visone – Analysis and Visualization of Social Networks

Jürgen Lerner

University of Konstanz

Egoredes Summer Course Barcelona, 6.–10. July, 2009

Outline

1. visone– people, purpose, history.
2. visone basics.
3. **Scene 1** Juan draws a network.
4. **Scene 2** José analyzes a personal network.
5. **Scene 3** Jorge explores a random network.

Each scene introduces a researcher who wants to perform a specific task.

- We're going to help him to do this step by step in visone; intermingled with notes on the methods applied.

Outline

1. visone– people, purpose, history.
2. visone basics.
3. **Scene 1** Juan draws a network.
4. **Scene 2** José analyzes a personal network.
5. **Scene 3** Jorge explores a random network.

Each scene introduces a researcher who wants to perform a specific task.

- We're going to help him to do this step by step in visone; intermingled with notes on the methods applied.

visone – people, purpose, history.

Developed by groups at

- University of Karlsruhe (Dorothea Wagner) and
- University of Konstanz (**Ulrik Brandes**).

Highly involved: Michael Baur, Martin Mader, Uwe Nagel.

Main purpose: making novel network analysis and visualization techniques available to social scientists.

- Ease of use: **visual** creation, interaction, exploration, analysis, and representation of networks.
- Current focus on small and medium size networks.
- Convenient handling of actor and tie attributes.

Started around 2001 – gets constantly extended, improved, . . .

visone – people, purpose, history.

Developed by groups at

- University of Karlsruhe (Dorothea Wagner) and
- University of Konstanz (**Ulrik Brandes**).

Highly involved: Michael Baur, Martin Mader, Uwe Nagel.

Main purpose: making novel network analysis and visualization techniques available to social scientists.

- Ease of use: **visual** creation, interaction, exploration, analysis, and representation of networks.
- Current focus on small and medium size networks.
- Convenient handling of actor and tie attributes.

Started around 2001 – gets constantly extended, improved, . . .

visone – people, purpose, history.

Developed by groups at

- University of Karlsruhe (Dorothea Wagner) and
- University of Konstanz (**Ulrik Brandes**).

Highly involved: Michael Baur, Martin Mader, Uwe Nagel.

Main purpose: making novel network analysis and visualization techniques available to social scientists.

- Ease of use: **visual** creation, interaction, exploration, analysis, and representation of networks.
- Current focus on small and medium size networks.
- Convenient handling of actor and tie attributes.

Started around 2001 – gets constantly extended, improved, ...

VISONE basics.

Getting it, installing it, running it.

Import, export.

Network model.

Overview of analysis and visualization methods.

visone– getting it, installing it, running it.

http://visone.info/

Requires Sun **Java** Runtime Environment (JRE) 6 (`http://java.sun.com/`)

Works on Windows, Linux, Unix, MacOS, ...

Running: either

- click on webstart visone or
- download `visone-2.5.jar` and execute it (e. g.) by double-clicking.

$\Rightarrow$ visone

visone – import, export.

Import and export **network**-formats

- GraphML – main format, in XML
- UCINET, Pajek, (Siena)
- adjacency matrix in `.txt` or `.csv`
- GML, LEDA GraphWin, (centering resonance analysis)

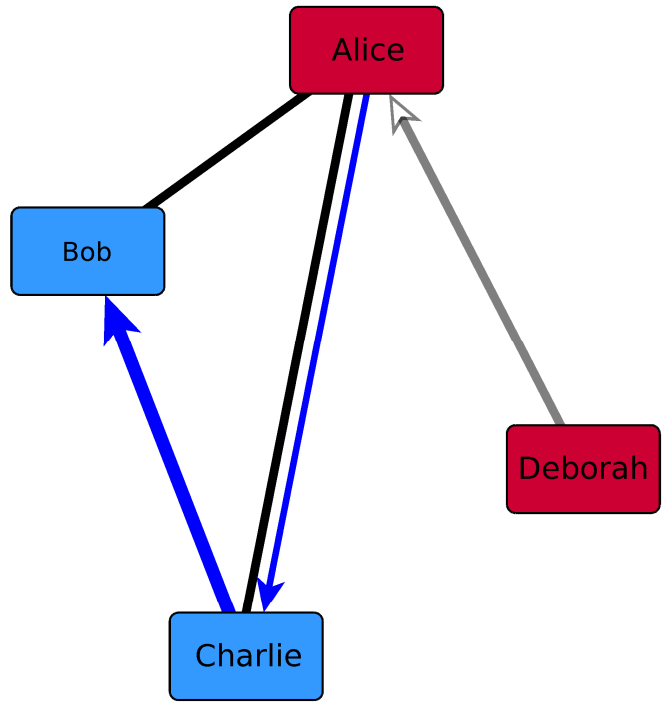
Export **image**-formats

- PNG, JPEG, GIF, BMP, SVG, PDF, ...
- printing

$\Rightarrow$ visone

visone– network model.

- mixed multigraph: (un-)directed, multiple ties
- actor and tie **attributes**: properties, numerical indicators, ...
- confirmed and un-confirmed ties



Confirmed and un-confirmed ties.

