Data Analytics SET10109

Graphs and Time

Lecturer: David Hunter

Course Content: Natalie Kerracher

Overview

- What is a graph?
- Some features of graph data
- Different ways to visually represent graphs and trees: considerations, advantages, disadvantages
- Visual approaches for the temporal dimension

Reading

Required

- Munzner, T. (2014). *Visualization Analysis and Design*.
  Chapter 9 – Arrange Networks and Trees

Recommended

- Von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., van Wijk, J. J., Fekete, J.-D., & Fellner, D. W. (2011). Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges. *Computer Graphics Forum*, *30*(6), 1719–1749. doi:10.1111/j.1467-8659.2011.01898.x

- Aigner, W., Miksch, S., Muller, W., Schumann, H., & Tominski, C. (2007). Visualizing time-oriented data—A systematic view. *Computers & Graphics*, *31*(3), 401-409. doi:10.1016/j.cag.2007.01.030

What is a graph?

Graph G= (V,E)

- V = a set of vertices (nodes)
- E = a set of edges (links)

Useful terminology:

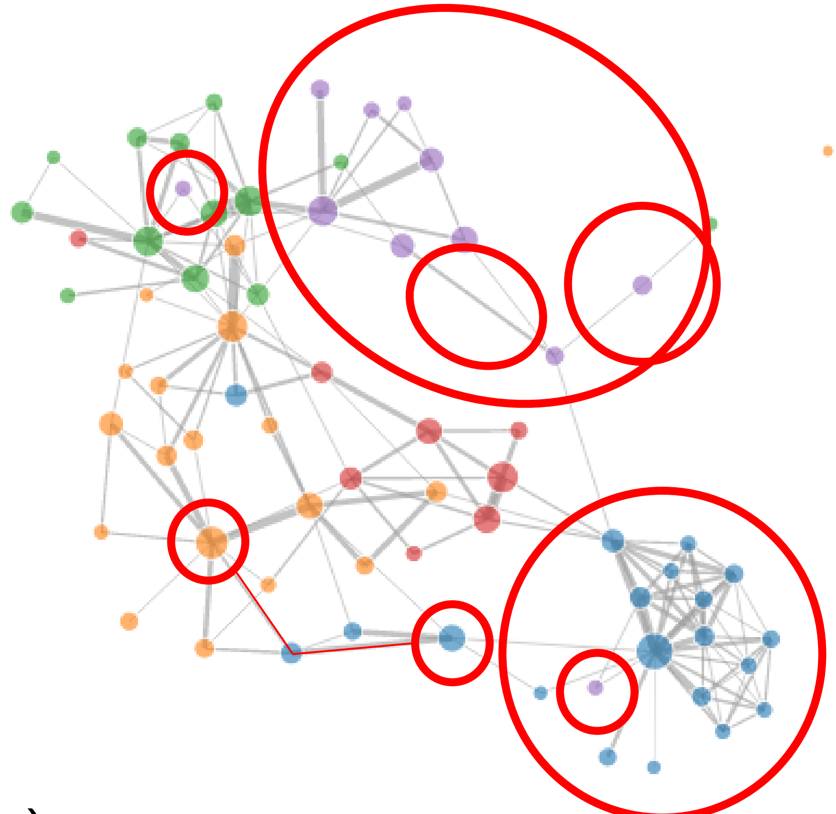
Clusters: based on connectivity

Groups: based on attributes

Paths: between two nodes

Attributes on nodes

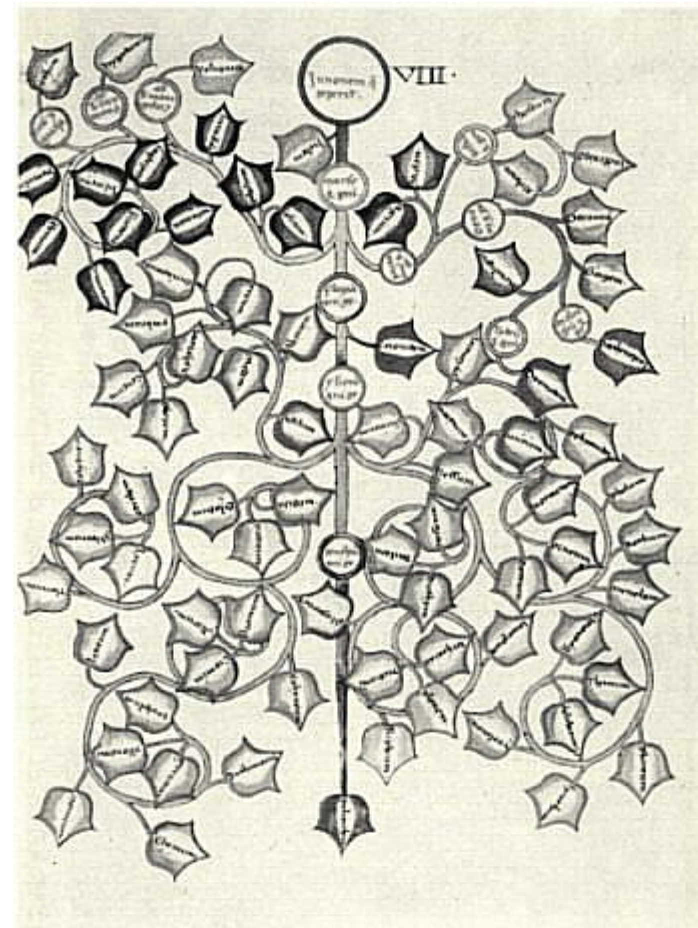
Weightings on edges (weighted graphs)



http://www.soc.napier.ac.uk/~cs22/test2/iidivis.html

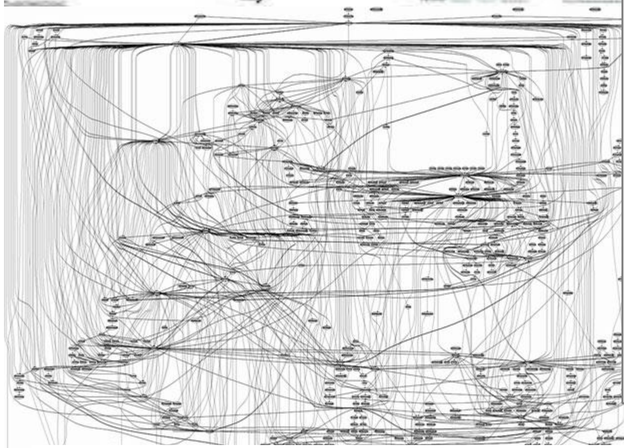
Graph Drawing: a Long History

- Graph drawing has been around for a *long* time ('Mill games', ancient Egypt)
- Euler (1707-1783) credited with originating graph theory in 1763
  - Did not make significant use of graph visualisation
- Wasn't until mid 1800s that graph drawings began to appear
  - E.g. illustrations in mathematical papers



Hand drawn family tree from the middle ages.
(Kruja et al, 2002, Figure 2)

Sources of graph data

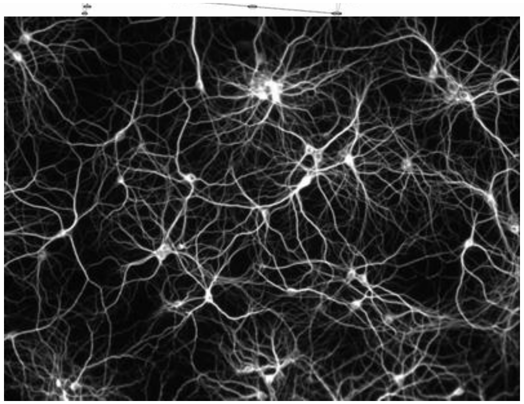


- Social networks
  - Organisational hierarchies
  - Family trees
  - Terrorist networks
  - Social media
- Computer networks
- Transportation networks
- Software systems
- Biological networks

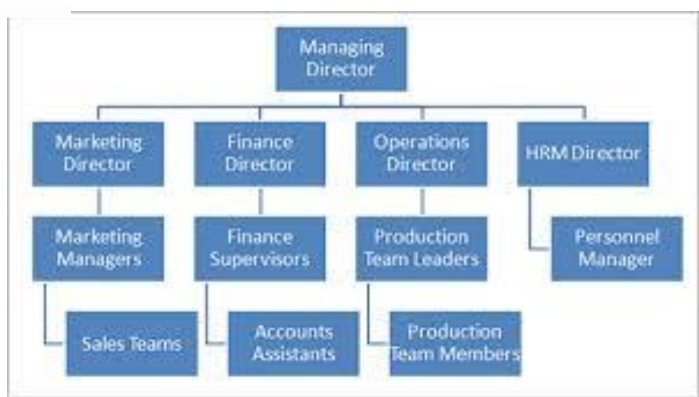
…

Paul Butler Facebook visualisation (2010)
Source http://on.fb.me/hy6dmb

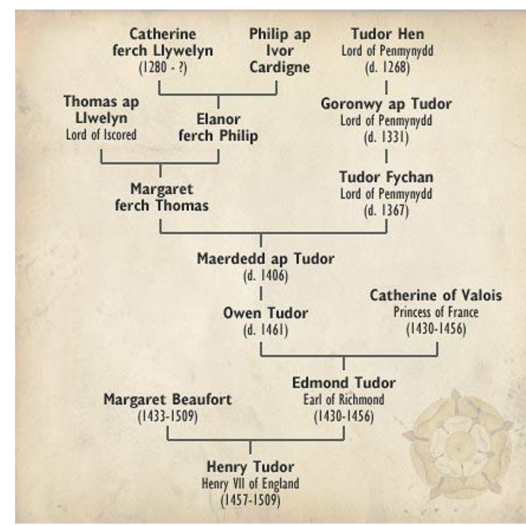
IIS call graph. Source: http://rixstep.com/1/1/20070206,00.shtml



Dissociated culture of rat hippocampal neurons
Paul De Koninck Laboratory - Universite Laval (2005)

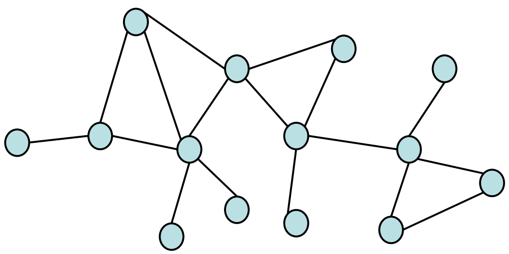


Organisation hierarchy source: www.tutor2u.net

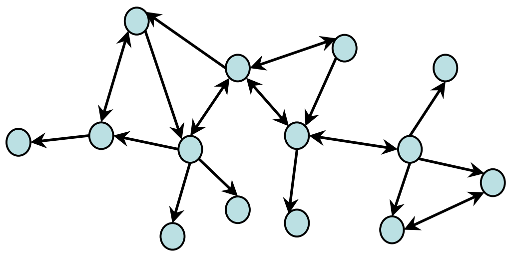


source http://www.bbc.co.uk/wales/history

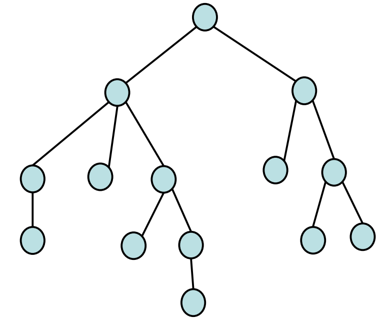
Graph structure: some considerations



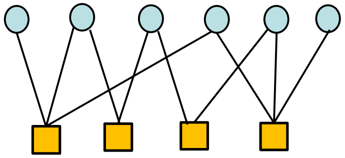
undirected graph



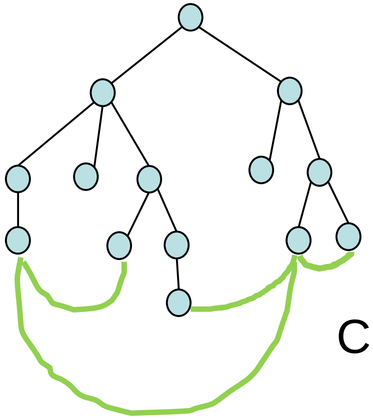
directed graph



Tree
(hierarchical structure)



bipartite graph



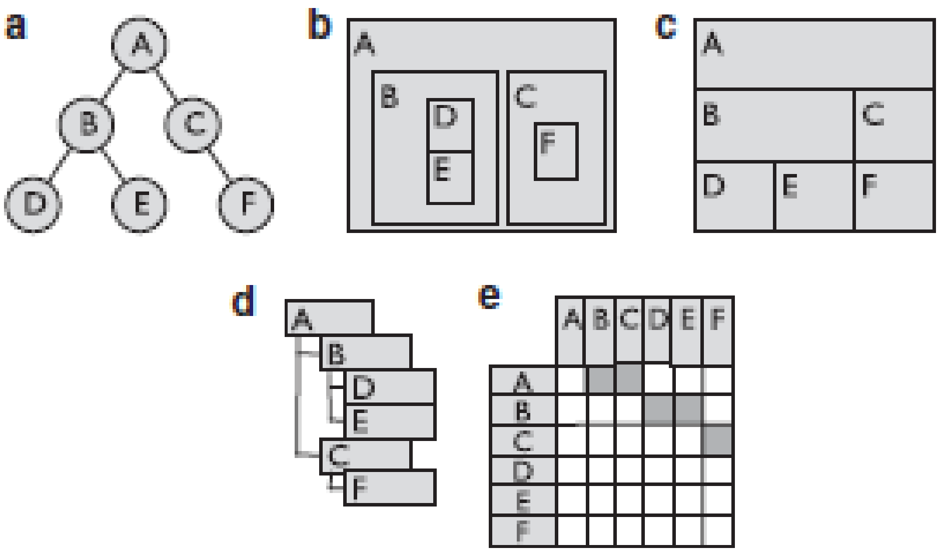
Compound graph

How can we visually represent graph data?

