Visualizing Networks with Spring Embedders: two-mode and valued graphs

Lothar Krempel
Max Planck Institute for the Study of Societies
Cologne,Germany
krempel@mpi-fg-koeln.mpg.de

August 5, 1999

1. What is a spring embedder ?
   (a) definition
   (b) models
   (c) properties

2. How to obtain solutions for two common social science data types
   (a) valued data
   (b) two-mode data

What is a spring embedder ?

A system where the nodes of a graph are replaced by steel rings and the edges with springs that exert mechanical forces.

✔ type of forces

- mechanical forces based on the Euclidean distance only
- electrical fields to impose placement constraints
- magnetic fields can be used to affect the overall appearance

✘ constraints

- iteration schemes
  - sequential
  - batch wise

- search strategies
  - deterministic
  - stochastic

- criteria for a good solution
  - low potential energy state of the total system
  - statistical fit ?
  - readability ?

Constraints

- uniform edge length (binary graph)
- minimal edge crossings
- nodes should not be to close
- angle of incident edges (angular resolution)
- nodes should not be to close to edges
- width of the layout should be optimal
- symmetry ?

Criteria to evaluate a solution

- low potential energy in the total system (all attached forces balance)

- fit to data

  - binary: all edges have the same length
  - valued: high correlation with raw data
    - metric: Pearson´s r
    - ordinal: Kendall tau

- Readability: a high resolution of the image to communicate additional content

Attraction and Repulsion for binary Graphs I

- The literature that has emerged around the basic scheme of a spring embedder has adopted a large body of field and force concepts and included additional order constraints to maximize the relational resolution of the drawings.

1. spring forces between adjacent nodes
2. electrical fields around single sources
3. gravitational fields
4. magnetic forces between magnetic poles

Attraction and Repulsion for binary Graphs II

In their recent book on graph drawing (Battista et al (1999)), p. 306) summarize different approaches to the force directed drawing of straight line graphs, with a generalized formula which describes all forces which are simultaneously attached to a single vertex u.

$$F(v) = \sum_{(u,v) \epsilon E} f_{uv} + \sum_{(u,v) \epsilon VxV} g_{uv}$$

- where f_{uv} is a spring force between adjacent Nodes u and v
- and g_{uv} is an electrical force proportional to the Euclidean distance of u and v and the zero length of a spring. The electrical force follows an inverse square law.

Attraction and Repulsion for binary Graphs III

- a desired length l_{uv} for the distance between two vertices u and v ,
- their current Euclidean distance $d(p_u, p_v)$,
- spring embedders differ in the way how they weight the current displacement $(d(p_u, p_v) - l_{uv})$ from a desired length l_{uv}
- and the way in which the squared distance enforces an additional spacing between all units.

The resulting x component of the Force $F(v)$ can be written as

$$F(v) = \sum_{(u,v) \epsilon E} k_{uv}^{(1)} (d(p_u, p_v) - l_{uv}) \frac{x_v - x_u}{d(p_u, p_v)} + \sum_{(u,v) \epsilon V x V} \frac{k_{uv}^{(2)}}{(d(p_u, p_v))^2} \frac{x_v - x_u}{d(p_u, p_v)}$$

A Simple Algorithm for a Spring Embedder

```
produce start-configuration
define krit
while fit > krit do
    for all  lines do
        compute Node_displacement for attractive
        forces
    end for
    for all  nodes do
        displace
    end for
    for all nodepairs do
        compute current Euclidean distance
        apply repulsive forces
    end for
    rescale to view image
    compute goodness of fit
end while
```

Properties I: Spring embedders with no repulsive forces: Barycenter Drawing

- Tutte (1960, 1963) has proven that any three-connected graph (a graph where the nodes have at least three links) can be drawn (without repulsive forces) by partitioning the nodes into two sets, where one set (with at least three nodes) is fixed and the nodes of the second set are allowed to move.