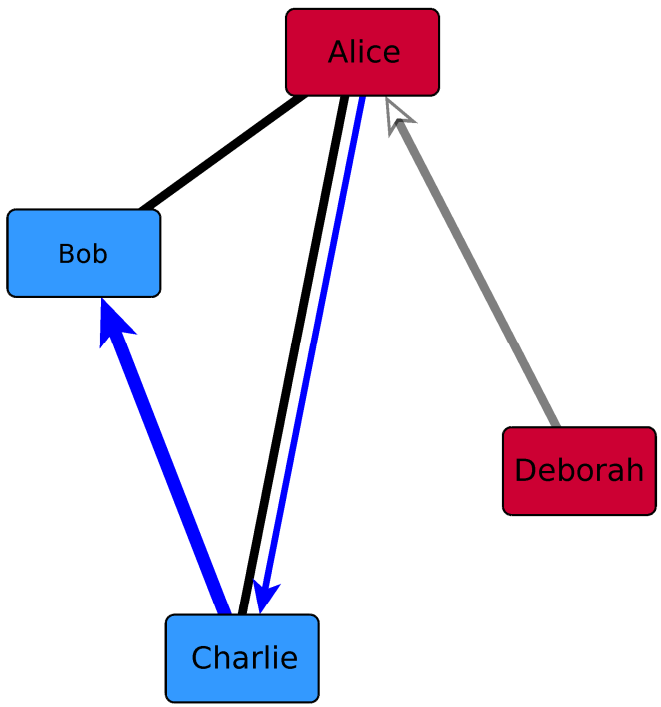
"Name your best friends."

Friendship (symmetric relation) encoded in black lines.

B names *A* and *A* names *B*.
C names *A* and *A* names *C*.

D names *A*, but *A* **does not** name *D*.

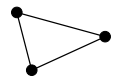
- inconsistencies often treated as "error"
- might encode valuable information
- do not "correct" it



Overview of visualization methods.

General purpose: draw networks to visualize structure.

- MDS, stress minimization, spring embedder, spectral, circular, random.

$\Rightarrow$ quick layout button 

Show properties: draw networks to emphasize attributes.

- **map attributes** to color, size, shape, width, . . .
- **centrality** layouts (draw central actors in the center)
- **status** layouts (draw high status actors on top)

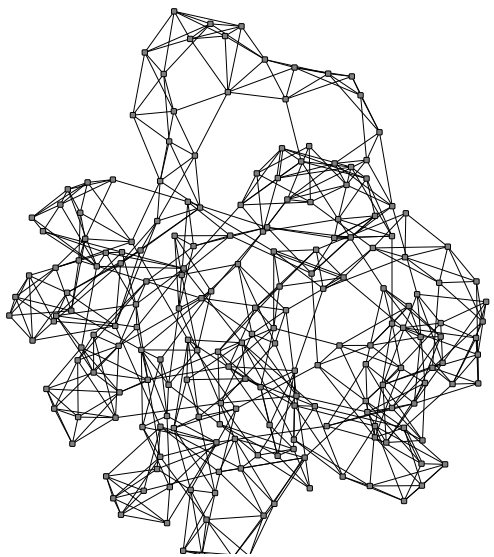
Geometric transformation

- rotate, translate, scale, reflect.

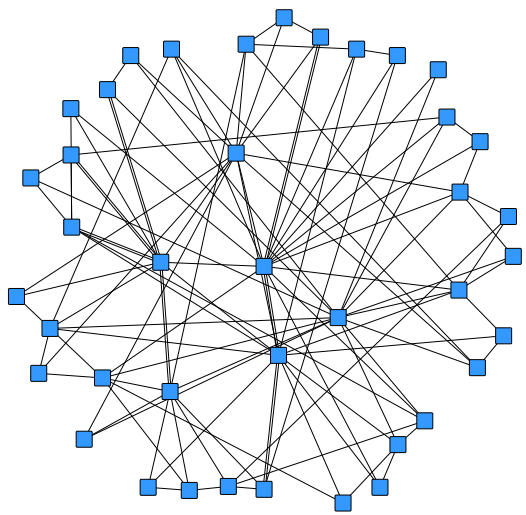
$\Rightarrow$ visone

Comparison of visualization methods.

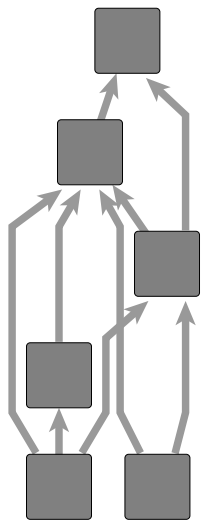
general purpose



centrality layout



status layout



Overview of analysis methods.

Indexing: computing properties of actors and ties.

- node and link **centrality**: degree, eigenvector, betweenness, closeness, status, page rank, ...
- **density**: clustering coefficient
- **distance** to selected actors

Grouping: computing groups of actors (in development)

- **clustering** (groups of densely connected actors)
- **role equivalence** (groups of similar actors)

Modeling (SIENA) (new feature)

$\Rightarrow$ visone

Outline

1. visone– people, purpose, history.
2. visone basics.

⇒ **Scene 1** Juan draws a network.

3. **Scene 2** José analyzes a personal network.
4. **Scene 3** Jorge explores a random network.

Scene 1 Juan draws a network for presentation.

Scenario: Juan writes a term paper on ***regular equivalence***.

For the presentation he wants to illustrate this concept on a small example network.

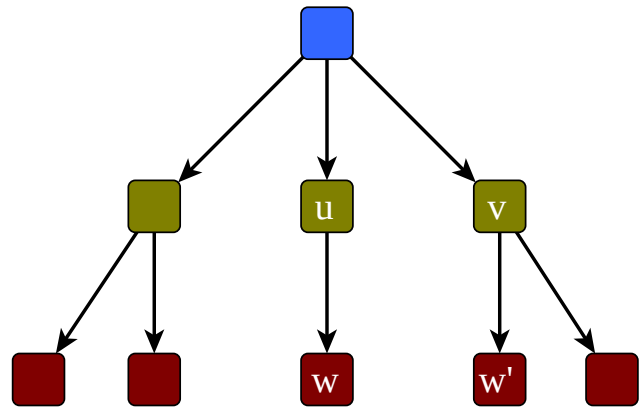
An network image helps understanding.