Problem: show connections, and thus relationships, clearly

Common relational layouts:

- Network (Graph) Data
  - Node-link diagrams
  - Matrix
- Tree / Hierarchy Data
  - Node-link diagrams
  - Adjacency layouts
  - Space-filling (treemaps, icicle plots)

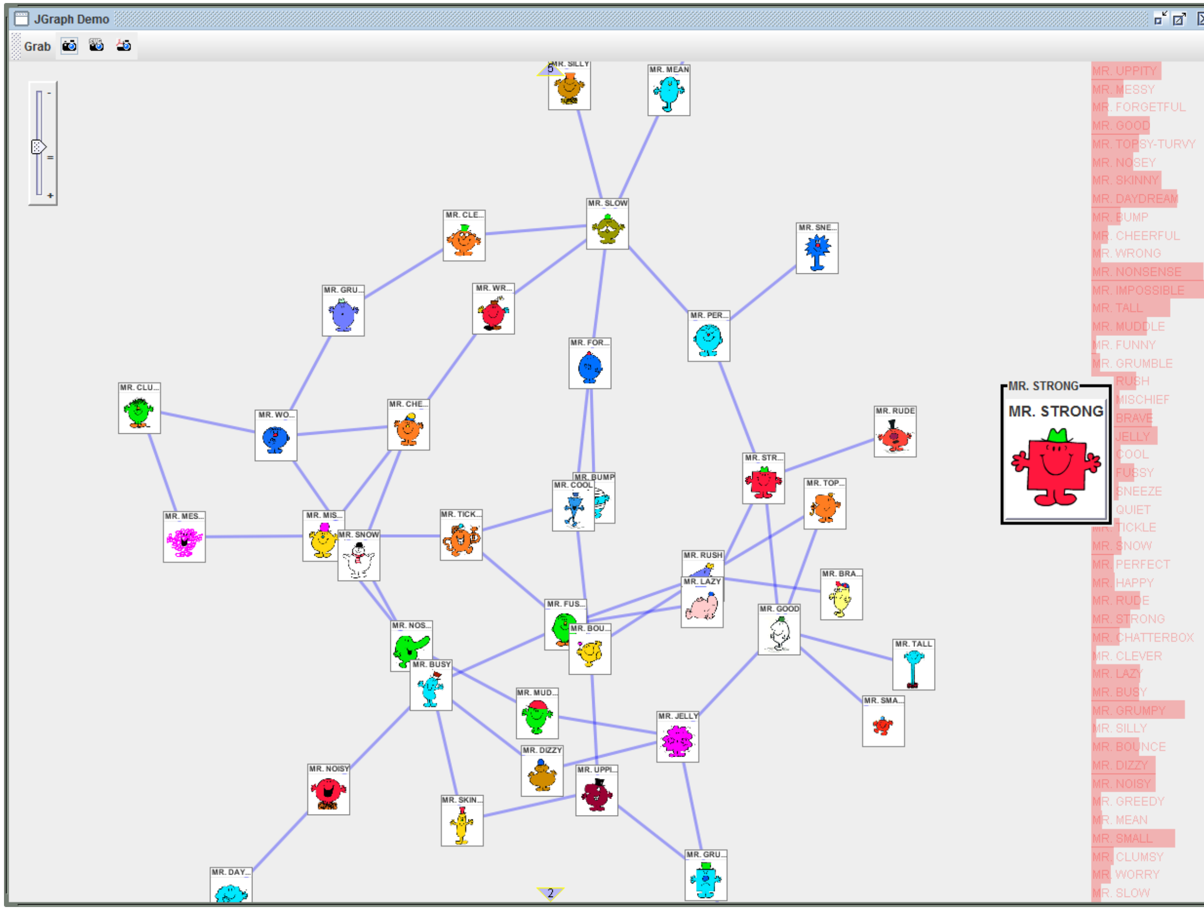


Graham & Kennedy (2010), Figure 3

- Useful resource: Manuel Lima's website http://www.visualcomplexity.com/vc/

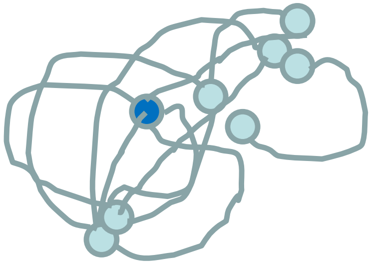
Network – Node-Link

- The most common form of showing a network is with a node-link diagram
- Nodes represent entities, edges represent the relationships between them
- Readily understood by viewers



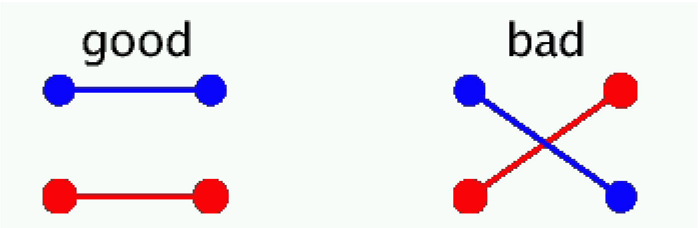
Graph Drawing Issues

Key Challenge: graph layout

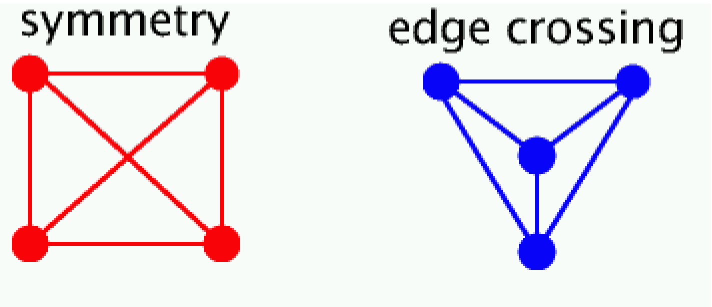


Optimise readability

Graph aesthetics:



Minimising edge crossings



Mutually incompatible criteria

From Munzner (2006)

Difficulties:

Some criteria mutually incompatible

Computation time and complexity (most criteria individually NP-hard)

Summary of aesthetic considerations

Node Metrics	Edge Metrics	Overall Layout Metrics
Cluster similar nodes	Minimize edge crossings	Maximize consistent flow direction
Distribute nodes evenly	Keep edge lengths uniform	Keep correct aspect ratio
	Minimize edge length (total and maximum)	Minimize area
Keep nodes apart from edges	Minimize edge bends	
Maximize node orthogonality		Maximize convex faces
Nodes should not overlap (except for nested nodes)	Keep edge bends uniform (angle/position)	Maximize global symmetry
	Maximize edge orthogonality	
	Maximize minimum edge angles	Maximize local symmetry

Source: Bennett et al. (2007)

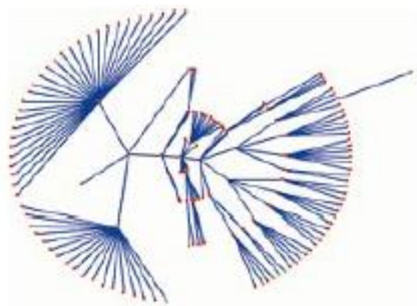
Layout algorithms

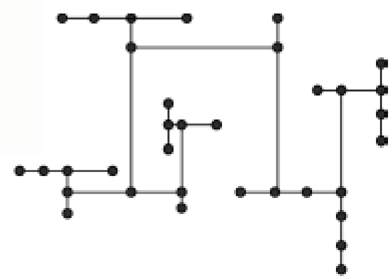
Edinburgh Napier UNIVERSITY

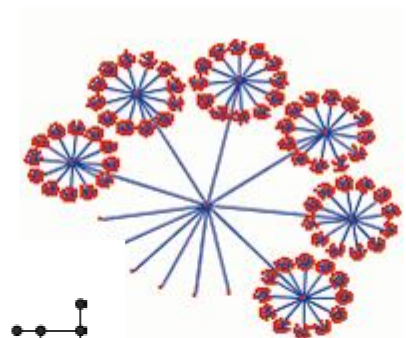
A key challenge for graph drawing is the computation of the lay out of nodes and edges

Categories of layout algorithms (classified according to type of node placement (von Landesberger et al. (2011)):

- Force directed
- Constraint-based
- Multiscale approaches
- Layered layouts
- Non-standard layouts







Images: Herman et al (2000)

Force directed algorithms are common: each node is pulled together into a final layout by its edges (see e.g. http://bl.ocks.org/mbostock/4062045)

See von Landesberger et al (2011) or Herman et al (2000) as a starting point if you're interested in layout algorithms…

Edinburgh Napier UNIVERSITY

Exercise: draw the following graph, trying to follow the principles of a "good layout"

9 nodes labelled 1-9

12 edges:

1-2	2-5	6-8
1-3	3-6	6-9
1-4	4-7	7-9
2-1	5-8	8-5

Exercise: draw the following graph, trying to follow the principles of a "good layout"

9 nodes labelled 1-9

12 edges:

1-2	2-5	6-8
1-3	3-6	6-9
1-4	4-7	7-9
2-1	5-8	8-5

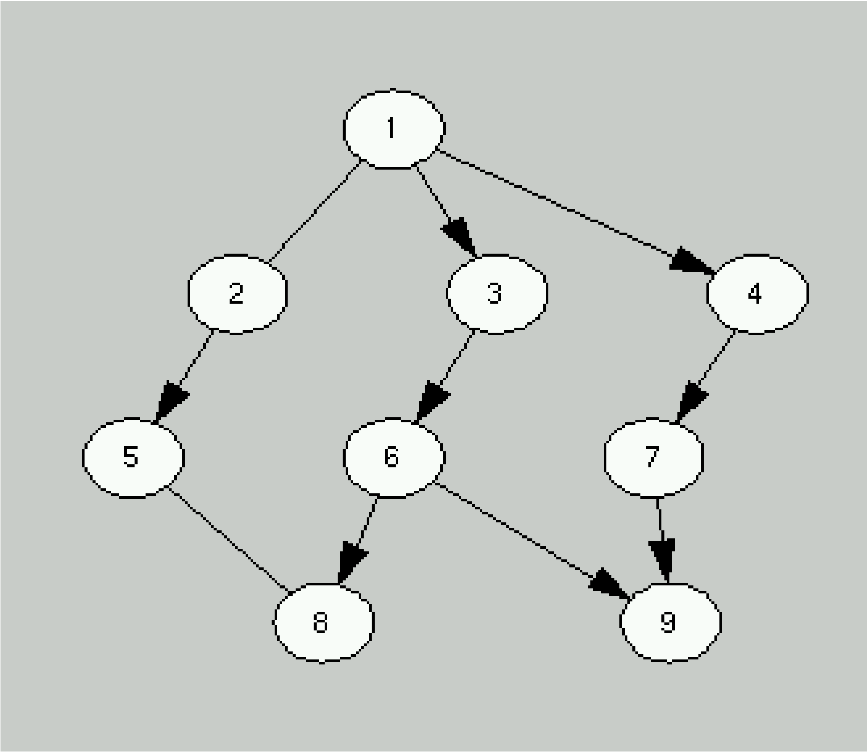
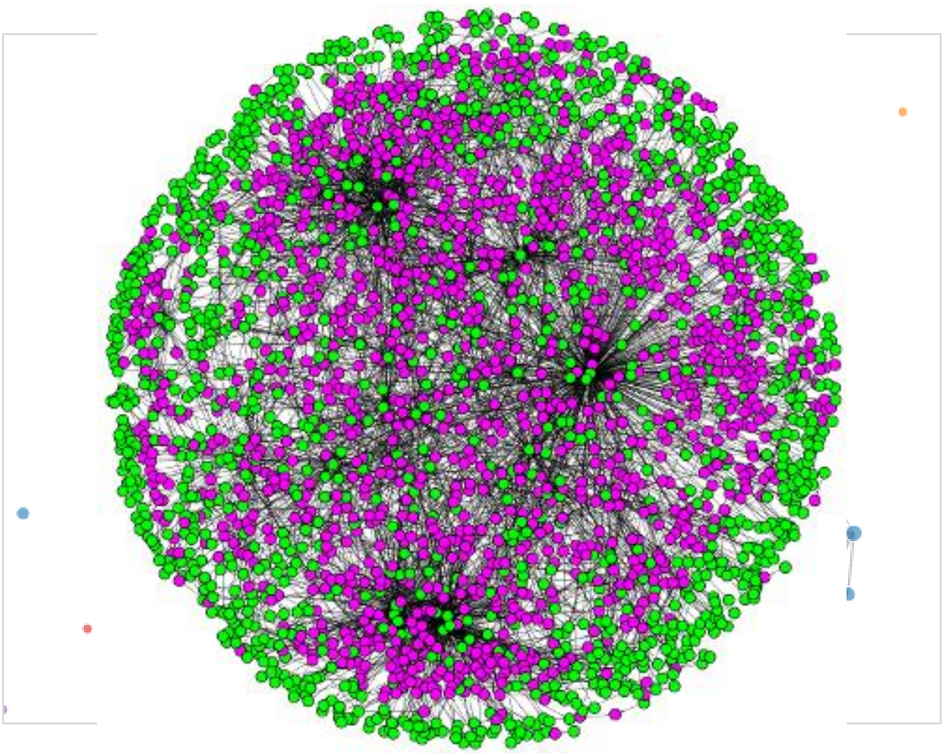


