Graphs and Networks 2

CS 7450 - Information Visualization
November 11, 2015
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

Interaction

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

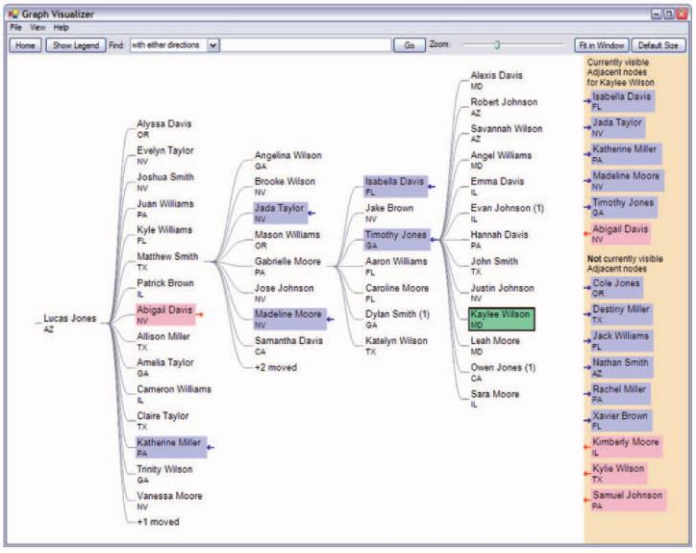
TreePlus

- 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

TreePlus

Video



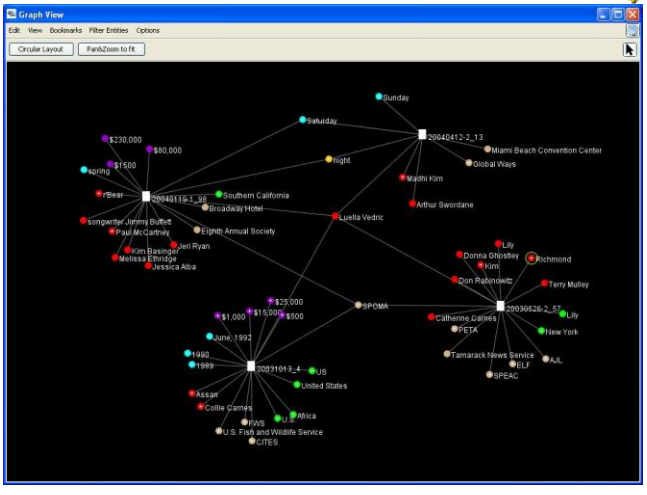
Green – current selection

Blue – vertices from current selection

Red – vertices to current selection

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

Recent Trends in GraphViz

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

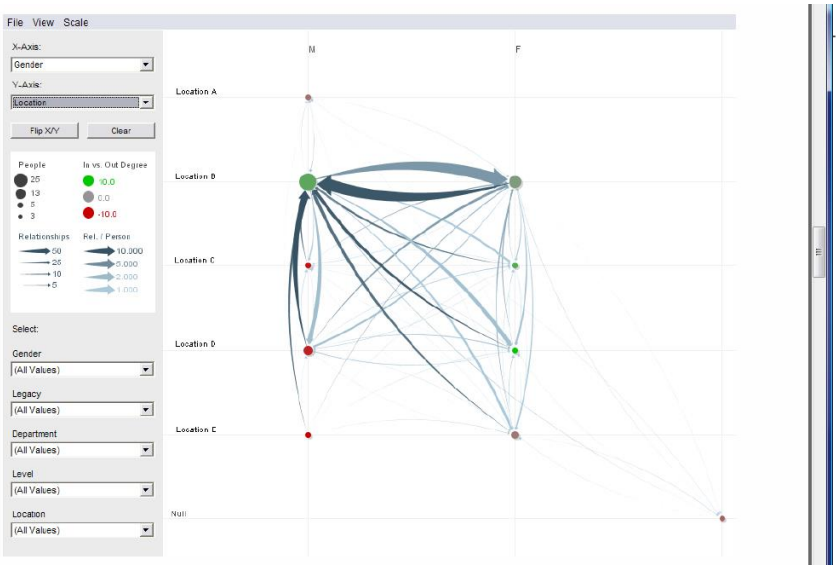
Attribute-based layout

- Largely driven by interest in social network analysis

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



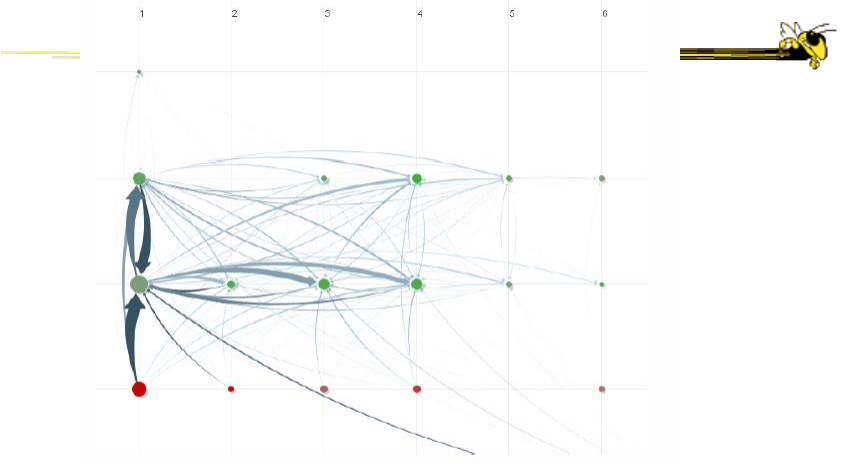


