

Graphs and Networks 2



CS 7450 - Information Visualization
November 16, 2016
John Stasko

Review



- Last time we looked at graph layout aesthetics and algorithms, as well as some example applications
- Today we look at more recent InfoVis network visualization systems & projects

Learning Objectives



- Understand how interaction can be used to assist network visualization
- Be able to explain concept of "attribute-based layout"
 - Provide examples of systems using technique
- List and define different network analysis metrics (eg., betweenness centrality)
- Explain different approaches to visualizing large networks
- Describe and explain contribution of following network visualization systems
 - TreePlus, PivotGraph, Semantic Substrates, CiteVis, Social Action, vanHam&Perer, NodeTrix, Ploceus

Interaction



- One of the key ways we move beyond graph layout to network visualization (InfoVis) is interaction with the graph

TreePlus



- Don't draw entire graph
- Have a focus vertex, then incrementally expand and show connections (min span tree) from there
- Interaction:
 - Single-click: show connections via highlight
 - Double-click: new focus vertex
 - Smooth animated change in focus
- "Plant a seed and watch it grow"

Lee et al
TVCG '06

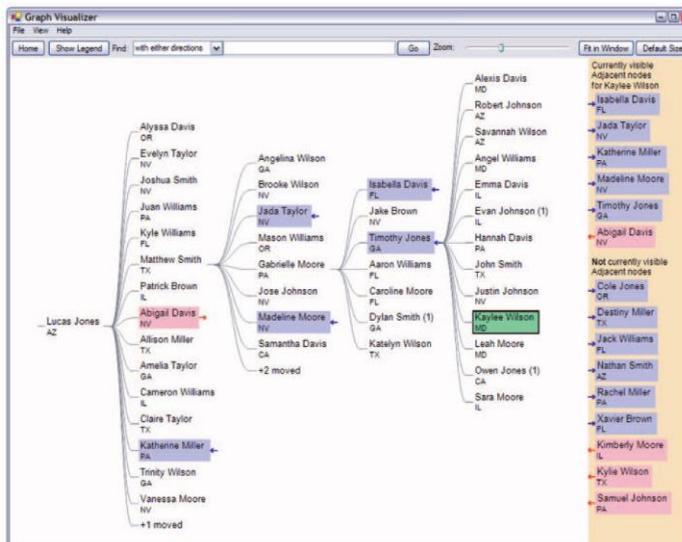
Fall 2016

CS 7450

5

TreePlus

Video



Green – current selection

Blue – vertices from current selection

Red – vertices to current selection

Fall 2016

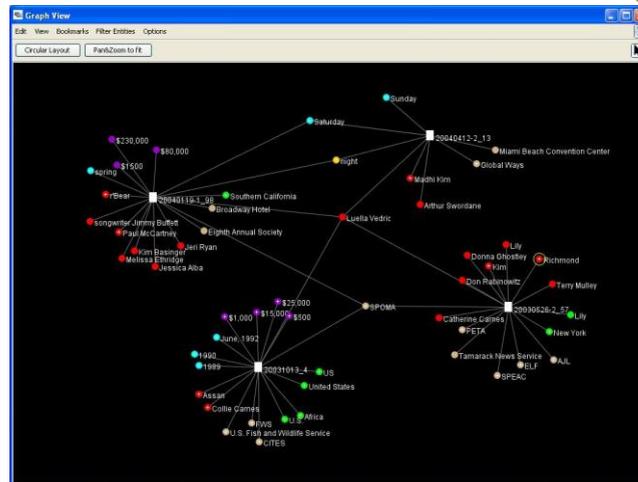
CS 7450

6

Jigsaw's Graph View



Don't draw everything, but allow the viewer to interactively explore (expand & compress) the graph



Stasko, Görg & Liu
Information Visualization '08
Fall 2016

CS 7450

Demo

7

Recent Trends in GraphViz



- Attributes of nodes influence geometric positioning
 - Not just some arbitrary layout
- Utilize graph statistical analysis too

Attribute-based layout

- Largely driven by interest in social network analysis

Fall 2016

CS 7450

8

PivotGraph



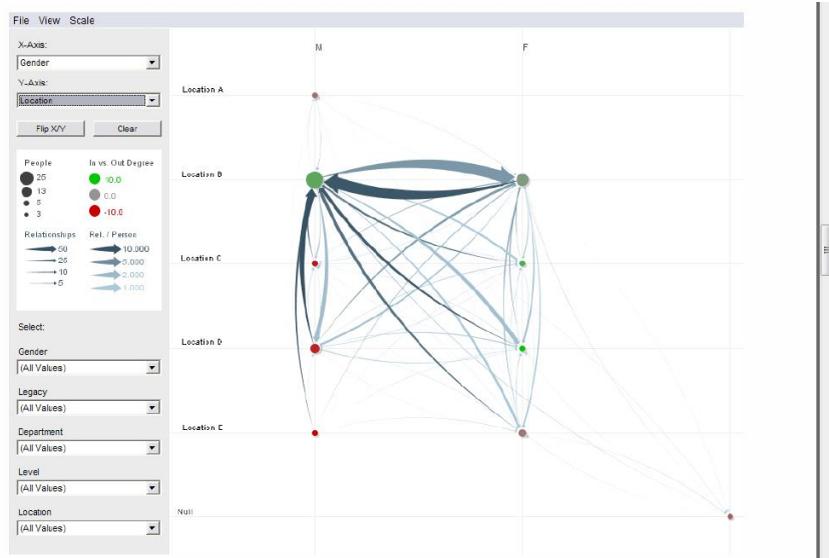
- Position nodes into a grid based on attributes
- Cluster on common node attributes
 - Put all A's together, all B's together, ...
- "Roll up" nodes
 - Draw edge from A to B depending on how many edges from some A to some B