Image: www.cmpe.boun.edu.tr

Graph Drawing Issues: Scale



Node-link diagram
"Hairball"
http://eagereyes.org/techniques/graphs-hairball

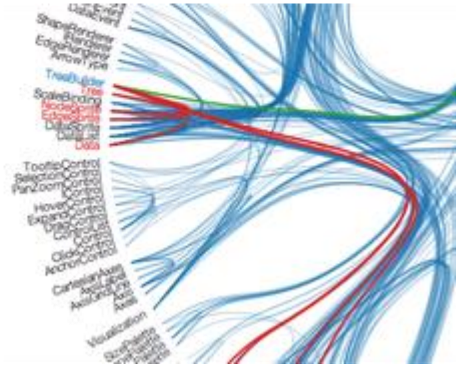
Problems of scale:

Computation
- Computationally expensive
- Time to render

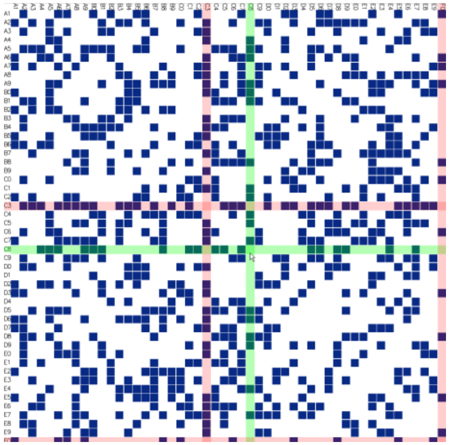
Readability:
- Leads to occlusion
- Makes interaction difficult
- Run out of screen space to draw nodes

Overcoming issues of scale

- Parallel processing on the GPU (for faster computation)
- Larger screen
- Reduction techniques
  - Edge bundling
  - Node bundling (clustering)
- Interaction techniques
- Use an alternative representation e.g. matrix



Edge bundling + interaction
http://mbostock.github.io/d3/talk/20111116/bundle.html

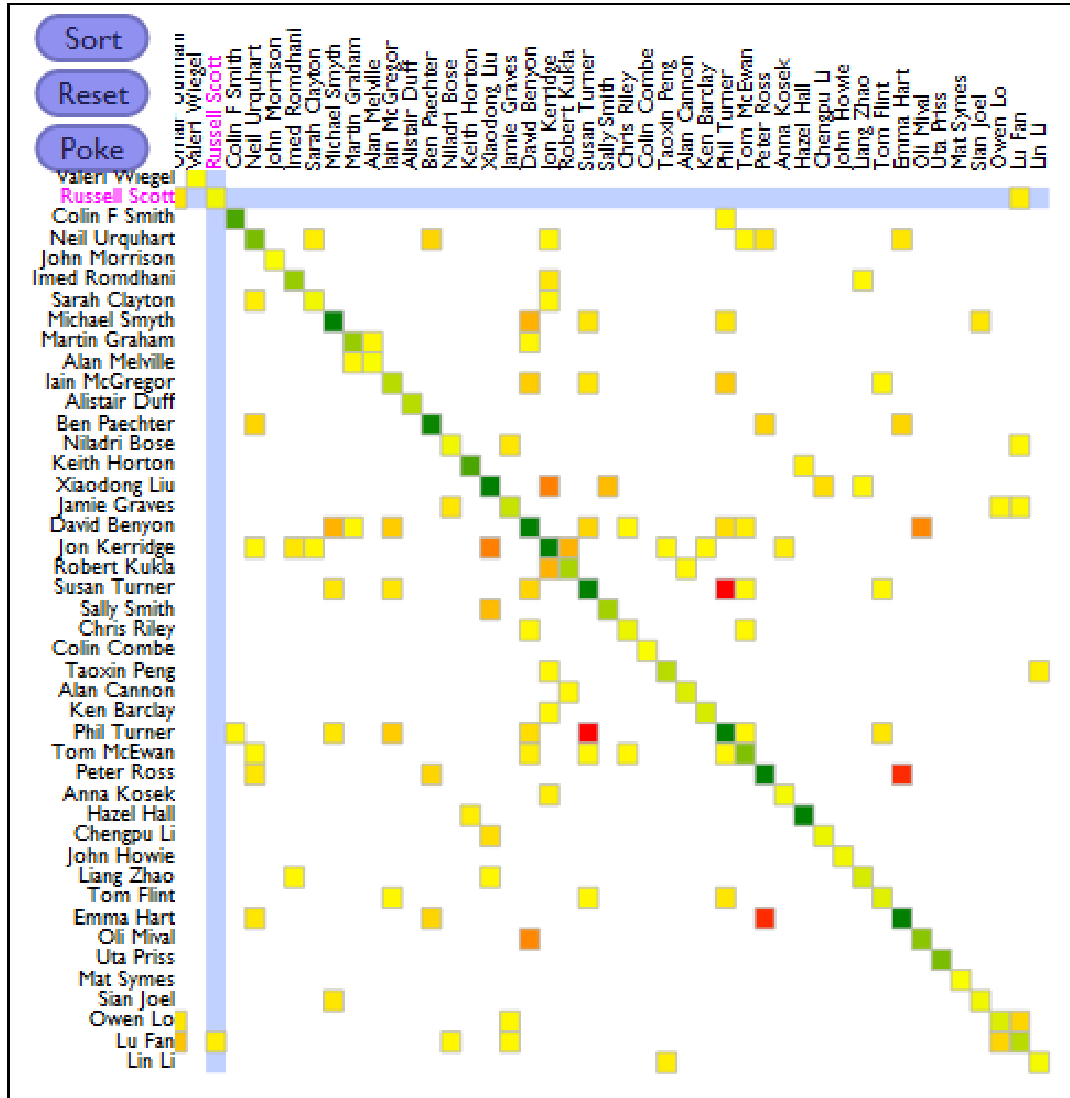


Matrix (Ghoniem et al. 2005)

Network – Matrix

Edinburgh Napier UNIVERSITY

- Entities (nodes) arranged on x/y axes
- Connections shown as marks between rows and columns



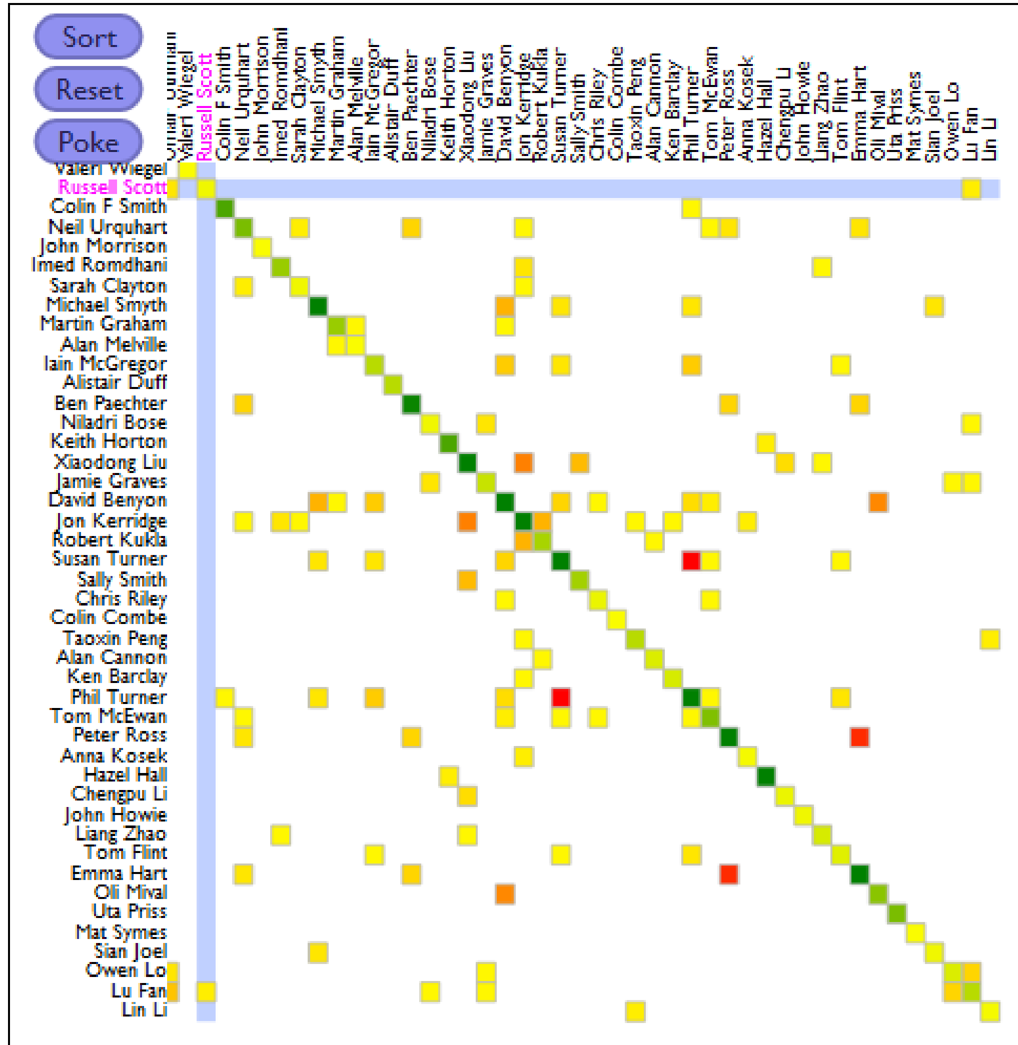
Network – Matrix

Advantages:

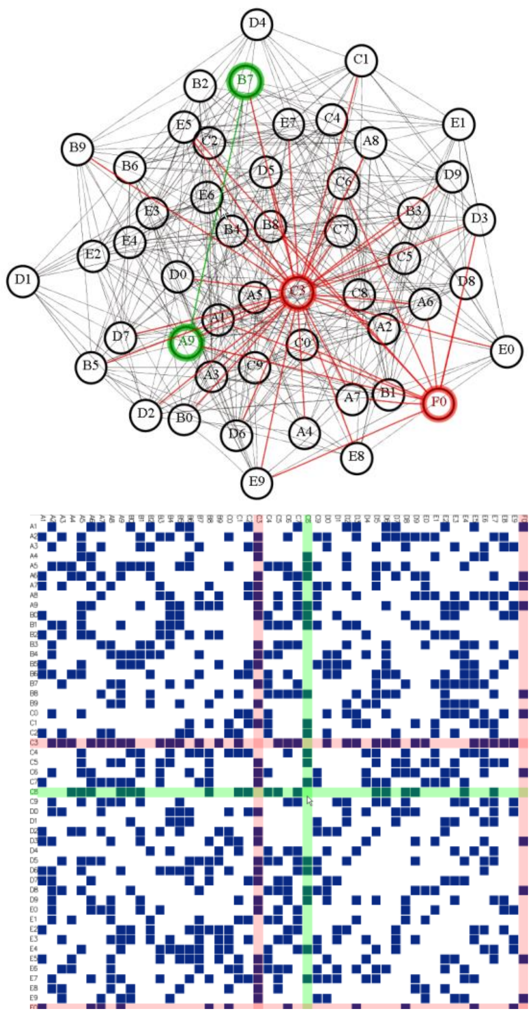
- Scale well
- No clutter, easy to calculate layout
- Outperform node-links on most tasks, where graph > 20 nodes (Ghoniem et al., 2005)

Disadvantages:

- Not intuitive
- Hard to follow paths
- Often sparse (space issues)
- Require ordering to show clusters



Node-link v Matrix summary



from Ghoniem, Fekete, & Castagliola (2005)

Node-link diagrams

\+ Intuitive

\+ Good for path-finding tasks

\+ Good for showing structure in sparse graphs

\- Do not scale well (readability, computation)