A vertex-coloring $c\colon V \to C$ is *regular* for a graph $G = (V, E)$ if

whenever two vertices u and v have the same color

and one of them is connected to another vertex $(u, w) \in E$

then v is connected to a vertext w' that has the same color as w



Wasserman/Faust (1994)

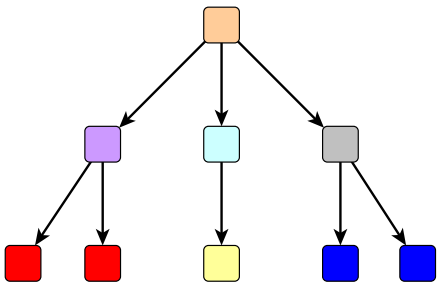
Such images can easily be created with visone.

Wrapping up.

- To create a small network that you have in mind, chose the new empty network option.
- Nodes and lines can be created in edit mode by use of the mouse – node and line templates can be defined.
- Node and tie indices are computed internally – can be mapped to visual characteristics.
- Export in various image formats is supported.

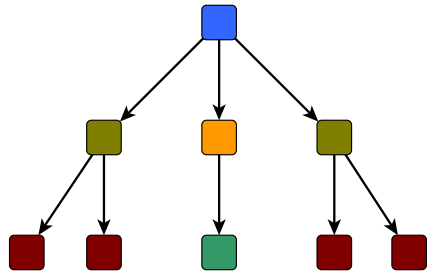
Intermezzo – different notions of role equivalence.

structural



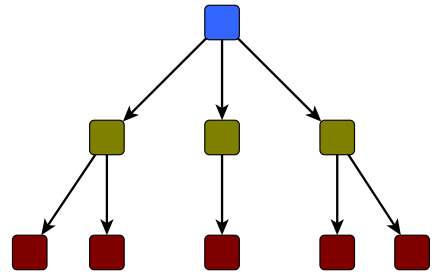
connected to the **same** others

exact regular



connected to the same number of **equivalent** others

regular



connected to **equivalent** others

Outline

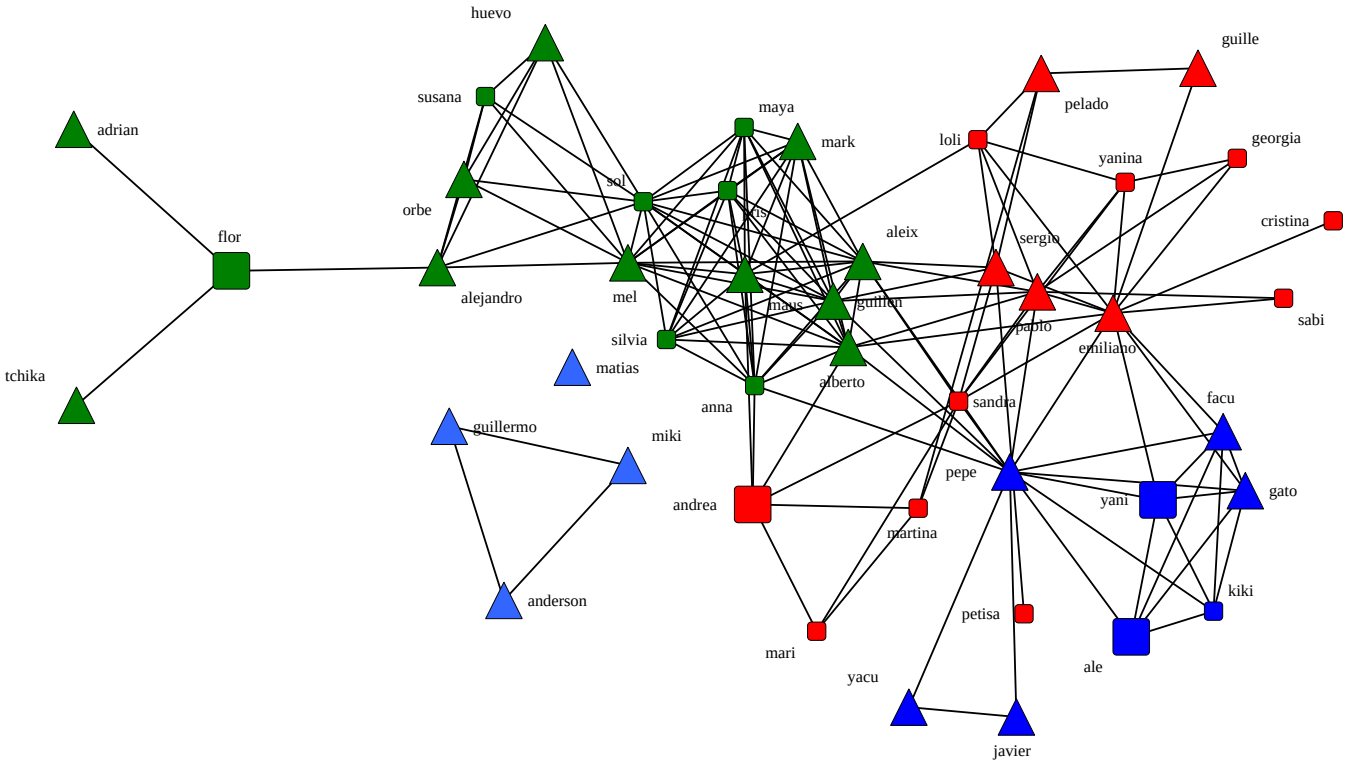
1. visone– people, purpose, history.
2. visone basics.
3. **Scene 1** Juan draws a network.

⇒ **Scene 2** José analyzes a personal network.

4. **Scene 3** Jorge explores a random network.

Scene 2 José analyzes a personal network.

Scenario José collected many personal networks containing various actor attributes.



$\Rightarrow$ visual exploration and analysis in Visone.

Wrapping up.