$$p_u = \frac{1}{degree_u} \sum_{(uv) \epsilon E} p_v$$

Properties I :no repulsive forces

- the solution is dependent of the start configuration

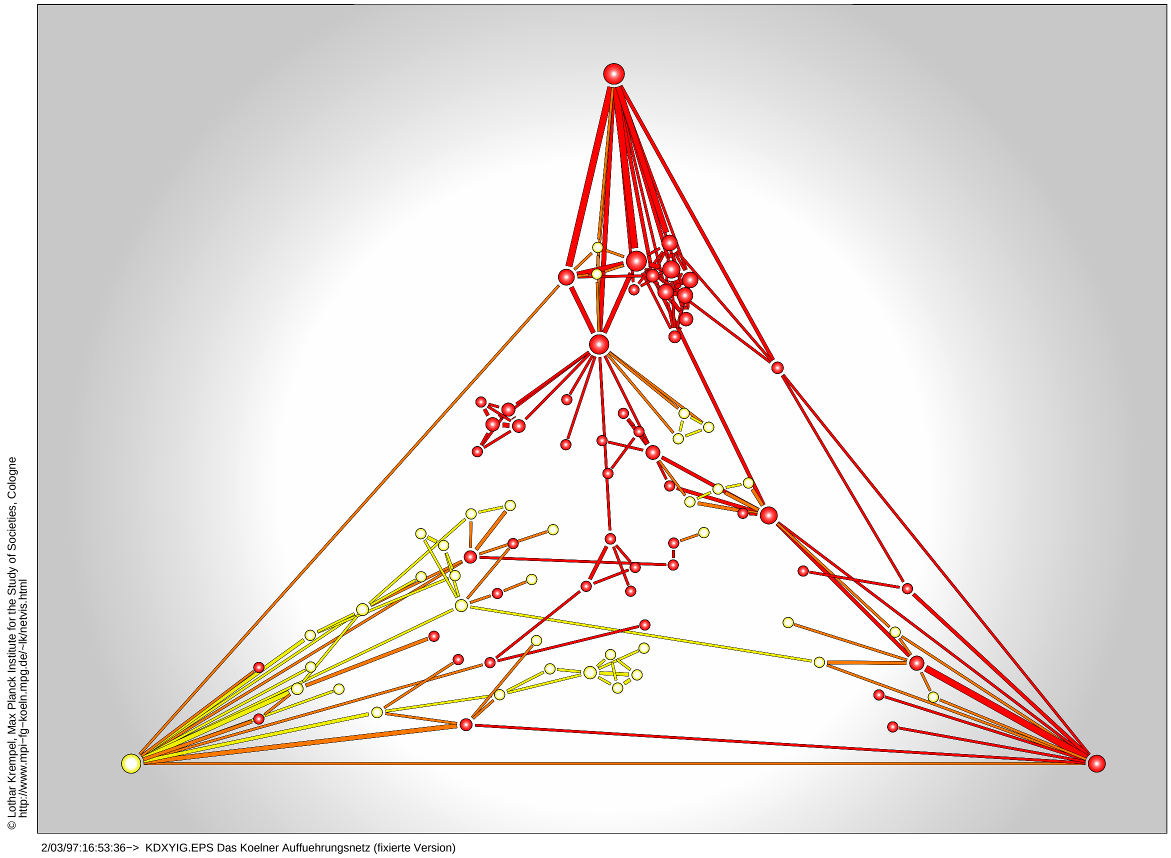


Figure 1: A barycentric drawing of co-performances of compositions from composers of new music in Cologne. Intimate knowledge allowed the researcher to choose two local (red) and a prominent external composer (yellow) to be fixed.
The graph is connected, but not three-connected. Nodes that are attached with single links are not placed very well. (Data courtesy of Dominik Sack)

