The image features three blue spheres of varying sizes. One large sphere is at the top right, a medium-sized one is in the center right, and another large sphere is at the bottom right. Two thin blue lines originate from the top left and extend diagonally across the page, one passing near the top sphere and the other near the bottom sphere.

Cytoscape for the Social Sciences

**Marci Brandenburg and
Justin Joque**

3/12/2015

Contents

Cytoscape Background	2
Installing Cytoscape	3
Workshop Data	3
Importing Data	3
Navigating Your Network	7
Menu Bar	7
Tool Bar	7
Control Panel	7
Network View Window	7
Overview Panel	7
Table Panel	7
Attribute Browser Panel	7
Docking and Undocking Panels	7
Layout	8
Importing Node Attribute Data	9
Viewing Attribute Data	10
Using Color to Represent Continents	11
Using Edge Attributes to Represent Migration Numbers	12
Export the Network Graph	13
Create a Legend	14
Filtering Data	15
Calculating Metrics	16
Graphing Large Networks	16

Cytoscape Background

Cytoscape

- Cytoscape = open source network visualization and analysis software
- Software “core” provides basic functionality
- Plugins/apps
 - Add extra functionality beyond the core
- Network (graph) with countries represented as nodes, and number of immigrants represented as links (edges) between nodes
- <http://www.cytoscape.org/>

Cytoscape

- Initially made public in 2002
- Latest version = 3.2
 - Today’s training on 3.2
- Cytoscape App Store:
<http://apps.cytoscape.org/>

Cytoscape Consortium

- Agilent Technologies
- Institute for Systems Biology
- Memorial Sloan-Kettering Cancer Center
- National Center for Integrative Biomedical Informatics (NCIBI)
- Institut Pasteur
- University of Toronto
- UCSD
- UCSF

Installing Cytoscape

1. Launch Firefox or your preferred browser
2. Go to <http://www.cytoscape.org/download.html>
3. Fill out the information and click 'Proceed to Download'
4. Download the 3.2 version, Windows 64-bit option
5. Run the installation file and choose the Desktop as your installation location


Workshop Data

We will be working with 2010 migration data that has been pre-compiled from the World Bank's "Global Bilateral Migration Database." Please download and unzip the data onto the Desktop of the machine you're working on.

1. These instructions and the link to the zip file of workshop data are available on GitHub - at <http://github.com/clarkdatalabs/>
2. Select *Cytoscape*
3. In the lower right corner click the *Download Zip* Button
4. Unzip all the files and move them to the desktop

Importing Data

To create a network, the data file to import must have at least 3 columns: source interaction, target interaction, and interaction type. A fourth column, containing an edge attribute, is optional.



	Source	Target	Edge Attribute	Interaction Type
1	source	dest	value	interaction
2	Afghanistan	Australia	21968	migration
3	Afghanistan	Belgium	1059	migration
4	Afghanistan	Canada	42100	migration
5	Afghanistan	Czech Republic	211	migration
6	Afghanistan	Denmark	12293	migration
7	Afghanistan	Finland	580	migration
8	Afghanistan	France	2958	migration
9	Afghanistan	Germany	79444	migration
10	Afghanistan	Greece	567	migration
11	Afghanistan	Hungary	606	migration
12	Afghanistan	India	8414	migration
13	Afghanistan	Iran Islamic Rep	1704199	migration
14	Afghanistan	Ireland	238	migration
15	Afghanistan	Italy	1382	migration
16	Afghanistan	Japan	710	migration
17	Afghanistan	Netherlands	30986	migration
18	Afghanistan	New Zealand	2346	migration

1. When opening Cytoscape, the Welcome to Cytoscape window appears.
2. Under Start New Session, click New/Empty Network
3. The Create New Network window now appears.
4. Leave default selections for all options and click OK.
5. Go to File -> Import -> Network -> File...
6. The Network file to load dialog box should appear.
7. Select the appropriate file that contains your data (for today: CytoSocSci_Data_2010.csv).
8. The Import Network From Table dialog box should appear.
9. Notice that the data from your file shows up in the bottom of the window.
10. Under the Advanced heading, check the box next to Show Text File Import Options.
11. Under Column Names, check the box next to Transfer first line as column names. This tells Cytoscape that the first row in your spreadsheet contains column headings.
12. Make sure the data at the bottom of the window is formatted correctly (if it is not, change the delimiter).

13. Under the Interaction Definition section, use the dropdown menus to select The Source Interaction, Interaction Type, and Target Interaction
 - a. Select Column 1 for Source Interaction, Column 4 for Interaction Type, and Column 2 for Target Interaction
 - b. Click on the Column 3 heading to designate it as an Edge Attribute (it should turn blue)
14. Click OK.