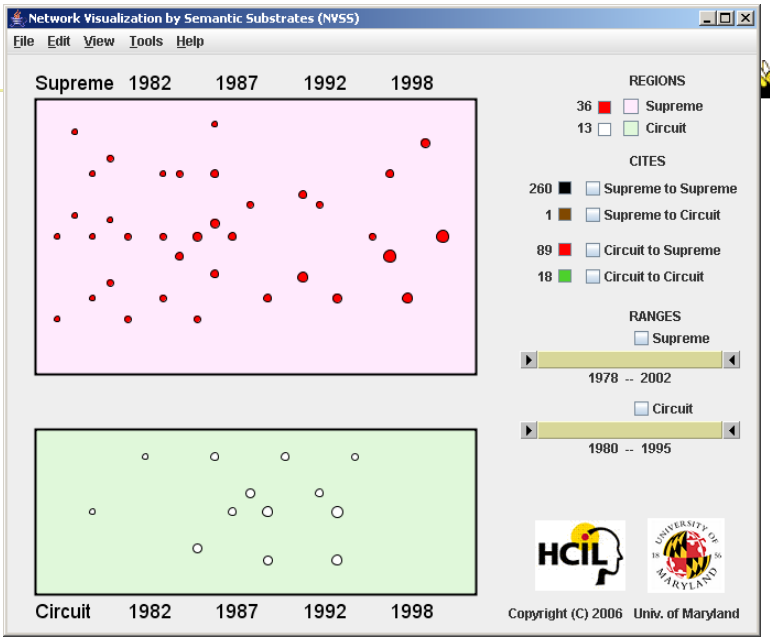
Figure 10. *Communication network of people in a large company. X-axis is division, y-axis is office geography. The division in the leftmost column has far more cross-location communication than the others.*

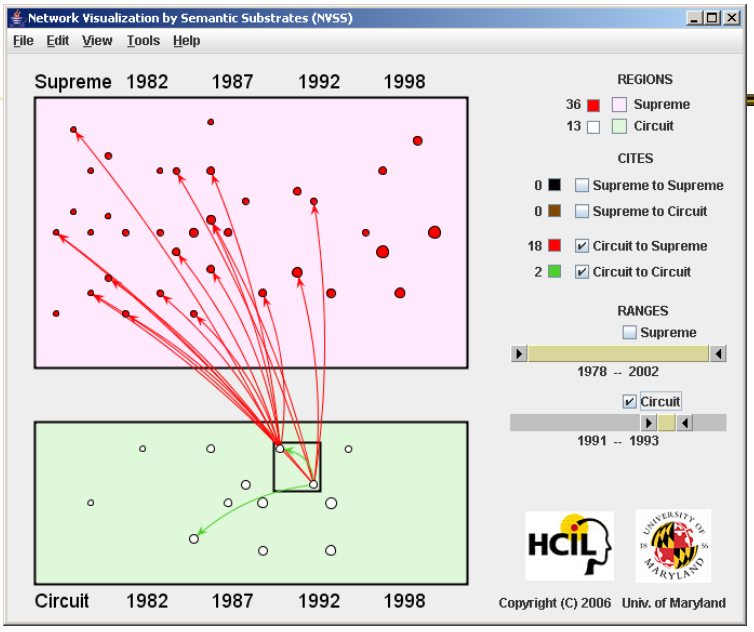
http://www.cs.umd.edu/hcil/nvss/

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

Shneiderman & Aris
TVCG (InfoVis) `06



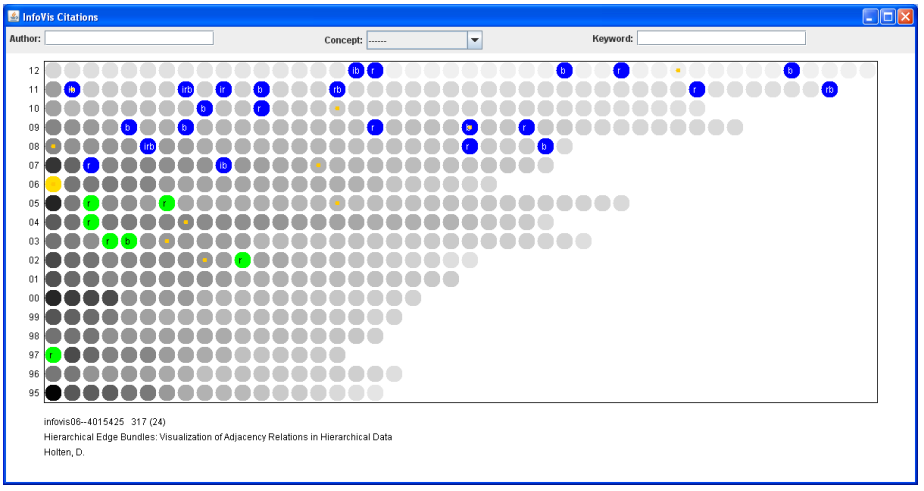


CiteVis

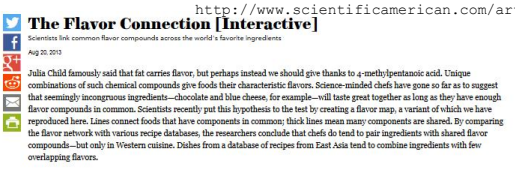
- Showing InfoVis Conference paper citation patterns
  - Papers are graph vertices
  - A cites B is graph edge
- Attribute-based layout
  - Year x Number of citations
- Uses color & interaction to show citations rather than drawn links

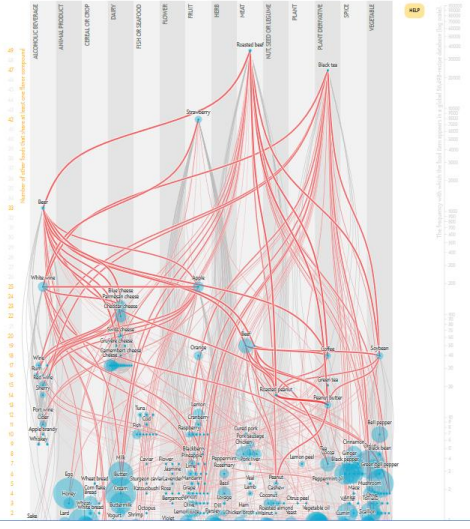
Stasko, Choo, Han, Hu, Pileggi, Sadana & Stolper
InfoVis poster '13

