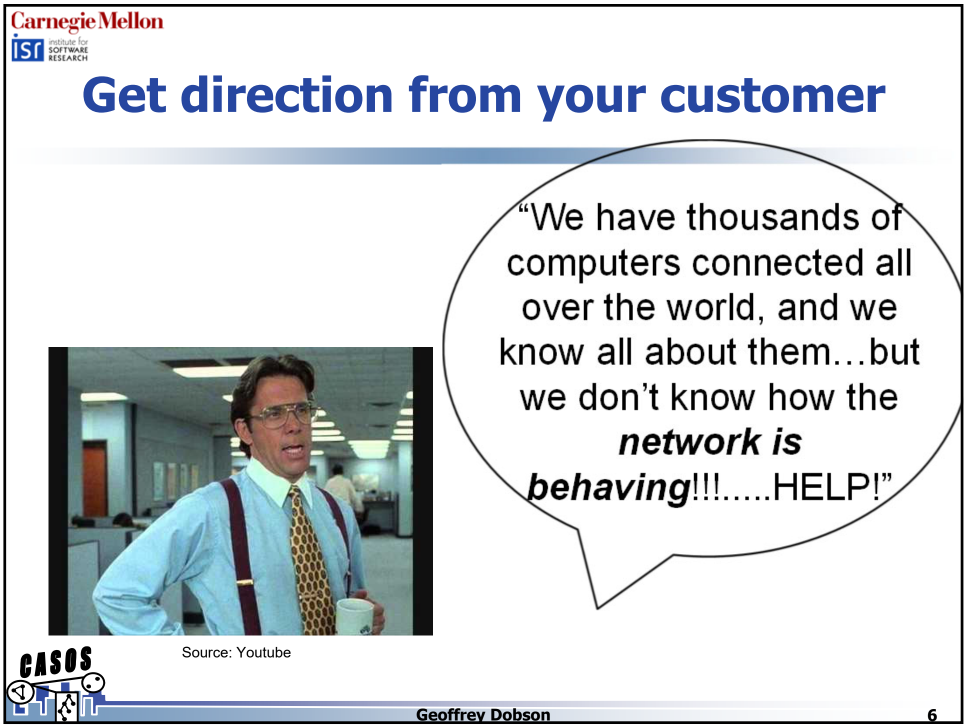
The BAE Systems' **Cyber & Communications Technologies Research (CCTR) Group** is a leading multidisciplinary organization dedicated to cyber, networking and communications R&D. CCTR is the cyber and communications R&D engine for all of BAE Systems North America, yet it runs like a boutique firm. It's the best of both worldscombining a small-company feel, in its own facility, with access to a large company's technology transition paths, customer base, and other resources. CCTR executes a variety of scientific and engineering R&D efforts. **Join us for amazing R&D challenges and work on revolutionary networking and communication technologies that make a direct real-world impact.**



Land a new job

Company	BAE Systems
Job Title	Senior Researcher, Network Science
Workcenter	Cyber Situational Awareness Cell
Job Description	Apply network science techniques and expertise to the Cyber Situational Awareness Cell of a multibillion dollar international corporation

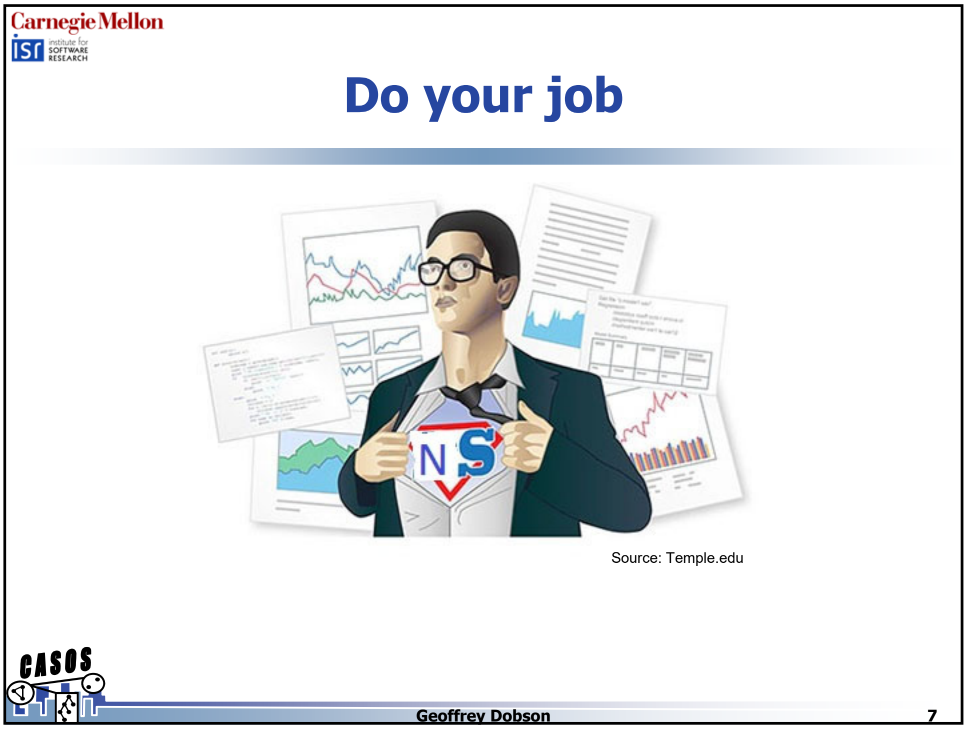
Source: Rutgers.edu



Get direction from your customer

"We have thousands of computers connected all over the world, and we know all about them…but we don't know how the ***network is behaving***!!!.....HELP!"