Import Network From Table

Select a Network Collection


Network Collection:

Mapping Column for New Network:

Mapping Column for Existing Network:

Interaction Definition

Source Interaction: Interaction Type: Target Interaction:

 Columns in BLUE will be loaded as EDGE ATTRIBUTES.

Advanced

☒ Show Text File Import Options

Text File Import Options


Delimiter: ☒ Tab ☐ Comma ☐ Semicolon ☒ Space ☐ Other

Preview Options: ☐ Show all entries in the file ☒ Show first entries.

Column Names: ☒ Transfer first line as column names Start Import Row: Comment Line:

Network Import Options: Default Interaction:

Preview

 **Excel™ Workbook** Left Click: Enable/Disable Column, Right Click: Edit Column

CytoSocSci_Data_2010

✓ source	✓ dest	✓ value	✓ interaction
Afghanistan	Australia	21968	migration
Afghanistan	Belgium	1059	migration

After the data has been imported, a network graph appears in the Cytoscape main view.

Moldova	Ghana	Ethiopia	Albania	Uruguay	Lao PDR	Tunisia	Saudi Arab
erroe Islands	Bangladesh	Rwanda	Hong Kong China	Chad	Botswana	Paraguay	Ukraine
Zimbabwe	St Kitts and Nevis	Congo Rep	China	Puerto Rico	Grenada	Suo Tom and Principe	Guyana
etherland	Dominican Republic	Liechtenstein	Marshall Islands	Nicaragua	Gambia	apua New Guinea	Liberia
Morocco	Uzbekistan	Azerbaijan	Egypt Arab Rep	Monaco	Palau	Vanuatu	Greenlan
Angola	French Polynesia	Al African Repu	Panama	San Marino	Bolivia	El Salvador	Qatar
imor-Leste	St Lucia	Latvia	Costa Rica	Spain	Cayman Islands	Georgia	Cameroo
Belize	Brazil	Maldives	Croatia	Sierra Leone	United Arab Emirates	Turkey	Bahrain

Navigating Your Network

Menu Bar

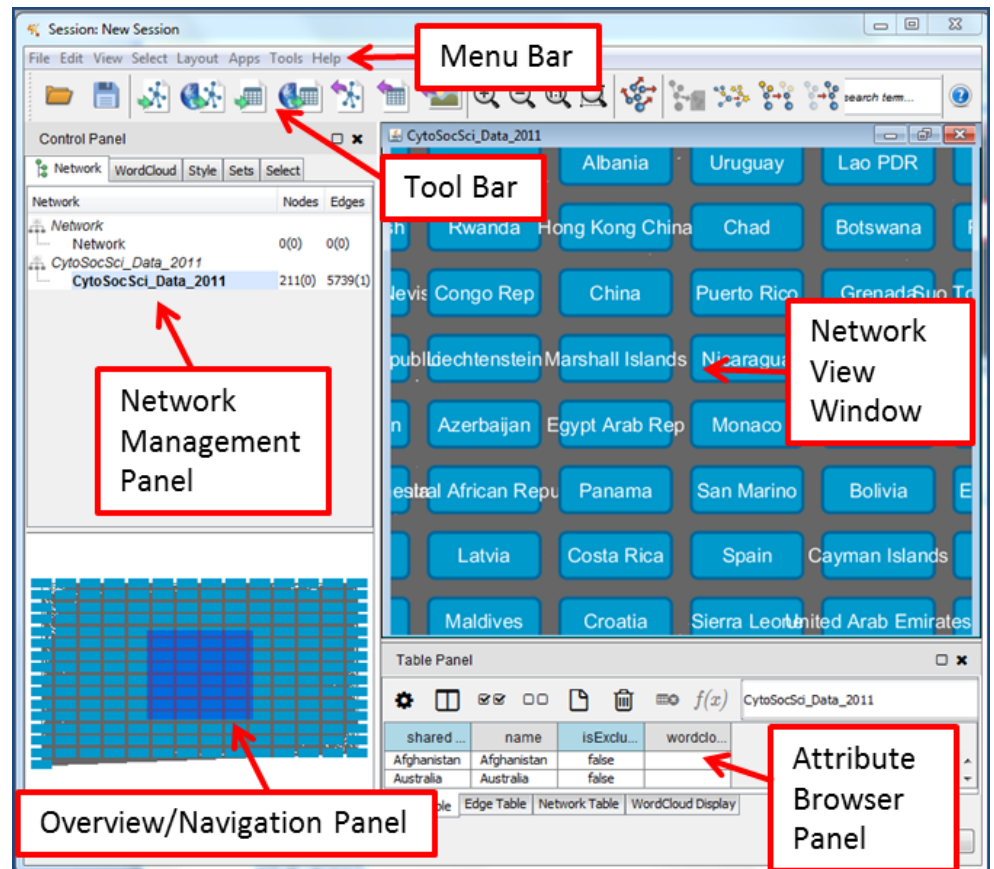
The Menu Bar contains all of the functions that you may need to use for your network in Cytoscape.

