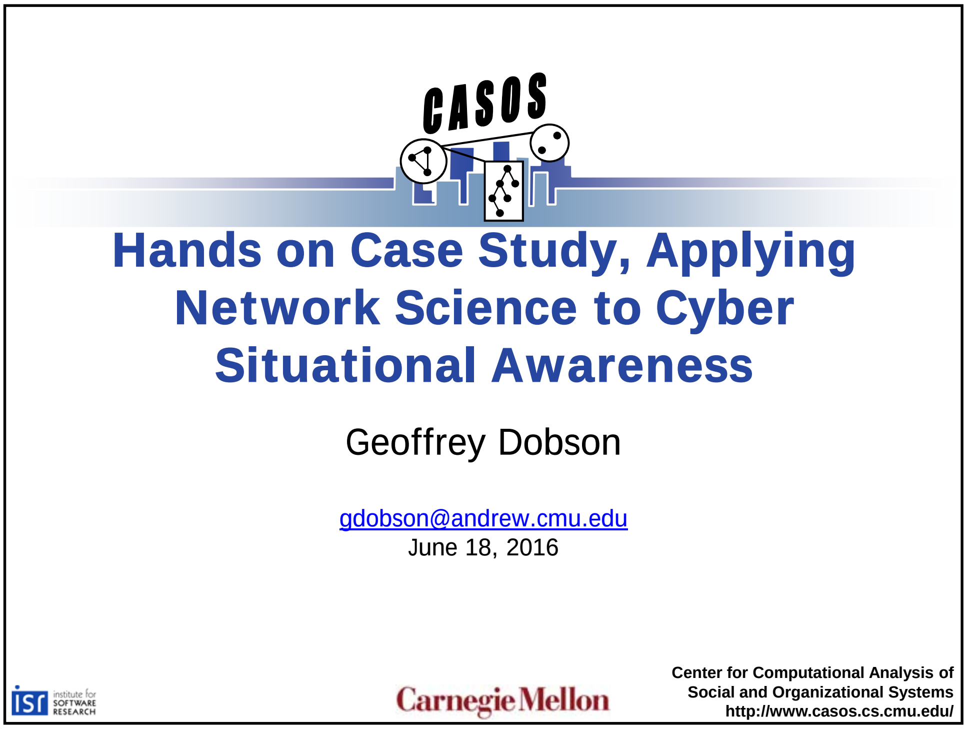


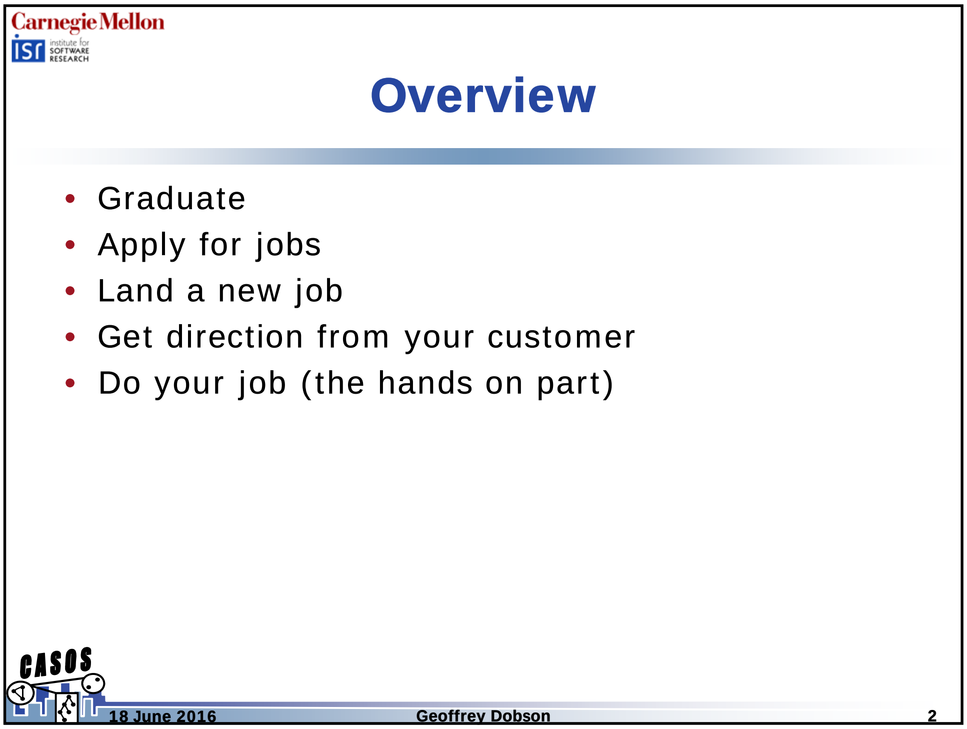
Hands on Case Study, Applying Network Science to Cyber Situational Awareness

