NodeXL for Social Network Analysis

CAR meeting, Raleigh, Feb 25 2011

Peter Aldhous, San Francisco Bureau Chief



peter@peteraldhous.com



What is NodeXL?

A free template for Excel 2007 and 2010 that makes Social Network Analysis easy for anyone familiar with basic spreadsheet functions.

What's really cool: NodeXL has been optimized for analyzing online social media – it includes built-in connections to query the APIs of Twitter, Flickr and YouTube, allowing you to draw networks of users and their activity.



NodeXL is Brought to You By...

Thanks especially to:

- Marc Smith, Connected Action, for Twitter network presented in these slides
- Ben Shneiderman, University of Maryland

Getting started

Download the template from <u>NodeXL site</u>, then open:

Home Insert Page Layout Formulas Data Review View Add-Ins Acrobat NodeXL Design	-lublu A												
	deba a la A												
Import * Import * <td< td=""><td>rkdook (f</td></td<>	rkdook (f												
A3 • (<i>f</i> _*													
A B C D E F G H I J M N O P Q	R												
1 Visual Properties Label Other Columns													
2 Vertex 1 Vertex 2 Color Vidth Style Opacity Visibility Label Color Size Columns Here													
3													
4 Vertex 1 Name													
6 the edge's first													
vertex.	_												
9													
10													
11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1													
12 12 12 14 14 14 14 14 14 14 14 14 14 14 14 14													
17 I I I I I I I I I I I I I I I I I I I	_												
18													
H + N Edges / Vertices / Groups / Group Vertices / Overall Matrics / *]													
Ready	I 100%												



Type in a simple edge list, and calculate graph metrics

ſ	0	(° - ∎-) =			Friends - Microsoft	Evcel	- X		Т	able Tools							
	Home	Insert	Page Layou	t Forr	nulas	Graph Metrics			XL	L	Design							
	Import -	1	Type: U	Indirected		Select the metrics to o workbook	Select the metrics to calculate and insert into the workbook			Graph Metrics						nns 👻	😰 Sa	imple W
		Show	* Layout: F	ruchterma	n * A	A About duplicate edges	hami	ic 📙	Grapi	n Metric	CS		Elements			ome Pag		
E	Prepare Data	• Graph (Automate		Co				Iters	Ľ	Aggr	egate C	Graph I	Metrics				egister
	Data		Graph	1		Dogree							Calcul	late selec	ted graph r	netrics a	ind	
	A2	- (fs	· Verte	x 1	V Degree		Detalls		_			insert	the resu	Its into the	workbo	ok.	
	А	В	С	D	E	✓ In-degree		<u>Details</u>	-		ſ	N	🎲 Frie	ends				Q
1			Visual Pro	perties		V Out-degree		Details			Other Co	lumn	Pre	ss F1 for	more help.			
									ont		Add You	r Own						
2	Vertex 1 💌	Vertex 2 🔽	Color 💌 V	Vidth 💌	Style 💌	Centrality					Columns	Here						
3	Ann	Bob				Betweenness a	nd closeness	<u>Details</u>										
4	Bob	Carol				centralities												
5	Carol	Dave				Eigenvector cer	ntrality	Details		_								
6	Ann	Carol					-		_	_								
7	Carol	Ed				Miscellaneous			-	_								
8	Ed	Frank				PageRank		Details	-	_								
9	Gary	Helen				Chustering and	Galant	Detaile					-					
10						Clustering coel	licient	Details		-								
12						Group metrics		<u>Details</u>		-								
13						Overall graph n	netrics	<u>Details</u>										
14										-								
15						Select All Des	elect All											
16										+								
17						Calculat	e Metrics	Cancel										
18																		
i A	← → → Eda	es Vertice	Groups	Group	Vertices	Overall Metrics					-0	4						
Re	ady		- oroupo	2 oroup	- Statess A													