Collected node and link attributes can be

- used to define classes of actors and ties;
- mapped to color, shape, label, size, . . .
- exported to attribute tables (`.csv`) for further analysis.

Network indices (e. g., centrality) again define attributes

- are computed internally and can then be treated as any other attributes (mapping to visual properties, . . .).

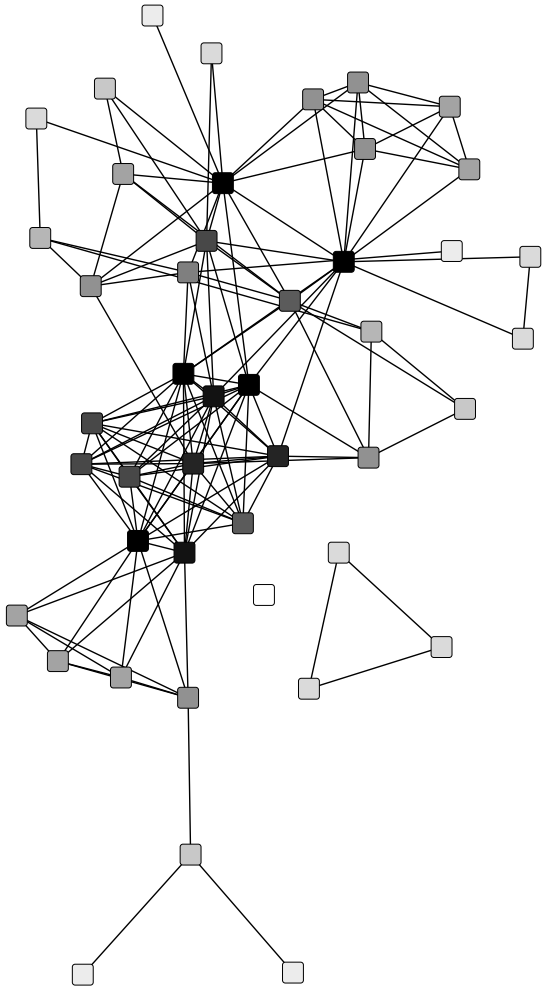
Intermezzo – centrality measures in networks.

degree indegree, outdegree

shortest path centralities
closeness, current flow closeness, betweenness, current flow betweenness, radiality, stress, eccentricity

feedback centralities (eigenvector)
eigenvector, hubs & authorities, PageRank, Katz' status

Degree centralities.



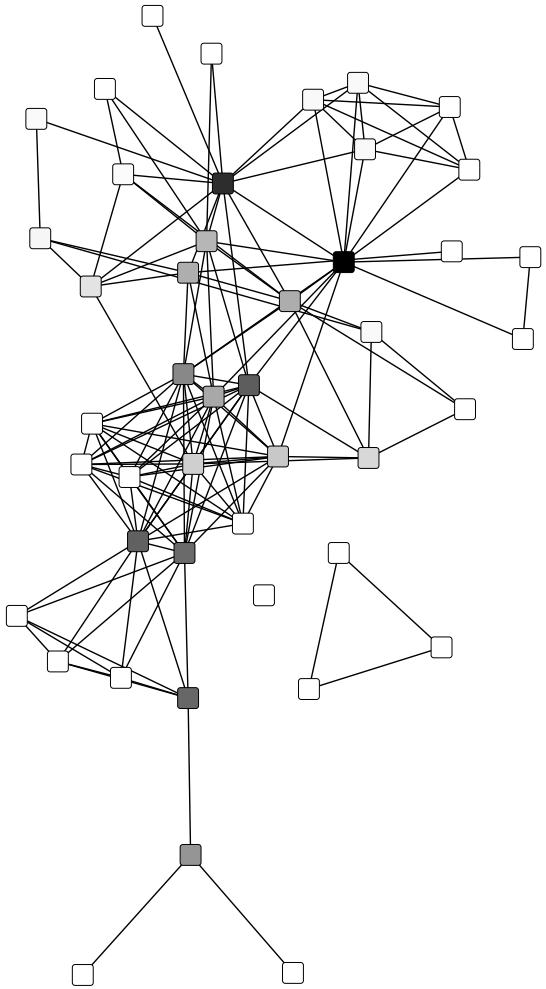
degree: number of lines connected to a node

indegree: number of *incomming* lines connected to a node

outdegree: number of *outgoing* lines connected to a node

$\Rightarrow$ visone

Shortest path centralities.



betweenness: being on shortest paths between alters

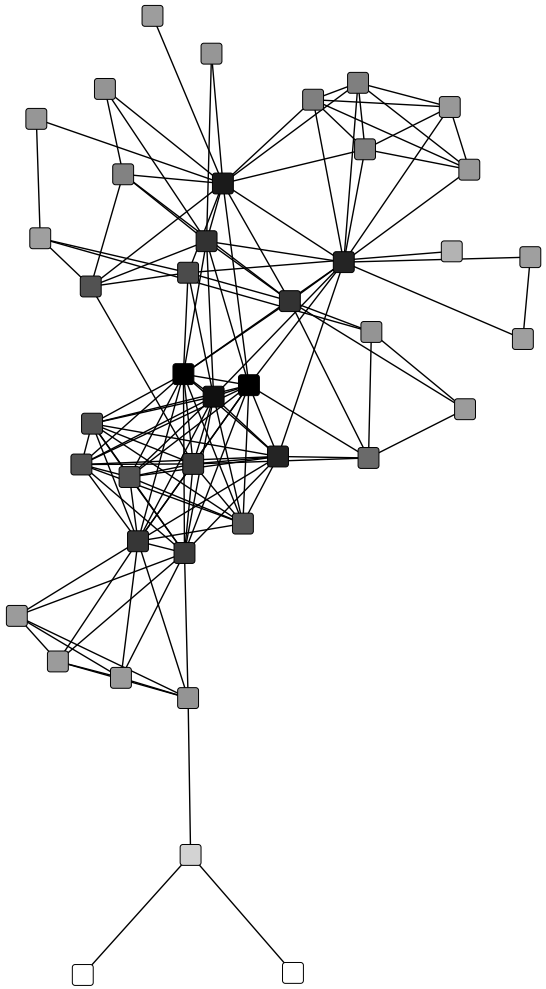
closeness: short distance to alters

current flow betweenness: high throughput of electric current
(also called **information centrality**)

current flow closeness: small potential difference when seen as an electric network