http://www.cc.gatech.edu/gvu/ii/citevis



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





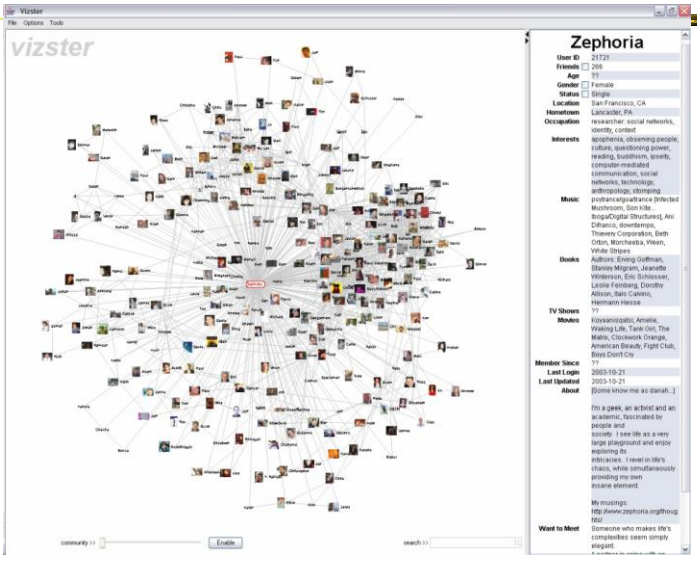
Flavor Connection

Vizster

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

Heer & boyd
InfoVis '05

Visualization



Combining Features

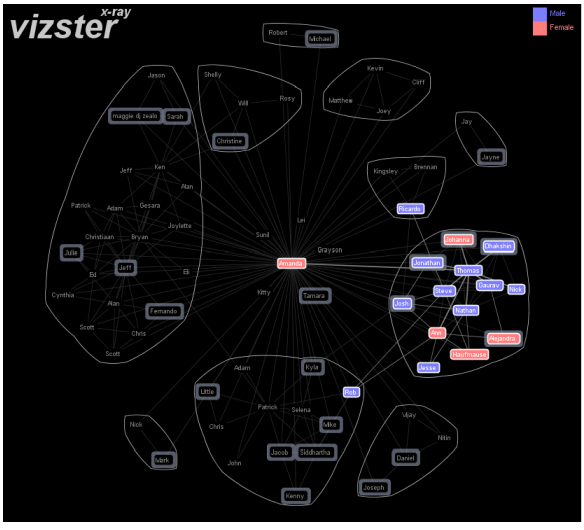
Video

Colors: Gender

Halo: Search for "student"