Now the vertices worksheet has been populated with data

6	Friends - Microsoft Excel Table Tools																		
ſ	<u> </u>	Home	Insert	Page Lay	out Forr	nulas Data		Review	View	Add-Ins	A	Acrobat	NodeX	L Design					
Import ▼ Import ™ Import ™ <t< td=""><td colspan="8">Image: Color Image: Color <td< td=""><td colspan="4">Metrics • Workbook Colu aph Images • Graph Element s • Show/Hide</td><td>Colu ents ide</td></td<></td></t<>							Image: Color Image: Color <td< td=""><td colspan="4">Metrics • Workbook Colu aph Images • Graph Element s • Show/Hide</td><td>Colu ents ide</td></td<>								Metrics • Workbook Colu aph Images • Graph Element s • Show/Hide				Colu ents ide
	$J10 - f_x$												_						
	A B C D E F G H I J K L																		
1	1 Graph Metrics Visual Properties																		
			_	In-	Out-	Betweennes	5S	Closeness	_	Eigenvecto	r		C	ustering			_	_	
2	2 V	ert 🌮 💌	Degree 💌	Degree 💌	Degree 💌	Centrality	•	Centrality		Centrality		PageRank		pefficient 🔽	Color	▼ S	hape 💌	Size 💌	Op
3	3 A	nn	2			0.0	000	(0.111	0	.202	0.9	55	1.000					
4	4 Bo	ob	2			0.0	000	(0.111	0	.202	0.9	55	1.000					
5	5 Ca	arol	4			8.0	000	(0.167	0	.278	1.8	379	0.167	1				
6	5 Da	ave	1			0.0	000	(0.100	0	.117	0.5	49	0.000					
7	7 Ec	d	2			4.0	000	(0.125	0	.142	1.0	60	0.000					
8	3 Fr	rank	1			0.0	000	(0.083	0	.060	0.6	i00	0.000					
9	G	ary	1			0.0	000	1	1.000	0	.000	1.0	000	0.000		_			
1	0 H	elen	1			0.0	000	1	1.000	0	.000	1.0	000	0.000					
1	1														V	erte	x Color		
1	2														Т	o se	lect an op	tional	
1	3														V	erte: lick	color, rig	ht- Select	
1	4														0	Color	on the rig	ht-	
15 click menu.																			
16																			
17																			
1	8																		
I	^	Edd	jes Verti	ces Grout	os / Group	Vertices / O	vera	Metrics	*					1	4				
R	eady							A											



Add more data: a new column for the gender of each friend

- -			Friends	- Microsoft	Excel			-	Table Tools					
Pa	ge Layout F	ormulas Data	Review	v View	Add	-Ins Acro	obat	NodeXL	Design					
	ype: Undirect ayout: Fruchter utomate Graph	ted • man • Autofil Column	Color Opac S A Visib	r r v city	′ertex Sha ′ertex Size dge Widt Properties	pe 👻 🌆 Bacl e 🛛 🏠 Scho h	kground emes	Dynamic Filters	Graph Metri Subgraph Ir Groups 👻 Analysis	ics • nages •	Workbook Co Graph Elemer Show/Hid	e	😰 Sample V 🚹 Home Pa ≕ Register	/orkbook ge
6	× ✓ Jx F													
	Н	I	J	K	L	M	N	0	Р	Q	R	S	AE	
			Visual Pr	roperties				•	Labels				Other Col	umns
or _	. —	Clustering		_	_	_	Image		_	Label Fill	Label			_
×	PageRank 💌	Coefficient 💌	Color 💌	Shape 💌	Size 💌	Opacity 💌	File 💽	Visibility	💌 Label 💌	Color	 Position 	Tooltip 💌	Sex	×
0.202	0.955	1.000									_		F	
0.202	0.955	1.000									_		М	
0.278	1.879	0.167									_		F	
0.117	0.549	0.000									_		М	
0.142	1.060	0.000									_		М	
0.060	0.600	0.000									_		М	
0.000	1.000	0.000									_		М	
0.000	1.000	0.000											F	
ices /	Groups / Gro	up Vertices 📈 Ov	erall Metric	cs 🦯 🞾 🗸	·									
														10

$\overline{\mathbf{F}}$

Incorporate attributes from the data into the network graph

		🖳 Autofill Columns											
	Frinzle Misserett Fred		How Autofill Columns works										
Home Insert Page Layout	Formulas Data Review View Add- ected	When Autofill is clicked, fill in these worksheet columns	e Options										
Export - Show	erman Autofill Opacity	Edge Color	▼ ()										
Data Graph & Automate Graph	Columns Columns Visibility = Edge Width Visual Properties	Edge Width	▼ 👂										
AB10 ▼ (× √ f _x F	Autofill Columns	Edge Opacity	▼										
A G H	Calculate and fill in the graph's	Edge Visibility	▼										
1 Eigenvector	clusterin values from other columns.	Edge Label	▼ 🔍										
2 Vertex 🔽 Centrality 🔽 PageRank	Coefficie 🎲 Friends	Vertex Color Sex	▼										
3 Ann 0.202 0.95 4 Bob 0.202 0.95	5 Press F1 for more help. 5 1.000	Vertex Shape	▼ 🔍										
5 Carol 0.278 1.87	9 0.167	Vertex Size: Betweenness	Cen 👻 📄										
6 Dave 0.117 0.54	9 000.0	Vertex Opacity	▼ 👂										
8 Frank 0.060 0.60	0 0.000	Vertex Visibility	▼ 👂										
9 Gary 0.000 1.00	0 0.000	Vertex Label Vertex	▼										
10 Helen 0.000 1.00	0.000	L Fill Color	▼										
12		Vertex Label Position	▼										
13		Vertex Tooltip Vertex	▼										
15		Vertex Layout Order	▼ 👂										
16 17		Vertex X	▼ 🔷										
18		Vertex Y	▼ 👂										
H + + H Edges Vertices Groups G	Image: Wertices Groups Group Vertices Overall Metrics Vertex Polar R												
Enter		Clear All Wederback Columns New											