Wattenberg
CHI '06

Fall 2016

CS 7450

9



Fall 2016

CS 7450

10

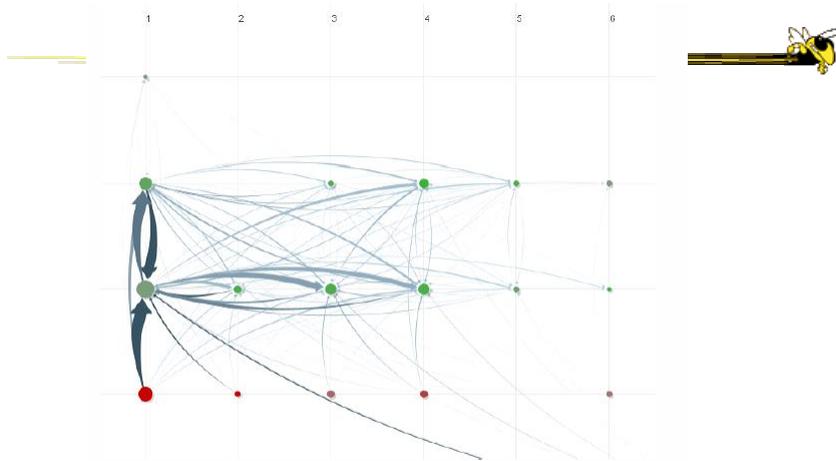
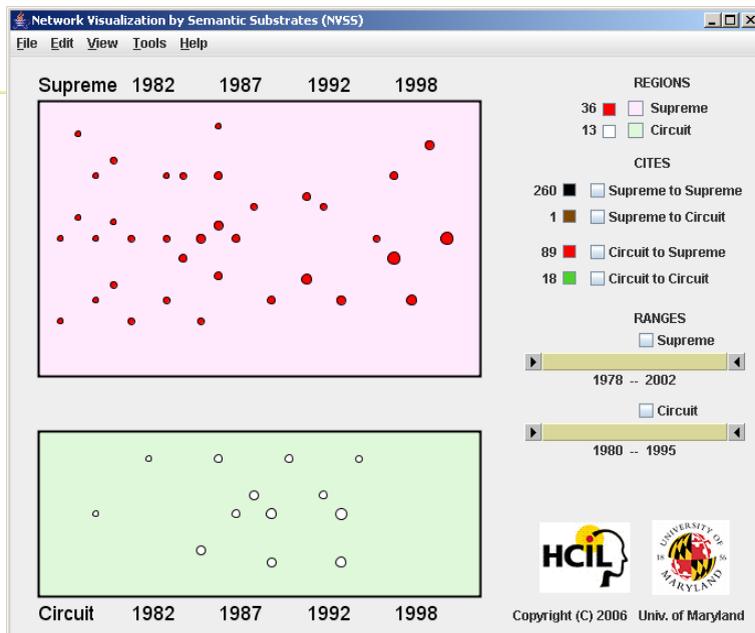


Figure 10. *Communication network of people in a large company. X-axis is division, y-axis is office geography. The division in the leftmost column has far more cross-location communication than the others.*

<http://www.cs.umd.edu/hcil/nvss/>

Semantic Substrates

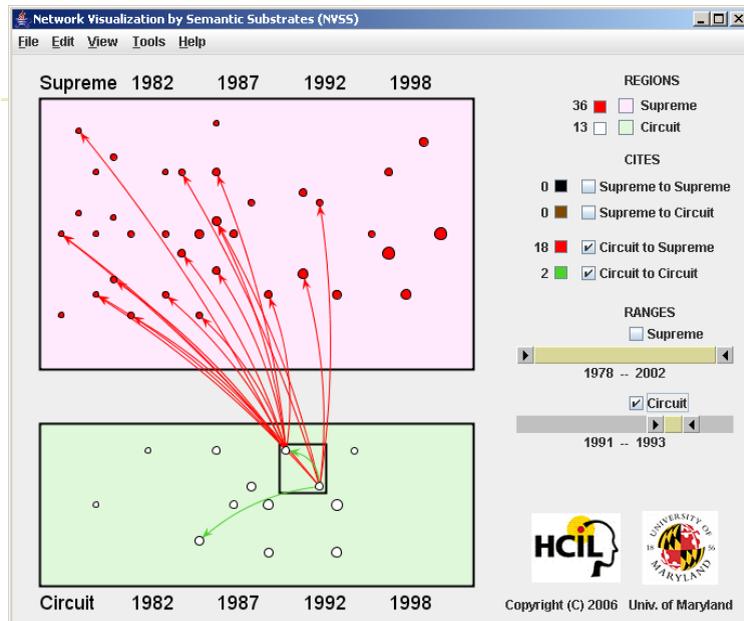
- Group nodes into regions
 - According to an attribute
 - Categorical, ordinal, or binned numerical
- In each region:
 - Position nodes according to some other attribute(s)
- Give users control of link visibility



Fall 2016

CS 7450

13



Fall 2016

CS 7450

Video

14

dotlink360

Connections between companies
(supplier, partner, etc.)