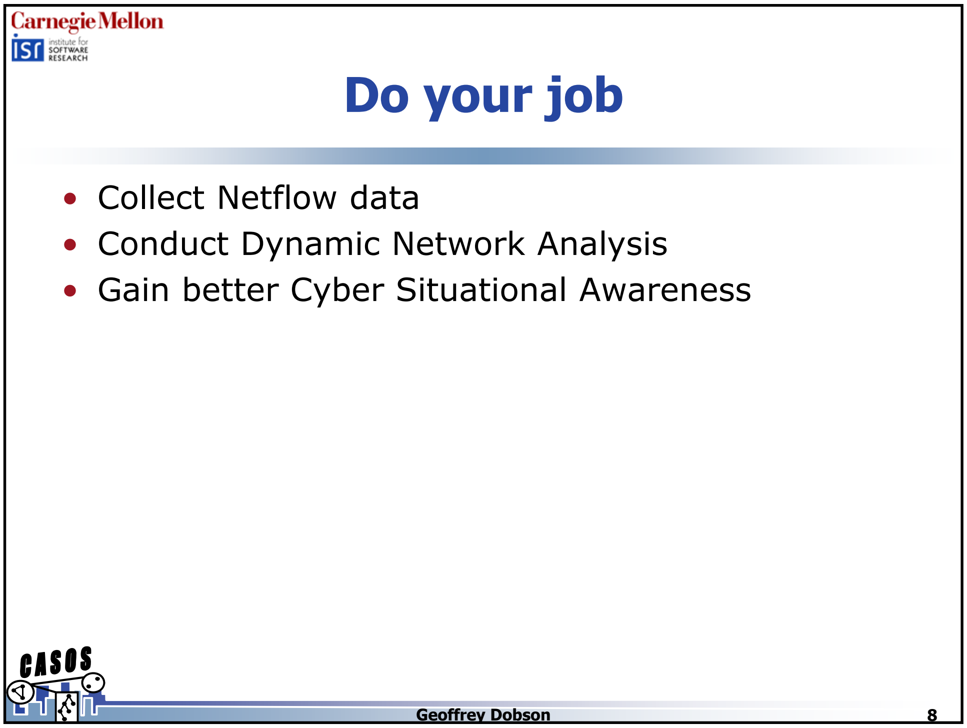
Source: Youtube



Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Do your job

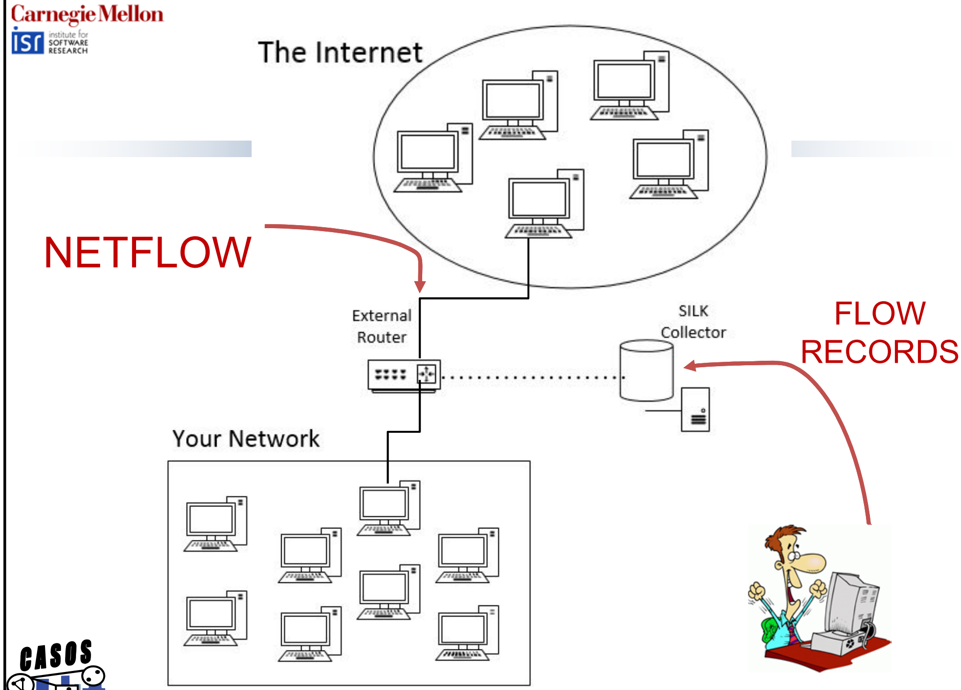
Source: Temple.edu



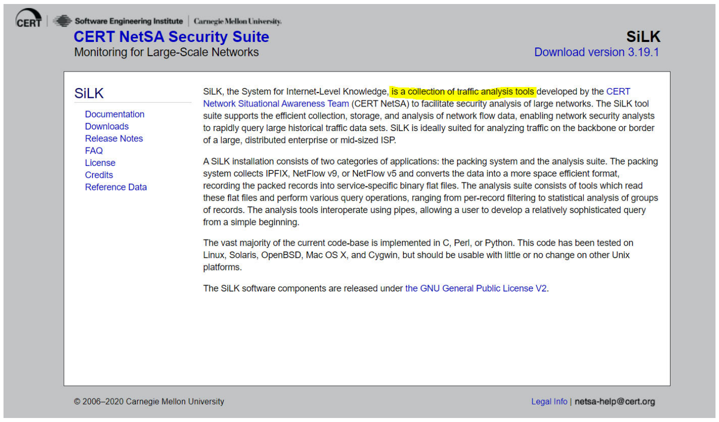
Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Do your job