Reset All Autofill Settings

Autofill

Close

x

The resulting friends network:



Citations in the hottest area of stem cell research





12 NewScientist | 12 june 2010

Import data from other formats and online social media

C. ,) 🖬 🤊 -	C- <) ≂		NodeXL	Graph [Read-	Only] - Mic	rosoft Ex	cel		Table Tools		
U	Home	Insert	Page Layout	Formulas	Data	Review	View	Add-Ins	Acrobat	NodeXL	Design		
B	import *		🚍 Type: Un	directed 🔹		ိ္နိ Color	P Verte	x Shape	Import from Twitt	ter Search I	Network	x pook C	olum
	Clear Node From UCIN From Graph From Graph	XL Workbook ET Full Matrix hML file hML files	: First DL File		T fill	Opacity S I Usibility	Verte:	x Size Width erties	Search for peop GOP Advanced searce Add an edge fo	le whose tw <u>ch help</u> or each	veets contain:	i Eleme	nts * de
(10) [01]	From Pajek From Open From Open	Matrix Work	book		_		Lab	el	✓ <u>F</u> ollows re ✓ " <u>R</u> eplies-t	elationship (to" relations " relationsh	(slower) hip in tweets ip in tweets	0	_
=	From Email From Flickr From Flickr	Network Related Tags User's Netwo	Network ork		ac	ity 🔽 Visibi	lity 🔽 Lab	el 🔽 (✓ Limit to 100 ▲ Add a Tweet) v pe	ople he Vertices worksheet		
	From Twitt	er Search Netv	work						Add statistic	columns to	the Vertices worksheet (slowe	er)	
-	From Twitte From YouTu From YouTu	er User's Netw ibe User's Net ibe Video Net	vork twork twork	Import from T Optionally o workbook, f	witter Sea lear the No then get th	r ch Network odeXL ne network of			Why this might r Request lifting o Your Twitter ac	not work: Tv of Twitter rat count	<u>witter rate limiting</u> te limiting		
	From Node	XL Workbook	Created on An	people who specified wo	se tweets (ord.	contain a	_		I don't have import only	e a Twittera ⁄a limited n	account. I will be able to umber of Twitter networks into	,	
12				NodeXLG	Graph				NodeXL.				
13				FIESS F1 I	or more ne	cip.			Thave a Tv authorized	NoteXL to	use my account to import		
14									Twitter net Web page.	works. Tak	e me to Twitter's authorization		
16									I have a Ty	witter accou	int. and I have authorized		
17									NodeXL to	use my acc	count to import Twitter		
18									networks.				
19											OK Cancel		
20													
21													
22													