Properties II :Spring embedders with attractive and repulsive forces

- the solution is independent of the start configuration

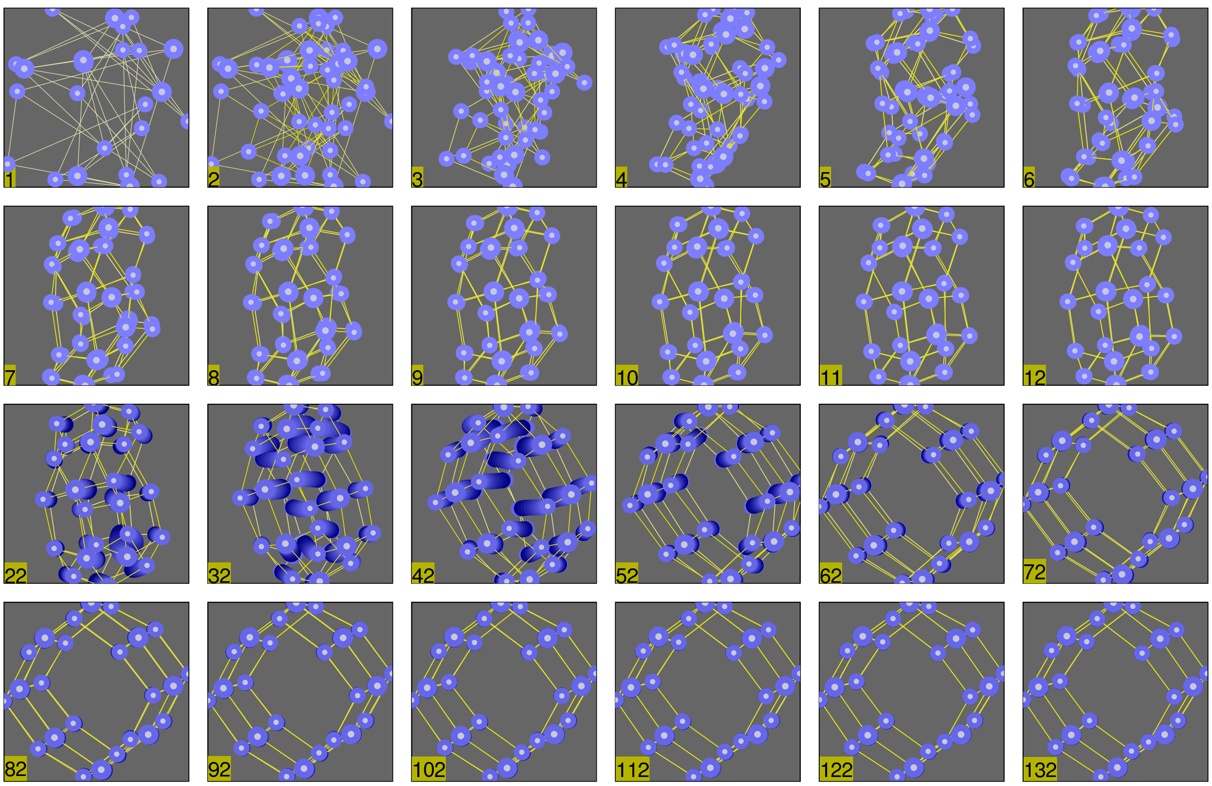


Figure 2: Animation of the embedding of a graph describing a torus with a Fruchterman & Reingold type spring embedder.

Three Groups of Algorithms

Spring Embedders differ in the way how forces are conceptualized and by the information that is used.

- forces that are a function of the current placement in the image only (Euclidean distance)
- forces that affect the overall appearance of a solution by additional field concepts
- forces that are a function of additional (structural) information derived from the data

Forces based on the Euclidean distance only I

Eades (1984) minimization: springs	
node attraction	$f_{a_{uv}} = k_a \log d_{uv}$
node repulsion	$f_{r_{uv}} = k_r / d_{uv}^2$

Fruchterman &Reingold (1994) minimization: annealing	
node attraction	$f_{a_{uv}} = d_{uv}^2 / k$
node repulsion	$f_{r_{uv}} = -k^2 / d_{uv}$

The models of Eades and Fruchterman & Reingold differ in their functions. The distance between of two nodes can be determined with the help of the resulting force $f_{res} = f_a + f_r$.. The attractive and repulsive functions for the Fruchterman and Reingold model are much steeper.