Tool Bar

The Tool Bar contains shortcut keys for the most popular Menu Bar functions.

Control Panel

The Control Panel contains tabs that allow you to specify functions for your network. The Network tab gives you access to the Network Management Panel where you select specific networks with which to work if you have multiple networks uploaded in a single session.



Network View Window

The Network View Window displays the networks you have uploaded or created.

Overview Panel

The Overview Panel allows you to navigate through portions of your network by dragging and moving the shaded square.


Table Panel

The Table Panel allows you control the information displayed in the Attribute Browser Panel.

Attribute Browser Panel

The Attribute Browser Panel allows you to look at attribute information for your nodes, edges and overall network.

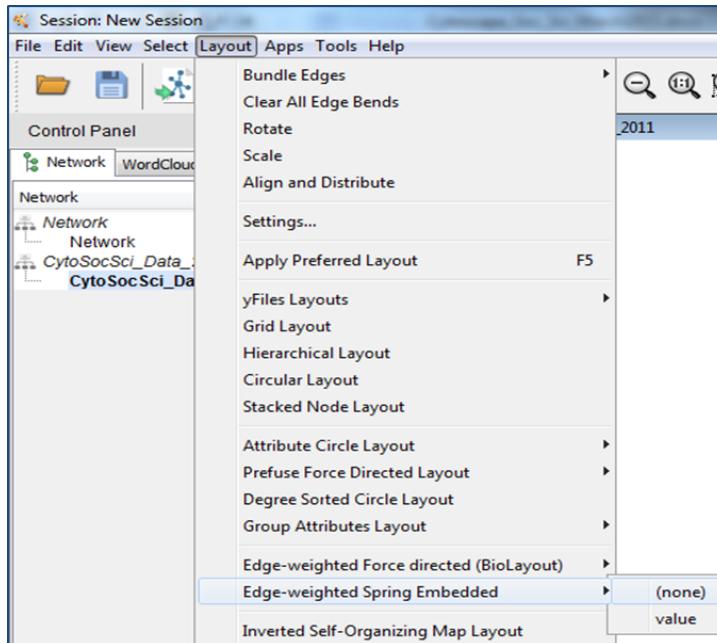
Docking and Undocking Panels

Use the Dock/Undock icon  in the Table and Control Panels to have these windows appear separately from or return to the Cytoscape window.

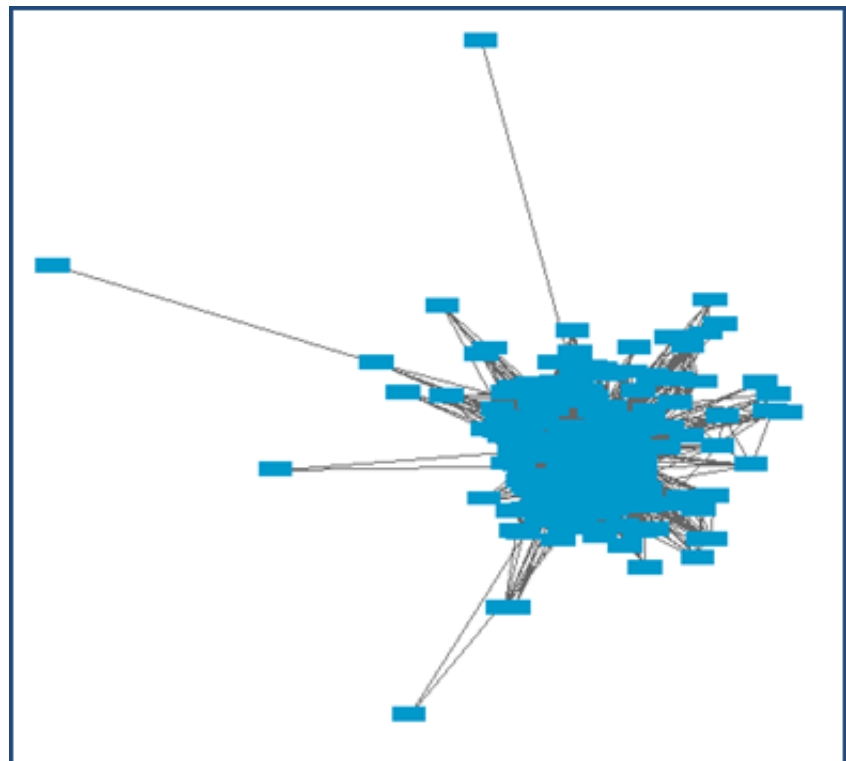
Layout

A variety of layouts are available through the Cytoscape software. The layout options are found in the Layout menu. An appropriate layout can be determined based on your data and needs.

