

COMM 645 HANDOUT – NODEXL BASICS

NodeXL: Network Overview, Discovery and Exploration for Excel. Download from nodexl.codeplex.com

Plugin for social media/Facebook import: socialnetimporter.codeplex.com

Plugin for Microsoft Exchange import: exchangespigot.codeplex.com

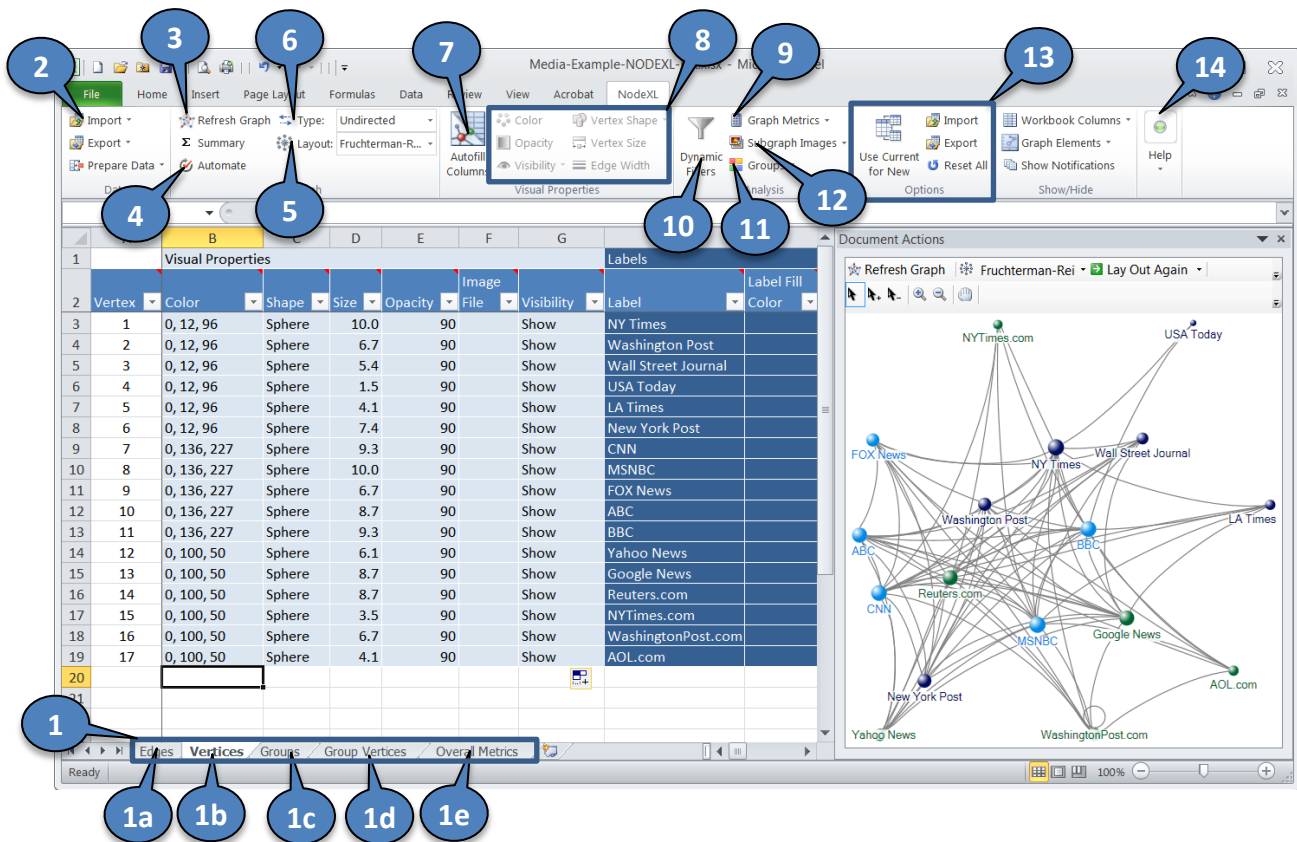
Plugin for Voston hyperlink network import: voston.anu.edu.au/node/13#VOSON-NodeXL

Note that NodeXL requires MS Office 2007 or 2010. If your system does not support those (or you do not have them installed), try using one of the computers in the PhD office.