Push Properties

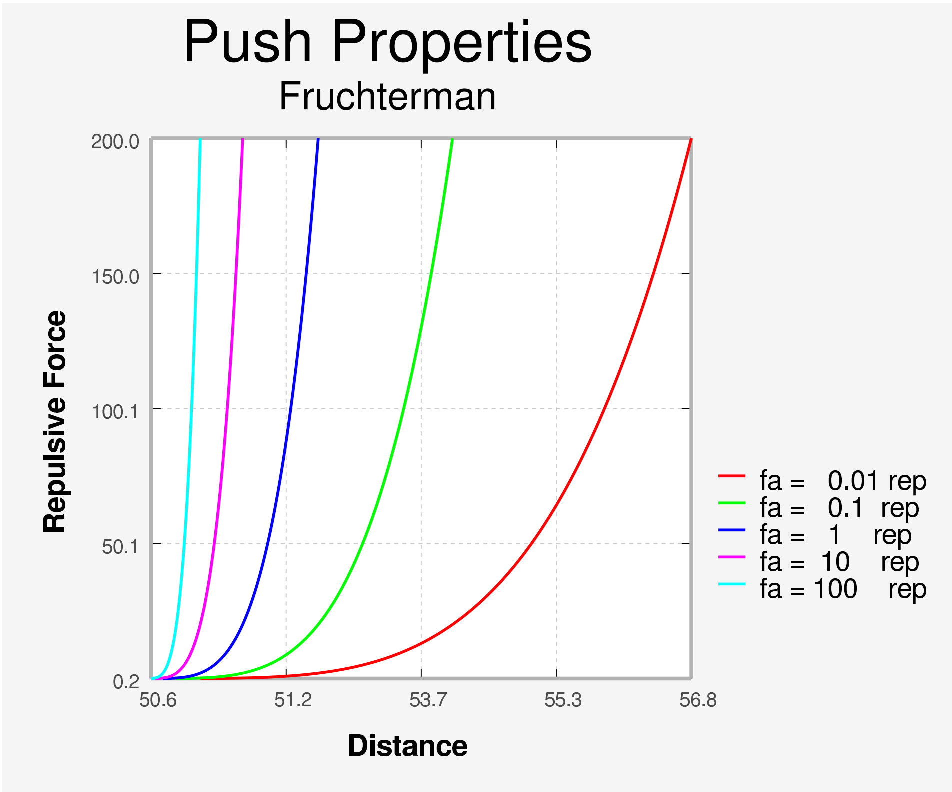


Figure 3: placement by increase of the repulsive force for attractive forces of different size.

Forces based on image distances only II

Davidson & Harrel (1996) minimization: annealing	
equal distribution	$a_{ij} = \frac{\lambda}{d_{ij}}$
distance to borders	$\frac{1}{r^2} + \frac{1}{l^2} + \frac{1}{t^2} + \frac{1}{b^2}$
edge length	$\frac{k}{d^2}$
node-edge distance	$h_{kl} = \frac{w}{g_{kl}^2}$

Coleman & Parker (1996) minimization: annealing	
node/node rep	$\frac{1}{d}$
edge length min	d^2
centripetal rep	$\frac{1}{d}$
node/edge rep	$\frac{1}{d}$
edge midpoints	$\frac{1}{d}$

Placement with structural information

- ✘ Two approaches that try to enhance the placements by adding additional structural information into the ordering process are especially noteworthy. These approaches point into the direction how information deficits of the classical model can be cured for more complex graphs.

Centrality measures (Freeman (1979)) seem to be very well suited to summarize the structure of a graph and to provide the necessary information that prevents artifacts as they can occur with the basic model for less regular and complex graphs.

Centrality Constraints

Frick & et al (1994)	
barycenter	$B = \frac{1}{N}\sum_{i=1}^{n} p(v_i)$
degree centrality	$\lambda(1 + degree(v))(B - p(v))$

Frick, Ludwig and Mehldau (1994) have presented an algorithm which explicitly takes the *degree centrality* of a node into account when constructing forces. The authors used an additional attractive force to the barycenter of the total configuration, which is weighted with the degree of a node.

Brandes &Wagner (1997) minimization: annealing	
closeness centrality	$c_3(d(x_v, \zeta)^2$ $-c_4(Cv + 1 - \max_{u \epsilon V}(C_u))$

Brandes & Wagner (1997) propose an attractive force to the center of a drawing that takes the *closeness centrality* of pairs of nodes into account. Closeness centrality characterizes a node by the sum of the minimal path lengths to all other nodes.