$\Rightarrow$ visone

Shortest path centralities.



betweenness: being on shortest paths between alters

closeness: short distance to alters

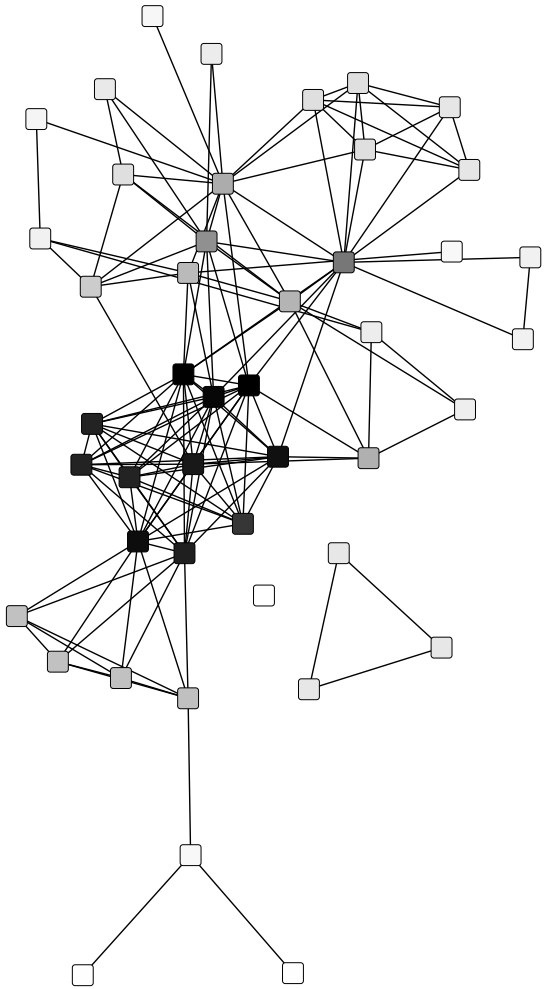
current flow betweenness: high throughput of electric current
(also called **information centrality**)

current flow closeness: small potential difference when seen as an electric network

$\Rightarrow$ visone

Feedback centralities.

Idea: node important if connected to (many) important others.



eigenvector of adjacency matrix

page rank: stable distribution in a random surfer model

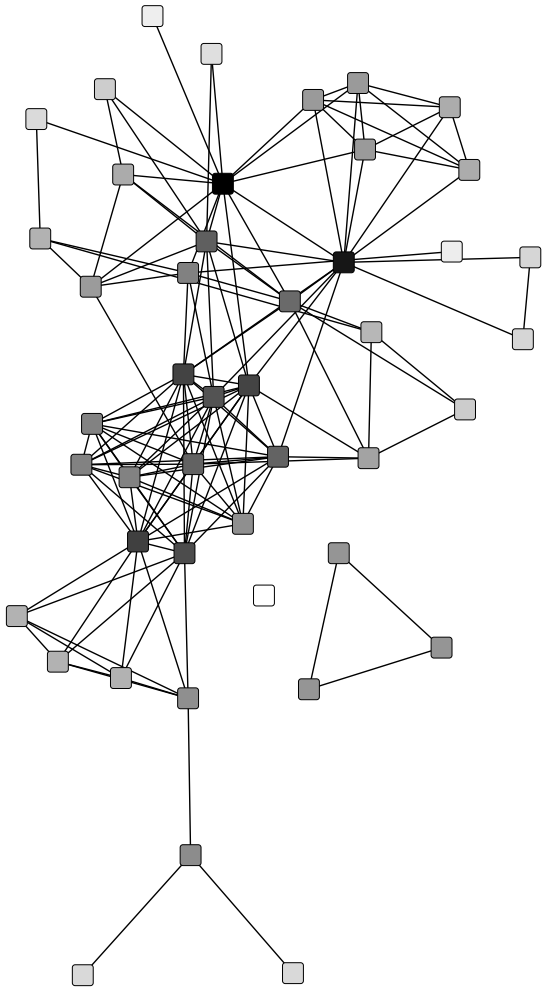
hubs & authorities: a strong hub points to many strong authorities;
a strong authority is pointed at by many strong hubs

Katz' status: nodes give status to nodes they point at – directly or indirectly

$\Rightarrow$ visone

Feedback centralities.

Idea: node important if connected to (many) important others.



eigenvector of adjacency matrix

page rank: stable distribution in a random surfer model

hubs & authorities: a strong hub points to many strong authorities;
a strong authority is pointed at by many strong hubs