1. Select Layout from the menu bar.
2. Select Edge-weighted Spring Embedded -> (none)



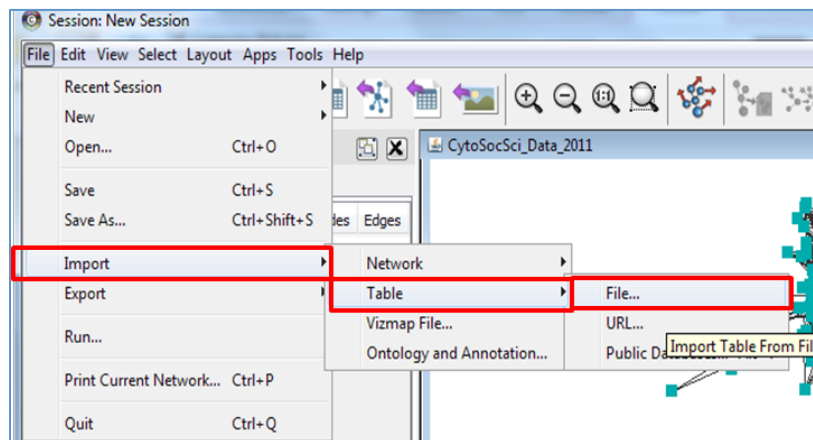
3. Based on the edge-weighted spring-embedded layout algorithm, the network graph is redrawn, as seen in the image to the right.



Importing Node Attribute Data

	Country	Continent
	A	B
1	countries_2010	Continent 2010
2	Afghanistan	Asia
3	Albania	Europe
4	Algeria	Africa
5	American Samoa	North America
6	Andorra	Europe
7	Angola	Africa
8	Antigua and Barbuda	North America
9	Argentina	South America
10	Armenia	Asia

1. Go to File -> Import -> Table -> File...



2. The Data Table file dialog box should appear on your screen.
3. Select the appropriate file that contains your data (for today: countries_continents.xlsx).
4. The Import Columns from Table dialog box should appear on your screen with your data showing at the bottom.
5. Next to Import Data as: use the dropdown menu to select Node Table Columns.
6. Under the "Advanced" heading, select "Show Text File Import Options." Several more options should now appear below this heading.
7. Under "Column Names," select "Transfer first line as column names."
8. Click OK.

Import Columns From Table

Target Table Data

Where to Import Table Data: To a Network Collection

Select a Network Collection

Network Collection: CytoSocSci_Data_2011

Key Column for Network: shared name

Importing Type

Import Data as: Node Table Columns

Advanced

☐ Show Mapping Options ☒ Show Text File Import Options ☒ Case Sensitive

Text File Import Options

Delimiter: ☒ Tab ☐ Comma ☐ Semicolon ☐ Space ☐ Other

Preview Options: ☐ Show all entries in the file ☒ Show first 100 entries.

Column Names: ☒ Transfer first line as column names Start Import Row: 1 Comment Line:

Refresh Preview

Preview

Excel™ Workbook

Left Click: Enable/Disable Column, Right Click: Edit Column

Sheet1

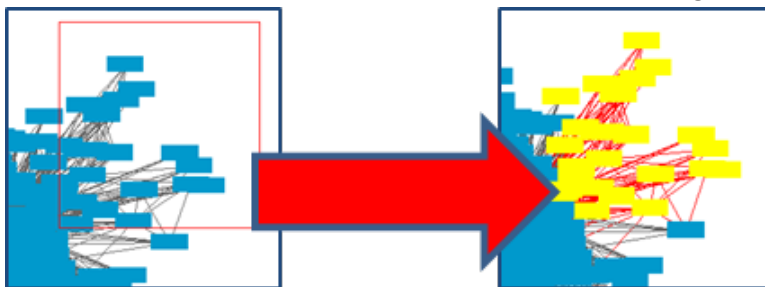
✓ countries_2010	✓ Continent_2010
Afghanistan	Asia
Albania	Europe
Algeria	Africa
American Samoa	North America
Andorra	Europe