Apply clustering algorithm

G) 🚽 🤊 - (× 📰 •	-) -	2010 - June	- NodeXL - (Collector - gop_201	.0-09-05_17-30-01 -	Microsoft Exce	k for	Table Tools	maine	dist'		
0	Home	Insert	Page Layout	Formula	as Data	Review Vi	ew Add-Ins	Acrobat	NodeXL	Design				
	Import * Export * Prepare Data *	Refresh Graph	Type: D Cayout: H Automate	irected arel-Koren	Autofi Column	Color Opacity s 🔊 Visibility *	P Vertex Shape ▼ 2 Vertex Size Edge Width	📓 Background	Dynamic Filters	 Graph N Subgrap Groups 	1etrics × oh Images ×	Workb	ook Columns * Elements *	Sa Maria Ho ⊡ Re
	Data		Graph			Visu	al Properties			Grou	o by Vertex At	tribute		1
	A2	•	(* <i>f</i> x	Vertex						Find	Connected Co	mponents		
4	А		В	С	D	E	F	G		Find	Clusters			L.
1	Vertex		Subgraph 🔽	Graph Met In- Degree 🔽	trics Out- Degree 🔽	Betweenness Centrality 🚽	Closeness Centrality 💌	Eigenvector Centrality	▼ Pagel	Clust Colla Colla	Find Cluster Group th clusters u looks at l	e graph's vo ising an alg now the ver	ertices into orithm that tices are	• 0
3	politico			133	2	63463.406	0.001	0.0	002	Expar Expar Select	nd connecte nd The clust color and refreshed	d to one an ers are indio I shape whe I. All the ve	other. ated by vertex en the graph is rtices in one	.0
4	scottwgraves			153	50	45379.121	0.001	0.0	07	Add S	example.	ight be blue ps checkbo	e disks, for x in the	.7
5	c0nservative1	L		146	93	34064.604	0.001	0.0	009	Remo	Show/Hi we menu mu the clust	de, Graph E ist be check ers to be sh	lements ribbon ed if you want own when the	.8
6	arizona_freed	lom		150	102	28549.034	0.001	0.0	009	5.740	The clust the Grou	refresned. er informati ps and Grou	on is stored in up Vertices	.8
7	conservative ²	132		135	74	20294.618	0.001	0.0	009	5.363	workshe	ets. June - Nod	eXL - Collector	8
8	kosmosnet			111	5	15681.759	0.001	0.0	05	4.104	Press F	010-09-05_1 1 for more 19	I7-30-01 help. Image	5.9
9	katyinindy			138	24	12879.665	0.001	0.0	07	4.763	0.07	72	Image	6.6
10	goodporkbad	pork		5	40	12209.131	0.001	0.0	001	2.238	0.04	10	Image	1.6

Eddes Vertices Groups Group Vertices

Twitter network for "GOP"

6		× •) ≠						Table Tools 201	10 - June - NodeXI	eXL - Collector - gop_2010-09-05_17-30-01 - Microsoft Excel	
C	Home	Insert	Page Layout	Formula	as Data	Review Vi	ew NodeXL	Design			🔞 – 🗖 🔀
	Import - Export - Prepare Data - Data	Refresh Graph	Type: D Layout: H Automate Graph	irected larel-Koren	Autofi	Color Opacity	Vertex Shape Vertex Size Edge Width	Background	Dynamic Filters Group	aph Metrics bgraph Images ▼ oups ▼ alvsis Show/Hide Help	Close Window
_	A2	· ((f _x	Vertex							*
	А		В	С	D	E	F	G	H 🔺	Document Actions	▼ x
1				Graph Me	trics		1	1		🖉 😽 Refresh Graph 🔯 Harel-Koren Fast Mul 🗸 🛐 Lay Out Again 🖌 🍸 Dynamic Filt	ers 🤫 Options
2	Vertex		Subgraph	In-	Out-	Betweenness	Closeness Centrality	Eigenvector Centrality	PageRank	▶ ▶ ▶ 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
3	politico			133	2	63463.406	0.001	. 0.002	2 8.24		
4	scottwgraves			153	50	45379.121	. 0.001	. 0.007	7 6.95!		
5	c0nservative1	L		146	93	34064.604	0.001	. 0.009	9 6.06		
6	arizona_freed	lom		150	102	28549.034	0.001	. 0.009	9 5.74		£
7	conservative4	132		135	74	20294.618	0.001	. 0.009	9 5.36:		
8	kosmosnet			111	5	15681.759	0.001	. 0.005	5 4.104		
9	katyinindy			138	24	12879.665	0.001	. 0.007	7 4.76		
10	goodporkbad	pork		5	40	12209.131	0.001	. 0.001	L 2.23		
11	larrysabato			49	3	9425.256	0.001	. 0.002	2 2.34:		~
12	terrisharpe			31	58	9320.256	0.001	. 0.001	L 2.23!		~
13	newsbusters			101	14	8653.937	0.001	. 0.004	4 3.79!		
14	whisper1111			74	37	8514.879	0.001	. 0.000) 3.67!		
15	slone			92	4	. 8418 530	0.001	0.005	5 3.17		

Learn more

Tutorial: Analyzing Social Media Networks: Learning By Doing With NodeXL

Book: Analyzing Social Media Networks with NodeXL: Insights From A Connected World



NodeXL for Social Network Analysis

CAR meeting, Raleigh, Feb 25 2011

Slides: <u>http://www.peteraldhous.com/CAR/Aldhous_CAR2011_NodeXL.pdf/</u>

Peter Aldhous, San Francisco Bureau Chief