Basole, et al
TVCG (InfoVis) '13

Fall 2016

CS 7450

15

CiteVis



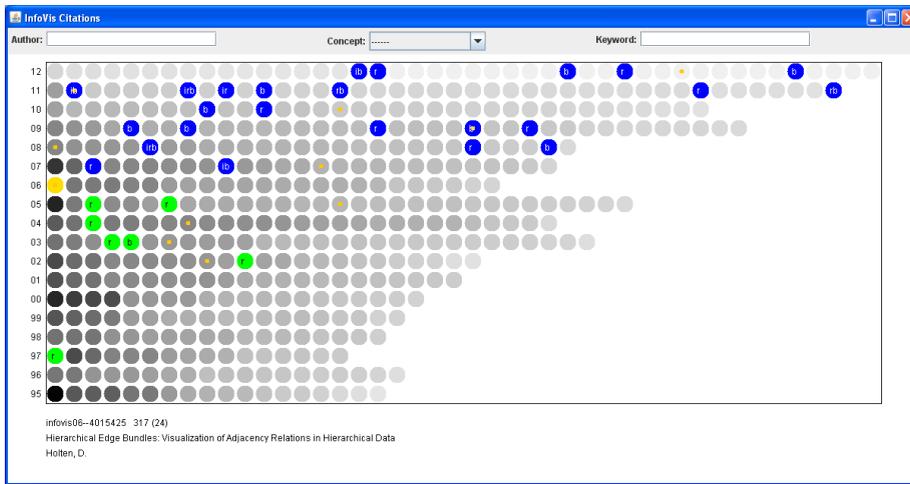
- Showing InfoVis Conference paper citation patterns
 - Papers are graph vertices
 - A cites B is graph edge
- Attribute-based layout
 - Year x Number of citations
- Uses color & interaction to show citations rather than drawn links

Stasko, Choo, Han, Hu, Pileggi, Sadana & Stolper
InfoVis poster '13

Fall 2016

CS 7450

16



Demo

Fall 2016

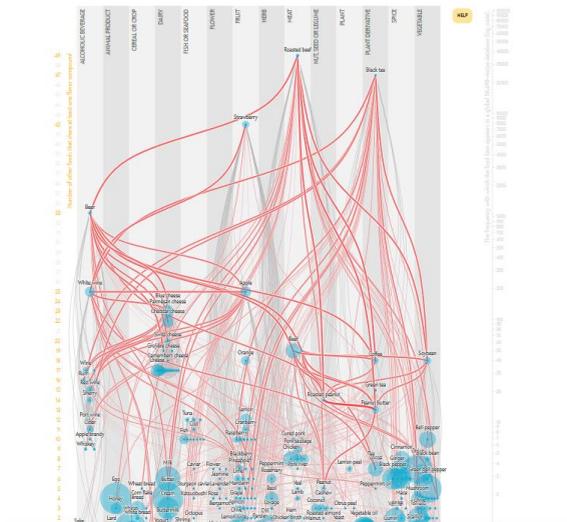
CS 7450

17

<http://www.scientificamerican.com/article/flavor-connection-taste-map-interactive/>

The Flavor Connection [Interactive]
 Scientists link common flavor compounds across the world's favorite ingredients
 Aug 20, 2013

Julia Child famously said that fat carries flavor, but perhaps instead we should give thanks to 4-methylpentanoic acid. Unique combinations of such chemical compounds give foods their characteristic flavors. Science-minded chefs have gone so far as to suggest that seemingly incongruous ingredients—chocolate and blue cheese, for example—will taste great together as long as they have enough flavor compounds in common. Scientists recently put this hypothesis to the test by creating a flavor map, a variant of which we have reproduced here. Lines connect foods that have components in common; thick lines mean many components are shared. By comparing the flavor network with various recipe databases, the researchers conclude that chefs do tend to pair ingredients with shared flavor compounds—but only in Western cuisine. Dishes from a database of recipes from East Asia tend to combine ingredients with few overlapping flavors.

Flavor Connection
(saw last time)

18

Vizster



- Visualize social networking sites like friendster, myspace, facebook
- Implementation
 - Crawled 1.5 million members (Winter 2003)
 - Written in Java using the *prefuse* toolkit (<http://prefuse.sourceforge.net>)
- Oppose Shneiderman's mantra. Instead: "Start with what you know, then grow."

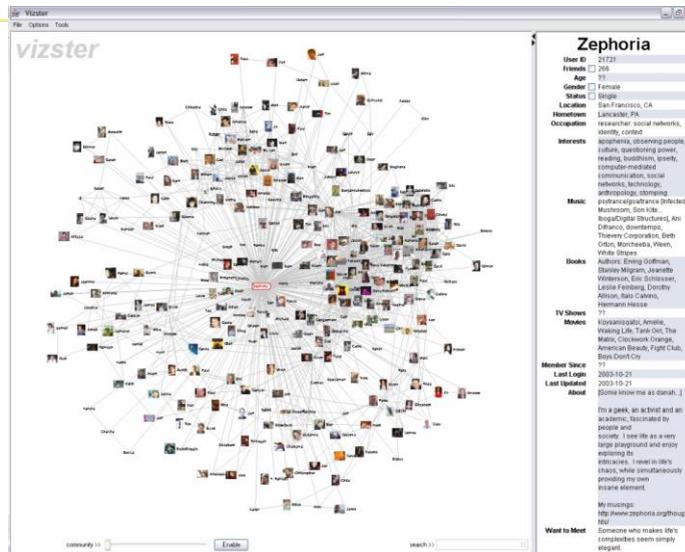
Heer & boyd
InfoVis '05

Fall 2016

CS 7450

19

Visualization



Fall 2016

CS 7450

20

Combining Features

Video

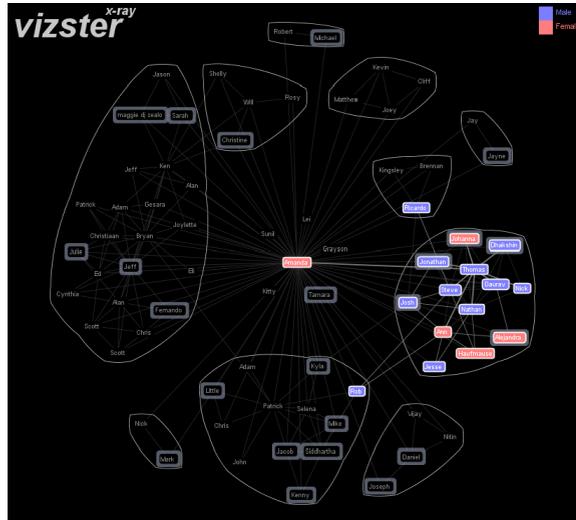


Colors: Gender

Halo: Search for "student"

