



## Locate Groups Report

Jeff Reminga

The CASOS Center  
COS Program, School of Computer Science, Carnegie Mellon  
Summer Institute 2020



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



## Locate Groups Report

- Load the Software Company dataset
- Select the software company meta-network in the manager
- Click the Generate Reports button
- Choose Locate Groups



June 2020

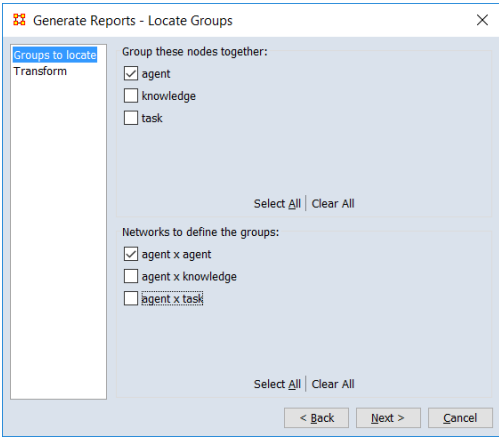
© 2020 CASOS, Director Kathleen M. Carley

2



Carnegie Mellon  
IST Institute for SOFTWARE RESEARCH

## Locate Groups Report...

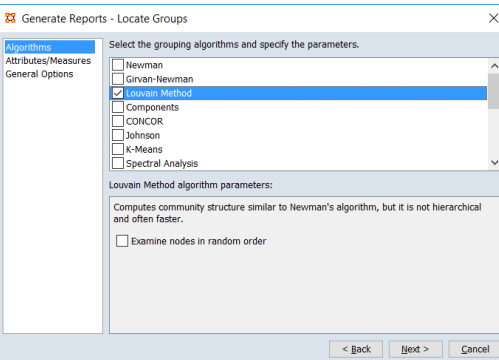


- Open the Generate Reports dialog
- Select the Locate Groups report
- Use "Group these nodes together:" to choose the composition of the groups (i.e. which nodes can be grouped together)
- Use "Networks to define the groups:" to select the network data used by the algorithm to create groups
- Here we choose to group Agent nodes using only network Agent x Agent

CASOS June 2020 © 2020 CASOS, Director Kathleen M. Carley 3

Carnegie Mellon  
IST Institute for SOFTWARE RESEARCH

## Locate Groups Report...



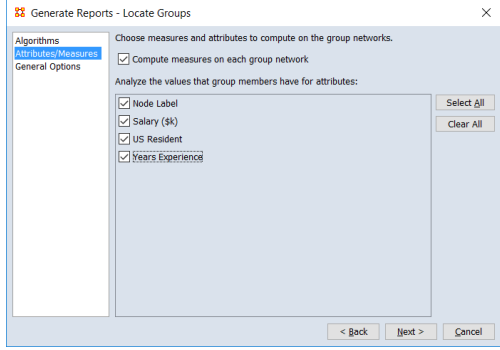
- The algorithms to choose from are determined by the choices from the previous panel:
  - nodes to group
  - networks to use
- Click on any algorithm to see its input parameters
- We will run the Louvain Method algorithm which finds groups of nodes that are densely connected

CASOS June 2020 © 2020 CASOS, Director Kathleen M. Carley 4



Carnegie Mellon  
IST Institute for SOFTWARE RESEARCH

## Locate Groups Report...

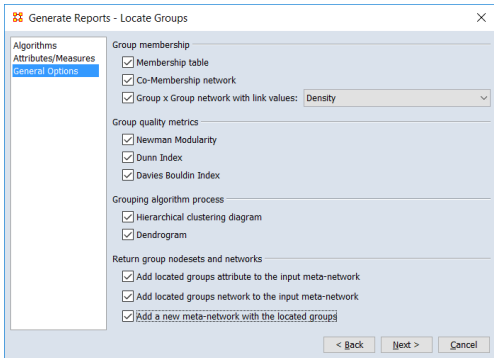


- Select Attributes/Measures
- Click to compute measures and report the statistics for the values of the group nodes
- Click the attributes to analyze
- The attribute values for the nodes of each group will be reported

CASOS June 2020 © 2020 CASOS, Director Kathleen M. Carley 5

Carnegie Mellon  
IST Institute for SOFTWARE RESEARCH

## Locate Groups Report...



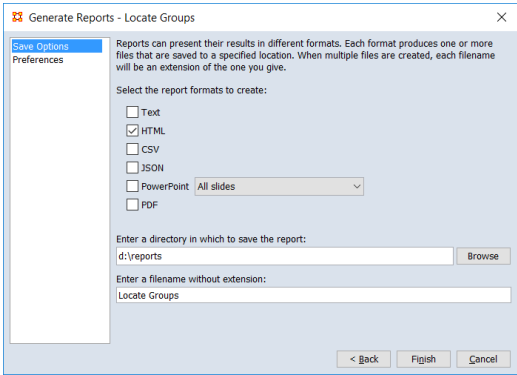
- Select to compute and output all General Options
- In particular: the last section will return the computed groups in three different ways:
  - As an attribute in the input meta-network nodes
  - As a collection of nodesets and networks in the input meta-network
  - As a separate meta-network

