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
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
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
dotlink360

Connections between companies (supplier, partner, etc.)

Basole, et al
TVCG (InfoVis) ’13

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

Demo

Flavor Connection
(saw last time)
**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](http://prefuse.sourceforge.net))
- Oppose Shneiderman’s mantra. Instead: “Start with what you know, then grow.”

Heer & boyd  
InfoVis ’05

---

**Visualization**

![Visualization Image]
Combining Features

Colors: Gender

Halo: Search for “student”

Highlight: Friends of selection

Blobs: Communities

Immersion

https://immersion.media.mit.edu/
Google+ Ripples

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

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

Note

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”
Graphs as Maps

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

Both Representations

Node-link

Map

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

http://www2.research.att.com/~yifanhu/MAPS/imap.html
Music Graph/Map
TV Shows
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.
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.

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

- 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/
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

GLO-STIX

- GLO-STIX: Graph-Level Operations for Specifying Techniques and Interactive Exploration

Encapsulate operations into an editing tool
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., betweeness 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

Upcoming

• Time series data

• Visual Analytics