Drawings using Distances I

Kamada and Tawai proposed to minimize the energy of a Hookean spring to order the total system: if the resting length of each of the springs is known, the energy is given by the squared difference of the Euclidian and target distances.

$$f_{(u,v)} = K_{u,v}(d_{(u,v)} - t_{(u,v)})^2$$

The authors chose the spring constant to be dependent on the distance so that springs with smaller displacements are stronger.

$$K_{u,v} = k/d^2_{(u,v)}$$

Kamada &Kawai minimization: steepest descent Newton Raphson	
desired width	$l_{uv} = geodesic(uv)$
weight	$k_{uv} = k/l^2_{uv}$
force	$f = k_{uv}(d_{uv} - l_{uv})^2$
Energy	$\sum\sum(d_{uv} - l_{uv})^2$

Drawings using Distances II

Cohen (1997):

The stress functions S_k, $k = 0, 1, 2$ are proportional to something of the form

$$D_k = \sum_{i<j} \frac{(d_{ij} - t_{ij})^2}{t_{ij}^k}$$

let vertex i have coordinates ($x_{i,}y_i$) so that

$$d_{ij} = \sqrt{(x_i - x_j)^2 (y_i - y_j)^2}$$

given the current vertex positions, an iteration consists of adding to each (x_i, y_i) the increment $(\Delta_{x_i,}\Delta_{y_i})$ where

$$\Delta_{x_i} = -\mu_i \frac{\partial D_k}{\partial_{k_i}} = -\mu_i \sum_{j:j \neq i} \frac{2(x_i - x_j)(1 - t_{ij}/d_{ij})}{t_{ij}^k}$$

Drawings using Distances III

The direct approach allows us also to introduce an additional mechanism into the computations which results in a similar flexibility that makes the simple spring embedders so useful to produce renderings with a high resolution: an additional spacing in the image can be enforced if we add a constant distance k to each of the desired distances as they are found in the data.

$$l_{uv} = k + t_{uv}$$

The gradient of the raw stress has a very intuitive meaning: it simply scales the spring with the ratio of the desired and the Euclidean distance, which becomes one if both are of the same length.

Once a minimum has been found by the algorithm, we can increase k, to enforce a minimal distance, while preserving the ranks of the distances.as in the simple spring embedders

Two-Mode Graphs: An Extension of Barycenter Drawing

Two-mode data are quite common in the analysis of social networks, when only the relations *between* two sets of nodes are observed. A classical study is that of Davis & Gardener (1932) who describe 18 women in the American South and their participation in 14 events.

We present a solution for this problem by using a modified barycenter approach: in each step of the iteration we fix one of the two sets of nodes and allow the other set to move into the direction of their barycenters.

	two-mode embedder
attraction	$f_a = k_a \frac{d_{uv}}{degree_u}$
repulsion	$f_r = k_r \frac{1}{d_{uv}^2}$

Two-Mode Graphs: Algorithm

1. for all links of Set A (rows)

   (a) compute attraction to the nodes of B to their barycenters
   (b) displace Nodes A
   (c) compute repulsion and displace nodes of A and B

2. for all links of Set B (cols)

   (a) compute attraction for nodes of B to their barycenters
   (b) displace nodes B
   (c) compute repulsion and displace nodes of A and B