Geoffrey Dobson

gdobson@andrew.cmu.edu

June 18, 2016

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



Overview

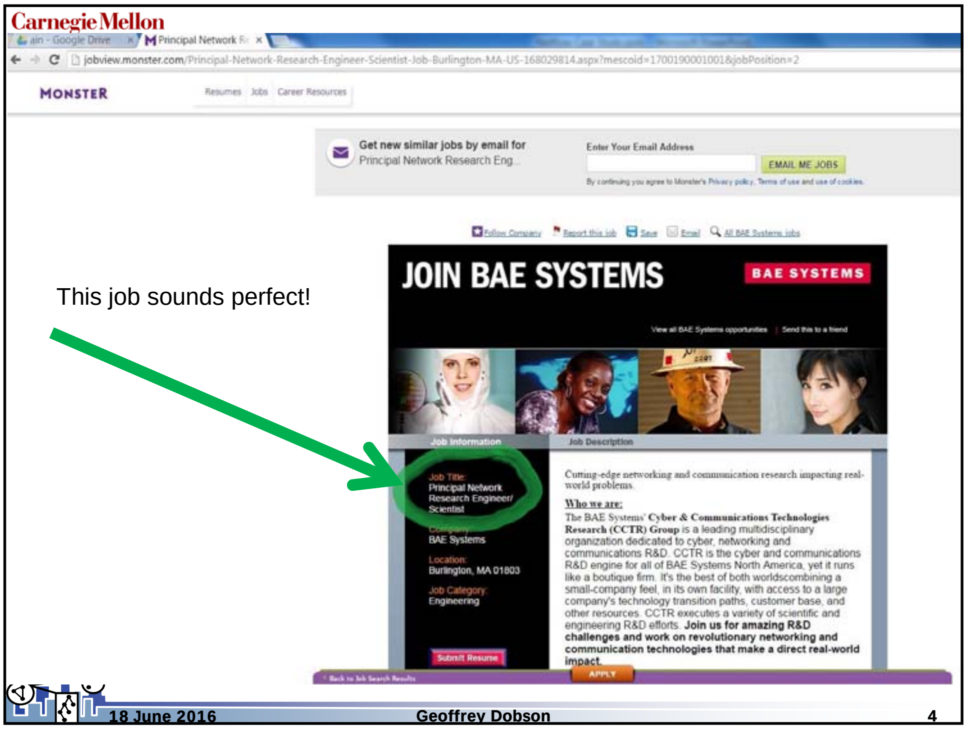
- Graduate
- Apply for jobs
- Land a new job
- Get direction from your customer
- Do your job (the hands on part)

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Graduate

✓ Ph.D.

Carnegie Mellon



jobview.monster.com/Principal-Network-Research-Engineer-Scientist-Job-Burlington-MA-US-168029814.aspx?mescoid=1700190001001&jobPosition=2

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

Enter Your Email Address

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

This job sounds perfect!