Matrix representation

\+ Ability to scale as no edges, so no path-crossing

\+ Outperform node links on most tasks

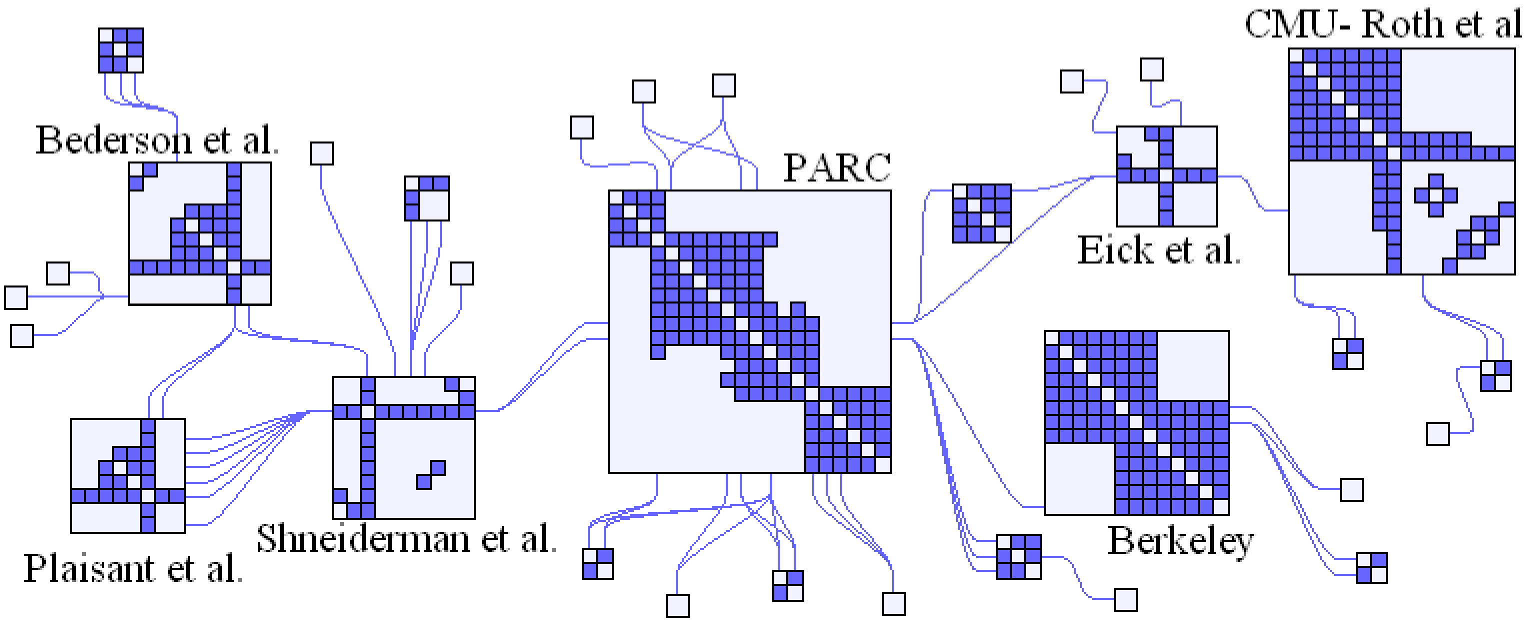
\- Less intuitive

\- Less good for sparse graphs

\- Not good for path finding tasks

Require ordering to show clusters

Node-link/Matrix Hybrid

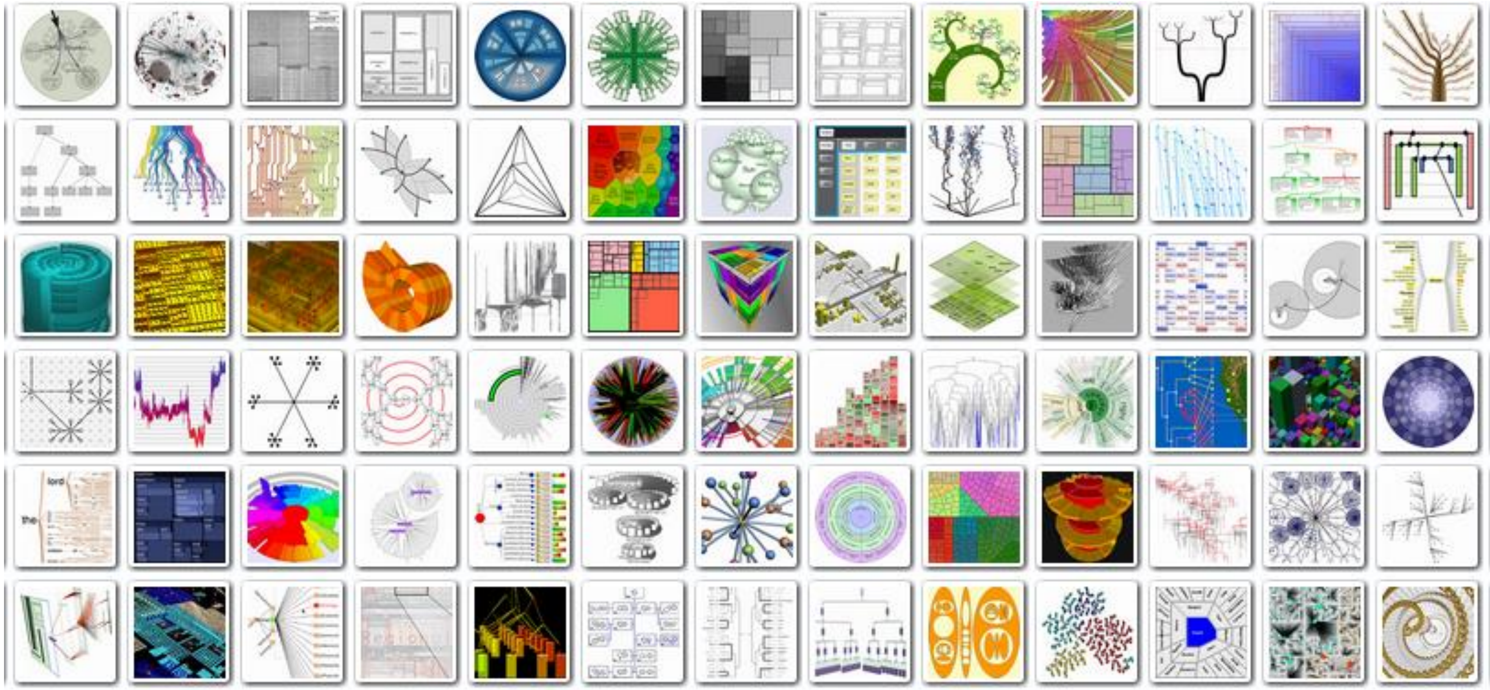


NodeTrix: Matrix/node-link hybrid (small world networks)
(Henry, Fekete, & McGuffin, 2007)

VISUAL REPRESENTATIONS OF HIERARCHICAL DATA (TREES)

Tree visualisation

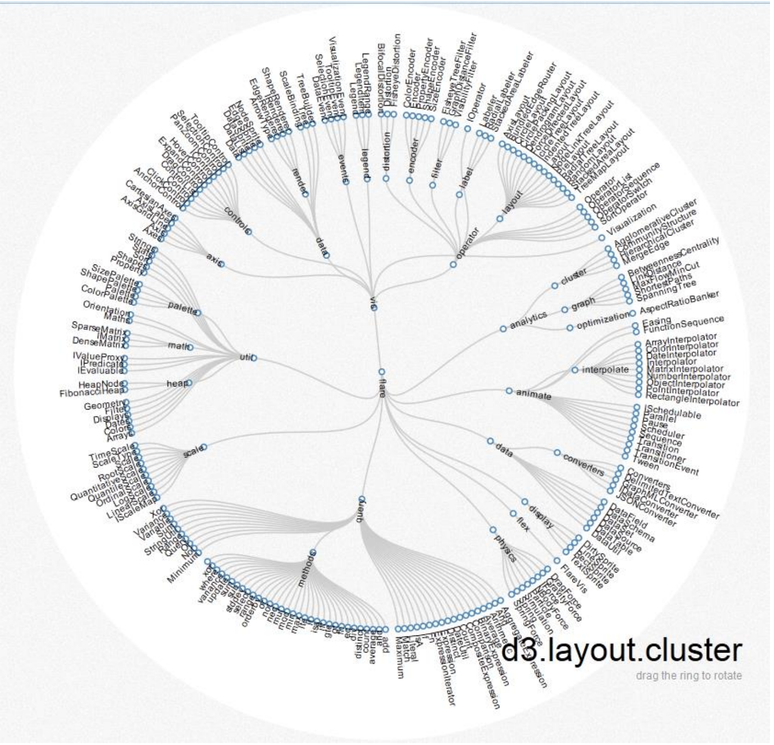
- Excellent resource: www.treevis.net
- Can search based on ‘three design axes’:
  - Dimensionality (2D, 3D, or hybrid)
  - Edge representation (explicit, implicit, or hybrid)
  - Node alignment (radial, axis-parallel, or free)

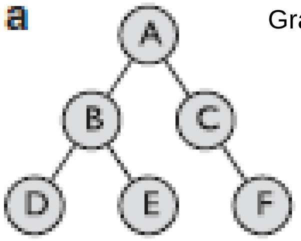


Screenshot: treevis.net

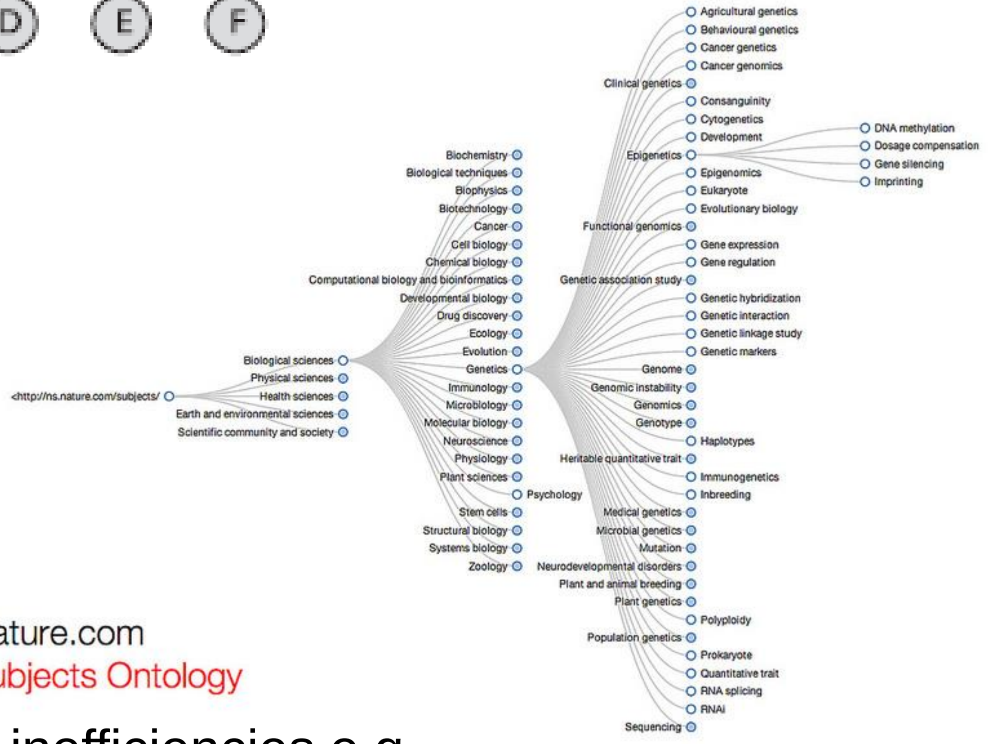
Tree – Node-Link

- Understandable but not very space efficient





Graham & Kennedy (2010)

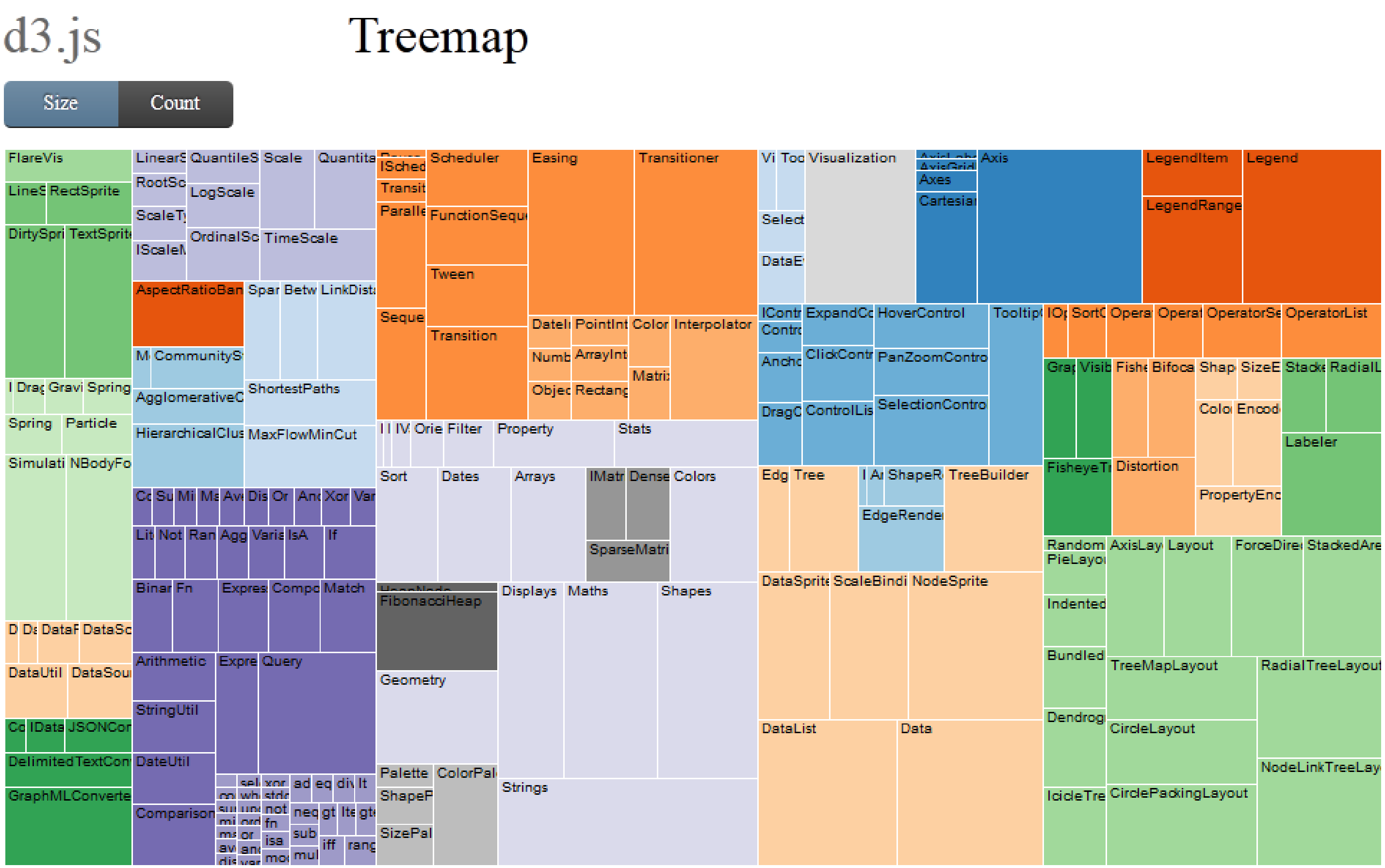


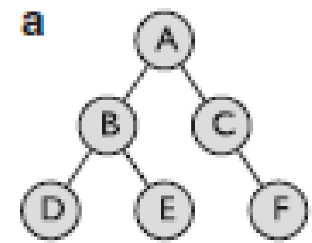
Nature.com
Subjects Ontology

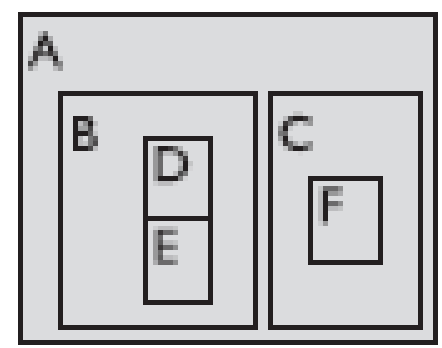
- Interaction to help overcome space inefficiencies e.g. http://bl.ocks.org/mbostock/4339083