- Collect Netflow data
- Conduct Dynamic Network Analysis
- Gain better Cyber Situational Awareness



SiLK



CERT | Software Engineering Institute | Carnegie Mellon University

CERT NetSA Security Suite

Monitoring for Large-Scale Networks

SiLK

Download version 3.19.1

SiLK

- Documentation
- Downloads
- Release Notes
- FAQ
- License
- Credits
- Reference Data

SiLK, the System for Internet-Level Knowledge, is a collection of traffic analysis tools developed by the CERT Network Situational Awareness Team (CERT NetSA) to facilitate security analysis of large networks. The SiLK tool suite supports the efficient collection, storage, and analysis of network flow data, enabling network security analysts to rapidly query large historical traffic data sets. SiLK is ideally suited for analyzing traffic on the backbone or border of a large, distributed enterprise or mid-sized ISP.

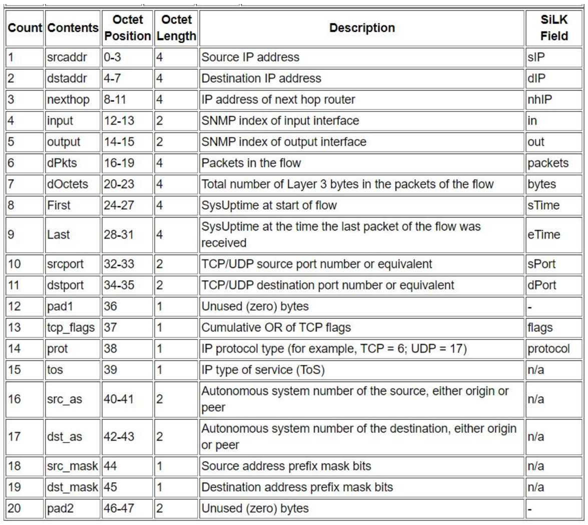
A SiLK installation consists of two categories of applications: the packing system and the analysis suite. The packing system collects IPFIX, NetFlow v9, or NetFlow v5 and converts the data into a more space efficient format, recording the packed records into service-specific binary flat files. The analysis suite consists of tools which read these flat files and perform various query operations, ranging from per-record filtering to statistical analysis of groups of records. The analysis tools interoperate using pipes, allowing a user to develop a relatively sophisticated query from a simple beginning.

The vast majority of the current code-base is implemented in C, Perl, or Python. This code has been tested on Linux, Solaris, OpenBSD, Mac OS X, and Cygwin, but should be usable with little or no change on other Unix platforms.

The SiLK software components are released under the GNU General Public License V2.

Legal Info | netsa-help@cert.org

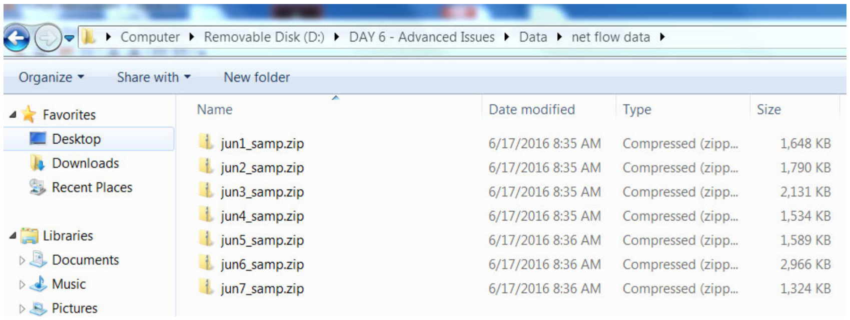
SiLK



Count	Contents	Octet Position	Octet Length	Description	SiLK Field
1	srcaddr	0-3	4	Source IP address	sIP
2	dstaddr	4-7	4	Destination IP address	dIP
3	nexthop	8-11	4	IP address of next hop router	nhIP
4	input	12-13	2	SNMP index of input interface	in
5	output	14-15	2	SNMP index of output interface	out
6	dPkts	16-19	4	Packets in the flow	packets
7	dOctets	20-23	4	Total number of Layer 3 bytes in the packets of the flow	bytes
8	First	24-27	4	SysUptime at start of flow	sTime
9	Last	28-31	4	SysUptime at the time the last packet of the flow was received	eTime
10	srcport	32-33	2	TCP/UDP source port number or equivalent	sPort
11	dstport	34-35	2	TCP/UDP destination port number or equivalent	dPort
12	pad1	36	1	Unused (zero) bytes	-
13	tcp_flags	37	1	Cumulative OR of TCP flags	flags
14	prot	38	1	IP protocol type (for example, TCP = 6; UDP = 17)	protocol
15	tos	39	1	IP type of service (ToS)	n/a
16	src_as	40-41	2	Autonomous system number of the source, either origin or peer	n/a
17	dst_as	42-43	2	Autonomous system number of the destination, either origin or peer	n/a
18	src_mask	44	1	Source address prefix mask bits	n/a
19	dst_mask	45	1	Destination address prefix mask bits	n/a
20	pad2	46-47	2	Unused (zero) bytes	-

Collect Netflow Data

1. Go Data -> net flow data

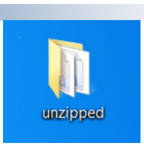


Name	Date modified	Type	Size
jun1_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	1,648 KB
jun2_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	1,790 KB
jun3_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	2,131 KB
jun4_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	1,534 KB
jun5_samp.zip	6/17/2016 8:36 AM	Compressed (zipp...	1,589 KB
jun6_samp.zip	6/17/2016 8:36 AM	Compressed (zipp...	2,966 KB
jun7_samp.zip	6/17/2016 8:36 AM	Compressed (zipp...	1,324 KB