Highlight: Friends of selection

Blobs: Communities



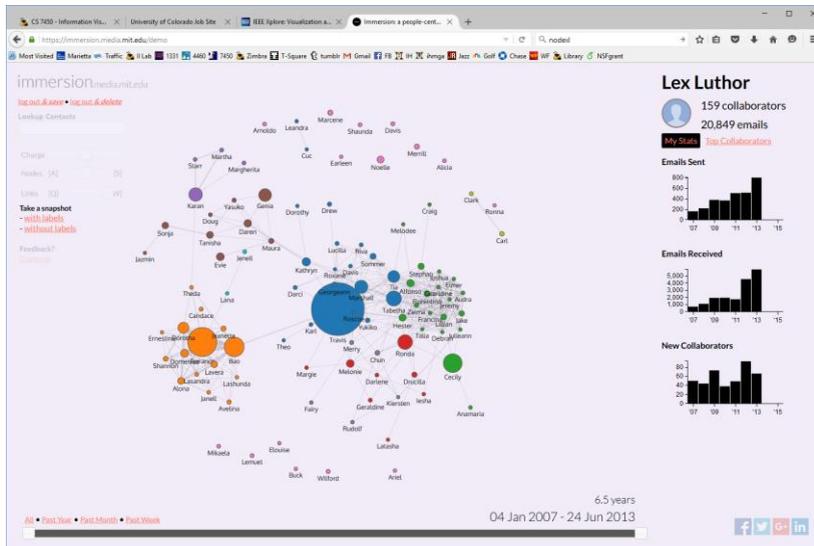
Fall 2016

CS 7450

21

<https://immersion.media.mit.edu/>

Immersion



Fall 2016

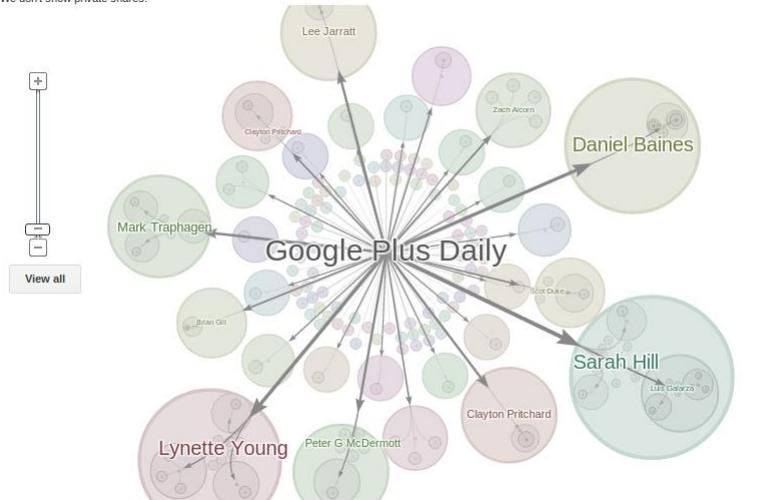
CS 7450

Demo

22

Google+ Ripples

Showing 241 public shares. 347 total shares.
We don't show private shares.



Fall 2016

CS 7450

Now defunct

23

<http://www.cs.umd.edu/hcil/socialaction/>

SocialAction

- Combines graph structural analysis (ranking) with interactive visual exploration
- Multiple coordinated views
 - Lists by ranking for analysis data
 - Basic force-directed layout for graph vis

Perer & Shneiderman
TVCG (InfoVis) '06

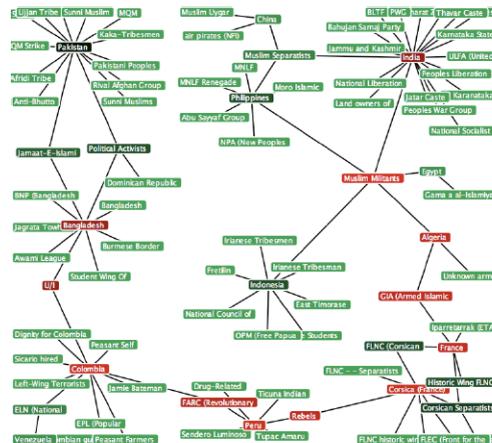
Fall 2016

CS 7450

24

Rank	Node	Type
2,316.00	Muslim Militants	Terrorist Group
2,436.90	Corsica (France)	Country
2,413.00	Colombia	Country
2,388.00	Peru	Country
2,280.50	France	Country
2,239.00	Algeria	Country
2,226.00	Hezbollah	Terrorist Group
2,214.00	GIA (Armed Islamic Group)	Terrorist Group
2,124.00	PARC (Revolutionary Armed For...	Terrorist Group
1,718.00	Bangladesh	Country
1,615.00	YPL	Terrorist Group
1,598.00	India	Country
1,063.00	Pakistan	Country
788.00	Osseken Separatists	Terrorist Group
704.00	RLNC (Corsican National Libera...	Terrorist Group
704.00	Historic Wing RLNC	Terrorist Group
657.00	Indonesia	Country
614.00	Political Activists	Terrorist Group
596.00	Philippines	Country
520.00	Jamaat-E-Islami	Terrorist Group
350.00	Muslim Separatists	Terrorist Group
276.00	ELN (National Liberation Army)	Terrorist Group
187.00	Venezuela	Country
187.00	China	Country
149.00	Egypt	Country
0.00	Dignity for Colombia	Terrorist Group
0.00	Jamie Bateman Canon Front	Terrorist Group
0.00	Sendero Luminoso	Terrorist Group
0.00	Jamaat-ul-Mujahideen	Terrorist Group
0.00	Timorese Students	Terrorist Group

(a) Ordered list of 97 nodes in the largest connected component of the terrorism network in 1996. The nodes are ranked according to their betweenness centrality.