TreeMap

- Enclosure encodes hierarchy
- Space-efficient but structure harder to interpret



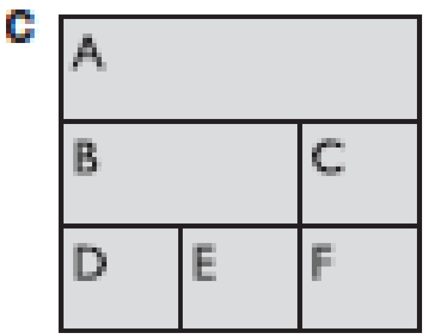




Graham & Kennedy (2010)

Tree – Adjacency layouts

- Halfway house between node-link and a treemap



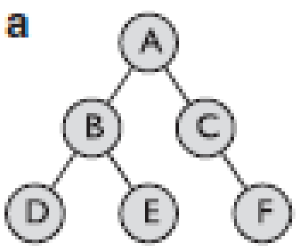
Graham & Kennedy (2010)

D3.js partition layout

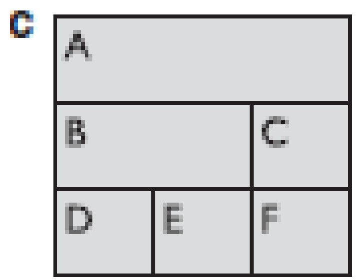
Tree layouts: summary

Edinburgh Napier UNIVERSITY

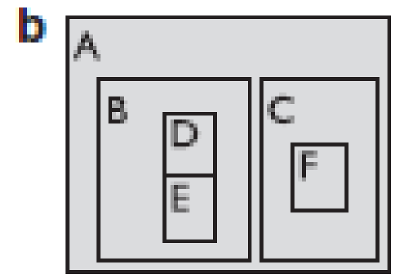
Node-link



Adjacency



Tree map



Graham & Kennedy (2010)

- More space
- More intuitive
- Structure easier to interpret

- Less space
- Less intuitive
- Structure harder to interpret
- Emphasises leaf nodes
- Cannot encode edge attributes

Graph vis tools

Easy to use graph visualisation tools:

- NodeXL (http://nodexl.codeplex.com/)
- Gephi (https://gephi.org/)

TIME

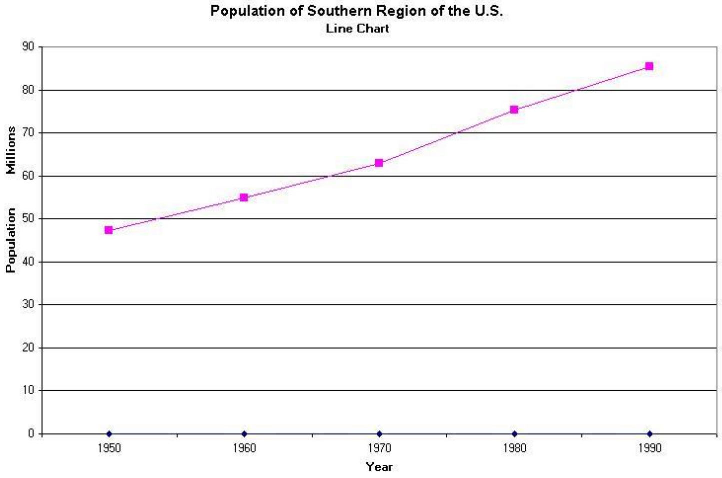
Visualising temporal data

- Useful resources: www.timeviz.net (survey) and HCIL's summary of temporal visualisation projects http://www.cs.umd.edu/hcil/temporalviz/

Image: www.timeviz.net

Visualising temporal data

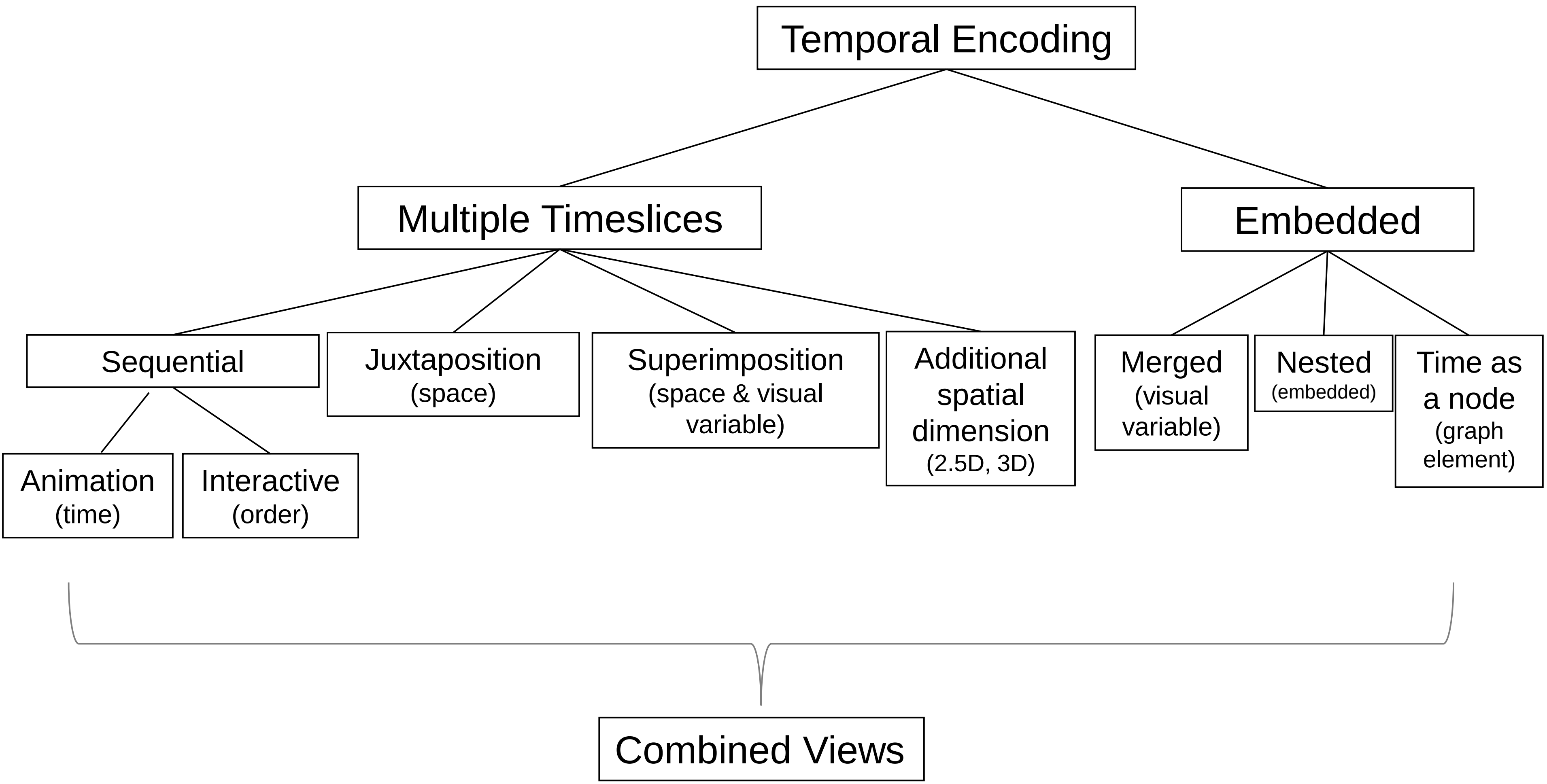
- Time is special – we can’t just treat it like any other dimension
- (unlike spatial data) it has no inherent mapping in space
- For tabular data over time, we can use e.g. line charts



Line chart image ciese.org

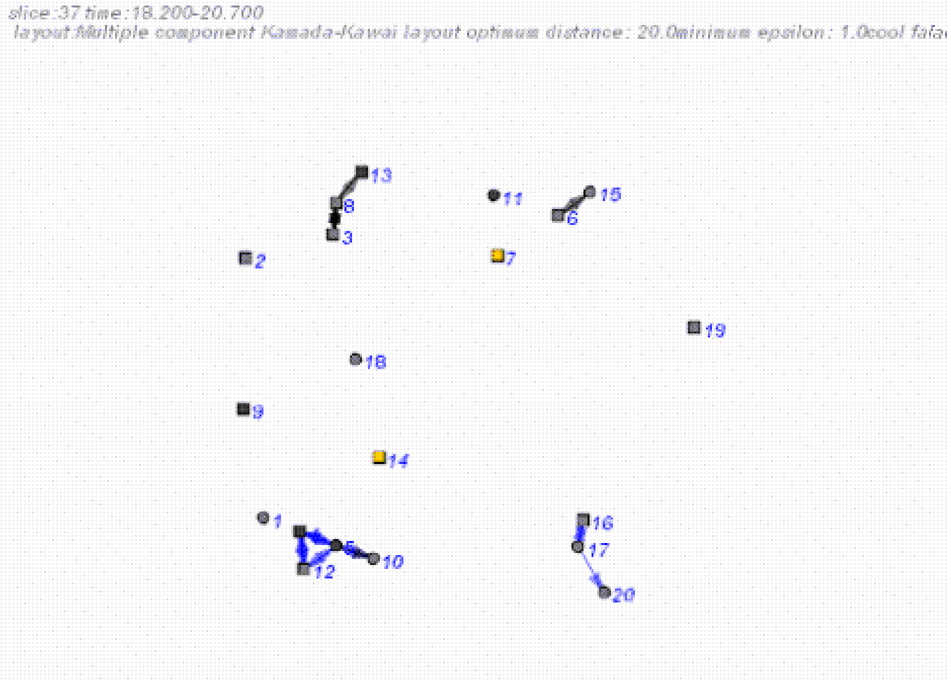
- Problem: relational, spatial, and multidimensional data over time
  - we use up two dimensions laying out the graph or showing the map

Encoding the temporal dimension: approaches



TIMESLICE APPROACHES

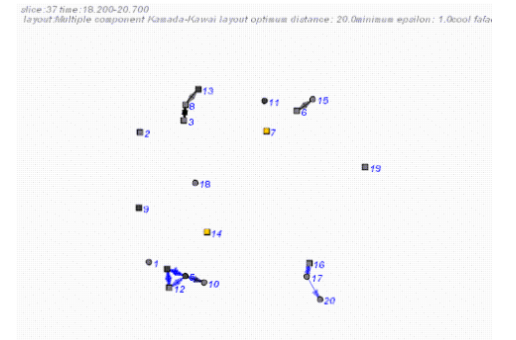
Sequential:
Mapping time to time (animation)



Bender-deMoll & Mcfarland’s (2006) animation of a classroom attention network, using SoNIA.

Available at http://www.cmu.edu/joss/content/articles/volume7/deMollMcFarland/images/cls33_10_16_96.gif

Sequential:
Mapping time to time (animation)



Advantages:

- An extra encoding channel
- People enjoy animation (even when it's not helping them)
- Reduces clutter – full screen space available
- More accurate for certain tasks (graph studies)

Disadvantages:

- Tasks take longer – need to view the whole sequence
- Cognitive overhead - comparison of timeslices need to be performed in memory
- Lack of interaction makes it difficult for the user to explore the data

Animation as an encoding channel

- Showing images sequentially to convey change over time

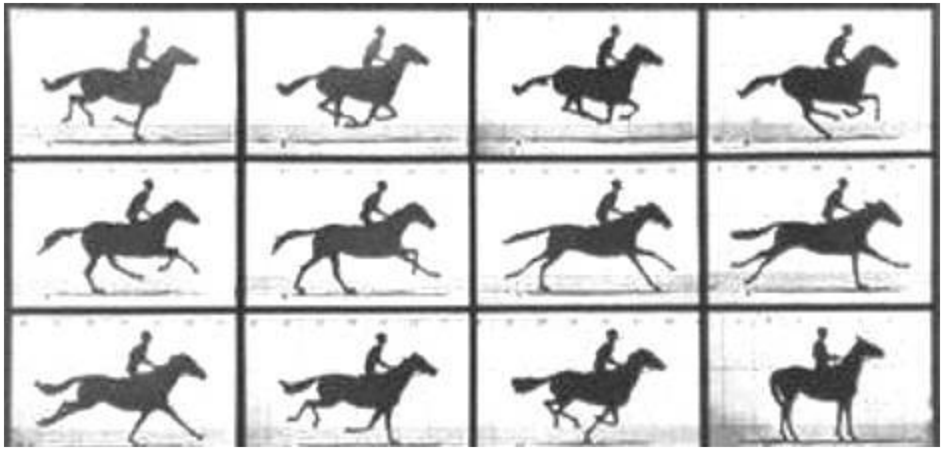
Animation

- Good:
  - Transitions (Heer & Robertson, 2007)
  - Simple comparison of adjacent items (Fekete, 2002)
  - (Narrated) storytelling (Robertson, Fernandez, Fisher, Lee, & Stasko, 2008) e.g. Gapminder
- Not so good:
  - In analysis e.g. making comparisons, exploring data (less accurate, slower)
  - When conveying complex systems (Tversky, Morrison, & Betrancourt, 2002)