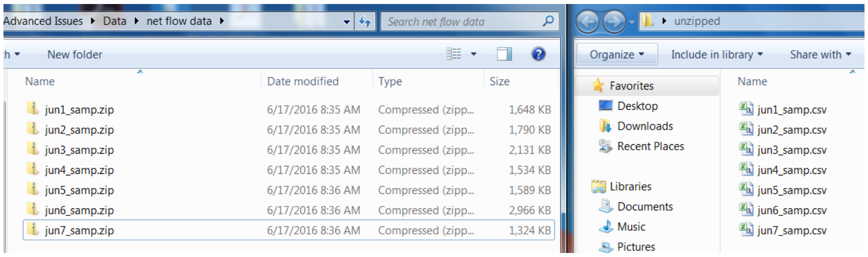
Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Collect Netflow Data

2. Create New Folder on Desktop called “Unzipped”



3. Go to Data drive and right click on each zip file, and extract to Unzipped Folder



Name	Date modified	Type	Size
jun1_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	1,648 KB
jun2_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	1,790 KB
jun3_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	2,131 KB
jun4_samp.zip	6/17/2016 8:35 AM	Compressed (zipp...	1,534 KB
jun5_samp.zip	6/17/2016 8:36 AM	Compressed (zipp...	1,589 KB
jun6_samp.zip	6/17/2016 8:36 AM	Compressed (zipp...	2,966 KB
jun7_samp.zip	6/17/2016 8:36 AM	Compressed (zipp...	1,324 KB

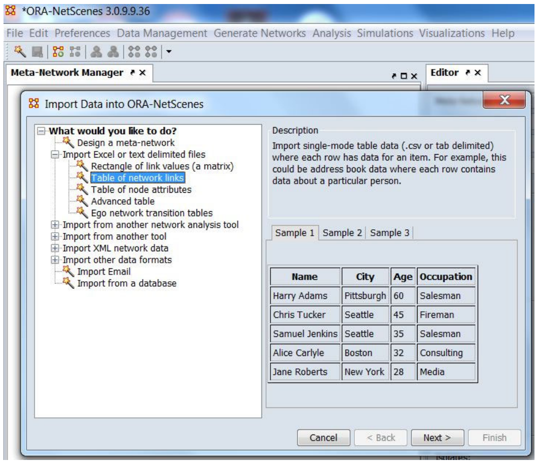
Name
jun1_samp.csv
jun2_samp.csv
jun3_samp.csv
jun4_samp.csv
jun5_samp.csv
jun6_samp.csv
jun7_samp.csv

CASOS

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Collect Netflow Data

4. Open Import Wizard and select Table of network links


*ORA-NetScenes 3.0.9.9.36
File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations Help
Meta-Network Manager
Editor
Import Data into ORA-NetScenes
Cancel
< Back
Next >
Finish

What would you like to do?
- Design a meta-network
- Import Excel or text delimited files
  - Rectangle of link values (a matrix)
  - Table of network links
  - Table of node attributes
  - Advanced table
  - Ego network transition tables
- Import from another network analysis tool
- Import from another tool
- Import XML network data
- Import other data formats
- Import Email
- Import from a database

Description

Import single-mode table data (.csv or tab delimited) where each row has data for an item. For example, this could be address book data where each row contains data about a particular person.

Sample 1 | Sample 2 | Sample 3

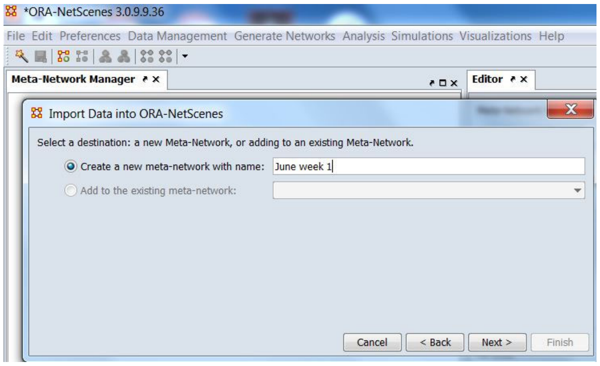
Name	City	Age	Occupation
Harry Adams	Pittsburgh	60	Salesman
Chris Tucker	Seattle	45	Fireman
Samuel Jenkins	Seattle	35	Salesman
Alice Carlyle	Boston	32	Consulting
Jane Roberts	New York	28	Media

CASOS

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Collect Netflow Data

5. Name the Meta Network


*ORA-NetScenes 3.0.9.9.36
File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations Help
Meta-Network Manager
Editor
Import Data into ORA-NetScenes
Cancel
< Back
Next >
Finish

Select a destination: a new Meta-Network, or adding to an existing Meta-Network.

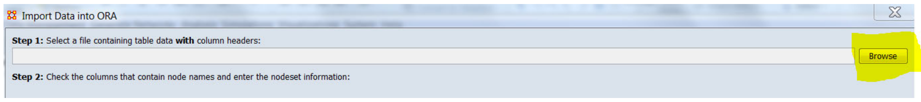
● Create a new meta-network with name: June week 1

○ Add to the existing meta-network:

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

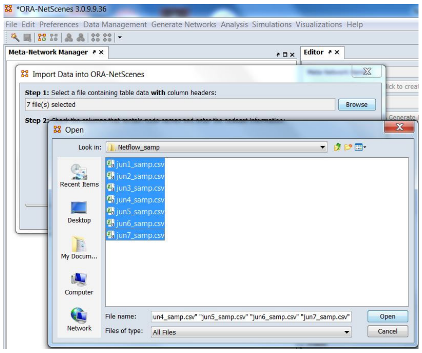
Collect Netflow Data

6. Browse to files



Step 1: Select a file containing table data **with** column headers:

Step 2: Check the columns that contain node names and enter the nodeset information:


*ORA-NetScenes 3.0.9.9.36
File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations Help
Meta-Network Manager
Editor
Import Data into ORA-NetScenes
Browse
Open
Recent Items
Desktop
My Docum...
Computer
Network
Open
Cancel

Step 1: Select a file containing table data **with** column headers:

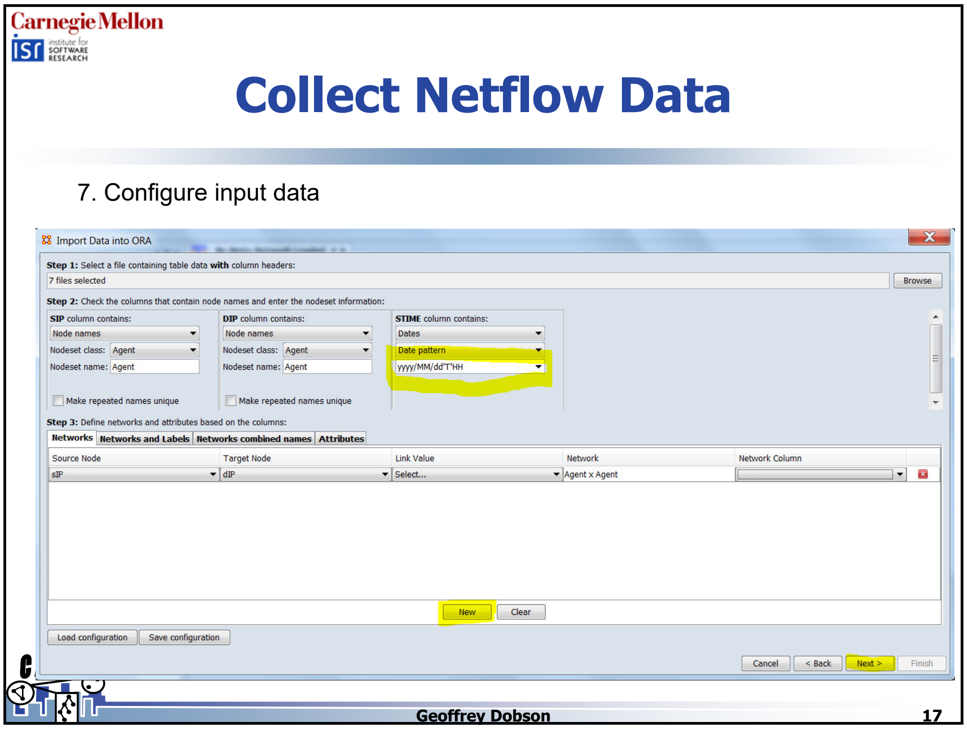
7 file(s) selected

Look in: Netflow_samp

jun1_samp.csv
jun2_samp.csv
jun3_samp.csv
jun4_samp.csv
jun5_samp.csv
jun6_samp.csv
jun7_samp.csv

File name: un4_samp.csv" "jun5_samp.csv" "jun6_samp.csv" "jun7_samp.csv"

Files of type: All Files



Collect Netflow Data

7. Configure input data

Import Data into ORA

Step 1: Select a file containing table data **with** column headers:

7 files selected

Step 2: Check the columns that contain node names and enter the nodeset information:

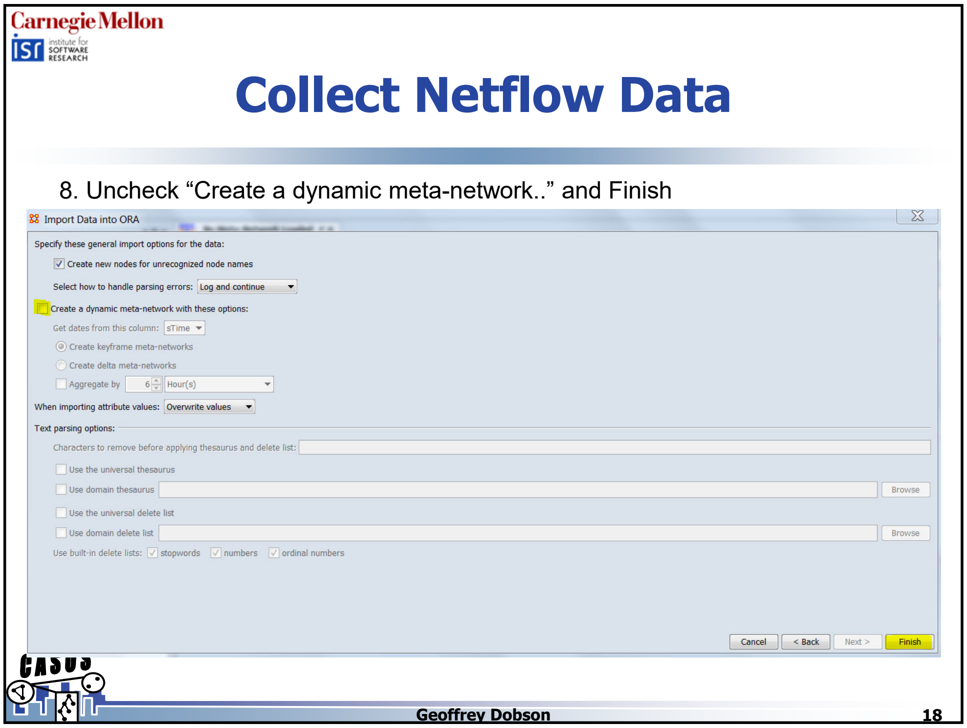
SIP column contains: Node names
Nodeset class: Agent
Nodeset name: Agent
□ Make repeated names unique