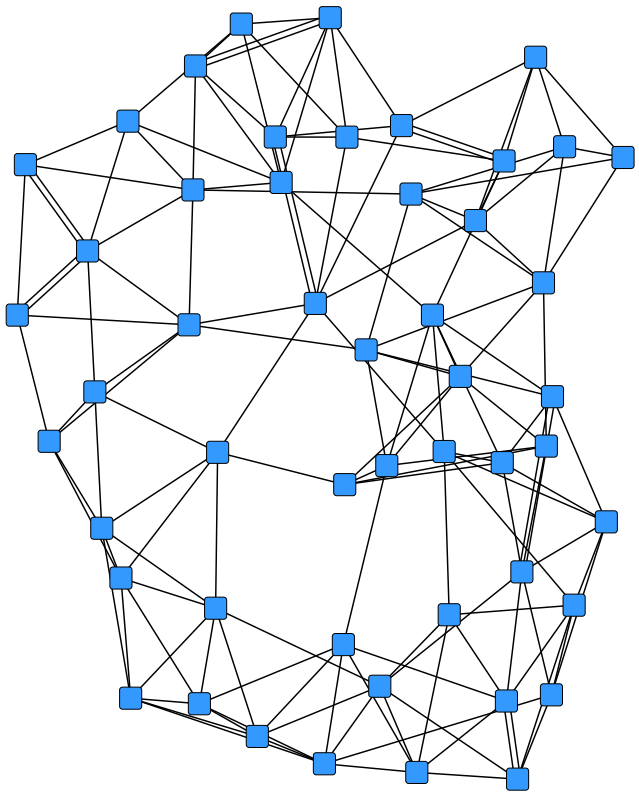
Katz' status: nodes give status to nodes they point at – directly or indirectly

$\Rightarrow$ visone

Intermezzo – network visualization methods.

- **stress minimization** – drawing short paths straight (MDS)
- **spring embedder** – equilibrium of physical forces
- **spectral** – minimizing edge lengths
- **classical MDS** – spectral approach to MDS
- **circular** – arange on circle, minimizing crossings
- **random** – *no comment*

Stress minimization.

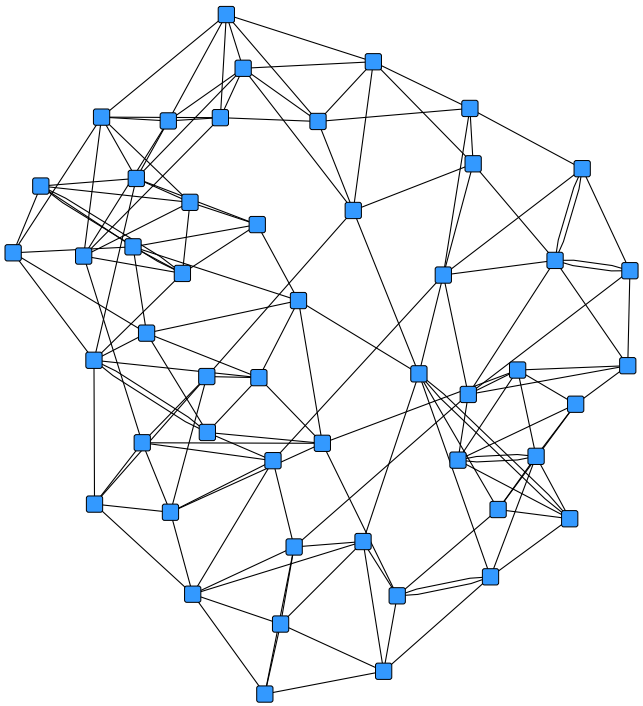


distance between nodes should correspond to graph distance
. . . especially for close nodes;

started by the quick layout button

$\Rightarrow$ visone

Spring embedder.

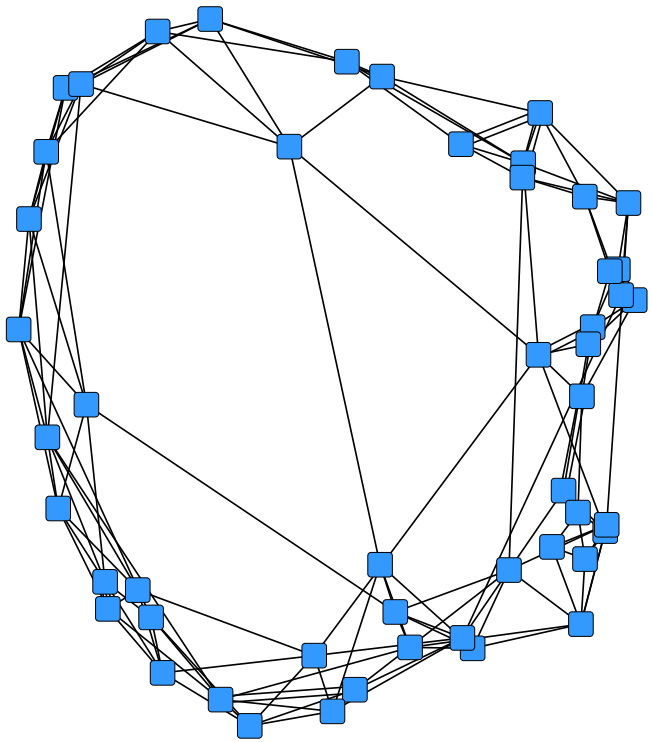


physical analogy: nodes repulse each other and edges have a preferred length (like springs);

layout determined by equilibrium state

$\Rightarrow$ visone

Spectral.



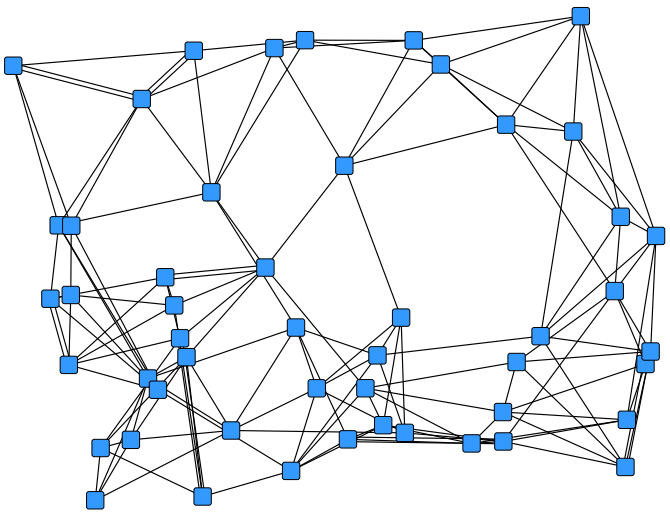
coordinates determined by eigenvectors of graph-matrices;

Laplacian: minimizing sum of squared edge lengths;

adjacency: draw nodes close if connected to the same others (structural similarity);

$\Rightarrow$ visone

Classical MDS.