CASOS June 2020 © 2020 CASOS, Director Kathleen M. Carley 6



Carnegie Mellon Institute for SOFTWARE RESEARCH

# Locate Groups Report...



- Choose HTML (or any other combination of report formats)
- Choose a filename and directory

**CASOS** June 2020 © 2020 CASOS, Director Kathleen M. Carley 7

Carnegie Mellon Institute for SOFTWARE RESEARCH

# Locate Groups Report...

## LOCATE GROUPS REPORT

Input data: software\_company  
Start time: Sat Jun 8 23:16:05 2019  
[Data Description](#)

### Parameters

Groups contain nodes from nodesets:	agent
Networks used to calculate groups:	agent x agent
Transformations applied:	None

### Clustering Overview

Global clustering coefficient (transitivity)	0.482
--	-------

- The HTML report shows at the top the parameters used to run it
- Then it gives the overall transitivity of the input network which indicates how clustered is the input network

**CASOS** June 2020 © 2020 CASOS, Director Kathleen M. Carley 8



**Carnegie Mellon**  
IST Institute for SOFTWARE RESEARCH

## Locate Groups Report...

**Louvain Method Clustering Algorithm**

The Louvain algorithm requires a symmetric network, and therefore the input network was symmetrized using the union method.

**Group Quality Metrics**

Newman Modularity	0.282	Higher values indicate a better clustering (i.e. more community structure).
Dunn Index	0.416	Higher values indicate a better clustering. It is defined as the ratio between the minimal inter-cluster distance to maximal intra-cluster distance.
Davies-Bouldin Index	2.243	Lower values indicate a better clustering. Each cluster is evaluated by the ratio of its internal spread to its distance to other clusters. This ratio is smallest for low spread and great distance.

The next report assesses the quality of the groups found by the algorithm.

**CASOS**  
June 2020 © 2020 CASOS, Director Kathleen M. Carley 9

**Carnegie Mellon**  
IST Institute for SOFTWARE RESEARCH

## Locate Groups Report...

**Group Membership**

Groups are composed of nodes from these nodesets: agent

Show 10 entries Search:

Group	Size	Members
1	7	Project Manager, Art Director, Technical Lead, Design Lead, Interactive Lead, Web Developer, Usability Engineer
2	5	Application Architect, Software Engineer 1, Software Engineer 2, Software Engineer 3, Software Engineer 4
3	4	Data Architect, Designer, Business Analyst 1, Business Analyst 2

The Group Membership section shows the groups found and lists the members (when there are fewer than 12 members per group).

The members are also shown via an attribute in the input meta-network.

**CASOS**  
June 2020 © 2020 CASOS, Director Kathleen M. Carley 10



**Carnegie Mellon**  
**IST** Institute for SOFTWARE RESEARCH

## Locate Groups Report...

**Betweenness Centrality**

These tables show how the measure values are distributed across each group.

Each table shows the min, mean, max, and std.dev values for members of each group. These group values are also compared with the values of the nodes as a whole (i.e. the population).

If the group mean measure value is greater than one standard deviation(s) above the population mean, then the group's row is colored red. The row is green if the group mean is within one standard deviation of the population mean. Finally, the row is colored blue if the group mean is less than one standard deviation(s) below the population mean.

Show 10 entries Search:

Group	Size	Min	Mean	Max	Std.dev
1	7	0	0.101	0.213	0.066
2	5	0	0.079	0.199	0.090
3	4	0.020	0.029	0.047	0.011

Showing 1 to 3 of 3 entries Previous 1 Next

	Size	Min	Mean	Max	Std.dev
All nodes	16	0	0.076	0.213	0.073

The measures tables compute the measures once on the input networks.

The measure values for the group nodes are reported as statistics.

**CASOS**  
 June 2020 © 2020 CASOS, Director Kathleen M. Carley 11

**Carnegie Mellon**  
**IST** Institute for SOFTWARE RESEARCH

## Locate Groups Report...

**Attribute US Resident (categorical) - Direct or Indirect**

This table shows the percentage of group members that directly or indirectly have each attribute value. A node indirectly has a value if it has a neighbor (in its group) with the value.

This is a text or numeric categorical attribute, such as 'Gender' or 'Years Experience'.

If the group mean measure value is greater than one standard deviation(s) above the population mean, then the group's row is colored red. The row is green if the group mean is within one standard deviation of the population mean. Finally, the row is colored blue if the group mean is less than one standard deviation(s) below the population mean.

Show 10 entries Search:

Group	Size	No	Yes	Herfindahl-Hirschman Index
1	7	0%	100%	1
2	5	100%	40%	0.388
3	4	75%	75%	0.250

Showing 1 to 3 of 3 entries Previous 1 Next

	Size	No	Yes	Herfindahl-Hirschman Index
All nodes	16	75%	75%	0.250

The attribute values for the group nodes are displayed in tables to show which groups have a high or low range of values.

**CASOS**  
 June 2020 © 2020 CASOS, Director Kathleen M. Carley 12