(b) Network visualization of the same 97 nodes, colored according to their ranking. The nodes with highest betweenness rankings, sometimes referred to as "gatekeepers", are painted red.

Figure 1.

Social Network Attributes



- **Bary center** – total shortest path of a node to all other nodes
- **Betweenness centrality** – how often a node appears on the shortest path between all other nodes
- **Closeness centrality** – how close a node is compared to all other nodes
- **Cut-points** – the subgraph becomes disconnected if the node is removed
- **Degree** – number of connections for node
- **HITs** – "hubs and authorities" measure
- **Power centrality** – how linked a node is to rest of network

Attribute Ranking



- Run these measures on all nodes and rank them
- Sort the rankings and show in lists and scatterplots
- Allow user to filter based on rankings
- Can aggregate rankings for cohesive subgroups of nodes

Fall 2016

CS 7450

27

Graph Visualization



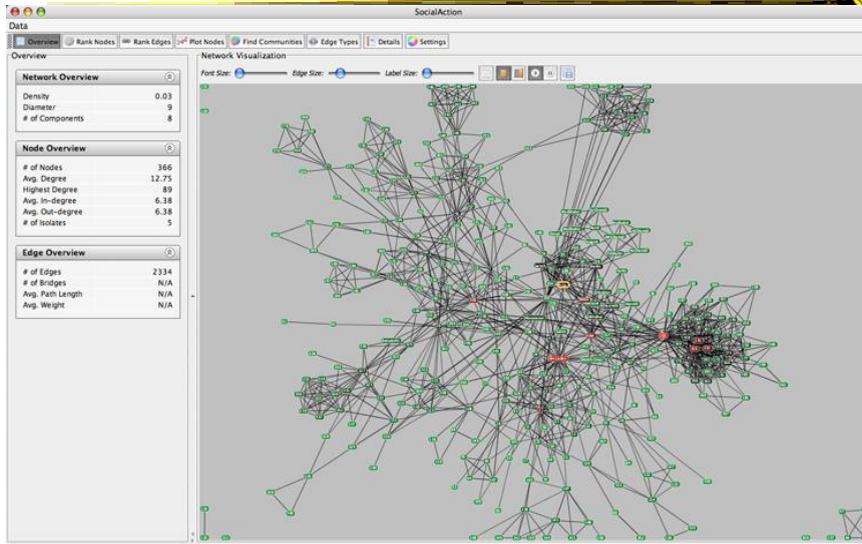
- Standard node-link
- Node positions remain constant across different metric views to promote comprehension
- Links can have types
- Coherent subgroups can be aggregated (like in Vizster)
 - Uses Newman's community identification algo

Fall 2016

CS 7450

28

Users begin with an overview of the entire social network. On the left side, overview statistics that describe the overall structure are presented. On the right, the network is visualized using a force directed algorithm.

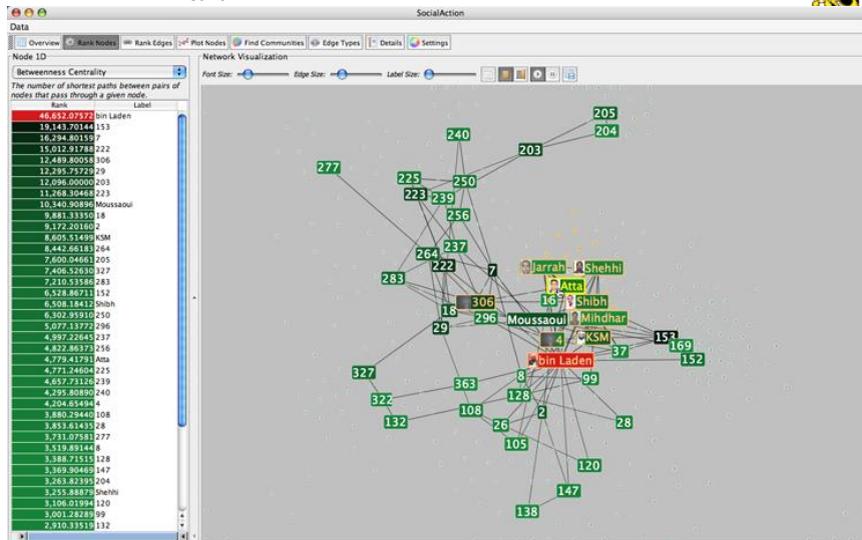


Fall 2016

CS 7450

29

The gatekeepers are found using a statistical algorithm. Users filter out the unimportant nodes using a dynamic slider which simplifies the visualization while maintaining the node positions and structure of the network.



Fall 2016

CS 7450

30

Implementation



- Jung
 - Network data structures and algorithms
- Prefuse
 - Graph drawing
- Piccolo
 - Scatterplot and Matrix views

Fall 2016

CS 7450

33

Comments



- One of my favorite recent InfoVis papers
- Not too innovative on the vis technique side, but wonderful application and synthesis of useful capabilities
- Actually, a very nice *visual analytics* example
- Good subsequent paper on case studies evaluation of it (on our later Eval day)

Fall 2016

CS 7450

34

Really Big Graphs



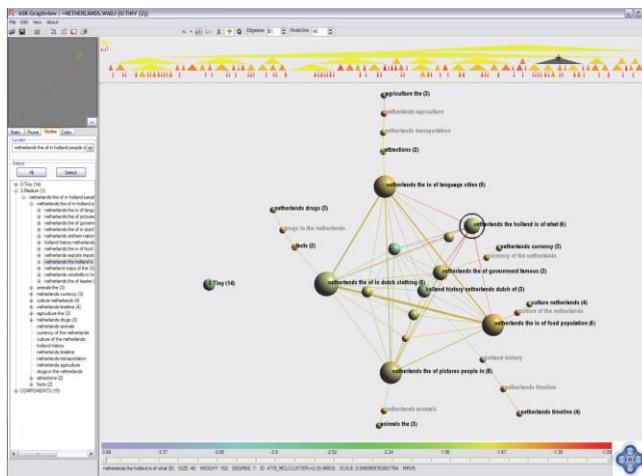
- May be difficult to keep all in memory
- Often visualized as “hairballs”
- Smart visualizations do structural clustering, so you see a high-level overview of topology