DIP column contains: Node names
Nodeset class: Agent
Nodeset name: Agent
□ Make repeated names unique

STIME column contains: Dates
Date pattern
yyyy/MM/dd'T'HH

Step 3: Define networks and attributes based on the columns:

Source Node	Target Node	Link Value	Network	Network Column
sIP	dIP	Select...	Agent x Agent	



Collect Netflow Data

8. Uncheck “Create a dynamic meta-network..” and Finish

Import Data into ORA

Specify these general import options for the data:

☑ Create new nodes for unrecognized node names

Select how to handle parsing errors: Log and continue

□ Create a dynamic meta-network with these options:

Get dates from this column: sTime

● Create keyframe meta-networks

○ Create delta meta-networks

□ Aggregate by 6 Hour(s)

When importing attribute values: Overwrite values

Text parsing options:

Characters to remove before applying thesaurus and delete list:

□ Use the universal thesaurus

□ Use domain thesaurus

□ Use the universal delete list

□ Use domain delete list

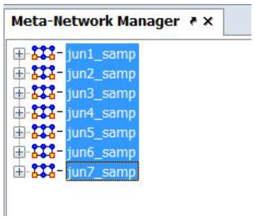
Use built-in delete lists: ☑ stopwords ☑ numbers ☑ ordinal numbers

Understand your data

- Describe your network data:
  - Undirected single mode network
  - Agent by Agent meta network
  - Bipartite graph
  - Flow records per day?
    - ~200,000
  - Links per day?
    - ~ 130,000
  - Nodes per day?
    - ~ 22,000

Perform Dynamic Network Analysis

1. Create a dynamic meta-network

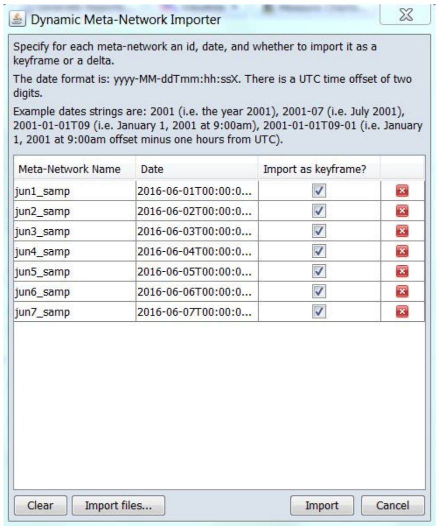


Meta-Network Manager

- jun1_samp
- jun2_samp
- jun3_samp
- jun4_samp
- jun5_samp
- jun6_samp
- jun7_samp

Perform Dynamic Network Analysis

2. Fill in Date field



Dynamic Meta-Network Importer

Specify for each meta-network an id, date, and whether to import it as a keyframe or a delta.

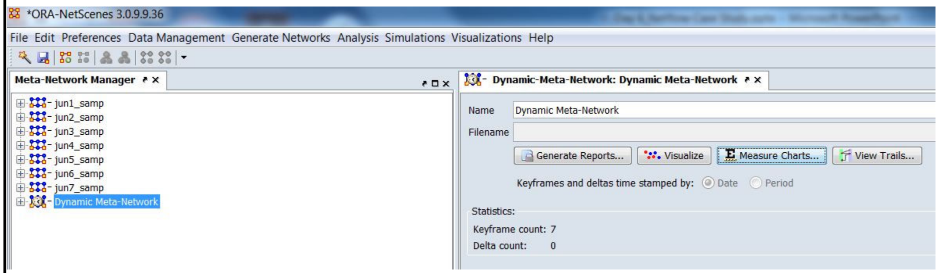
The date format is: yyyy-MM-ddTmm:hh:ssX. There is a UTC time offset of two digits.

Example dates strings are: 2001 (i.e. the year 2001), 2001-07 (i.e. July 2001), 2001-01-01T09 (i.e. January 1, 2001 at 9:00am), 2001-01-01T09-01 (i.e. January 1, 2001 at 9:00am offset minus one hours from UTC).

Meta-Network Name	Date	Import as keyframe?	
jun1_samp	2016-06-01T00:00:0...	☑	☒
jun2_samp	2016-06-02T00:00:0...	☑	☒
jun3_samp	2016-06-03T00:00:0...	☑	☒
jun4_samp	2016-06-04T00:00:0...	☑	☒
jun5_samp	2016-06-05T00:00:0...	☑	☒
jun6_samp	2016-06-06T00:00:0...	☑	☒
jun7_samp	2016-06-07T00:00:0...	☑	☒

Perform Dynamic Network Analysis

3. Click Measure Charts



*ORA-NetScenes 3.0.9.9.36
File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations Help
Meta-Network Manager
jun1_samp
jun2_samp
jun3_samp
jun4_samp
jun5_samp
jun6_samp
jun7_samp
Dynamic Meta-Network
Dynamic-Meta-Network: Dynamic Meta-Network
Generate Reports...
Visualize
Measure Charts...
View Trails...