Job Information

Job Title:
Principal Network Research Engineer/ Scientist

Company:
BAE Systems

Location:
Burlington, MA 01803

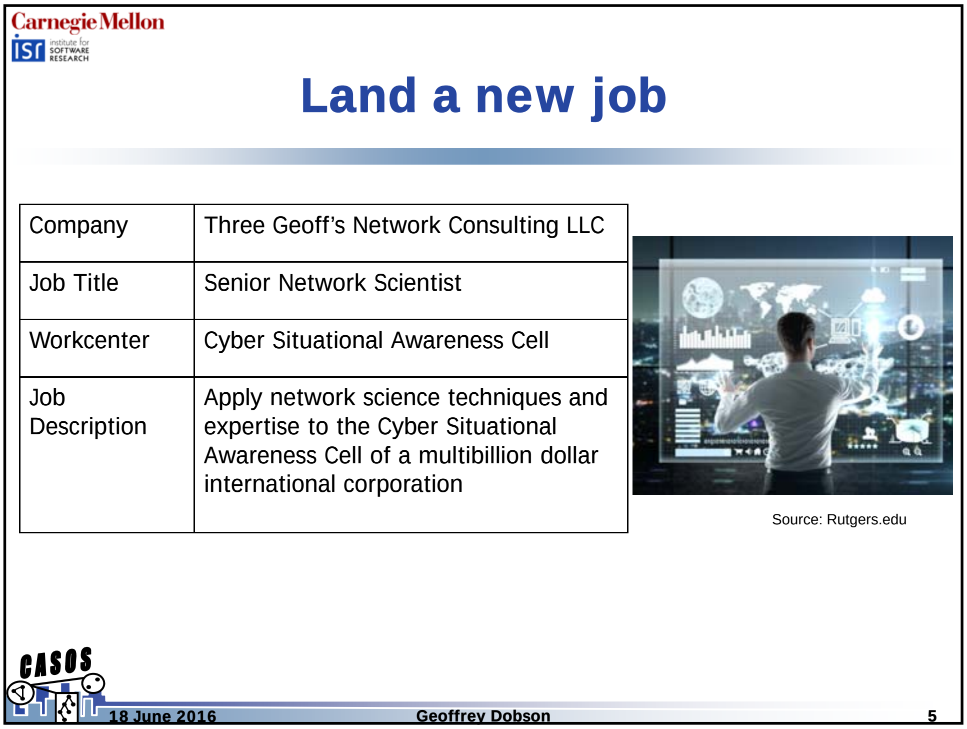
Job Category:
Engineering

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.**

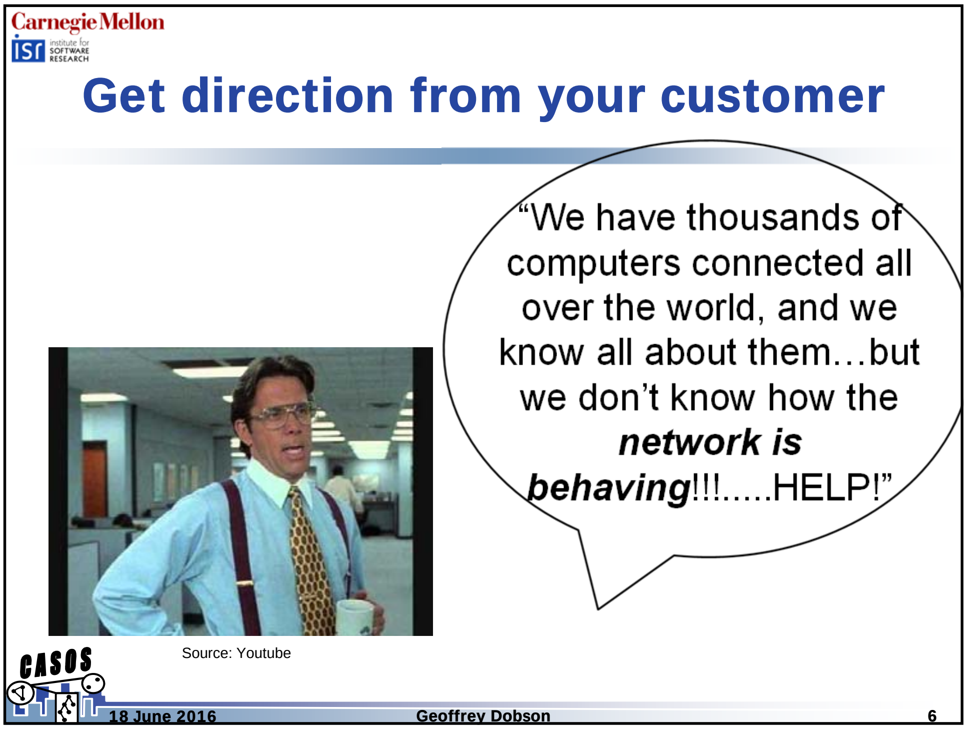


Carnegie Mellon

Land a new job

Company	Three Geoff’s Network Consulting LLC
Job Title	Senior Network Scientist
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