OK Cancel

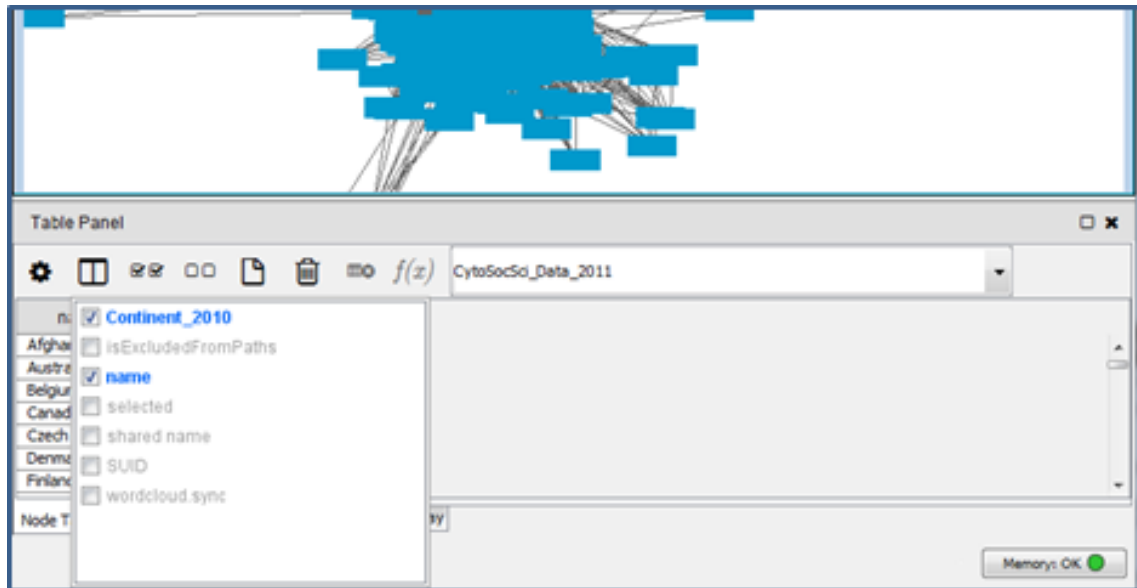
Viewing Attribute Data

In order to view attribute data about your network, select the nodes and edges of interest

1. Click on individual nodes or edges to highlight them OR
2. Left mouse click and draw a box around the nodes and edges of interest.



- Highlighted nodes and edges are the ones whose data will be displayed in the Attribute Browser Panel.
- To choose only specific attributes to view, click on the Show Column icon  in the Table Panel and then check the boxes of the specific attributes that you would like to view. For this example, make sure the box next to Continent 2010 is checked.



- Click outside the box and the Data Panel will refresh with the attribute data displayed.

Additional icons:

 : Use this icon to hide all attribute columns

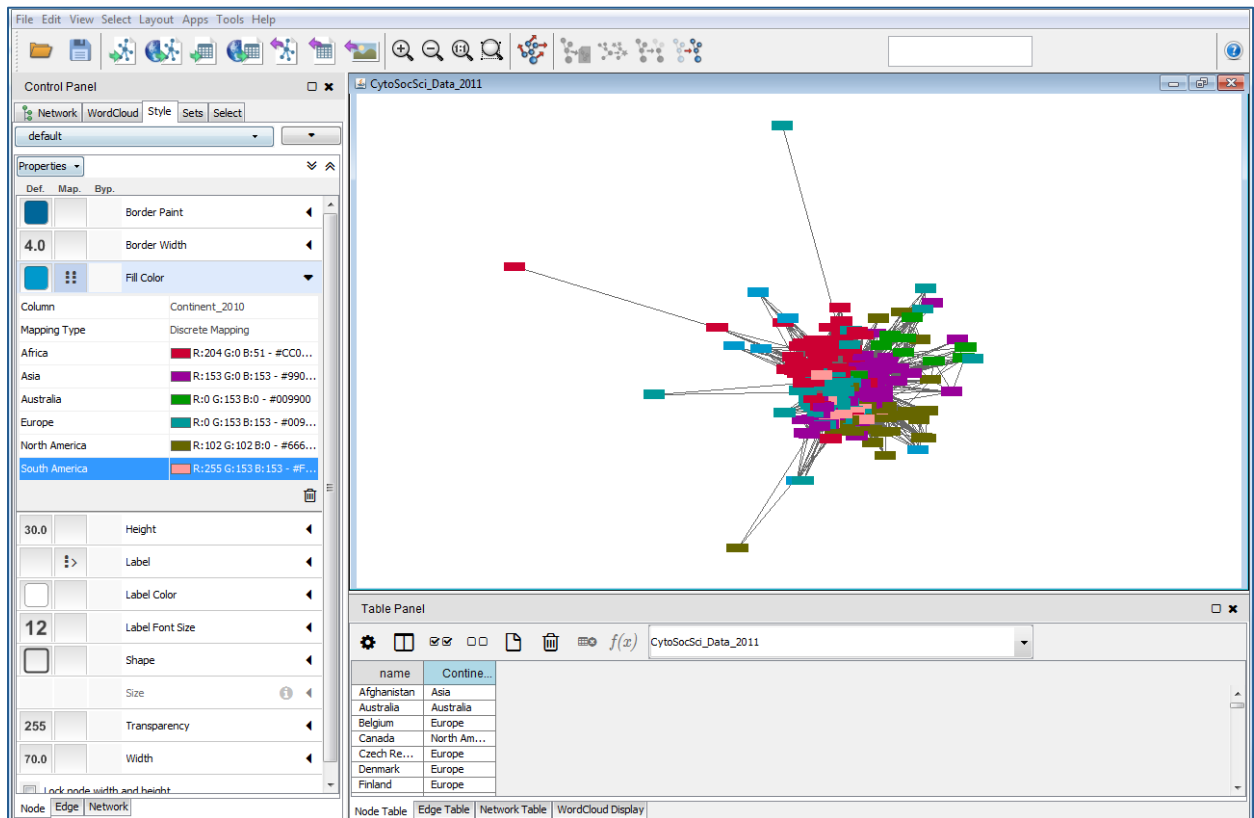
 : Use this icon to show all attribute columns