Name: Dynamic Meta-Network

Filename:

Keyframes and deltas time stamped by: ● Date ○ Period

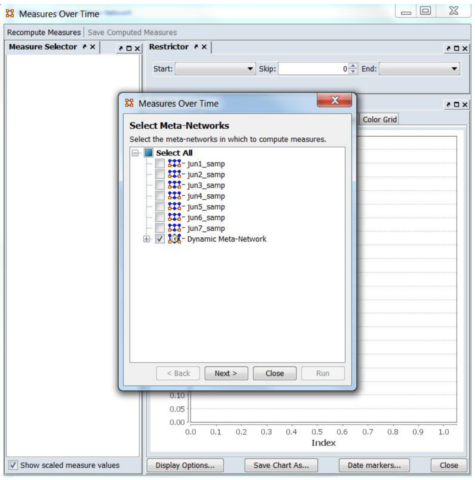
Statistics:

Keyframe count: 7

Delta count: 0

Perform Dynamic Network Analysis

4. Select the Dynamic Meta Network



Measures Over Time

Select Meta-Networks

Select the meta-networks in which to compute measures.

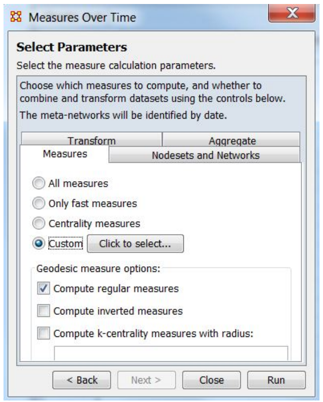
- Select All
  - □ jun1_samp
  - □ jun2_samp
  - □ jun3_samp
  - □ jun4_samp
  - □ jun5_samp
  - □ jun6_samp
  - □ jun7_samp
  - ☑ Dynamic Meta-Network

< Back | Next > | Close | Run

☑ Show scaled measure values

Perform Dynamic Network Analysis

5. Select Custom: Density and Network Centralization, Total Degree, Click Run


Measures Over Time

Select Parameters

Select the measure calculation parameters.

Choose which measures to compute, and whether to combine and transform datasets using the controls below. The meta-networks will be identified by date.

○ All measures

○ Only fast measures

○ Centrality measures

● Custom Click to select...

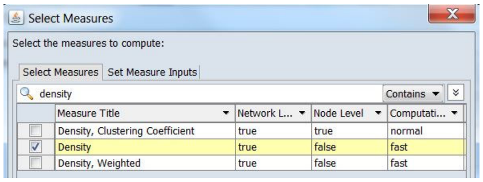
Geodesic measure options:

☑ Compute regular measures

□ Compute inverted measures

□ Compute k-centrality measures with radius:

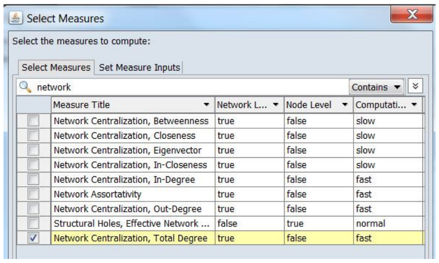
< Back | Next > | Close | Run


Select Measures

Select the measures to compute:

density Contains

	Measure Title	Network L...	Node Level	Computati...
□	Density, Clustering Coefficient	true	true	normal
☑	Density	true	false	fast
□	Density, Weighted	true	false	fast


Select Measures

Select the measures to compute:

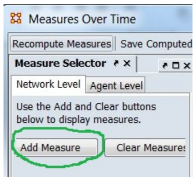
network Contains

	Measure Title	Network L...	Node Level	Computati...
□	Network Centralization, Betweenness	true	false	slow
□	Network Centralization, Closeness	true	false	slow
□	Network Centralization, Eigenvector	true	false	slow
□	Network Centralization, In-Closeness	true	false	slow
□	Network Centralization, In-Degree	true	false	fast
□	Network Assortativity	true	false	fast
□	Network Centralization, Out-Degree	true	false	fast
□	Structural Holes, Effective Network ...	false	true	normal
☑	Network Centralization, Total Degree	true	false	fast

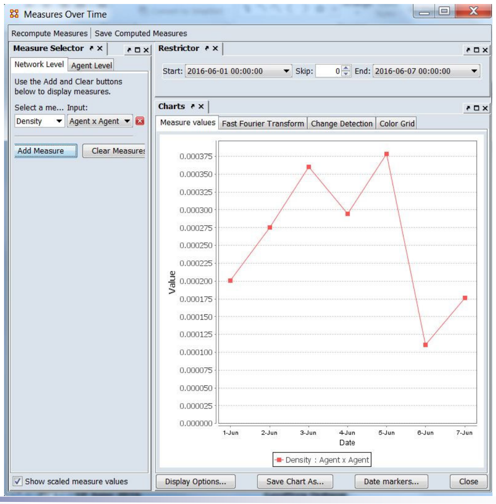
Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Perform Dynamic Network Analysis

6. Add Measure, then view various results



Use the Add and Clear buttons below to display measures.



Use the Add and Clear buttons below to display measures.

Select a me... Input: Density Agent x Agent

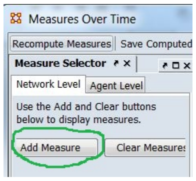
Start: 2016-06-01 00:00:00 Skip: 0 End: 2016-06-07 00:00:00

☑ Show scaled measure values

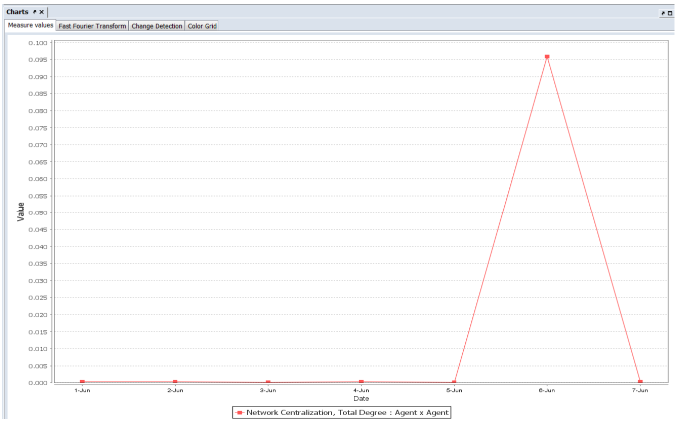
Carnegie Mellon
isr institute for SOFTWARE RESEARCH