Carnegie Mellon

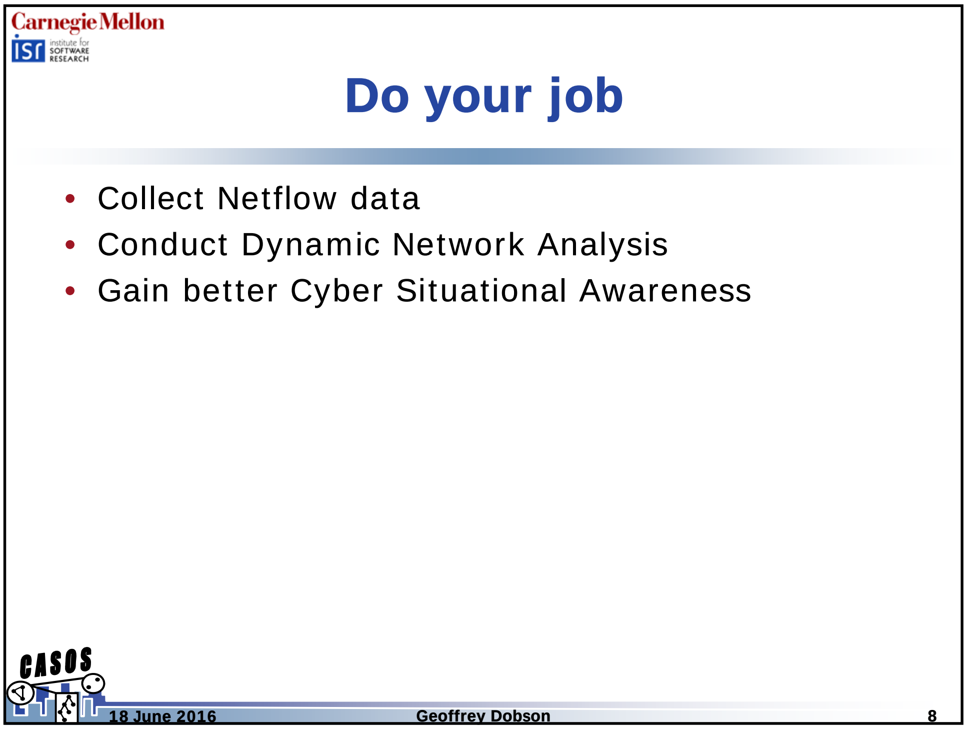
Get direction from your customer

“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

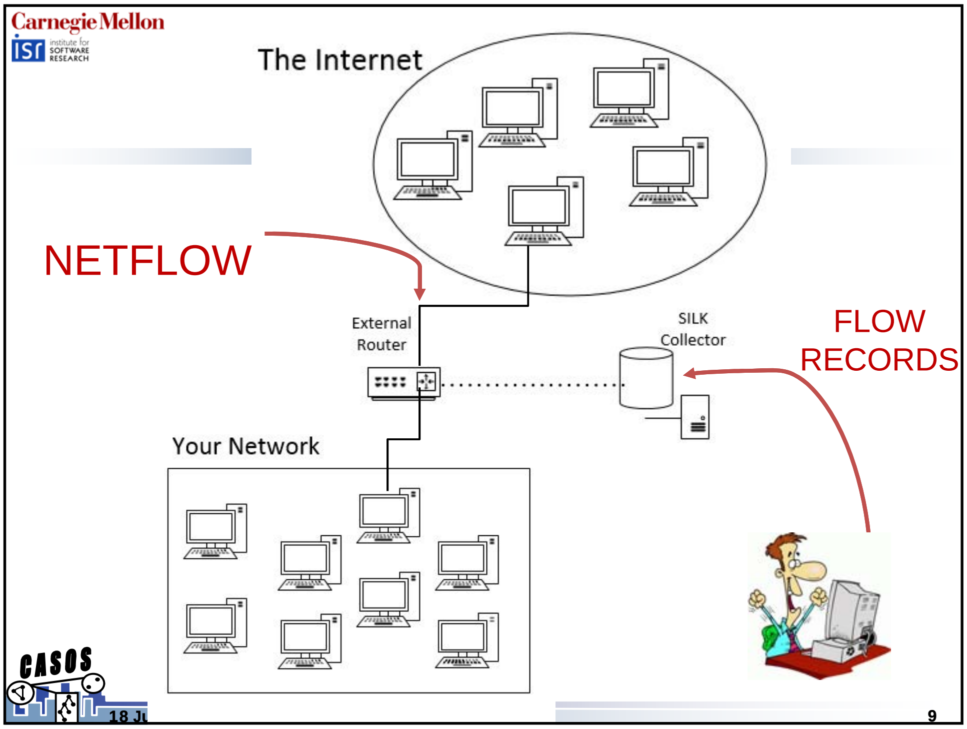
Do your job

Source: Temple.edu



Do your job

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



Carnegie Mellon
isr institute for SOFTWARE RESEARCH

The Internet

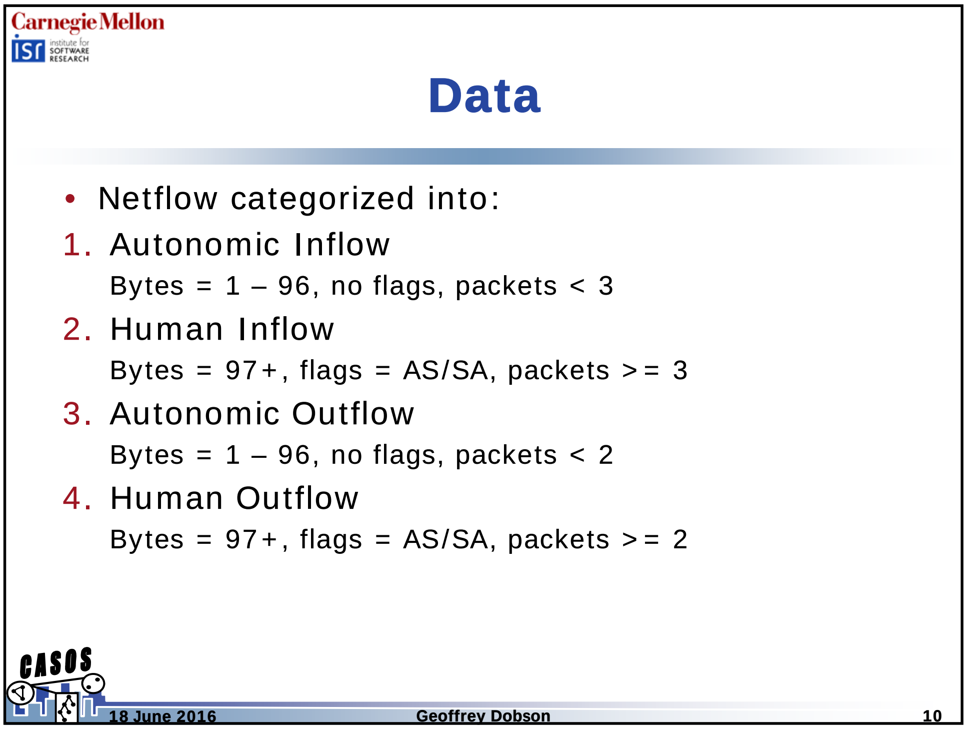
NETFLOW

External Router

SILK Collector

FLOW RECORDS

Your Network



Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Data

- Netflow categorized into:

1. Autonomic Inflow

   Bytes = 1 – 96, no flags, packets < 3

2. Human Inflow

   Bytes = 97+, flags = AS/SA, packets >= 3

3. Autonomic Outflow

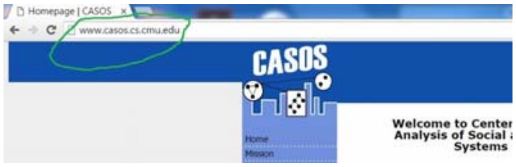
   Bytes = 1 – 96, no flags, packets < 2

4. Human Outflow

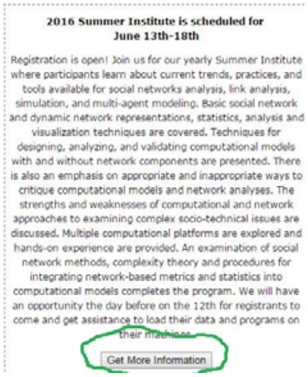
   Bytes = 97+, flags = AS/SA, packets >= 2

Collect Netflow Data

1. Go to casos.cs.cmu.edu



2. Click Get More Information

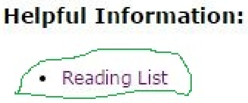


2016 Summer Institute is scheduled for June 13th-18th

Registration is open! Join us for our yearly Summer Institute where participants learn about current trends, practices, and tools available for social networks analysis, link analysis, simulation, and multi-agent modeling. Basic social network and dynamic network representations, statistics, analysis and visualization techniques are covered. Techniques for designing, analyzing, and validating computational models with and without network components are presented. There is also an emphasis on appropriate and inappropriate ways to critique computational models and network analyses. The strengths and weaknesses of computational and network approaches to examining complex socio-technical issues are discussed. Multiple computational platforms are explored and hands-on experience are provided. An examination of social network methods, complexity theory and procedures for integrating network-based metrics and statistics into computational models completes the program. We will have an opportunity the day before on the 12th for registrants to come and get assistance to load their data and programs on their machines.