Using Color to Represent Continents

Six continents are represented in our data set: North America, South America, Africa, Asia, Europe, and Australia. Visual styles can be used to easily see each country's continent.

- In the Control Panel, select the Style Tab.
- Make sure the Node tab is selected at the bottom of the Style Tab window.
- Click on the arrow to the right of Fill Color.
- To the right of Mapping Type, double click on --select value-- and select "Discrete Mapping" in the dropdown list.
- To the right of Column, double click on --select value-- and select "Continent_2010" in the dropdown list.
- Double click on the box next to Africa and click on the ellipsis button.
- Select a red color and click OK. The countries located on Africa will now be colored red.

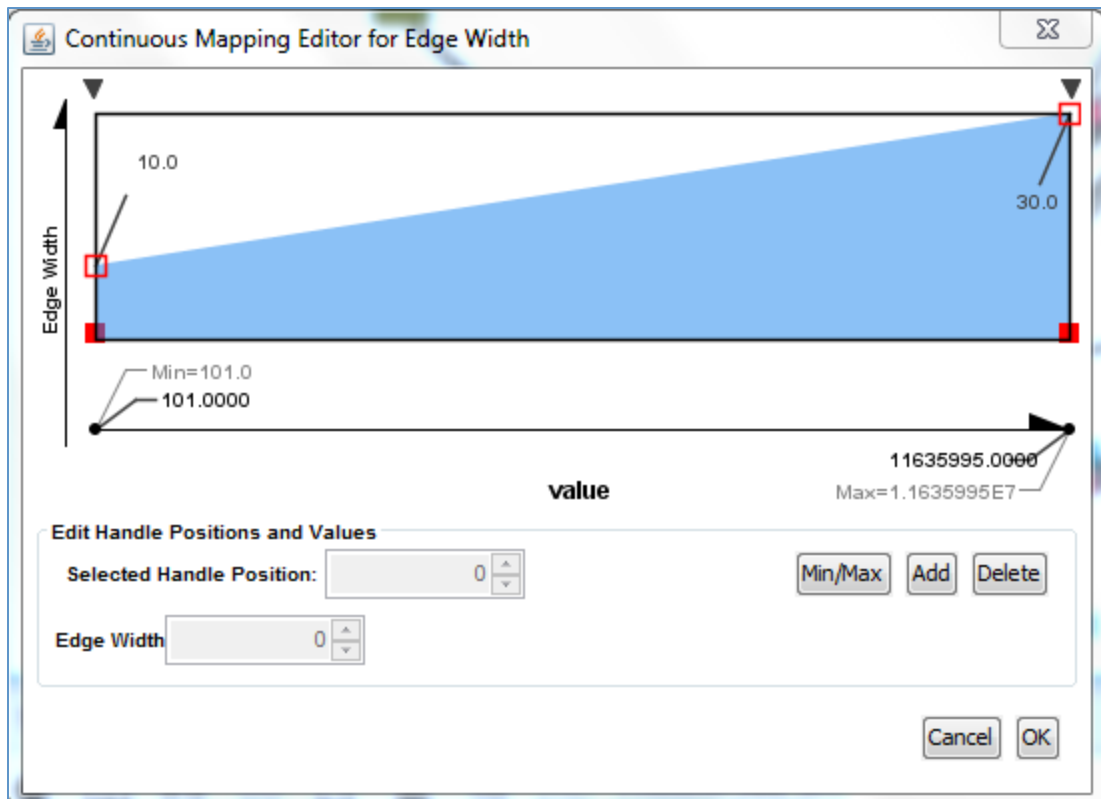
8. Double click on the box next to the value Asia and click on the ellipsis button.
9. Select a purple color and click OK. The countries located on Asia will now be colored purple.
10. Repeat the above steps for each country, assigning a different color to each one.



Using Edge Attributes to Represent Migration Numbers

Visual styles can also be used to control the edge appearances based on the data.