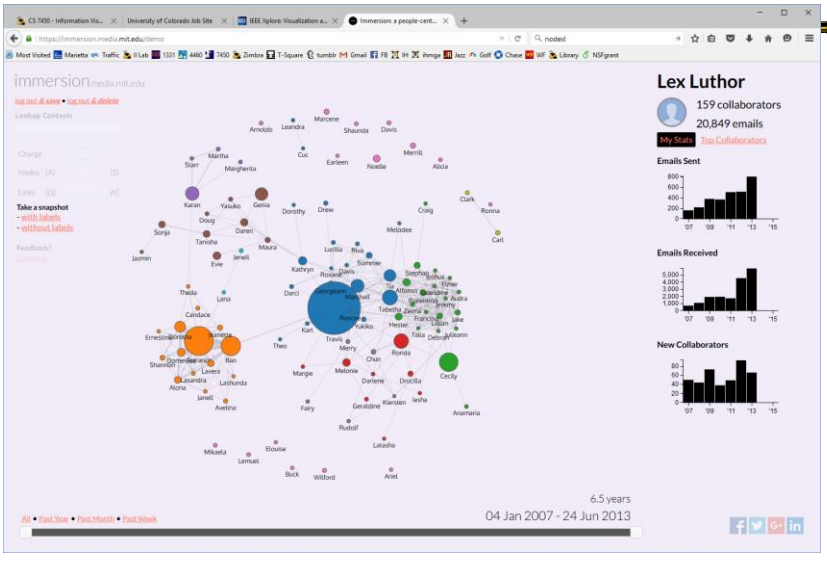
Highlight: Friends of selection

Blobs: Communities

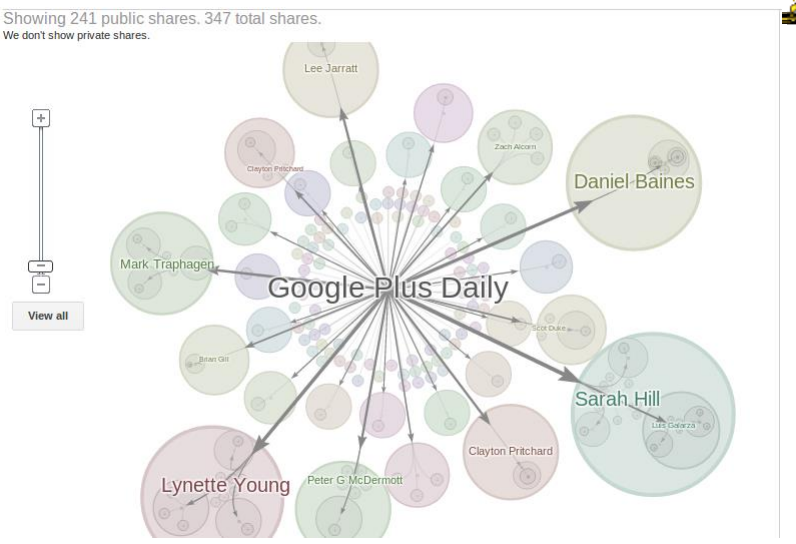


https://immersion.media.mit.edu/

Immersion



Google+ Ripples



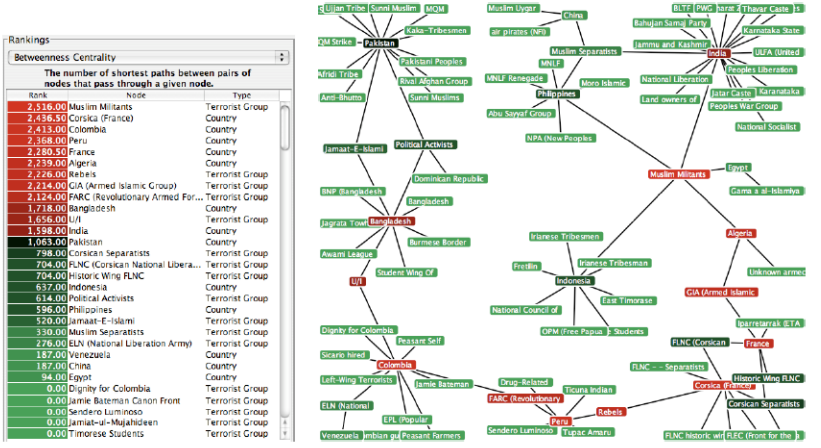
Now defunct

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



(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

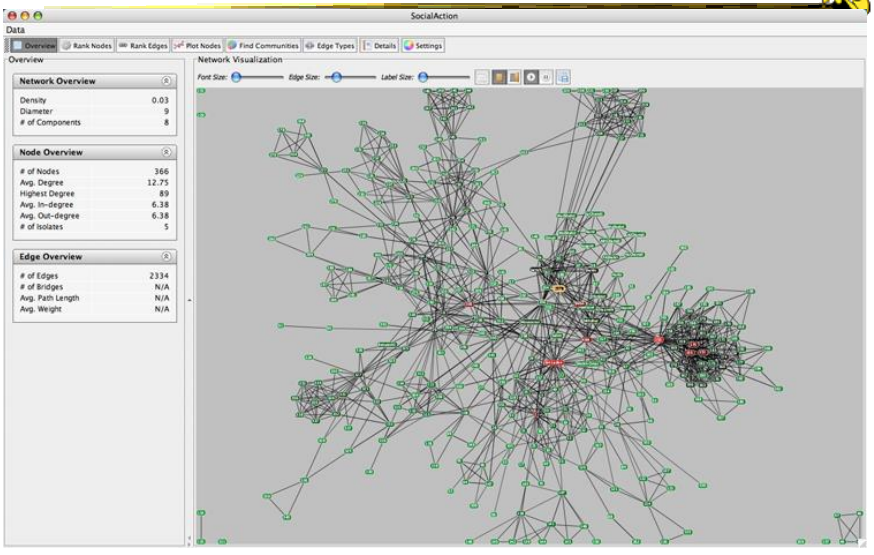
Attribute Ranking

- 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

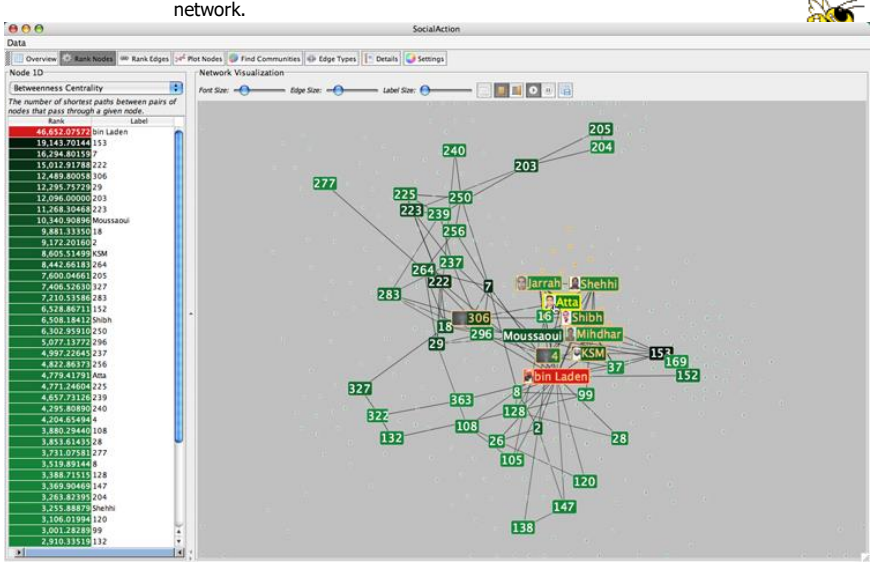
Graph Visualization

- 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

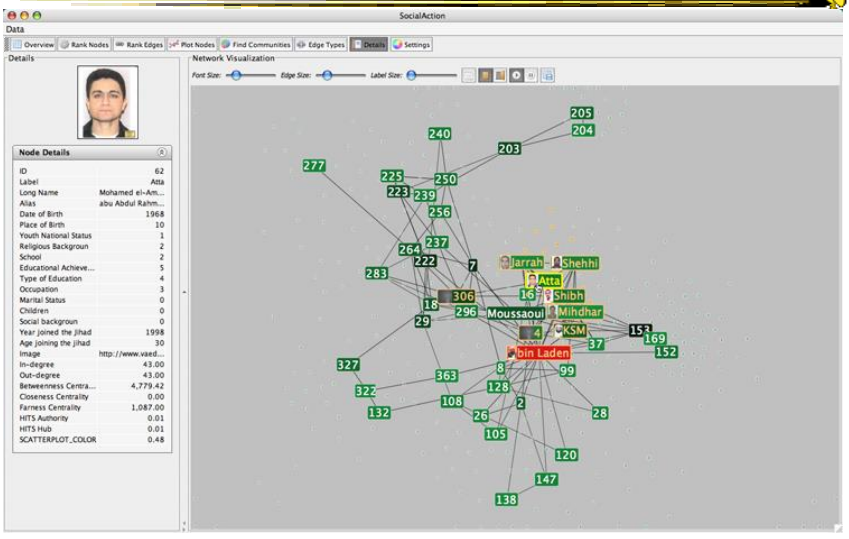
Users begin with an overview of the entire social network. On the left side, overview statistics that describe the overall structure are presented. On the right, the network is visualized using a force directed algorithm.



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.