Collect Netflow Data

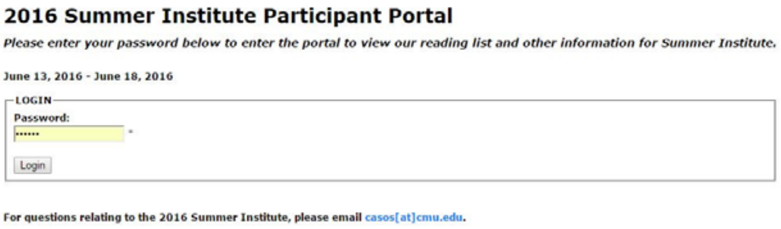
3. Click Reading List



Helpful Information:

- Reading List

4. Login with “si2016”



2016 Summer Institute Participant Portal

Please enter your password below to enter the portal to view our reading list and other information for Summer Institute.

June 13, 2016 - June 18, 2016

LOGIN

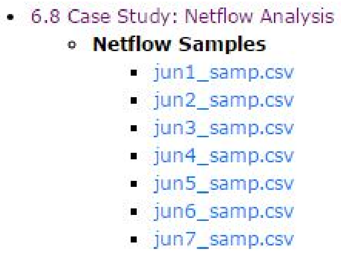
Password: ••••••

For questions relating to the 2016 Summer Institute, please email casos[at]cmu.edu.

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Collect Netflow Data

5. Download Netflow samples

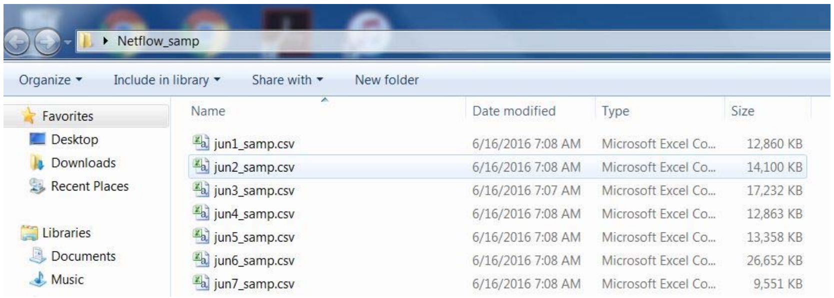


- 6.8 Case Study: Netflow Analysis
  - **Netflow Samples**
    - jun1_samp.csv
    - jun2_samp.csv
    - jun3_samp.csv
    - jun4_samp.csv
    - jun5_samp.csv
    - jun6_samp.csv
    - jun7_samp.csv

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Collect Netflow Data

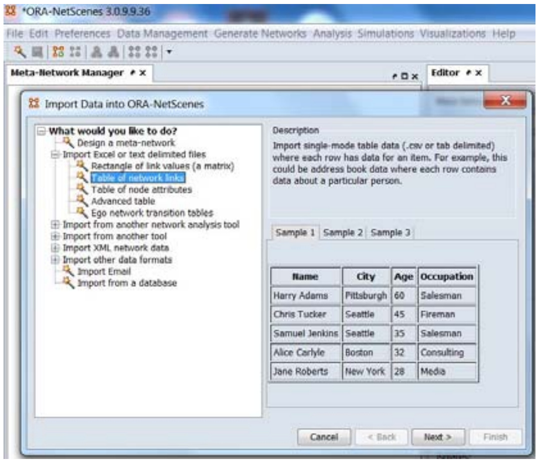
6. Unzip all to directory on Desktop



Name	Date modified	Type	Size
jun1_samp.csv	6/16/2016 7:08 AM	Microsoft Excel Co...	12,860 KB
jun2_samp.csv	6/16/2016 7:08 AM	Microsoft Excel Co...	14,100 KB
jun3_samp.csv	6/16/2016 7:07 AM	Microsoft Excel Co...	17,232 KB
jun4_samp.csv	6/16/2016 7:08 AM	Microsoft Excel Co...	12,863 KB
jun5_samp.csv	6/16/2016 7:08 AM	Microsoft Excel Co...	13,358 KB
jun6_samp.csv	6/16/2016 7:08 AM	Microsoft Excel Co...	26,652 KB
jun7_samp.csv	6/16/2016 7:08 AM	Microsoft Excel Co...	9,551 KB

Collect Netflow Data

7. Open Import Wizard and select Table of network links

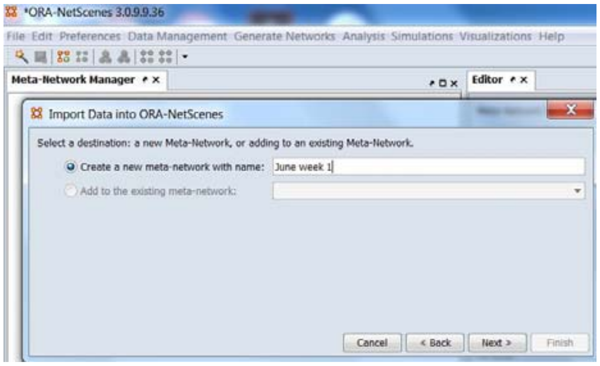

*ORA-NetScenes 3.0.9.9.36
File Edit Preferences Data Management Generate Networks Analysis Simulations Visualizations Help
Meta-Network Manager
Editor
Import Data into ORA-NetScenes
What would you like to do?
Design a meta-network
Import Excel or text delimited files
Rectangle of link values (a matrix)
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
Sample 1
Sample 2
Sample 3
Cancel
< Back
Next >
Finish

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.