1. Make sure you are in the Style Tab.
2. Click on the Edge tab within the Style tab window.
3. Click on the arrow to the right of Width.
4. To the right of Column, double click on --select value-- and select "value" in the dropdown list.
5. To the right of Mapping Type, double click on --select value-- and select "Continuous Mapping" in the dropdown list.
6. Double click on the image to the right of Current Mapping.



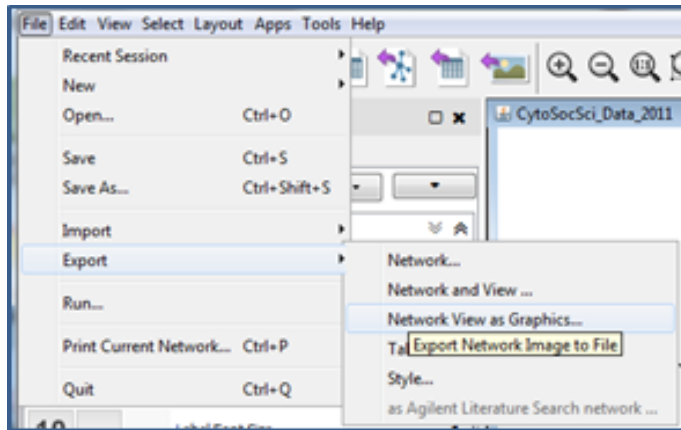
By dragging the arrows on the left and right it is possible to change the nature of the continuous mapping. Break points can also be added with the Add button and moved with the arrows on the top. These allow you to change the rate of change (i.e. so you can approximate a logarithmic relationship between value and line width).

7. Under the View menu choose: show graphic detail (this will allow you to see the different line weights even when zoomed out – though be aware that for large networks it takes significantly more memory and time to draw).

Note: there is also an option to set the edge opacity in the Style's Edge tab. To do this globally, click the leftmost box (Def. column) next to Transparency. An Edge Transparency box appears in which you can change the number. This can make it easier to see the variable widths of the edges.


Export the Network Graph

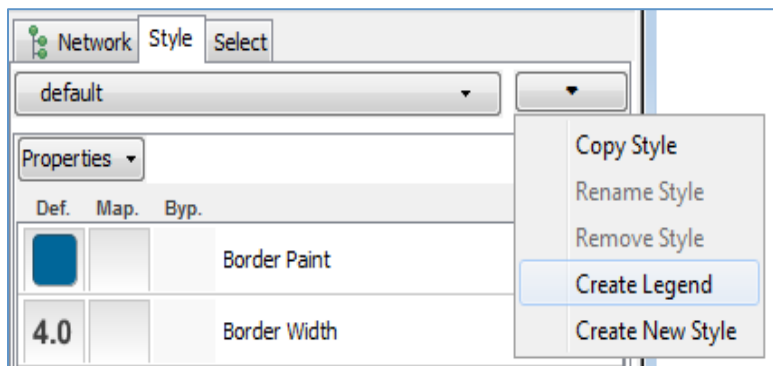
1. Click on the File Menu, select Export, and then Network View as Graphics...