3. repeat

Two-Mode Graphs:Example

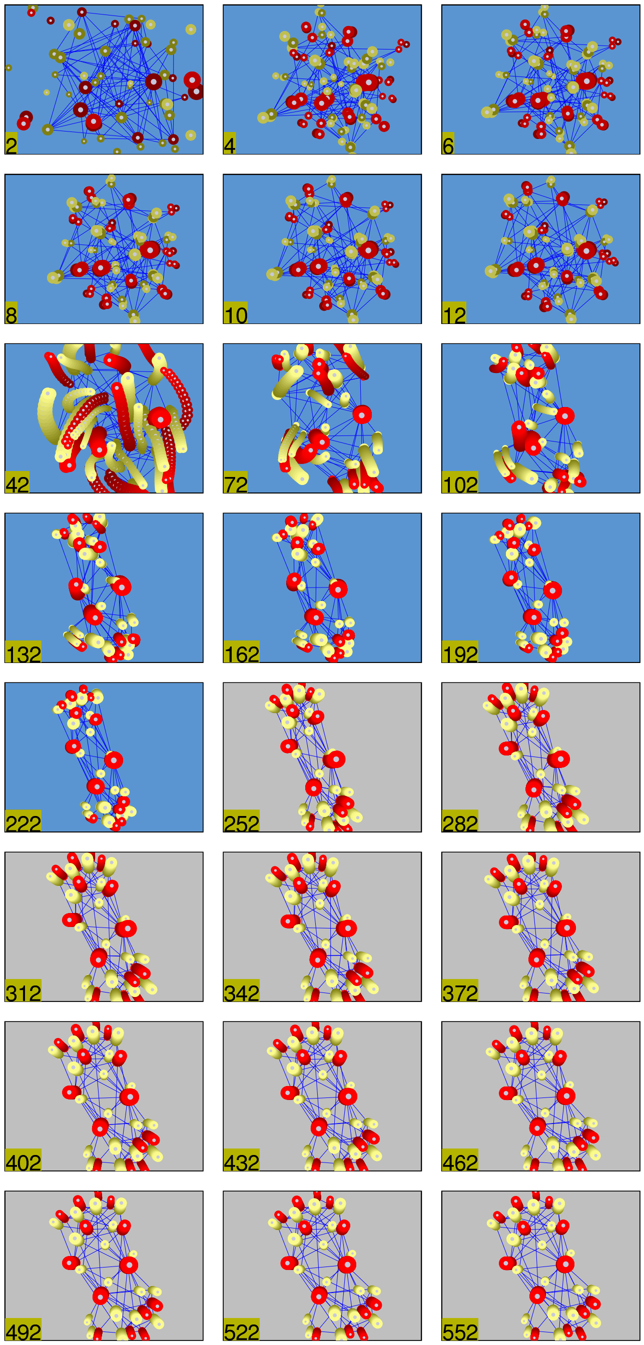