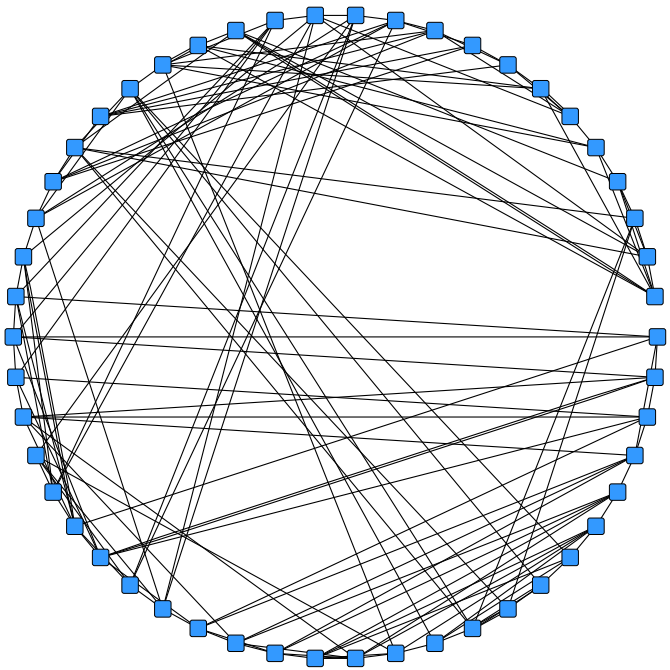


distance between nodes should correspond to graph distance;

best two-dimensional representation (computed by spectral means)

$\Rightarrow$ visone

Circular.



nodes arranged on a circle;

ordered such that crossings are reduced;

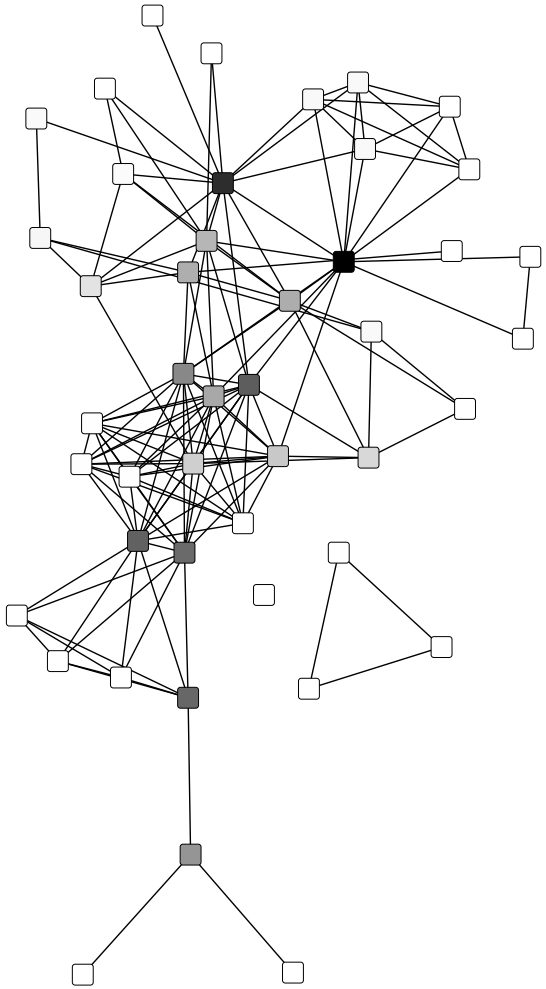
$\Rightarrow$ visone

Outline

1. visone– people, purpose, history.
2. visone basics.
3. **Scene 1** Juan draws a network.
4. **Scene 2** José analyzes a personal network.

$\Rightarrow$ **Scene 3** Jorge explores a random network.

Scene 3 Jorge explores random networks.



Scenario Jorge computed betweenness centrality on some network (dark nodes: more central).

He's wondering whether the observed centralization is higher than one would expect randomly.

$\Rightarrow$ random network generation in visone.

Wrapping up.

- Random networks from various distributions can be generated via the create network option.
  - **(uniform)** just varying density
  - **(preferential** attachment) skewed degree distribution
  - **(small world)** local clustering, small diameter, little degree variance
  - **(planar)** rare for social networks
- Node attributes (e. g., centrality) are computed internally
  (a) mapping of attributes to visual characteristics
  (b) export of attribute tables (`.csv`) for further analysis