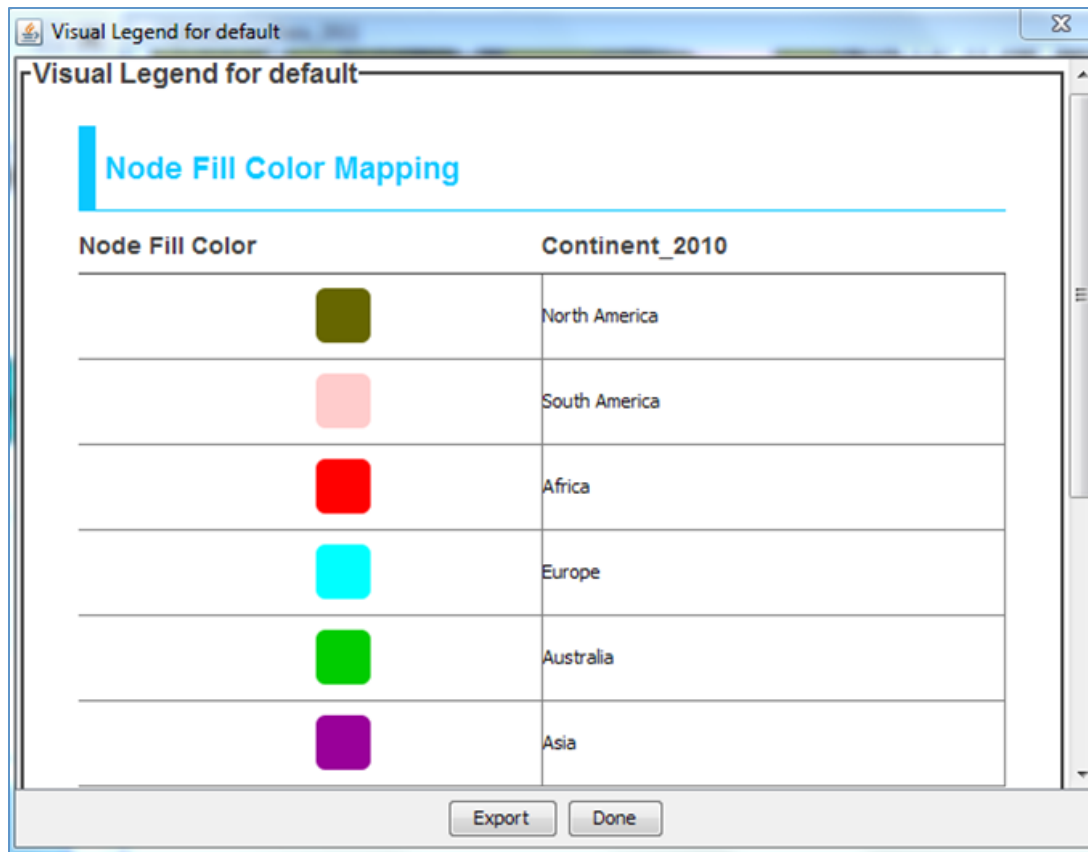
2. The Export Network View as Graphics window should now appear.
3. Use the Save Image As: button to determine where the file will be saved.
4. Use the dropdown menu next to Select the export file format to determine the format, such as pdf, in which the network will be saved.
5. Click OK.
6. The portion of the network graph in the main view alone (not the entire Cytoscape window or network) will be saved and can then be used in publications or presentations.

Create a Legend


1. In the Control Panel, select the Style Tab.
2. At the top right of the tab, click on the dropdown arrow button .
3. Click on Create Legend.

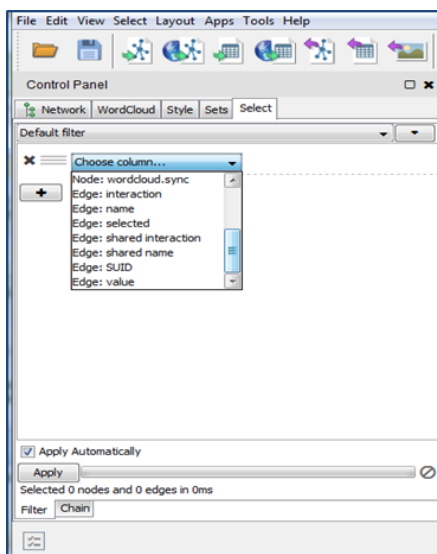


4. The Visual Legend for default window should now appear, providing the legend for your network.
5. Click Export to save the legend.



Filtering Data

1. Click on Select in the Control Panel.
2. Click  to create a new filter. Select Column Filter.
3. Using the Choose column dropdown menu, select Edge: value.



4. In the text box enter 0 and 1000 (so it reads “is between 0 and 1000 inclusive”)
5. If the Apply Automatically option is checked, the main network should show all edges between 0 and 1000 as being selected.
6. Under the Select menu at the top choose Edges -> Invert Edge Selection.
7. Under the Edit menu at the top choose Delete Selected Nodes and Edges (to remove those that fall below your filter).

Calculating Metrics

In-Degree: The number of edges pointing towards a node

Out-Degree: The number of edges leaving a node (in and out degree are the same for each node on an undirected graph)

Shortest Path: The lowest number of edges that must be traversed to get from one node to another

Diameter: The longest shortest path in a network

1. Choose Tools -> NetworkAnalyzer -> Network Analysis -> Analyze Network.
2. If prompted choose to treat the network as either directed or undirected (depending on the network)
3. After metrics have been calculated you can view the results and choose to save them (n.b. even if you do not choose to store them the calculated values will be attached to nodes and edges and can be used for visualization purposes)

Note that the default network analysis package does not account for weighted edges. Exactly how you would deal with weighted edges depends on what they represent and using an external plug-in is most likely the best way to handle this.

Graphing Large Networks

Large networks may require more memory than default to graph without crashing. By launching your java virtual machine with a larger stack size and more memory Cytoscape is capable of graphing networks with hundreds of thousands of nodes and edges. Cytoscape can be started with more memory from the command line with these options:

-Xmx1GB -Xss10M

-Xmx sets the amount of memory and -Xss sets the stack size (10M is usually sufficient for solving most ‘out of memory’ errors for layouts and more will slow down some processes). More information is available: http://wiki.cytoscape.org/Cytoscape_User_Manual/Launching_Cytoscape.