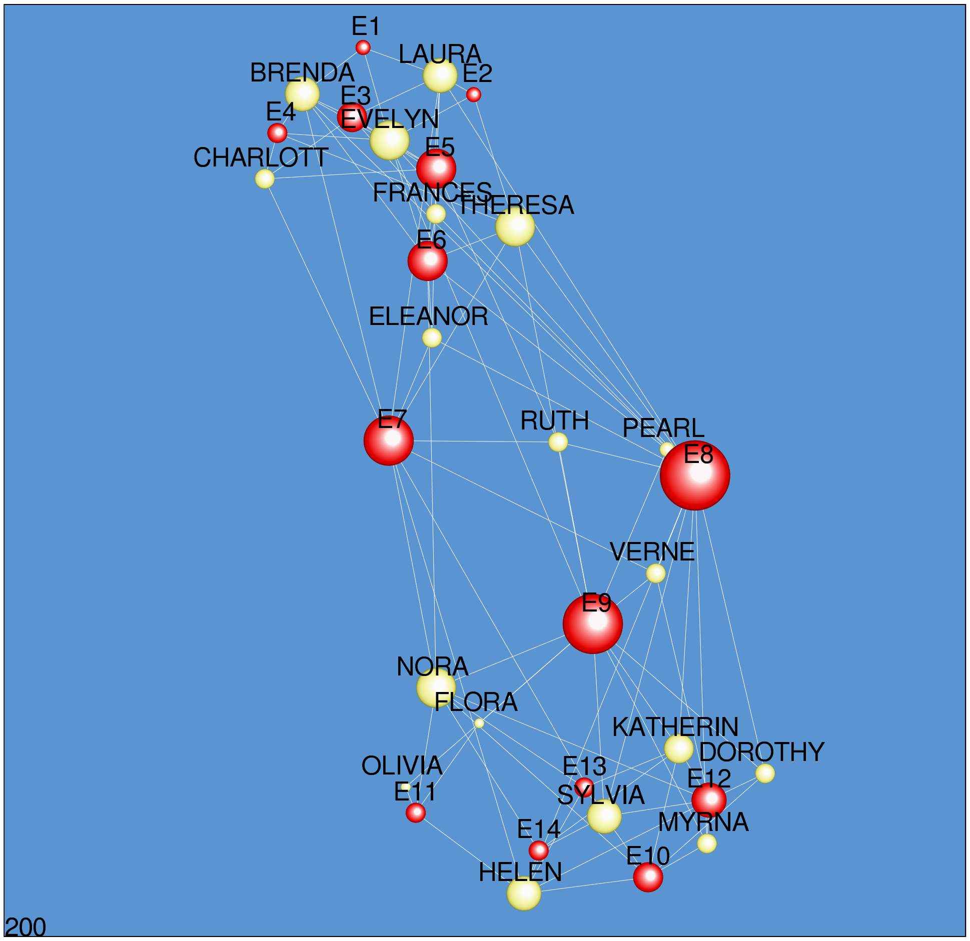
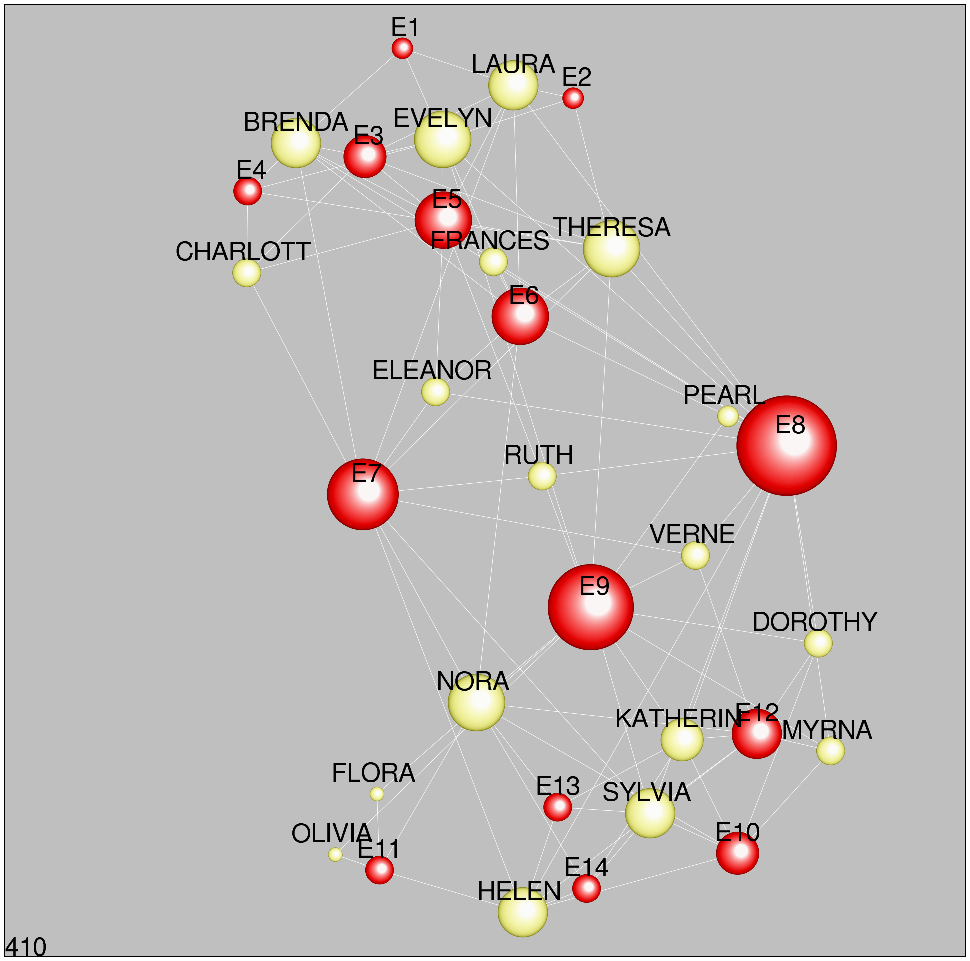