Major sections within NodeXL:

- **Edges Tab:** Edge list (*Vertex 1 = source, Vertex 2 = destination*) and attributes (Fig.1→1a)
- **Vertices Tab:** Nodes and attribute (nodes can be imported from the edge list) (Fig.1→1b)
- **Groups Tab:** Groups of nodes defined by attribute, clusters, or components (Fig.1→1c)
- **Groups Vertices Tab:** Nodes belonging to each group (Fig.1→1d)
- **Overall Metrics Tab:** Network and node measures & graphs (Fig.1→1e)

Figure 1: The NodeXL Interface



After you install the NodeXL template, a new NodeXL tab will appear in your Excel interface. The following features will be available in it:

Fig.1 → 1: Switch between different data tabs. The most important two tabs are "Edges" and "Vertices".

Fig.1 → 2: Import data into NodeXL. The formats you can use include GraphML, UCINET DL files, and Pajek .net files, among others. You can also import data from social media: Flickr, YouTube, Twitter, Facebook (requires a plugin), or a hyperlink networks (requires a plugin).

Fig.1 → 3: Refresh the graph visualization after changing some of its properties.

Fig.1 → 4: Automate and run a series of tasks (each also available in other parts of the interface).

Fig.1 → 5: Select the graph layout you'd like to use for the network visualization.

Fig.1 → 6: Are the links in your data directed or undirected?

Fig.1 → 7: Use "Autofill Columns" to change the visual characteristics of nodes and edges based on attributes or network measures. For example, you can change the color, shape, size, and opacity of nodes based on an attribute (e.g. color by gender) or a structural property (e.g. size by in-degree centrality). The "Options" button in the "Autofill Columns" menu will allow you to tweak the settings for the specific attribute/property you are interested in visualizing.

Fig.1 → 8: Change the visual characteristics (color, size, shape, line width, opacity, etc.) of selected nodes or edges. To use images for nodes (e.g. the profile pictures of social media users), you need to select shape "Image". Then add URLs / local file locations for the image file associated with each node to a column named "Image File" (URLs will be added automatically when you use NodeXL to collect Twitter/etc. social network data).

Fig.1 → 9: Calculate various node and network-level metrics. Those include node in-degree, out-degree, betweenness, closeness, and eigenvector centrality; network reciprocity, clustering coefficient, modularity, and so on. Once calculated, node metrics will appear as new columns in your "Vertices" tab (Fig.1→1b). Graph metrics will appear in the "Overall Metrics" tab (Fig.1→1e).

Fig.1 → 10: Select specific nodes/edges based on different attributes and structural characteristics (e.g. select nodes with specific levels/ranges of degree, closeness centrality, attributes, etc.) The nodes that are not included in the selection will be hidden from the graph visualization.

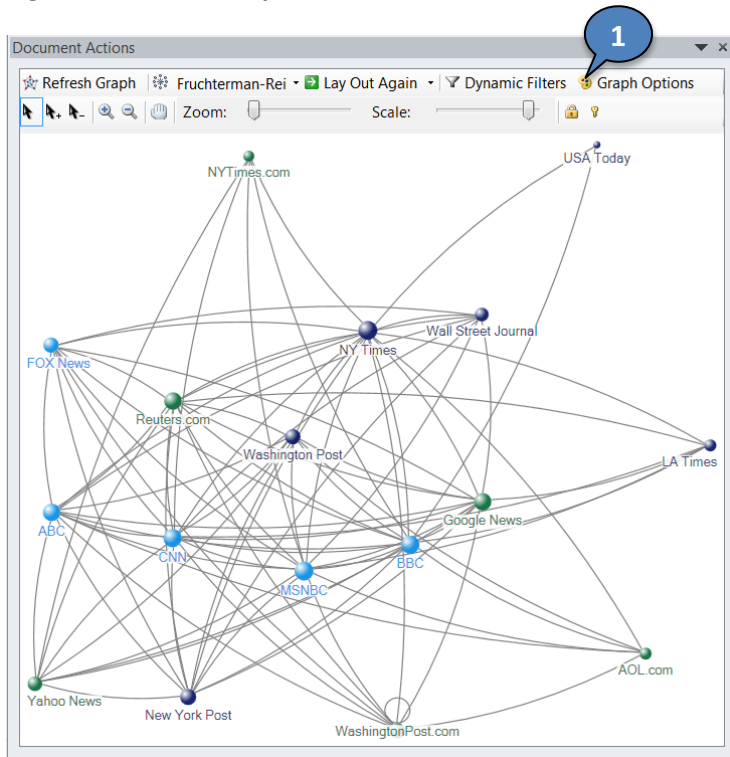
Fig.1 → 11: Create thumbnails/subgraph images for each vertex in the data. You can select how many steps out from your focal node will be included in the subgraphs.