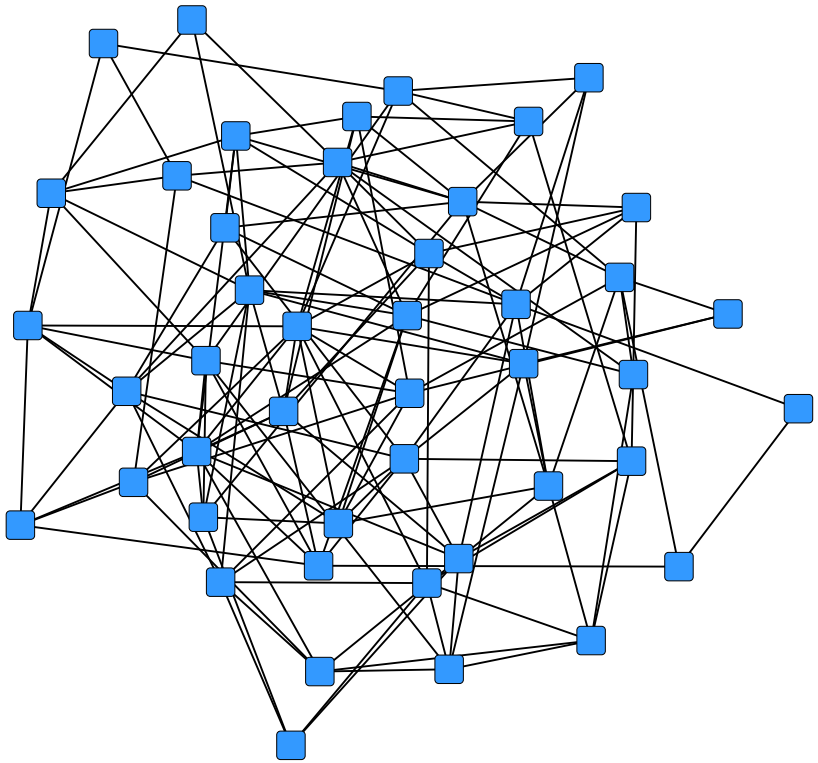
Intermezzo – random graph distributions.

uniform random graphs no structure at all

preferential attachment “the rich get richer”

small world locally clustered, small diameter

Uniform random graphs $G(n, p)$



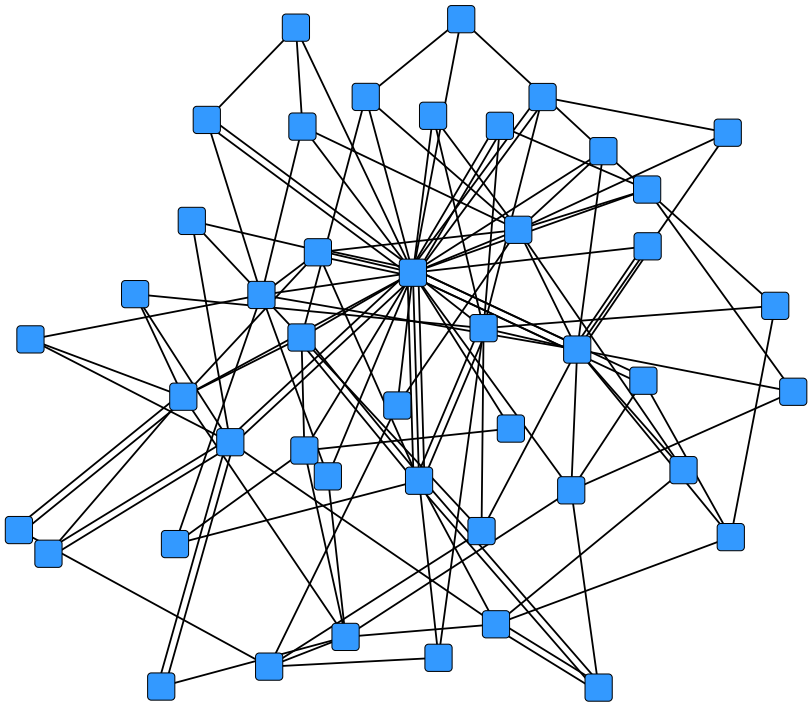
Generating process

1. fix a number n of nodes;
2. independently include edges $\{1, 2\}, \{1, 3\}, \ldots, \{n-1, n\}$ with the same probability p.

Very little variation in degrees, no clusters, small diameter, ...

$\Rightarrow$ visone

Preferential attachment.



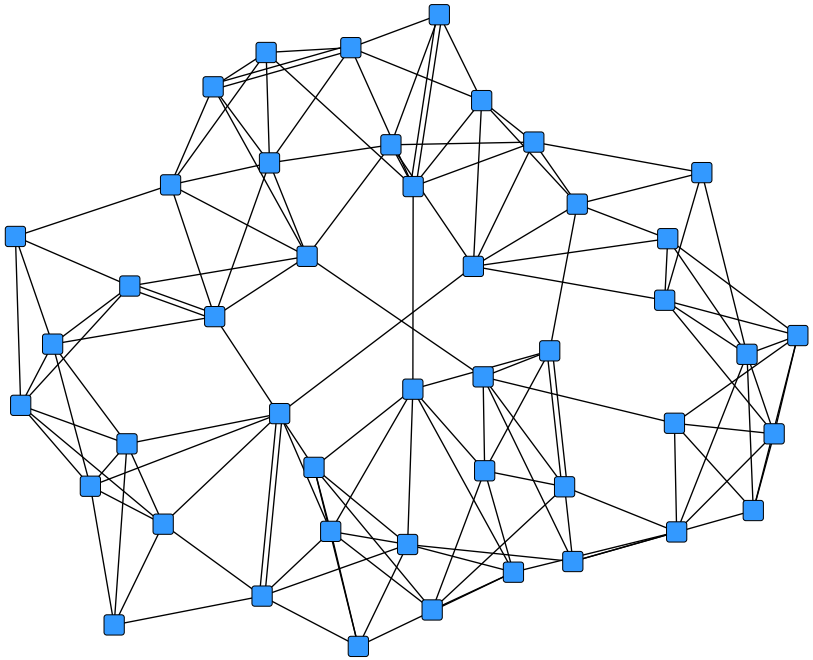
Generating process

1. insert nodes $v_1, \ldots, v_n$ one by one;
2. each node v_i choses d neighbors in $\{v_1, \ldots, v_{i-1}\}$
3. . . . with probability proportional to their current degree.

Enormous variation in degrees, small diameter, . . .

$\Rightarrow$ visone

Small worlds.



Generating process

1. circularly connect nodes with the k next neighbors
2. rewire each edge with probability p

Little variation in degrees, local clustering, small diameter if p is sufficiently large, . . .

$\Rightarrow$ visone

That’s it!