What’s in a graph?
Why Visualize Graphs?

- Quickly “get a sense” of a graph’s structure, something almost impossible from raw data.

- Surfaces patterns and structures that might be difficult to discover algorithmically (unless we know exactly what we are looking for).

- Many networks are naturally physical.
What questions might we ask?

- How does knowledge disseminate in online communities?
- What kinds of natural groupings (or communities) exist in co-author networks?
- How are two networks similar? How are they different?
- Which entities in a social network might be terrorists?
What questions might we ask?

- How does knowledge disseminate in online communities?
- What kinds of natural groupings (or communities) exist in co-author networks?
- How are two networks similar? How are they different?
- Which entities in a social network might be terrorists?
- Distributions, patterns, trends, gaps & outliers
The “ideal” graph vis toolkit

- Enable rapid exploration and hypothesis formulation
  - Extensive interactive features such as rollup, drill-down, layout adjustment, filters, selection
  - Fast calculation of statistics and metrics (centrality, clustering, distributions over features etc.)
  - Comparison of random variables with their expected values.
- Express confidence intervals over graph elements.
Putting things into Perspective
Putting things into Perspective
Putting things into Perspective
Size $\propto (\text{Expressible Context})^{-1}$

Source: http://mkweb.bcgsc.ca/linnet/
Size $\propto (\text{Expressible Context})^{-1}$
Problems with Scale

- **Pixel limitations** – there’s only so much we can fit on a screen, after all.
- **Storage limitations** – we have only so much storage space.
- **Processing limitations** – many algorithms $O(n^2)$ or worse.
- **Theoretical limitations** – some models and metrics lose their meaning as we shift focus from local to global
Node-link diagram limitations

- Nodes, edges are given either **equal** or **arbitrarily biased** representation.
Source: http://mkweb.bcgsc.ca/linnet/
Node-link diagram limitations

- Nodes, edges are given either equal or arbitrarily biased representation.

- Layout artifacts induce false positives and negatives in pattern detection.
Node-link diagram limitations

- Nodes, edges are given either *equal* or *arbitrarily biased* representation.
- Layout artifacts induce false positives and negatives in pattern detection.
- Missing data is not well represented.
Node-link diagram limitations

- Nodes, edges are given either equal or arbitrarily biased representation.
- Layout artifacts induce false positives and negatives in pattern detection.
- Missing data is not well represented.
- Incorporating multiple attributes is hard.
What do we do?!
Hybridize current methods
NodeTrix

Henry et al. ‘07
MatLink

(a) Node-Link (NL)  (b) Matrix (MAT)  (c) MatLink  (d) Zoom on MatLink

Henry & Fekete ‘10
Use domain knowledge to inform layout.
Internet Traffic
Language Genealogy
TimeNets

Kim, Card & Heer, '10
Time-dependent Interactions

Saraiya et al. ‘05
Semantic Substrates

Schneiderman & Aris, ‘06
Infer global properties through aggregation
Node Rollup

Node and Link Diagram

PivotGraph Roll-up
Edge Bundling

Holten ‘06
Communication density

Leskovec & Horwitz '08
Infer local properties through motif-finding and aggregation
Ego-network Motifs

Adamic et al. ‘08
Ego-network Motifs

Adamic et al. ‘08
Signed motif enumeration

Leskovec et al. ‘10
Motif Filtering

von Landesberger et al. ‘09
Incorporate Multiple Dimensions
PivotGraph

Wattenberg ‘06
Hive Plots

Kryzwinski, http://mkweb.bcgsc.ca/linnet/
Enable Graph Comparison
A Network Diff (Oesprey)
ManyNets

Freire et al. ‘10
GraphPrisms
GraphPrisms
In Sum

- There is no right way to visualize a graph!
- Always explore from multiple perspectives
- Exploit domain knowledge in layout if possible
- Global and local-scale (motif) aggregation often lends new insights
- Don’t assume something is significant without testing it statistically.
Discussants

- Andrea
- Sanjay
- Nikil