Animation: hard to perceive accurately

- Perception of motion may not be accurate (Tversky et al. 2002)

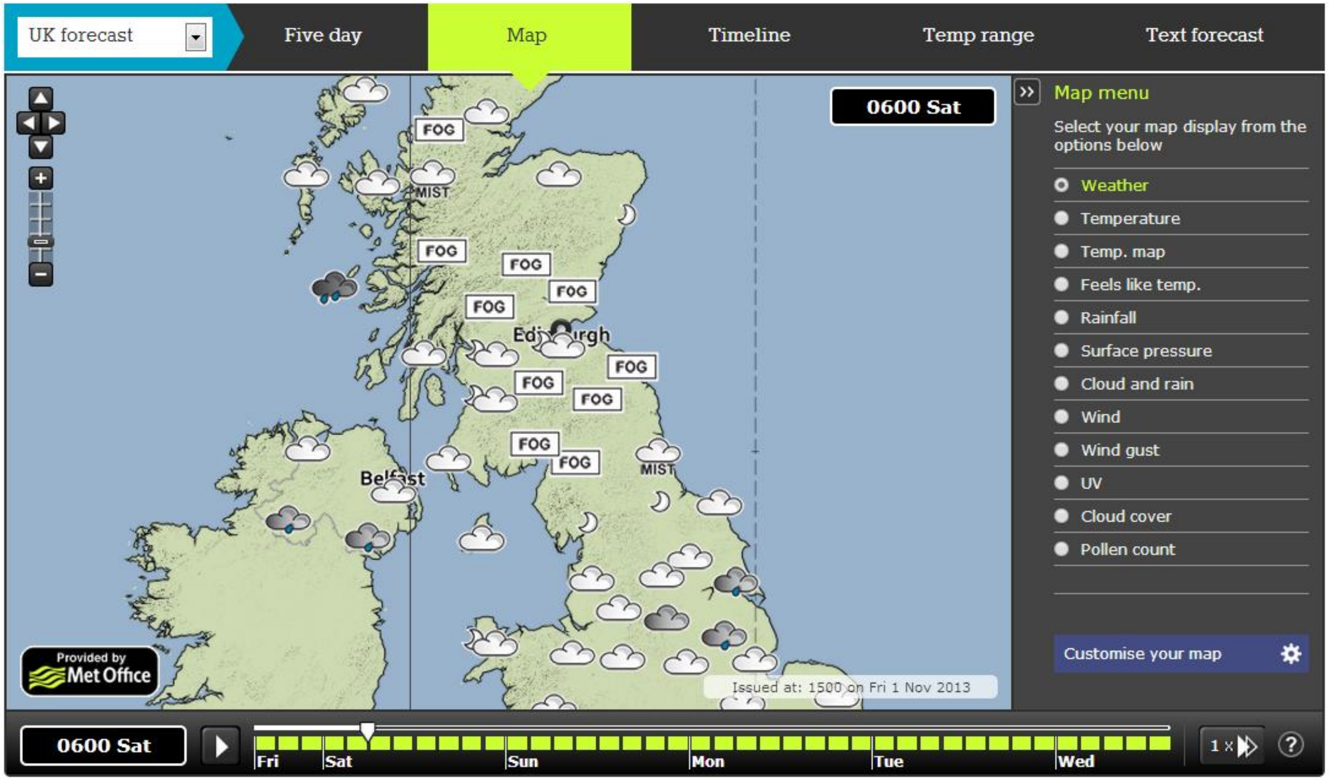
Stubbs (1724 - 1806)



Muybridge, 1878

Animation: cognitive burden

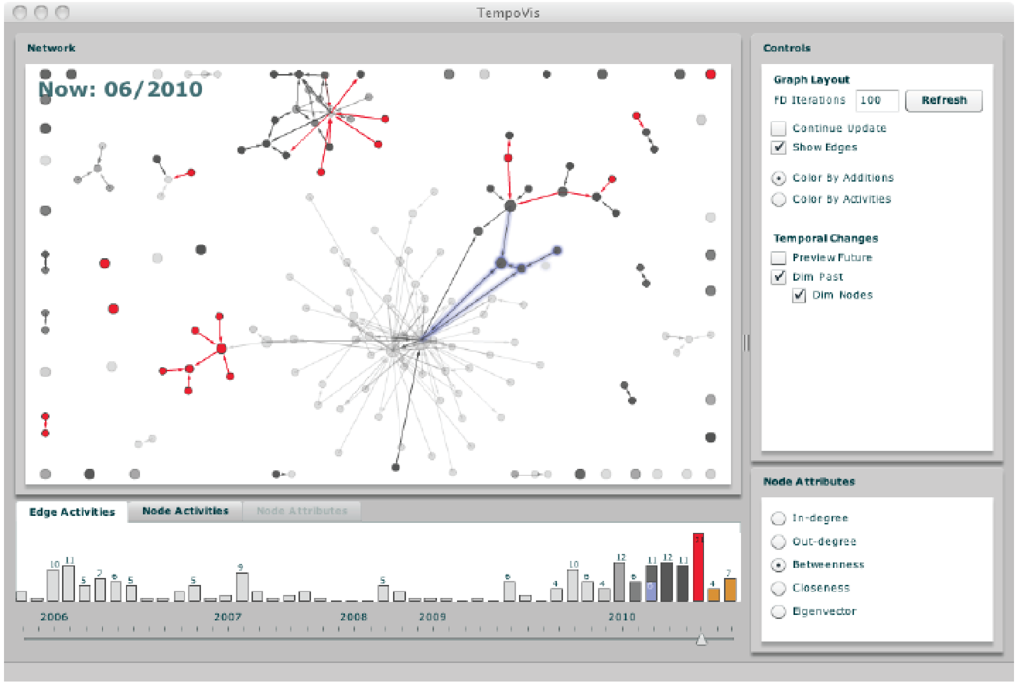
Test: which representation is easier animation or timeline?



http://www.metoffice.gov.uk/public/weather/forecast/map

Sequential:
Interactive approach (slideshow)

Edinburgh Napier
UNIVERSITY



Ahn et al.'s (2010) TempoVis prototype interface –use of timeslider interaction to navigate through different timepoints

Mapping time to space 1: Juxtaposition (small multiples)

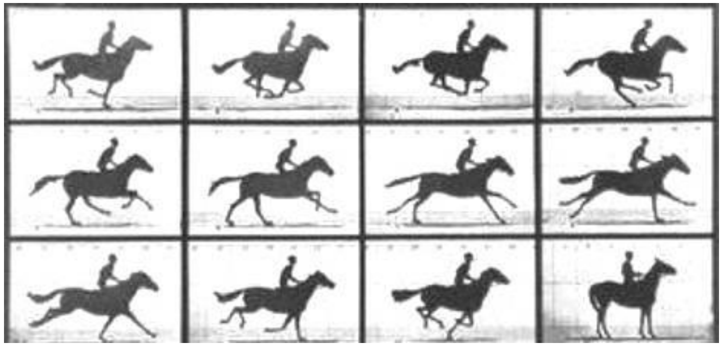
Edinburgh Napier UNIVERSITY

Sequential vs Juxtaposed views

Edinburgh Napier UNIVERSITY

- Showing images sequentially to convey change over time
  - Advantages:
    - Extra encoding channel
    - More screen space -> clutter reduction (Ellis & Dix, 2007)
    - More “fun”/“exciting” than static approaches (Robertson et al., 2008)
    - Can spot local changes between timeslices

  - Disadvantage:
    - Cognitive overhead – comparison performed in memory
    - Tasks take longer (Robertson et al., 2008; Archambault et al., 2011)
    - Less accurate; confusing (Robertson et al., 2008)

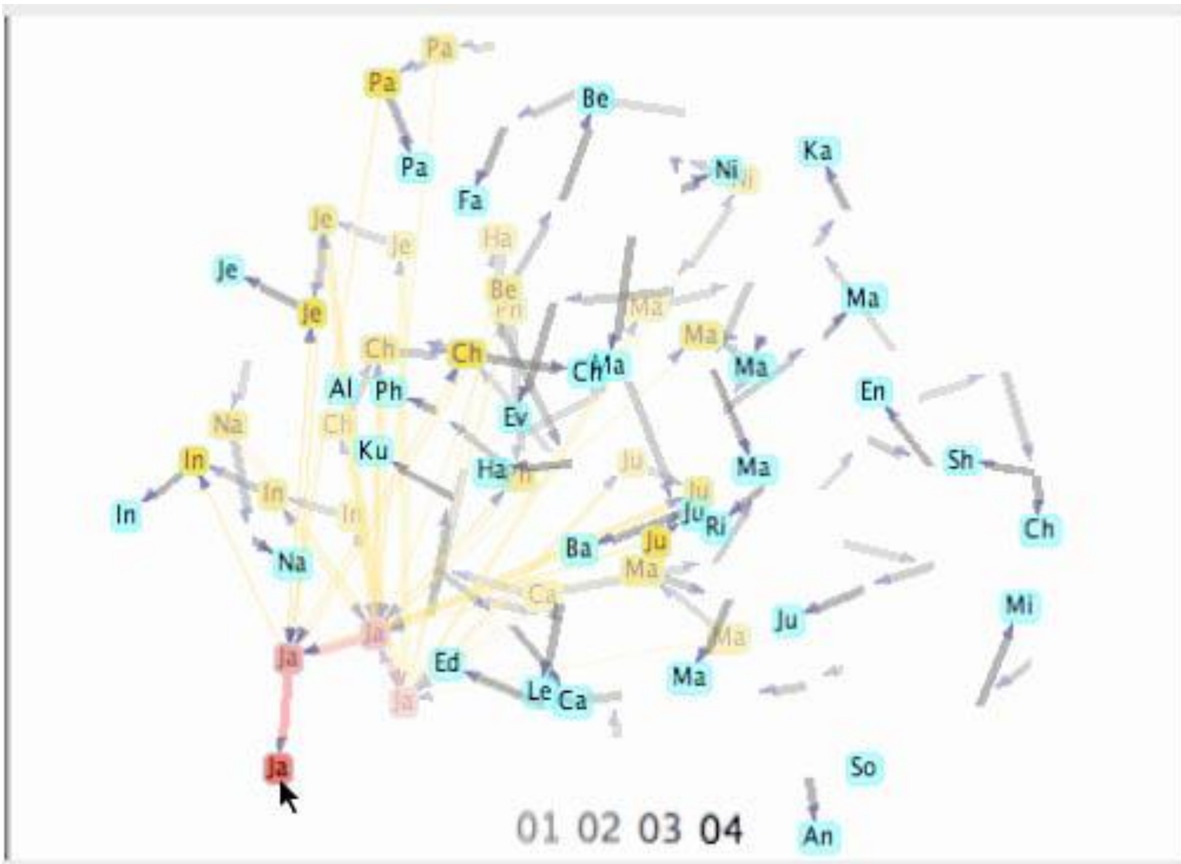


- Animation
  - Good:
    - Transitions (Heer & Robertson, 2007)
    - Simple comparison of adjacent items (Fekete, 2002)
    - Storytelling (Robertson et al., 2008)
  - Not so good:
    - In analysis, when conveying complex systems (Tversky, Morrison, & Betrancourt, 2002)

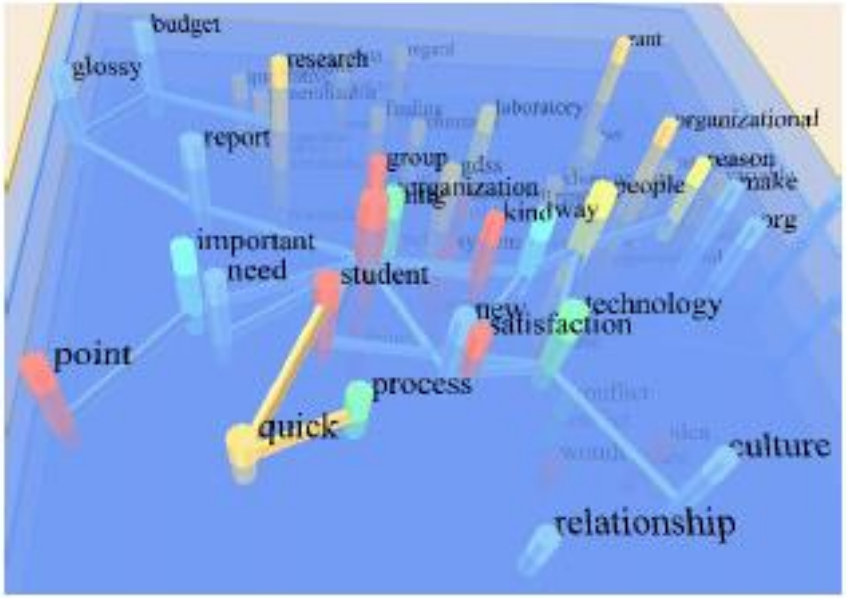
Mapping time to space & visual variable : Superimposition

- Four timeslices overlaid (each node appears multiple times)
- Opacity used to distinguish time (most recent = opaque, older = more transparent
- Red is highlight on mouse-over (shows single node).