Fall 2016

CS 7450

35

ASK-GraphView



Note

Uses clustering algorithms to construct a hierarchy

Fall 2016

CS 7450

Abello, van Ham & Krishnan
TVCG (InfoVis) '06

36

Alternate Big Graph Approach



- Show some of the details, rather than high level structure
- Allow users to focus on particular nodes
- Adapt DOI algorithm from trees to graphs
- Rely heavily on interaction
- Different paradigm: "Search, show context, expand on demand"

van Ham & Perer
TVCG (InfoVis) '09

Fall 2016

CS 7450

37

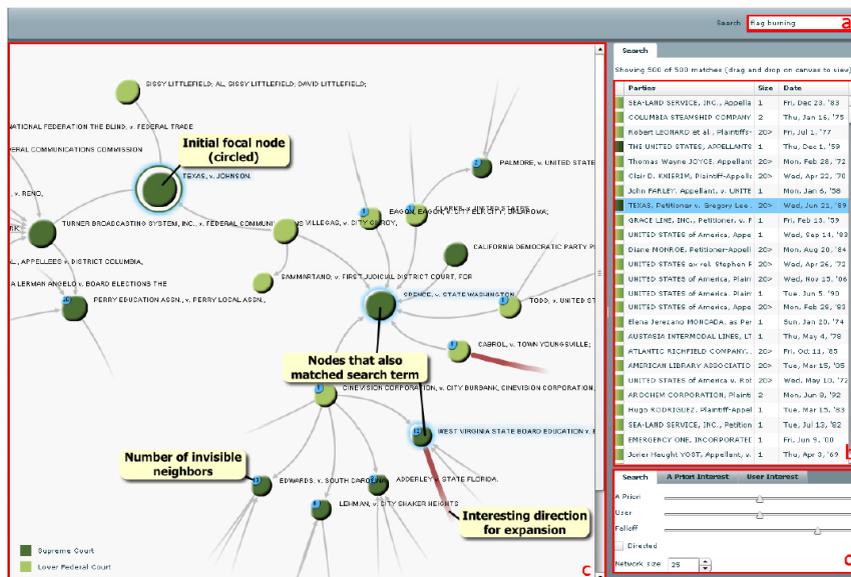


Fig. 3. Basic user interface layout. A user types a query in the searchbox (a) which yields a number of hits presented in tabular form (b). One of these hits can then be dragged to the main screen (c) which shows the subgraph centered on that node. Other nodes that matched the user's search are highlighted in blue. Users can adapt the balance between different components of the DOI function and the size of the subgraph in a separate panel (d).

Fall 2016

CS 7450

38

Graphs as Maps



- Represent a large graph as a map
- Maintain inherent structure and relationships between nodes
- Follow standard cartographic representations