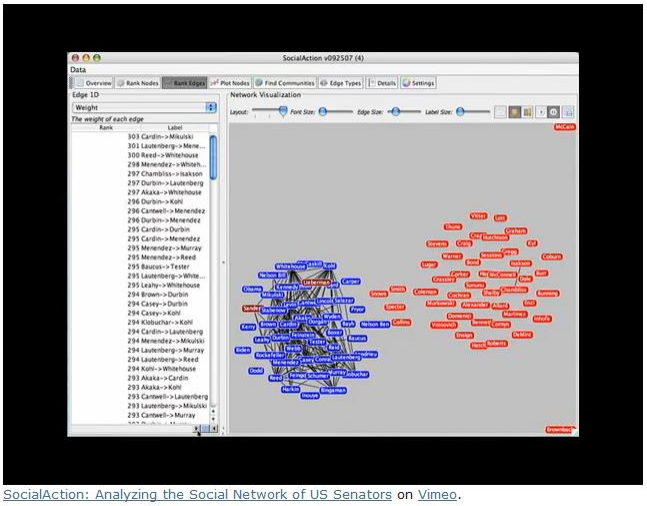


Labels are always given priority so users can understand what the data represents. When user selects a node, neighbors are highlighted and details appear on the left. In order to protect sensitive information, node labels have been anonymized except for those individuals publicly identified in the Zacarias Moussaoui trial.



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

Senate Voting Patterns



Video

SocialAction: Analyzing the Social Network of US Senators on Vimeo.

Implementation

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

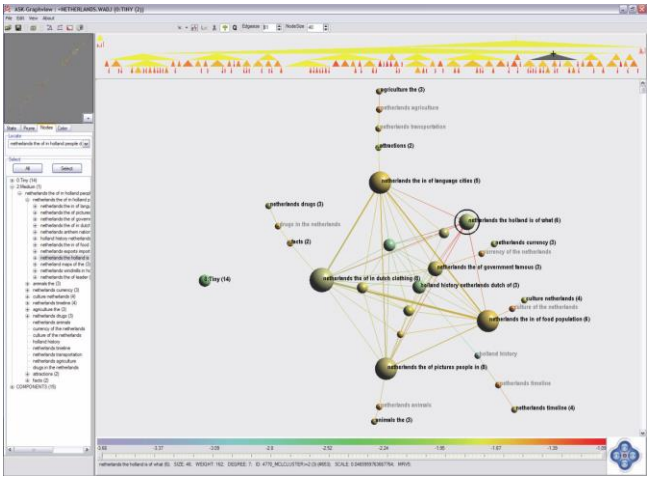
Comments

- 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)

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

ASK-GraphView



Uses clustering algorithms to construct a hierachy

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

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

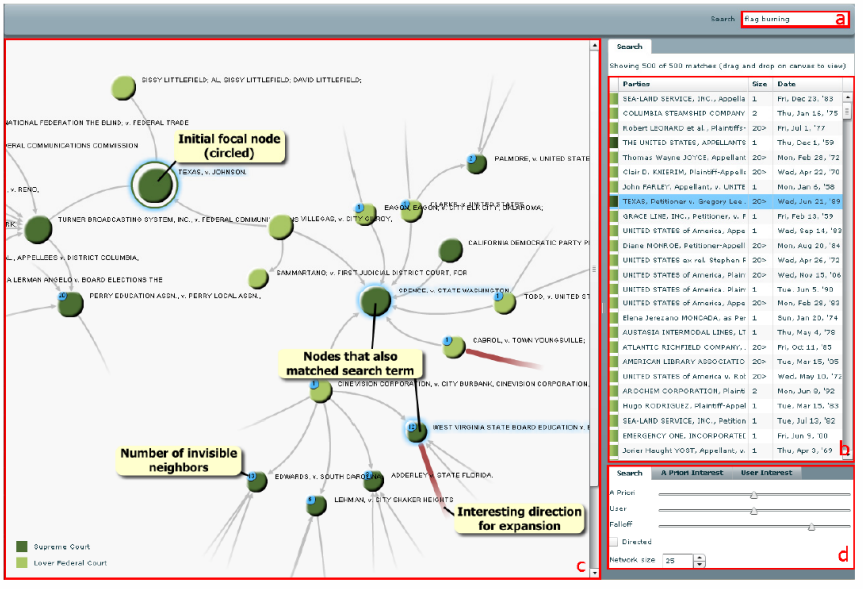


Fig. 3. Basic user interface layout. A user types a query in the searchbox (a) which yields a number of hits presented in tabular form (b). One of these hits can then be dragged to the main screen (c) which shows the subgraph centered on that node. Other nodes that matched the user's search are highlighted 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).