Fig.1 → 12: Group nodes based on their attributes, or the clusters/components they belong to. The groups will appear in the "Groups" and "Group Vertices" tabs (Fig.1→1c, 1d).

Fig.1 → 13: Import/Export/Reset the options you have set across NodeXL workbooks.

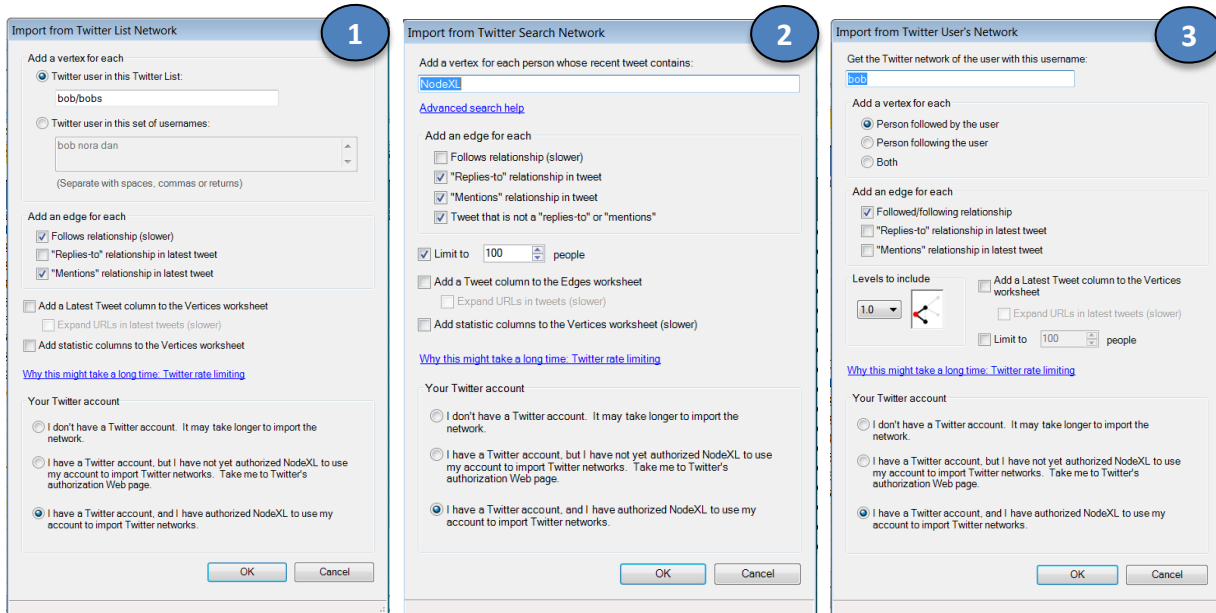
Fig.1 → 14: Get help!

Figure 2: NodeXL Graph Pane



Most of the options in the Graph Pane are similar to the ones available in the NodeXL tab. One additional menu here can be accessed from the Graph Options button (Fig.2 → 1). The menu will allow you to set multiple visual properties affecting all nodes and edges in the graph visualization (e.g. straight or curved edges, shadows for nodes, etc.).

Figure 3: Importing Twitter Data



Importing Twitter Data

To collect Twitter data using NodeXL, you have to select one of the three available options from the import menu (Fig.1 → 2):

- 1) **Twitter network based on a list of users (Fig.3 → 1).**
(Types of relationships include following or mentions/replies in latest tweets)
- 2) **Twitter network based on a search term/hashtag (Fig.3 → 2).**
(Types of relationships include following or mentions/replies in tweets. You can also opt to include in the analysis tweets that do not contain mentions/replies).
- 3) **Twitter networks centered around a certain user (Fig.3 → 3).**
(Vertices include people followed by and/or following the user. Types of relationships include following and mentions/replies).
"Levels to include" tells NodeXL how many steps out from the focal user it needs to go. One step captures all following/followers. Including 1.5 levels also captures the relationships among those following/followers. The maximum steps out you can go is 2.

Note that Twitter limits the number of requests you can make to its system per hour. The limits are stricter if NodeXL tries to access Twitter without using your account. If you link your account and authorize NodeXL to use it, the limit is a bit higher and you can collect more data per hour before reaching it.