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

Collect Netflow Data

8. 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.

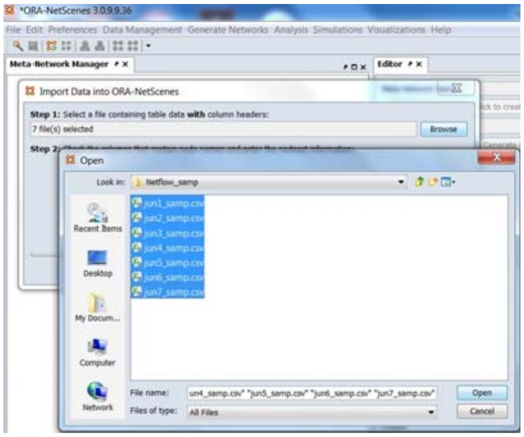
● Create a new meta-network with name: June week 1

○ Add to the existing meta-network:

Carnegie Mellon

Collect Netflow Data

9. Browse to files


*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

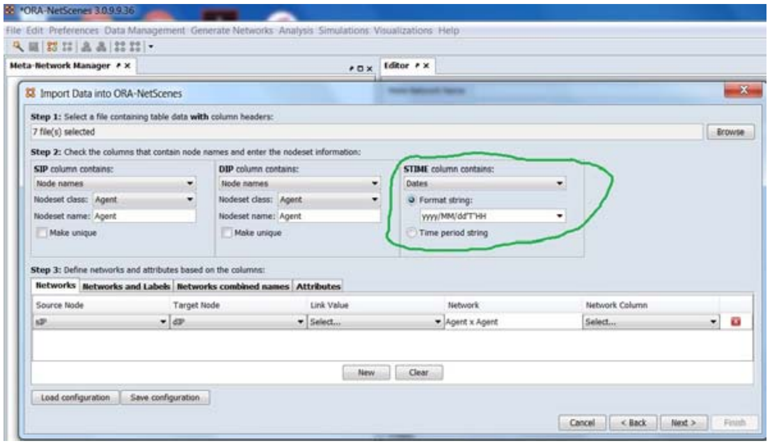
File name: un4_samp.csv" "jun5_samp.csv" "jun6_samp.csv" "jun7_samp.csv"

Files of type: All Files

Carnegie Mellon

Collect Netflow Data

10. Configure input data



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

7 file(s) 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 unique

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

STIME column contains: Dates
● Format string: yyyy/MM/dd'T'HH
○ Time period string

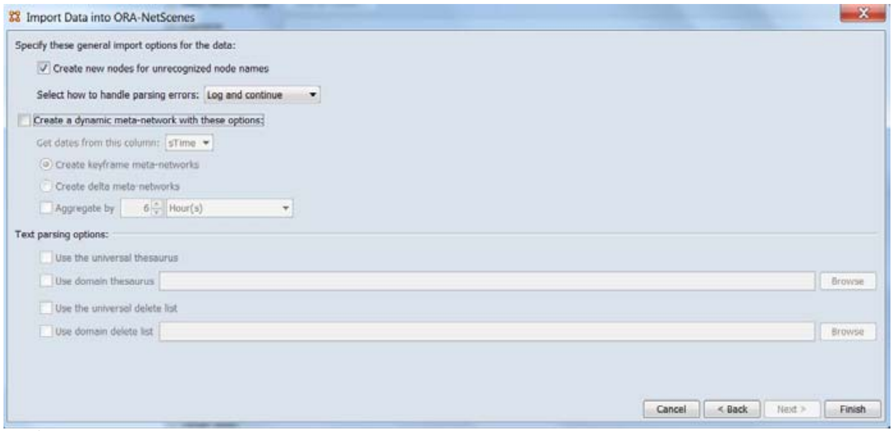
Step 3: Define networks and attributes based on the columns:

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

Carnegie Mellon

Collect Netflow Data

11. Uncheck “Create a dynamic meta-network..”



Import Data into ORA-NetScenes

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)

Text parsing options:

□ Use the universal thesaurus

□ Use domain thesaurus

□ Use the universal delete list

□ Use domain delete list

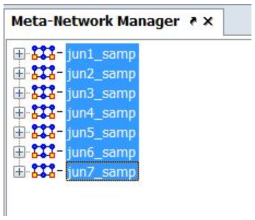
Carnegie Mellon

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?
    - ~ 130,394
  - Nodes per day?
    - ~ 22,032

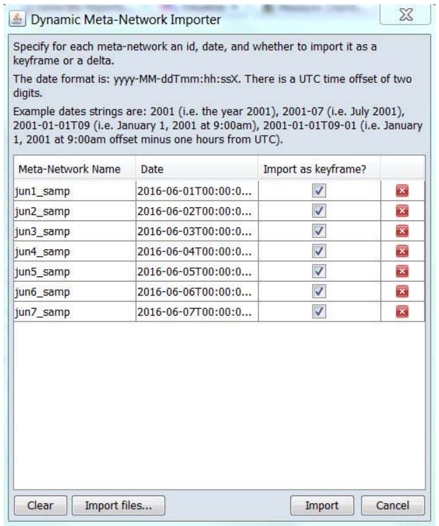
Perform Dynamic Network Analysis

1. Create a dynamic meta-network



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.