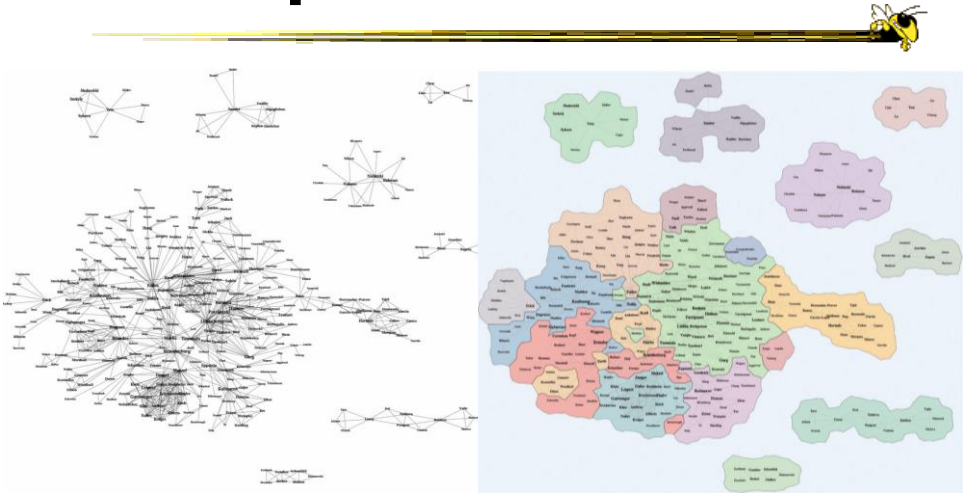
Graphs as Maps

- 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

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

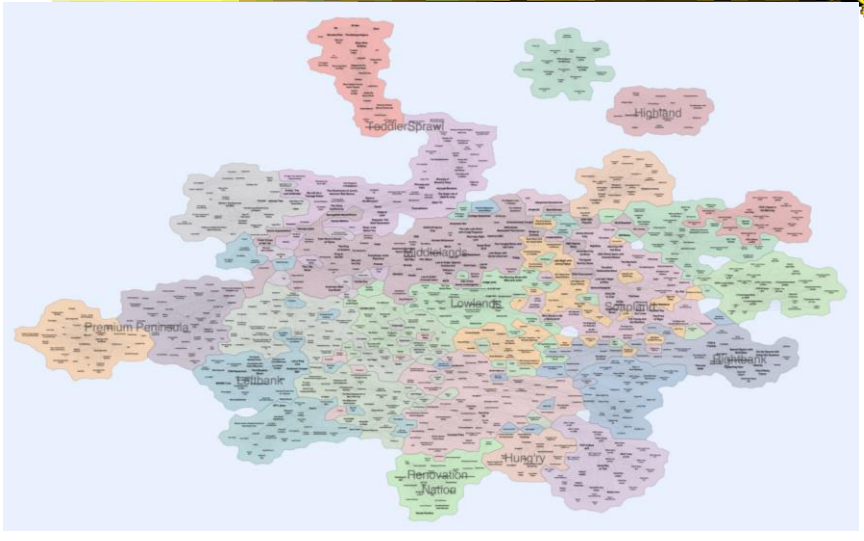
Both Representations

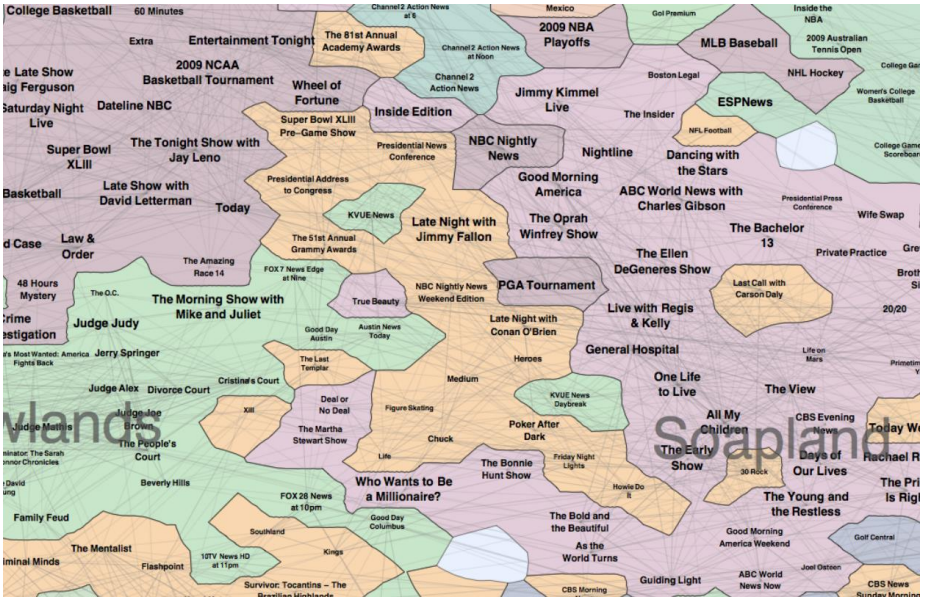


Node-link Map

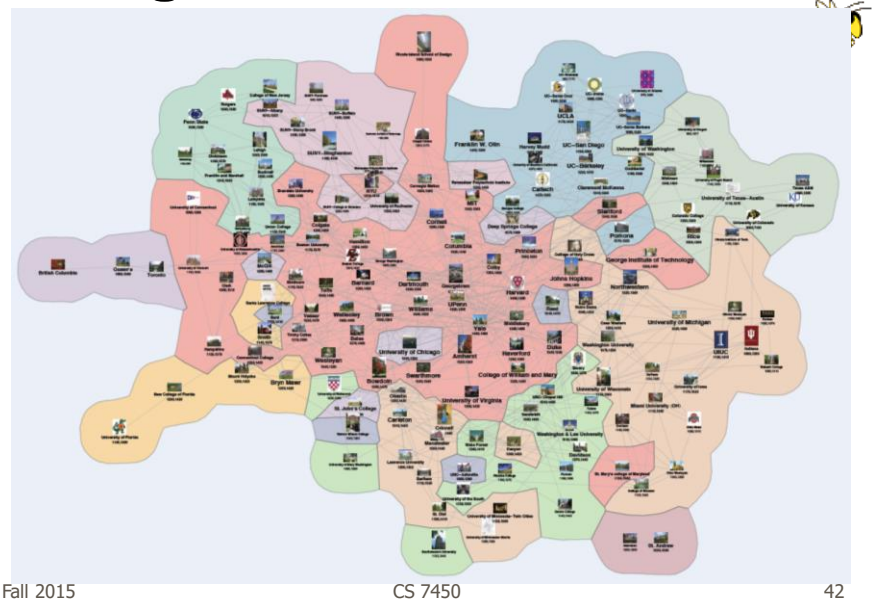
Music Graph/Map

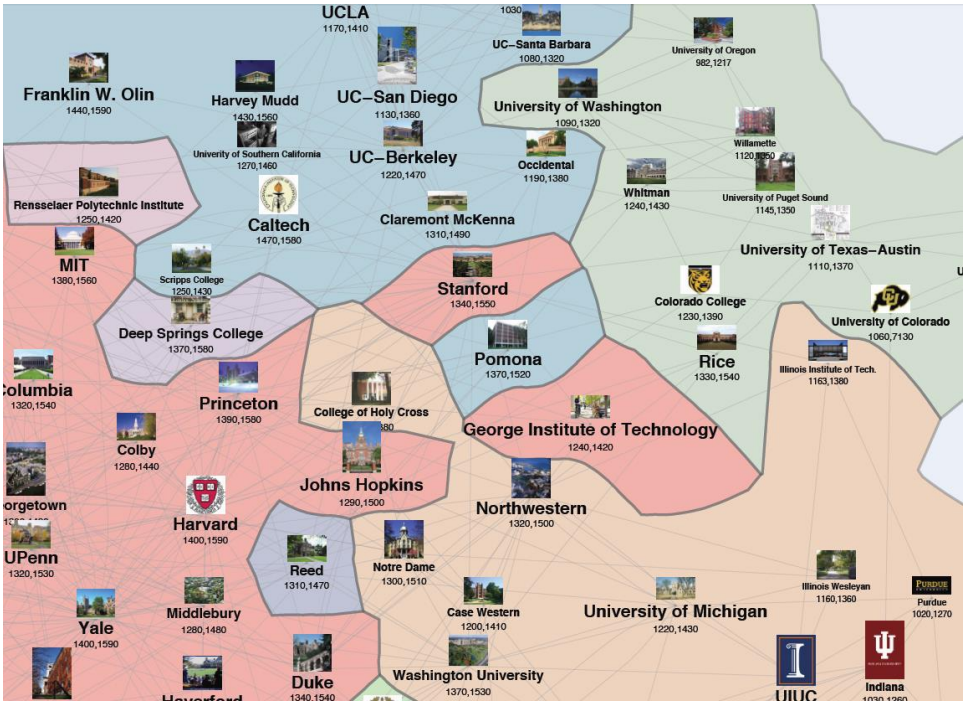
TV Shows




College Basketball
60 Minutes
Entertainment Tonight
The 81st Annual Academy Awards
2009 NBA Playoffs
MLB Baseball
2009 Australian Tennis Open
2009 NCAA Basketball Tournament
Wheel of Fortune
Jimmy Kimmel Live
NHL Hockey
ESPNews
Dateline NBC
Inside Edition
The Insider
Super Bowl XLIII
The Tonight Show with Jay Leno
NBC Nightly News
Nightline
Dancing with the Stars
Late Show with David Letterman
Today
Good Morning America
ABC World News with Charles Gibson
Law & Order
Late Night with Jimmy Fallon
The Oprah Winfrey Show
The Bachelor 13
The Ellen DeGeneres Show
The Morning Show with Mike and Juliet
PGA Tournament
Judge Judy
Live with Regis & Kelly
Jerry Springer
General Hospital
Judge Alex
Divorce Court
Cristina's Court
One Life to Live
The View
The People's Court
Deal or No Deal
The Martha Stewart Show
Poker After Dark
All My Children
The Early Show
Days of Our Lives
Beverly Hills
Who Wants to Be a Millionaire?
The Bonnie Hunt Show
The Young and the Restless
FOX 26 News at 10pm
Family Feud
The Bold and the Beautiful
As the World Turns
Good Morning America Weekend
The Mentalist
Flashpoint
Guiding Light
ABC World News Now
Survivor: Tocantins – The

Colleges





Drawing Graphs Better

- Can we do clever "tricks" to make dense graphs more readable?