Gansner, Hu & Kobourov
IEEE CG&A (PacificVis) '10

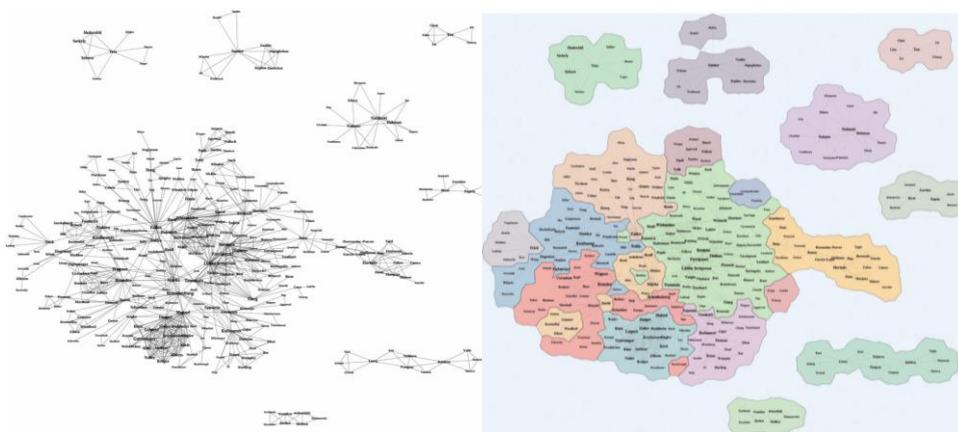
Fall 2016

CS 7450

39

<http://www2.research.att.com/~yifanhu/MAPS/imap.html>

Both Representations



Node-link

Map

Fall 2016

CS 7450

40

Music Graph/Map



Fall 2016

CS 7450

41

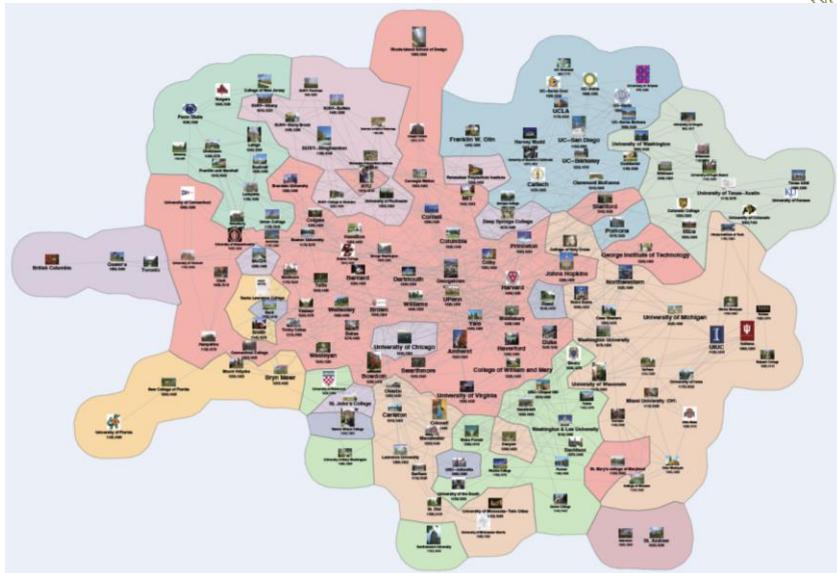


Fall 2016

CS 7450

42

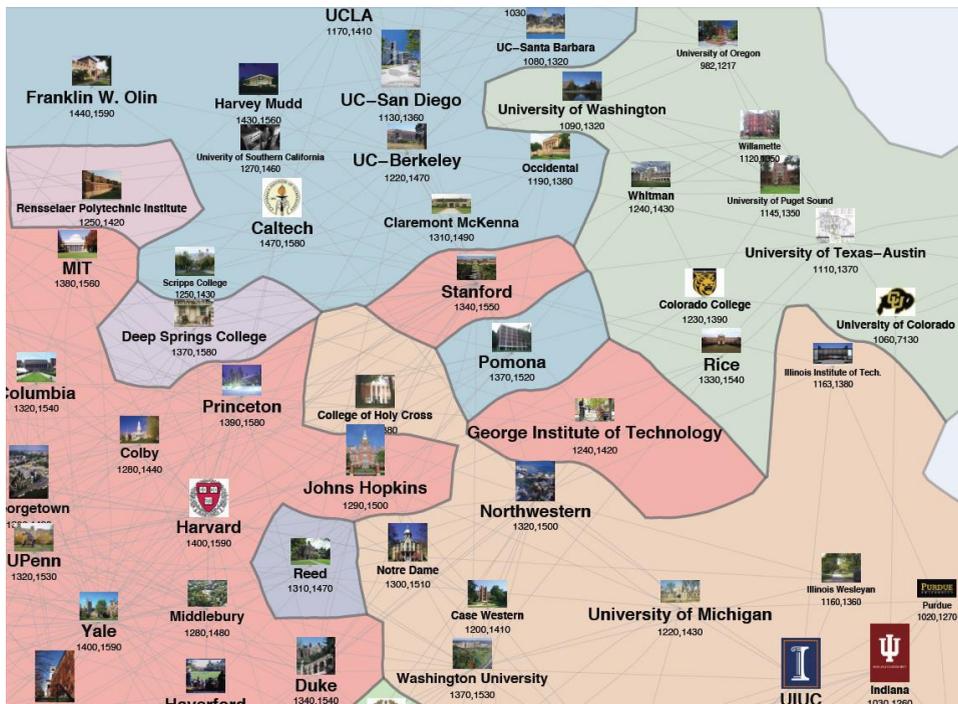
Colleges



Fall 2016

CS 7450

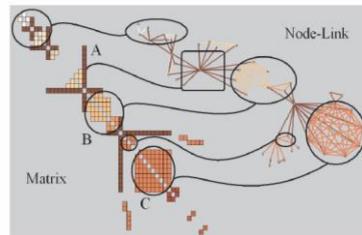
45



Matrix Representations



- There has been renewed interest in matrix representations of graphs recently
- I think the regularity, symmetry, and structure of a matrix are a win – people understand them well, but they don't scale up really well



Fall 2016

CS 7450

47

MatrixExplorer



- Provides matrix view in combination with node-link and various operations for gaining different perspectives

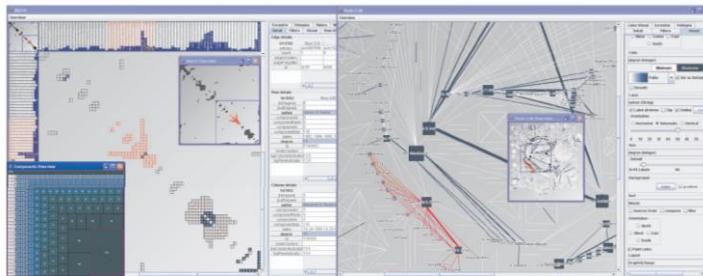


Fig. 1. MatrixExplorer showing two synchronized representations of the same network: matrix on the left and node-link on the right.

Fall 2016

CS 7450

Henry & Fekete
TVCG (InfoVis) '06

48

Node Reordering



Extremely important operation with matrix representations

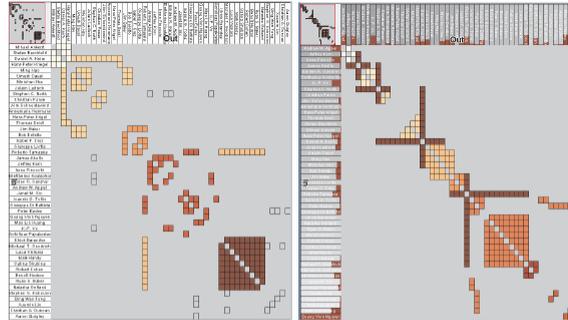


Fig. 6. Initial order (left) and TSP order (right). Colors represent clusters found by the user. Clusters are different in the two representations. Users found more clusters with TSP order. Headers red indicators (right) represents the distance between adjacent rows/columns.