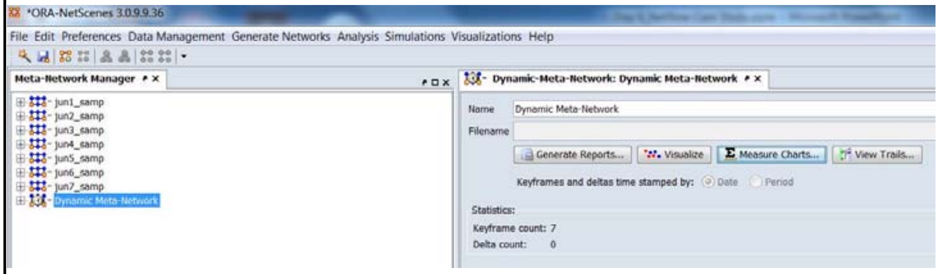
Example dates strings are: 2001 (i.e. the year 2001), 2001-07 (i.e. July 2001), 2001-01-01T09 (i.e. January 1, 2001 at 9:00am), 2001-01-01T09-01 (i.e. January 1, 2001 at 9:00am offset minus one hours from UTC).

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...	☑	☒

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Perform Dynamic Network Analysis

3. Click Measure Charts



Name: Dynamic Meta-Network

Filename:

Keyframes and deltas time stamped by: ● Date ○ Period

Statistics:

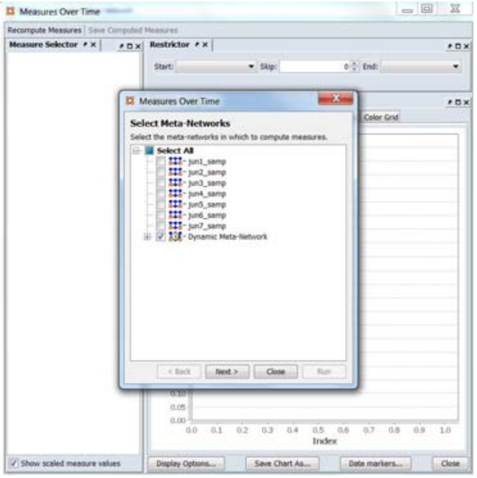
Keyframe count: 7

Delta count: 0

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

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.

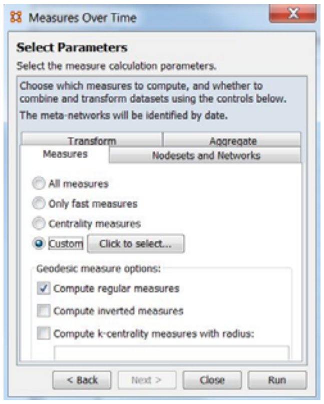
- Select All
  - □ jun1_samp
  - □ jun2_samp
  - □ jun3_samp
  - □ jun4_samp
  - □ jun5_samp
  - □ jun6_samp
  - □ jun7_samp
  - ☑ Dynamic Meta-Network

☑ Show scaled measure values

Carnegie Mellon

Perform Dynamic Network Analysis

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



Select Parameters

Select the measure calculation parameters.

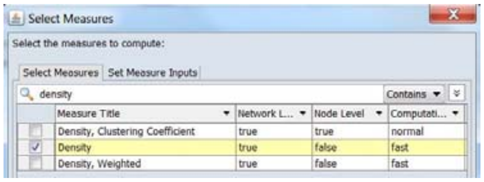
Choose which measures to compute, and whether to combine and transform datasets using the controls below.

The meta-networks will be identified by date.

○ All measures
○ Only fast measures
○ Centrality measures
● Custom

Geodesic measure options:

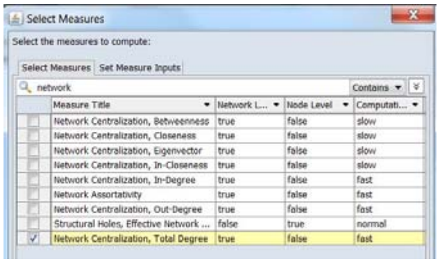
☑ Compute regular measures
□ Compute inverted measures
□ Compute k-centrality measures with radius:



Select the measures to compute:

density

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



Select the measures to compute:

network

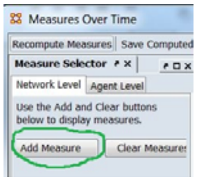
	Measure Title	Network L...	Node Level	Computati...
□	Network Centralization, Betweenness	true	false	slow
□	Network Centralization, Closeness	true	false	slow
□	Network Centralization, Eigenvector	true	false	slow
□	Network Centralization, In-Closeness	true	false	slow
□	Network Centralization, In-Degree	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

CASOS

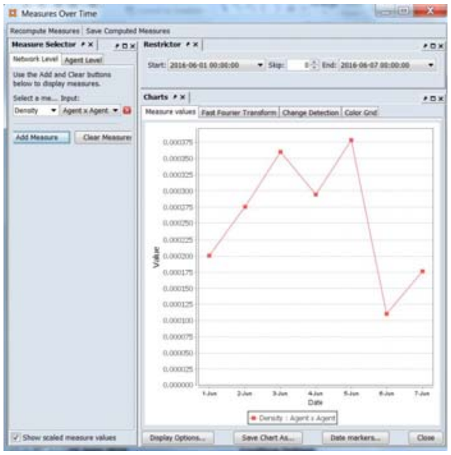
Carnegie Mellon

Perform Dynamic Network Analysis

6. Add Measure, then view various results



Use the Add and Clear buttons below to display measures.