Hierarchical Edge Bundles

- Bundle edges that go from/to similar nodes together
  - Like wires in a house
- Uses B-spline curves for edges
- Reduces the clutter from many edges

Holten
TVCG (InfoVis) '06

Example

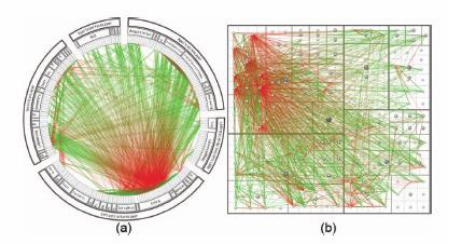


Fig. 11. A software system and its associated call graph (caller = green, callee = red). (a) and (b) show the system without bundling using a radial and a squarified treemap layout (node labels disabled), respectively. (a) and (b) mainly show hot spots; the actual connectivity information is more difficult to discern due to visual clutter.

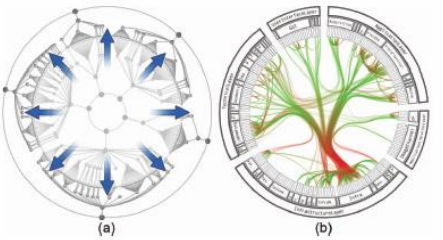


Fig. 12. Radial layout construction. (a) A radial tree layout is used for the inner circle and subsequently mirrored to the outside; (b) the inner layout is hidden and its structure is used to guide the adjacency edges. An icicle plot based on the mirrored layout is used to show the hierarchy.

Example

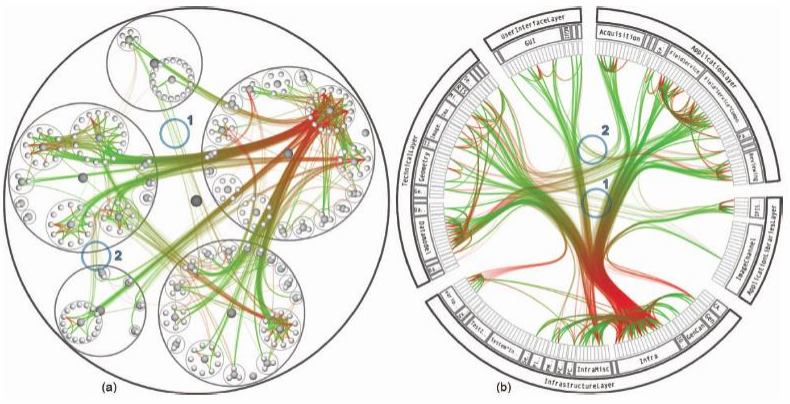
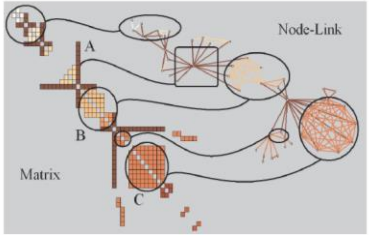


Fig. 13. A software system and its associated call graph (caller = green, callee = red). (a) and (b) show the system with bundling strength $\beta = 0.85$ using a balloon layout (node labels disabled) and a radial layout, respectively. Bundling reduces visual clutter, making it easier to perceive the actual connections than when compared to the non-bundled versions (figures 2a and 11a). Bundled visualizations also show relations between sparsely connected systems more clearly (encircled regions); these are almost completely obscured in the non-bundled versions. The encircled regions highlight identical parts of the system for (a), (b), and figure 15.