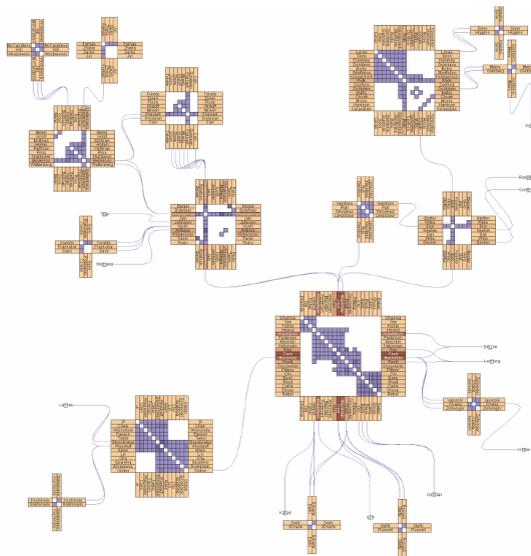
Fall 2016

CS 7450

49

NodeTrix

Hybrid of matrix and node-link



Henry & Fekete
TVCG (InfoVis) '07

Fall 2016

CS 7450

50

Simplifying Input



- Make it easier to input graphs and then explore them

Fall 2016

CS 7450

51

<http://nodexl.codeplex.com/>

NodeXL



NodeXL: Network Overview, Discovery and Exploration for Excel

CodePlex Open Source Community

72 people are following this project

Download

NodeXL Excel Template, version 1.0.1.051
Date: Wed Feb 9 2011 at 3:00 AM
Downloads: 2488
Ratings: No Ratings

You could win an HTC Surround* or an Ultra Light Laptop!

Write a Windows Phone 7 article or share a Windows Phone 7 application.
Enter Now

Activity: 30 All days
Page Views: 5660
Visits: 1876
Downloads: 2488

Fall 2016

CS 7450

52

Characteristics



- Plug-in for MS Excel
- Includes many network layout and network analysis metrics
- Data import:
 - List out vertices and edges in Excel columns
 - Native importers for email, Twitter, YouTube, etc.

Smith et al
C&T '09

Fall 2016

CS 7450

53

Non-Network Data?



- But what if you don't have vertex-edge data to begin?
 - May just have tabular data from spreadsheet or database
- Still may want to explore data modeled as a graph
 - Consider DB of NSF grants (PIs, institution, PM, amount, ...)
 - Look for clusters, patterns, connections, ...

Fall 2016

CS 7450

54

Ploceus



- Framework and system for modeling and visualizing tabular data as network
- Allow user to model data as graph interactively through direct manipulation
 - What are vertices, edges, edge weights, ...
- Visualizes graph on-the-fly (different layouts and network metrics)
- Advanced ops (project, aggregate, slice-n-dice) can be specified interactively too

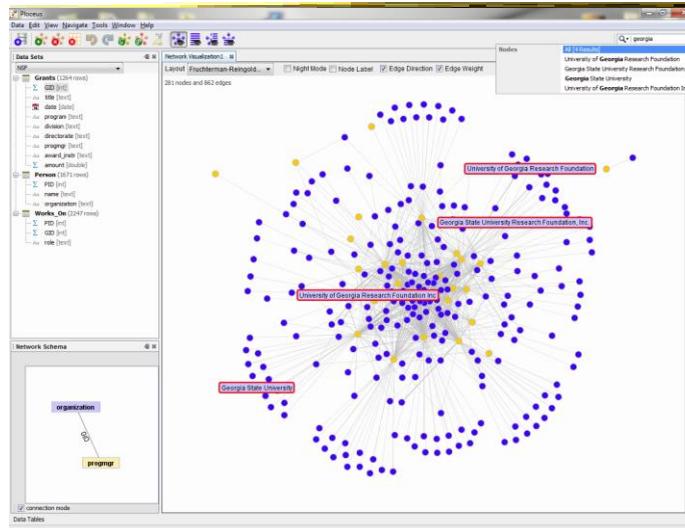
Fall 2016

CS 7450

55

<http://www.cc.gatech.edu/gvu/ii/ploceus/>

Ploceus



Video

Fall 2016

CS 7450

56

GLOs



- In all these network visualizations, can the differences/changes/operations be summarized into a small set?
 - Align, distribute, size, hide, clone, aggregate, position, substrate, etc.
- Can then use the operations to transform from one representation to another

Stolper, et al.
TVCG (InfoVis) '14

Fall 2016

CS 7450

57

GLO-STIX

Video



• *GLO-STIX: Graph-Level Operations for Specifying Techniques and Interactive exploration*

The screenshot shows the GLO-STIX interface. On the left is a 'Select GLOs' panel with a tree view of operations: Positioning Nodes (Align Left, Align Center, Align Right, Align Top, Align Middle, Align Bottom, Evenly Distribute on X, Evenly Distribute on Y, Evenly Distribute on X by Category, Evenly Distribute on Y by Category, Substrate on X by Gender, Substrate on Y by Category, Evenly Distribute on x Within Substrates, Position on X Relatively by Betweenness Centrality, Position on Y Relatively by Degree), Modifying Element Properties, Cloning, Aggregation, and Modifying Display Properties. The 'Applied GLOs' panel on the right lists: Substrate on X by Gender, Substrate on Y by Category, Aggregate by Gender and Category, Show All Links, Links to Curved, Size Nodes by Count, Show X Axis, and Show Y Axis. The main window displays a network graph with nodes of various sizes and colors (brown, purple, red, green, orange, blue) and edges. The graph is plotted on a coordinate system with X and Y axes. Below the main graph are two smaller thumbnail views of the graph.

Encapsulate operations into an editing tool

Fall 2016

CS 7450

58

Learning Objectives



- Understand how interaction can be used to assist network visualization
- Be able to explain concept of "attribute-based layout"
 - Provide examples of systems using technique
- List and define different network analysis metrics (eg., betweenness centrality)
- Explain different approaches to visualizing large networks
- Describe and explain contribution of following network visualization systems
 - TreePlus, PivotGraph, Semantic Substrates, CiteVis, Social Action, vanHam&Perer, NodeTrix, Ploceus

Fall 2016

CS 7450

59

Upcoming



- Time series data
- Visual Analytics

Fall 2016

CS 7450

60