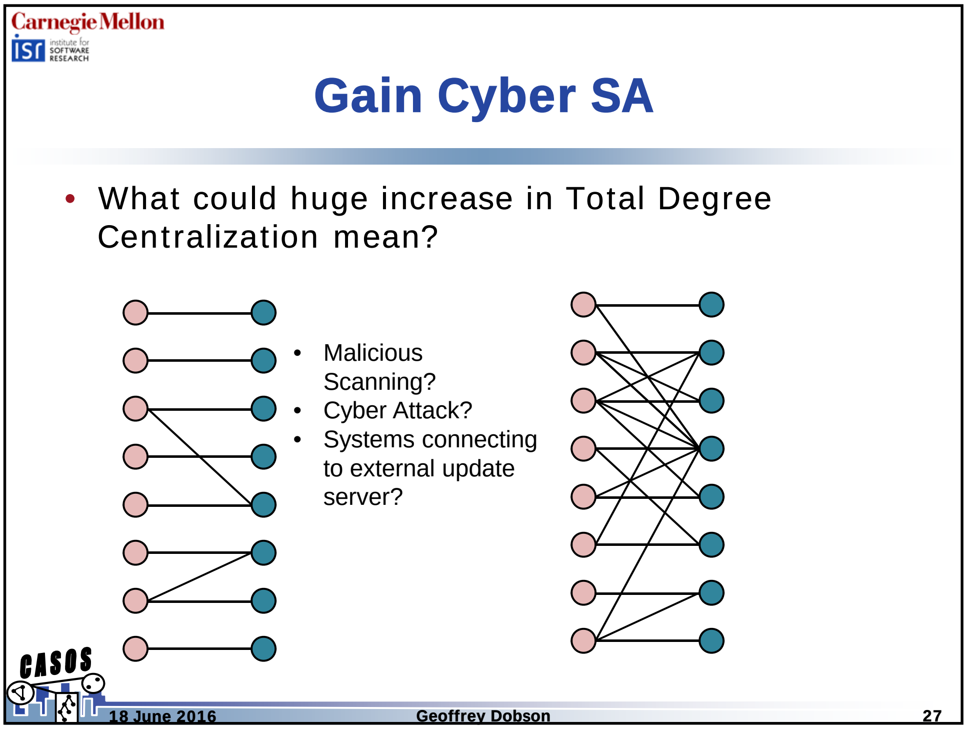
Use the Add and Clear buttons below to display measures.

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

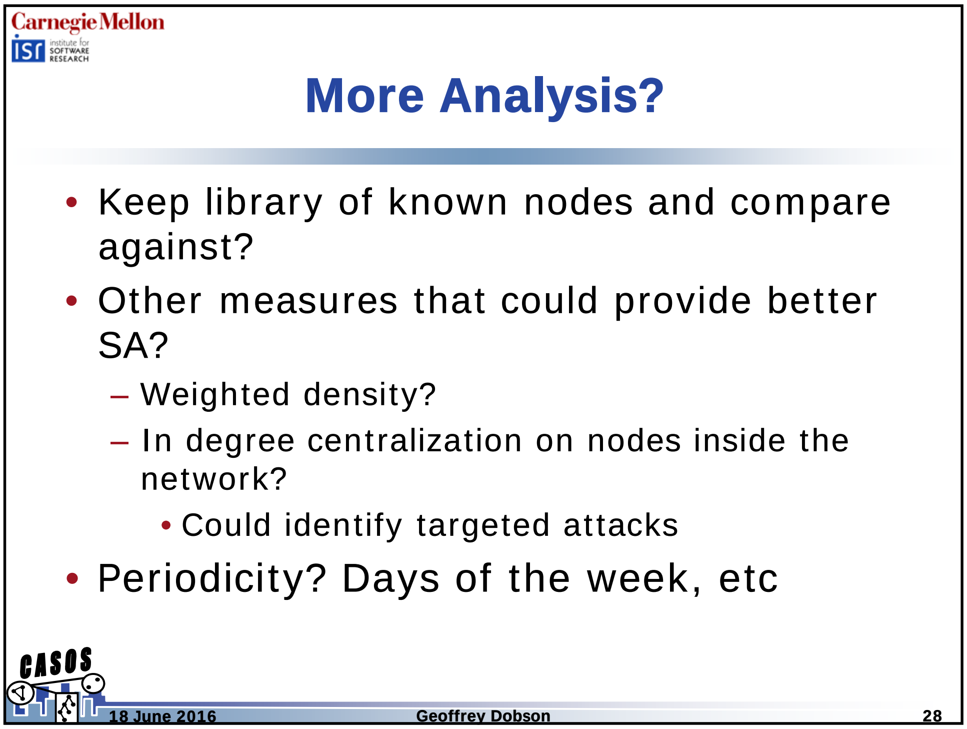
CASOS



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