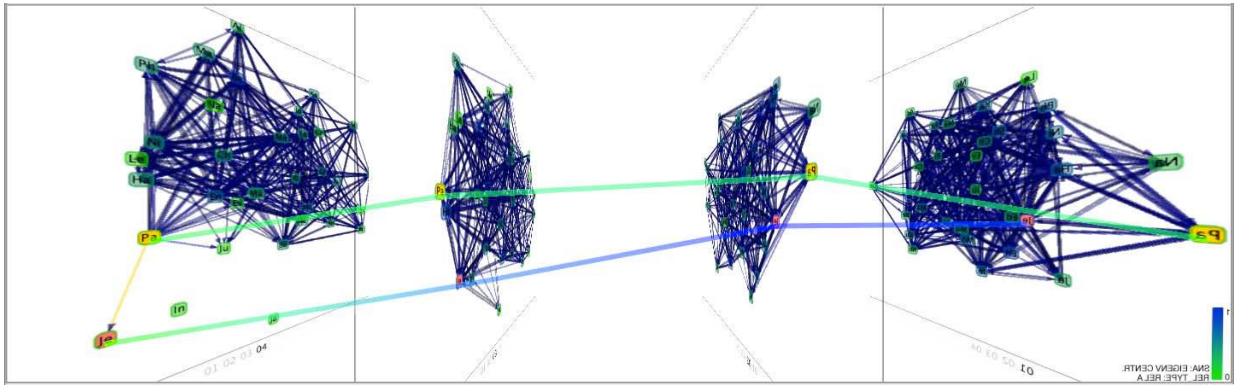
Federico et al (2011), fig 3.



Mapping time to an additional spatial dimension (2.5D, 3D)



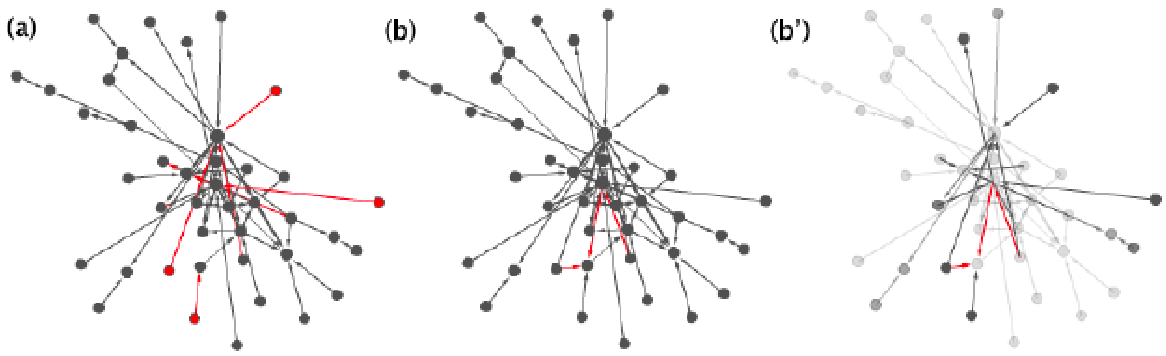
3D: Brandes and Corman (2003)



2.5D: Federico et al. (2011)

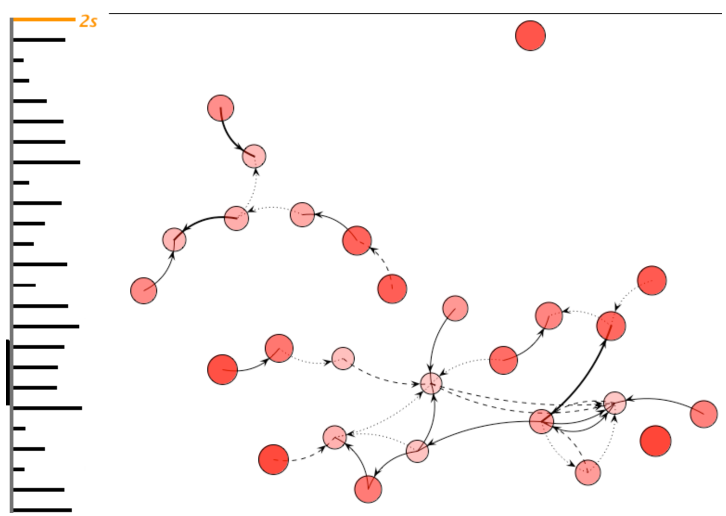
EMBEDDED APPROACHES

Mapping time to a visual variable: merged



In (b'), older activities painted in lower intensities to show their age

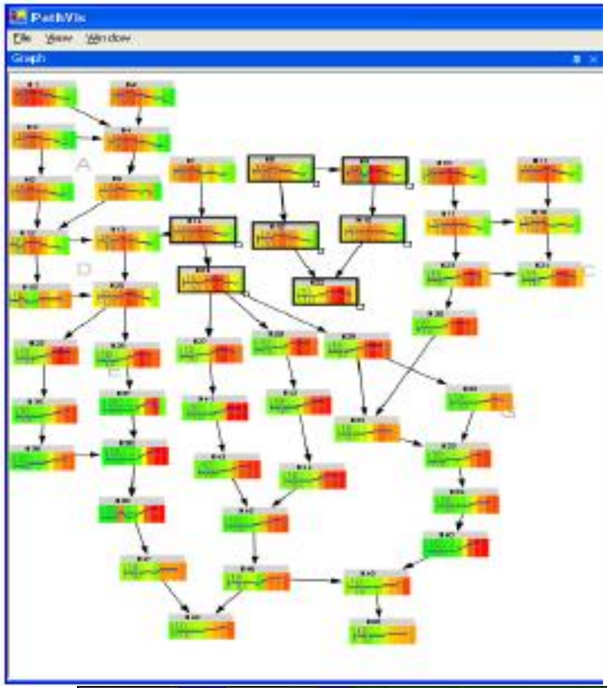
from Ahn, Taieb-maimon and Sopan (2010), fig 3, p6



Line length encodes the amount of time which has passed between changes

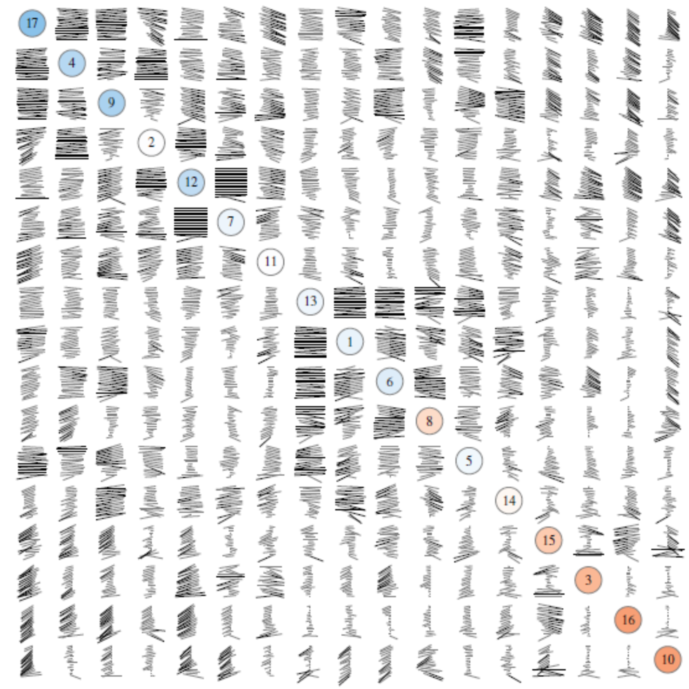
Shannon, Quigley, & Nixon (2010).

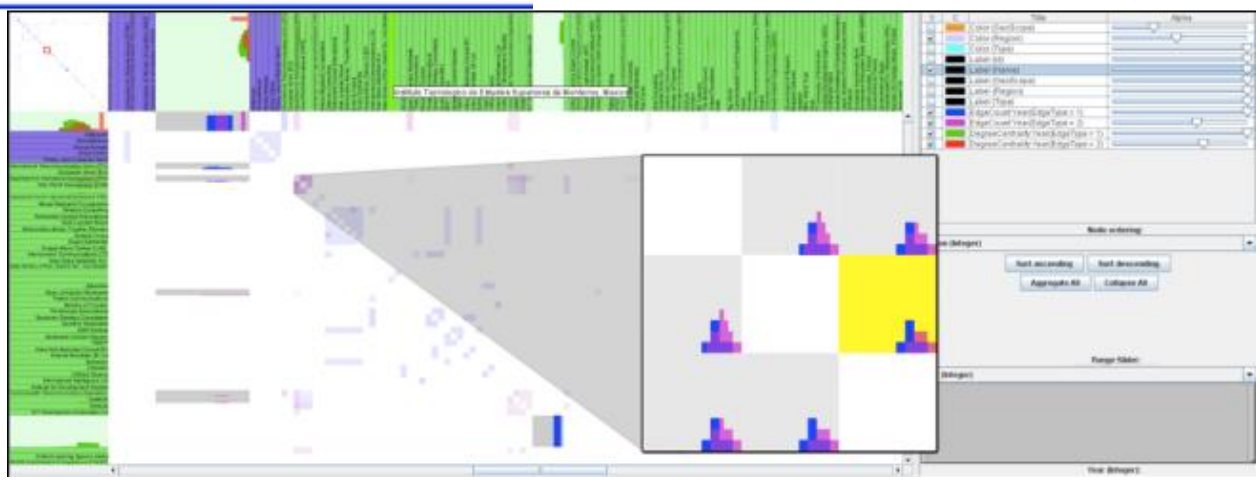
Nested time



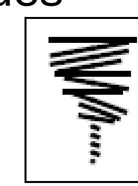
heat maps embedded in node link graph vertices (Saraiya et al., 2005);

histograms embedded in matrix (Yi et al., 2010)



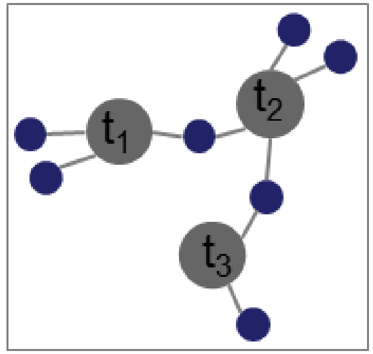


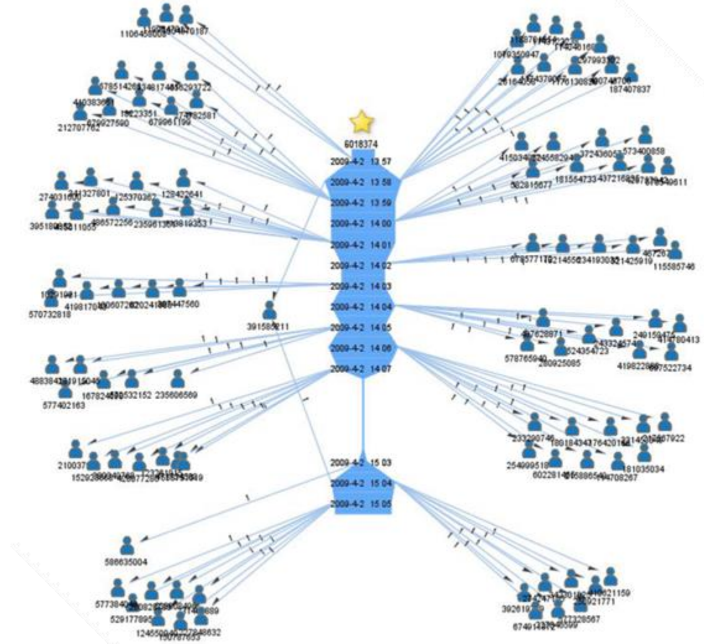
Gestaltmatrix (gestalt lines embedded in a matrix) (Brandes & Nick, 2011)



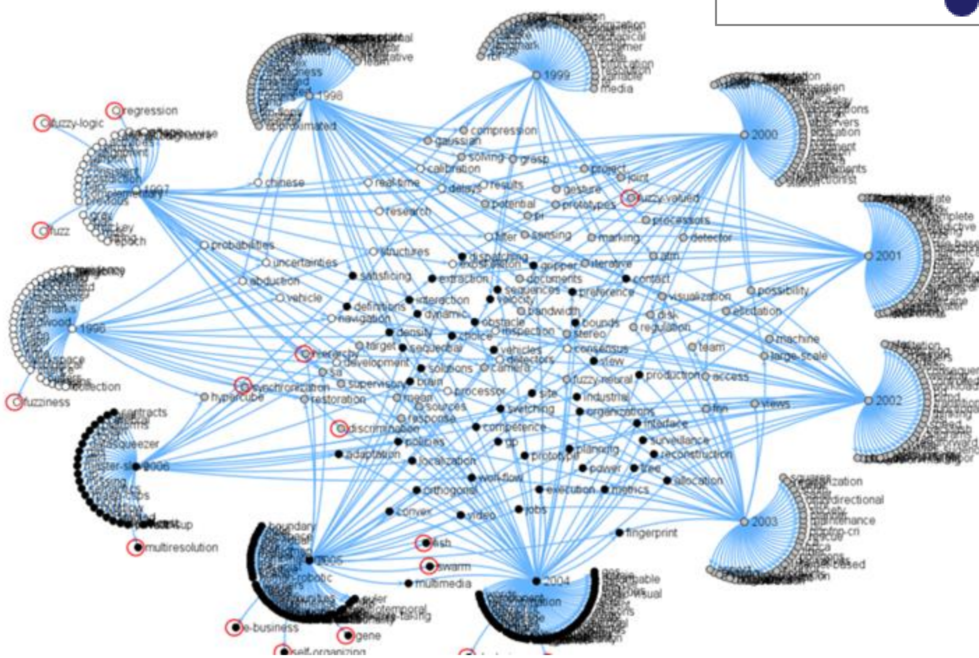
gestalt-based

Embedded: Time as a node

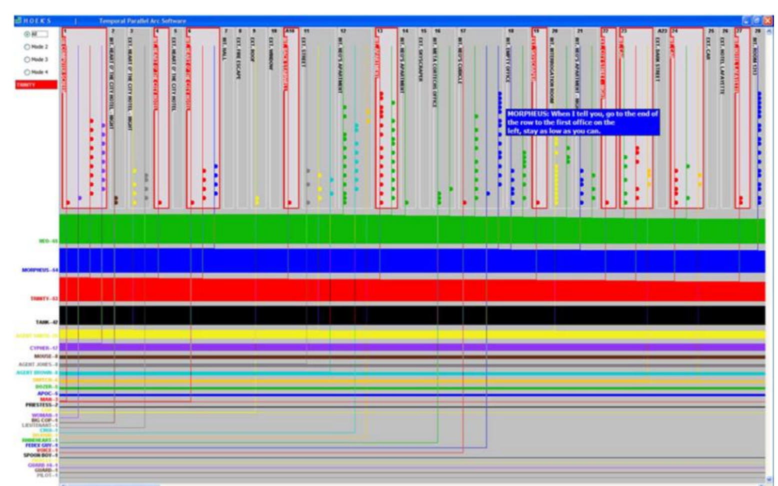




1.5D (Shi, Wang, & Wen, 2011, Figure 1)