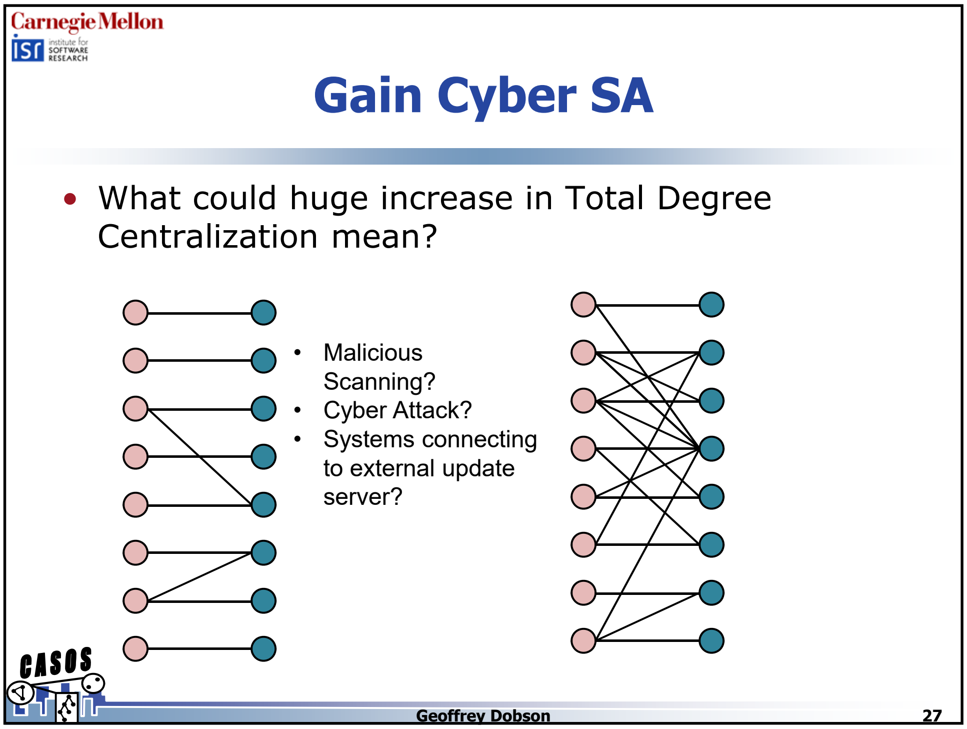
Perform Dynamic Network Analysis

6. continued



Use the Add and Clear buttons below to display measures.

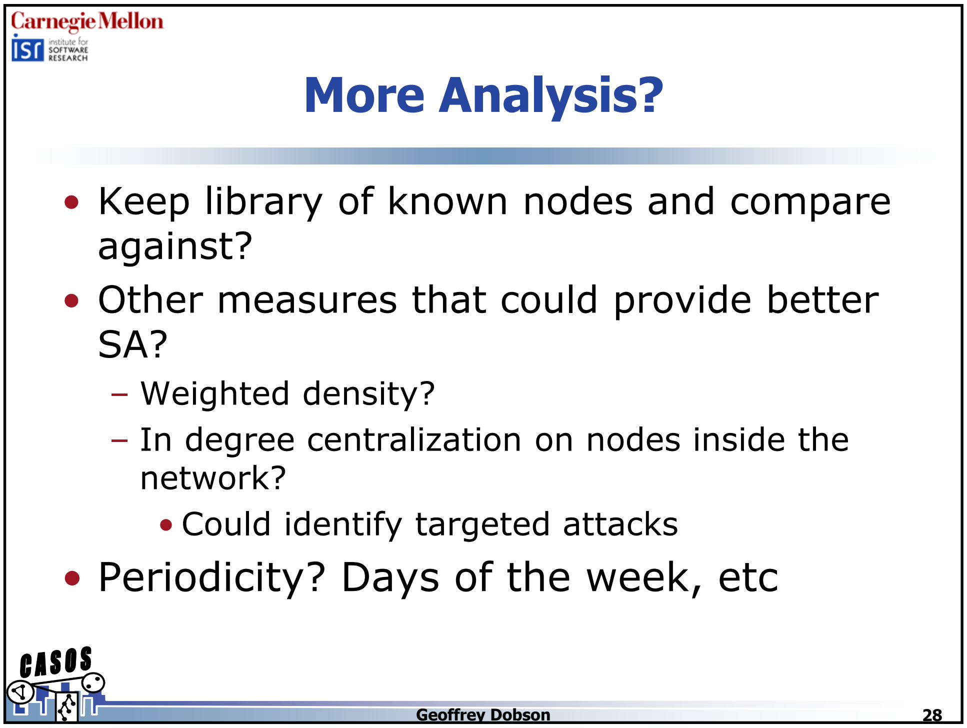




Gain Cyber SA

- What could huge increase in Total Degree Centralization mean?

- Malicious Scanning?
- Cyber Attack?
- Systems connecting to external update server?



More Analysis?

- Keep library of known nodes and compare against?
- Other measures that could provide better SA?
  - Weighted density?
  - In degree centralization on nodes inside the network?
    - Could identify targeted attacks
- Periodicity? Days of the week, etc