Figure 4: Animation of the iterations of a two-mode barycenter method with an additional repulsive force component.

Figure 5: Two solutions for the Davis & Gardener dataset illustrating the ordering capabilities of a two-mode barycenter approach blended with a repulsive force component



(a) Davis Gardener after 200 iterations



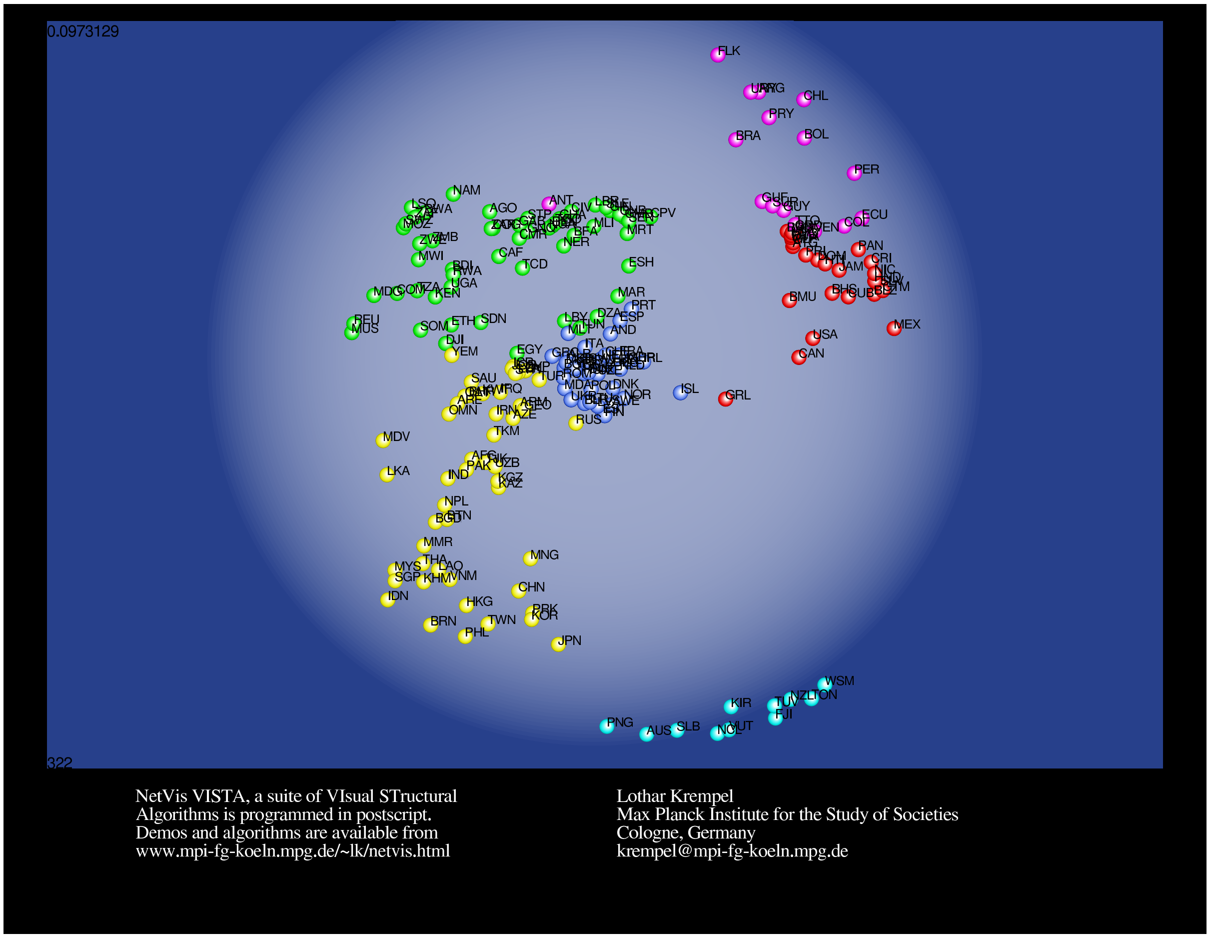
(b) as above after 200 additional iterations increasing the repulsive constant

A Metric Example: Ordering the Capitals of the World

To evaluate the potential of the two approaches to metric data, we use the distances of geographic locations which were read from a sphere. The grand circle distances between 194 capitals (their airline distances) are used to supply the algorithm with empirical target distances.

This is a problem for which there is no exact solution. as it involves to map a surface of a sphere onto a two-dimensional plane. Geographers have however established a number of projection conventions by which this problem is usually handled.

A Metric Example: Ordering the Capitals of the World



A world map reconstructed from the airline distances between 194 capitals, computed with a simple metric spring embedder. The Pearson correlation of the data with the image distances is r = .9641.

The solution allows to identify the continents (North America: red; South America: magenta, Africa (green), Europe (blue) and Oceania (cyan))[1]

[1]The location of the Antilles (ANT) is miscoded in the original data with a latitude and longitude of 0. Ok ?

Resume

- ✔ Force directed placement provides a tool box of ordering principles, which can be employed to show network properties of interest.

- ✔ The algorithmic concept of a Spring Embedder is a powerful framework that can be extended to handle BIPARTITE and VALUED graphs. Our simple approach to a metric problem can yield a fit that is comparable to multidimensional statistical procedures but allows at the same time to generate solutions which preserve the rank order of the distances.

- ✔ The latter can be used to enhance the overall resolution and readability of images of network drawings. This is often necessary when one wants to render additional substantive information on to network drawings.