Thiel, Dill, Kötter, & Berthold (2007, Figure 3)



TIPAD - Parallel Arc Diagrams (Hoek, 2011, Figure 14)

COMBINING TECHNIQUES

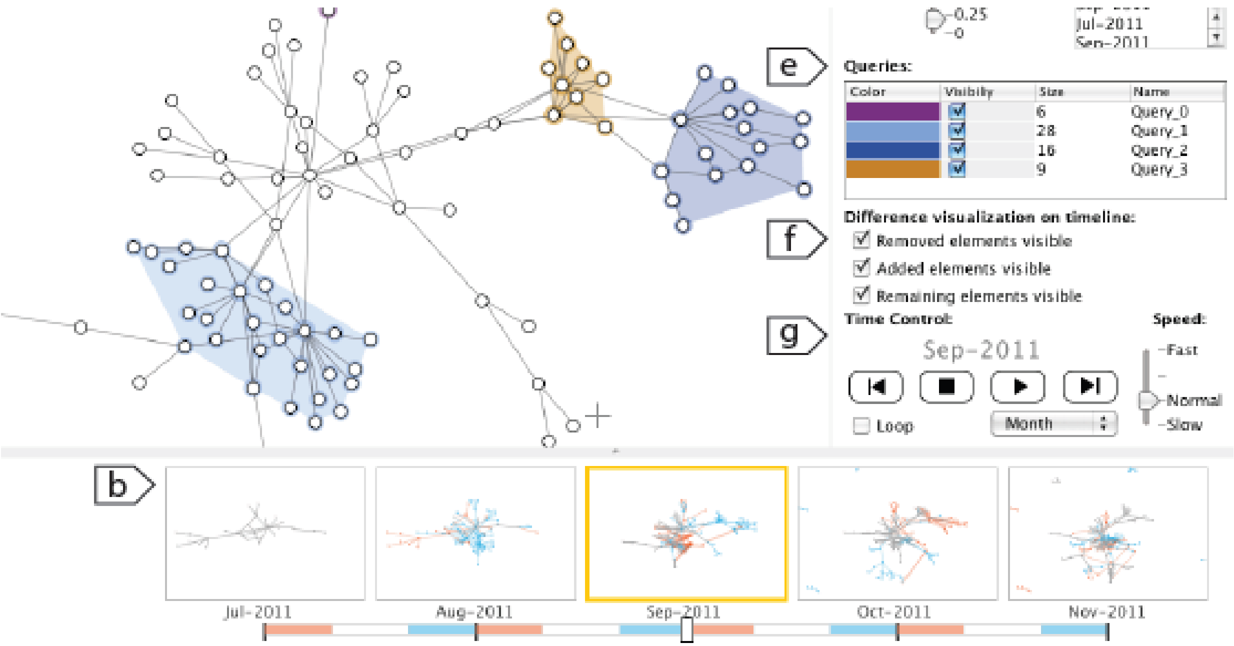
Combining Techniques

- Multiple views maximise insight and balance strengths and weaknesses of individual techniques

-> combine different temporal and/or graph structural encodings

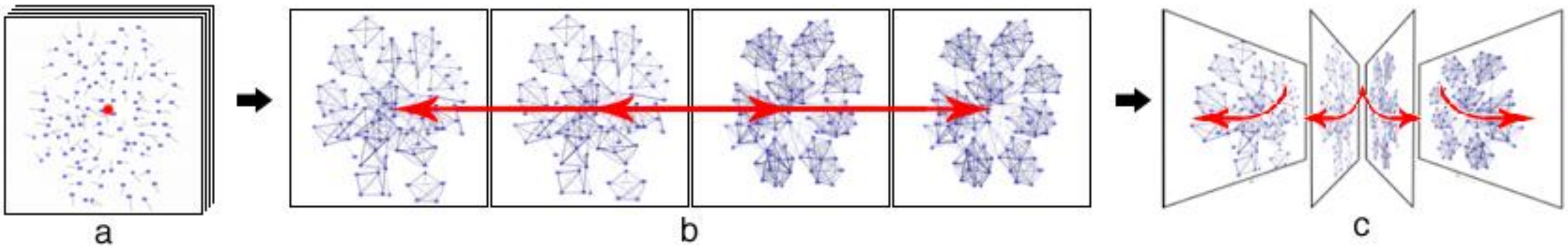
- Examples from literature:
  - Show different techniques in same screen space (MCV)
  - Allow user to select and switch between views
  - Allow users to select different views 'in situ'

Combined approaches: multiple view in same screen space



Animation and small multiples – GraphDiaries
(Bach, 2012)

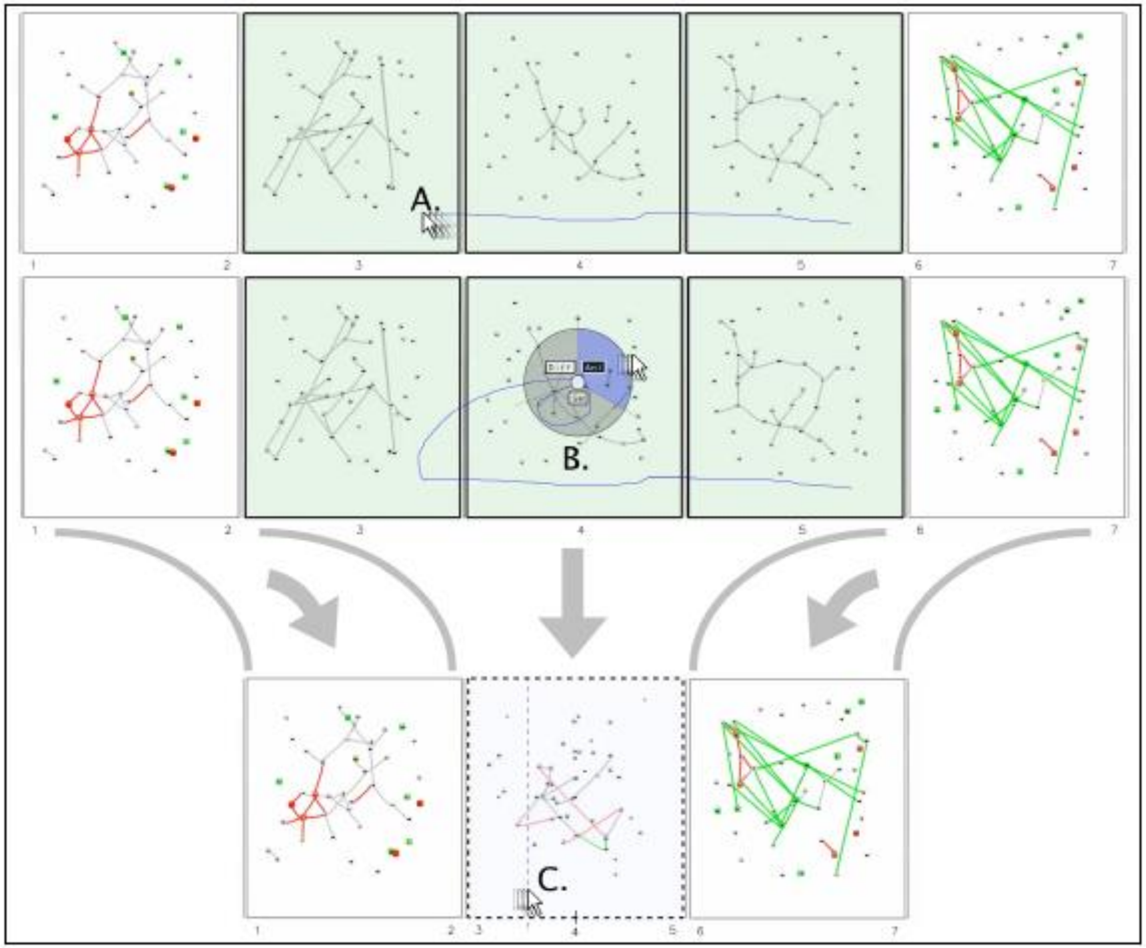
Combined approaches: user selects and switches between views



(a) superimposition, (b) juxtaposition, (c) 2.5D (Federico et al. 2011)

Interaction/transitioning technique designed to help orientate users when switching between views

Combined approaches: user selects views 'in situ'



DiffAni (Rufiange and McGuffin, 2013, Fig. 5: "Interactively converting three small multiple tiles into a single animation tile")

System offers juxtaposed, sequential and difference map (merged) views.

SUMMARY

Take away

What you should take away from this lecture…

- What a graph is
- The different techniques for representing graph data
- The advantages and disadvantages of each
- The approaches available for representing temporal data and the advantages and disadvantages of each

References

- Ahn, J.-wook, Taieb-maimon, M., & Sopan, A. (2010). Temporal Visualization of Social Network Dynamics : Prototypes for Nation of Neighbors. *Proc. Of Social Computing, Behavioral-Cultural Modeling and Prediction conference*, 309-316. Retrieved from http://cgis.cs.umd.edu/localphp/hcil/tech-reports-search.php?number=2010-28
- Bach, B., Pietriga, E., & Fekete., J.-D. (2012). Temporal Navigation in Dynamic Networks. *IEEE Vis Week. 2012.*
- Bender-deMoll, S., & Mcfarland, D. A. (2006). The Art and Science of Dynamic Network Visualization. *JoSS*, *7*(2). Retrieved from http://www.stanford.edu/group/sonia/papers/DNV_JOSS.pdf
- Bennett, C., Ryall, J., Spalteholz, L., & Gooch, A. (2007). The Aesthetics of Graph Visualization. *Computational Aesthetics in Graphics, Visualization, and Imaging*, 1-8. Retrieved from http://www.thechiselgroup.org/files/uploads/publications/submission1022.pdf
- Brandes, U., & Corman, S. R. (2003). Visual Unrolling of Network Evolution and the Analysis of Dynamic Discourse. *Information Visualization*, *2*(1), 40-50.
- Brandes, U., & Nick, B. (2011). Asymmetric relations in longitudinal social networks. *IEEE transactions on visualization and computer graphics*, *17*(12), 2283-90. doi:10.1109/TVCG.2011.169
- Federico, P., Aigner, W., Miksch, S., Windhager, F., & Zenk, L. (2011). A Visual Analytics Approach to Dynamic Social Networks. *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies (i-KNOW ’11)*. doi:10.1145/2024288.2024344
- Ghoniem, M., Fekete, J.-D., & Castagliola, P. (2005). On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. *Information Visualization*, *4*(2), 114-135. doi:10.1057/palgrave.ivs.9500092
- Graham, M., & Kennedy, J. (2010). A survey of multiple tree visualisation. *Information Visualization*, *9*(4), 235-252. doi:10.1057/ivs.2009.29

References (cont.)

- Henry, N., Fekete, J.-D., & McGuffin, M. J. (2007). NodeTrix: a hybrid visualization of social networks. *IEEE transactions on visualization and computer graphics*, *13*(6), 1302-9. doi:10.1109/TVCG.2007.70582
- Herman, I., Melancon, G., & Marshall, M. S. (2000). Graph visualization and navigation in information visualization: A survey. *Visualization and Computer Graphics, IEEE Transactions on*, *6*(1), 24-43. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=841119&tag=1
- Kruja, E., Marks, J., Blair, A., & Waters, R. (2002). A Short Note on the History of Graph Drawing. *GRAPH DRAWING Lecture Notes in Computer Science*, *2265/2002*(602-606), 272-286. doi:10.1007/3-540-45848-4_22
- Munzner (2006) 15 views of a node link graph http://www.cs.ubc.ca/~tmm/talks/nih06/
- Saraiya, P., Lee, P., & North, C. (2005). Visualization of Graphs with Associated Timeseries Data. *IEEE Symposium on Information Visualization 2005* (pp. 225-232). Minneapolis, MN, USA: IEEE.
- Shannon, R., Quigley, A., & Nixon, P. (2010). Showtime : Increasing viewer understanding of dynamic network visualisations. *Proceedings of Advanced Visual Interfaces, AVI2010*. Rome: ACM Press. Retrieved from http://rossshannon.com/publications/softcopies/Shannon2010Showtime.pdf
- Yi, J. S., Elmqvist, N., & Lee, S. (2010). TimeMatrix: Analyzing Temporal Social Networks Using Interactive Matrix-Based Visualizations. *International Journal of Human-Computer Interaction*, *26*(11-12), 1031-1051. doi:10.1080/10447318.2010.516722
- von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., van Wijk, J. J., Fekete, J.-D., & Fellner, D. W. (2011). Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges. *Computer Graphics Forum*, *30*(6), 1719–1749. doi:10.1111/j.1467-8659.2011.01898.x