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



MatrixExplorer

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

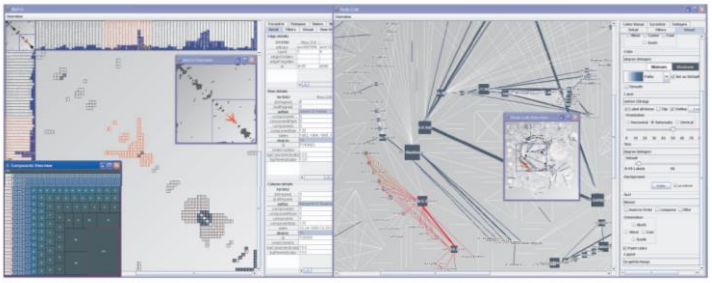


Fig. 1. MatrixExplorer showing two synchronized representations of the same network: matrix on the left and node-link on the right.

Node Reordering

Extremely important operation with matrix representations

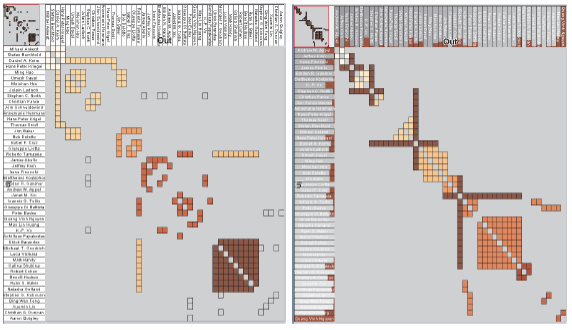
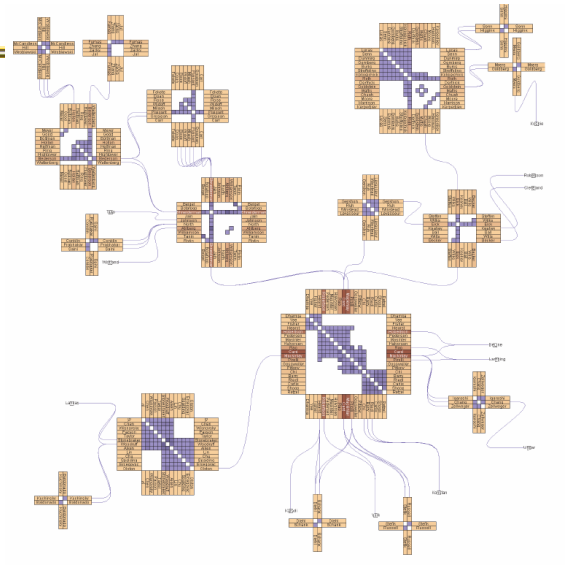


Fig. 6. Initial order (left) and TSP order (right). Colors represent clusters found by the user. Clusters are different in the two representations. Users found more clusters with TSP order. Headers red indicators (right) represents the distance between adjacent rows/columns.

NodeTrix

Hybrid of matrix and node-link



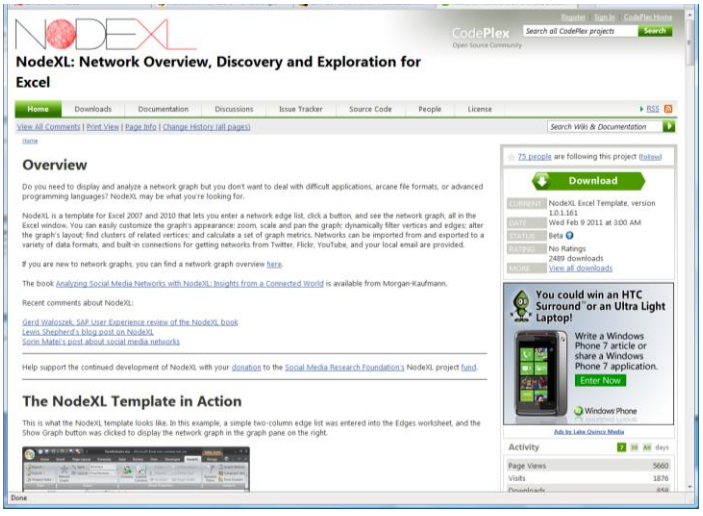
Henry & Fekete
TVCG (InfoVis) '07

Simplifying Input

- Make it easier to input graphs and then explore them

NodeXL

http://nodexl.codeplex.com/



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

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, …

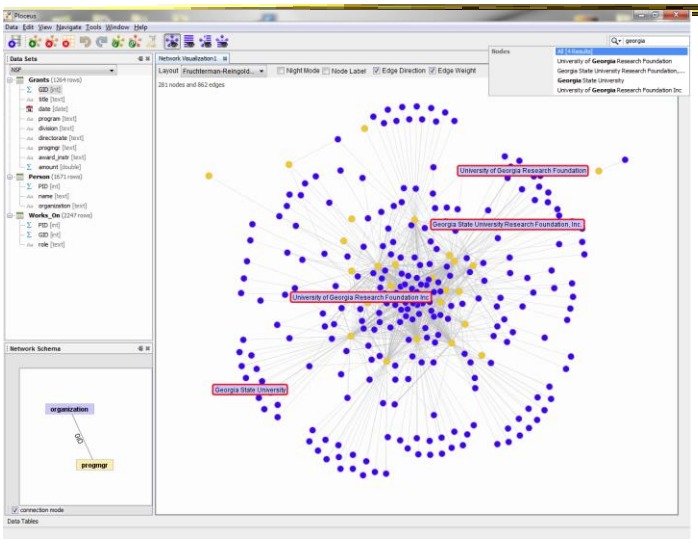
Ploceus

Liu, Navathe, Stasko
VAST '11, *Information Visualization* '14

- 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

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

Ploceus



Video

http://marvl.infotech.monash.edu/webcola/

cola.js

Graph Visualization Resource

- Very nice overview & survey
  - Herman et al, *IEEE TVCG* '00
  - but a little dated now

HW 5

- Examples
- Grading rubrics

Informal HW

- Download Jigsaw (and Java 8 if needed) onto your laptop and bring to class on Monday
  - Will email URL to use

Upcoming

- Visual Analytics
  - Reading
    - Keim et al '08
    - Stasko et al '08
- Visual Perception
  - Reading
    - Stone '06