Analysis and visualization with visone

Jürgen Lerner

University of Konstanz

Egoredes Summerschool Barcelona, 21.–25. June, 2010

About visone.

Visone is the Italian word for *mink*. In Spanish *visón*.

visone is a software for the **visual** creation, transformation exploration, analysis, and representation of network data.

Outline.

Introducing the visual graph editor.

Analysis and visualization of networks in visone.

Advanced attribute management.

task assignment

Dynamic networks.

Outline.

visone's visual graph editor.

In this trail you'll learn to use the graphical interface of visone

- how networks can be manually created in visone;
- changing the graphical appearance of nodes and edges;
- how attributes (e. g., age, gender, type of relationship) can be added to actors and ties;
- using templates and creating a legend;
- selecting actors and ties with certain characteristics;
- exporting image files suitable for publication.

Networks in visone.

A network in visone encodes three types of information.

- **Structure**: sets of nodes and edges
  (*who is in the network*, *who is connected to whom*).
- **Attributes**: properties of actors and ties, including
  - external properties such as age, gender, behavior, type of relationship, ...
  - network analytic measures such as centrality, group membership, ...
- **Graphical appearance**: color, shape, coordinates, ...

Cheat sheet: which file format should I choose?

You have to distinguish between

(I) saving/exporting the whole network

- Whenever possible, use **GraphML**. That's the only format which saves everything (structure, attributes, and graphics).
- Export in Ucinet or Pajek format (`dl` or `net`) is supported.
- For many other software, including Excel, SPSS, STATA, R, S-Plus, ... $\Rightarrow$ adjacency matrices in `csv`.
  Note: doesn't save graphics nor attributes.

(II) saving/exporting an image of the network

- If possible use vector-graphics: `pdf`, `eps`, `svg`.
  They give better quality and need smaller file size.
- Otherwise use pixel-based graphics: `png`, `jpg`, ...

Cheat sheet: which file format should I choose?

You have to distinguish between

(I) saving/exporting the whole network

- Whenever possible, use **GraphML.** That's the only format which saves everything (structure, attributes, and graphics).
- Export in Ucinet or Pajek format (`dl` or `net`) is supported.
- For many other software, including Excel, SPSS, STATA, R, S-Plus, ... $\Rightarrow$ adjacency matrices in `csv`.
  Note: doesn't save graphics nor attributes.

(II) saving/exporting an image of the network

- If possible use vector-graphics: `pdf`, `eps`, `svg`.
  They give better quality and need smaller file size.
- Otherwise use pixel-based graphics: `png`, `jpg`, ...

Cheat sheet: which file format should I choose?

You have to distinguish between

(I) saving/exporting the whole network

- Whenever possible, use **GraphML.** That's the only format which saves everything (structure, attributes, and graphics).
- Export in Ucinet or Pajek format (`dl` or `net`) is supported.
- For many other software, including Excel, SPSS, STATA, R, S-Plus, ... $\Rightarrow$ adjacency matrices in `csv`.
  Note: doesn't save graphics nor attributes.

(II) saving/exporting an image of the network

- If possible use vector-graphics: `pdf`, `eps`, `svg`.
  They give better quality and need smaller file size.
- Otherwise use pixel-based graphics: `png`, `jpg`, ...

Cheat sheet: which file format should I choose?

You have to distinguish between

(I) saving/exporting the whole network

- Whenever possible, use **GraphML.** That's the only format which saves everything (structure, attributes, and graphics).
- Export in Ucinet or Pajek format (`dl` or `net`) is supported.
- For many other software, including Excel, SPSS, STATA, R, S-Plus, ... $\Rightarrow$ adjacency matrices in `csv`.
  Note: doesn't save graphics nor attributes.

(II) saving/exporting an image of the network

- If possible use vector-graphics: `pdf`, `eps`, `svg`.
  They give better quality and need smaller file size.
- Otherwise use pixel-based graphics: `png`, `jpg`, ...

Cheat sheet: which file format should I choose?

You have to distinguish between

(I) saving/exporting the whole network

- Whenever possible, use **GraphML.** That's the only format which saves everything (structure, attributes, and graphics).
- Export in Ucinet or Pajek format (`dl` or `net`) is supported.
- For many other software, including Excel, SPSS, STATA, R, S-Plus, ... $\Rightarrow$ adjacency matrices in `csv`.
  Note: doesn't save graphics nor attributes.

(II) saving/exporting an image of the network

- If possible use vector-graphics: `pdf`, `eps`, `svg`.
  They give better quality and need smaller file size.
- Otherwise use pixel-based graphics: `png`, `jpg`, ...

Cheat sheet: which file format should I choose?

You have to distinguish between

(I) saving/exporting the whole network

- Whenever possible, use **GraphML.** That's the only format which saves everything (structure, attributes, and graphics).
- Export in Ucinet or Pajek format (`dl` or `net`) is supported.
- For many other software, including Excel, SPSS, STATA, R, S-Plus, ... $\Rightarrow$ adjacency matrices in `csv`.
  Note: doesn't save graphics nor attributes.

(II) saving/exporting an image of the network

- If possible use vector-graphics: `pdf`, `eps`, `svg`.
  They give better quality and need smaller file size.
- Otherwise use pixel-based graphics: `png`, `jpg`, ...

Exercise: Visone's visual graph editor.

Draw the social network of (some of) the **participants of this course**, as far as you are aware of it.

- Include characteristics of the actors.
- Include different types of relationships.
- Encode the actors' characteristics in graphical variables.
- Don't forget yourself.
- *Don't take this too seriously* ;-)

Export an image of your network.

You might do this in groups of two or three to foster discussion.

Outline.

Analysis and visualization of networks in visone.

In this trail you'll learn how analysis and visualization work hand in hand in visone

- importing networks and attributes from `csv` tables;
- how to layout a network to highlight its structure;
- computing network analytic measures such as centrality or group membership
- how to visualize these;
- exporting the computed values in `csv` tables.

Introduction to the dataset.

Personal networks collected by interviewing $\approx$ 500 migrants in Catalonia and Florida. (http://www.egoredes.net/)

From each respondent (**ego**)

1. **(questions about ego)** country of origin, years of residence, skin-color, health, language skills, ...
2. **(alters)** 45 people known to the respondent
3. **(questions about alters)** origin, country of residence, ...
4. **(ties)** which alters know each other

Here, analyze one personal network of a migrant from the Dominican Republic to the USA.

Overview: analysis of networks in visone.

visone distinguishes between three different levels of analysis.

1. **(indexing)** computation of node-level / edge-level properties.
   - centrality (importance) of nodes and edges;
   - clustering coefficient;
   - distance to selected nodes;

2. **(grouping)** computation of dense subgroups or partitioning the network into groups.

3. **(siena)** interface to the `RSiena` software.
   - data preparation and model specification in visone;
   - visualization of the estimated model and its fit to the observed network;

Cheat sheet: which visualization should I take?

What for?

Layout the network to give a clear picture of its structure.

- Usually: use the quick layout button.
- For comparison with other software, visone offers other algorithms (MDS, spring embedder, spectral, circular, . . .).
- Link routing bends edges to reduce crossings.
- Label placement reduces overlap.

Highlight node and edge properties.

- Attributes (external or network analytic) can be mapped to graphical variables (size, color, label, coordinates, . . .).
- For displaying importance: centrality and status layout.

Geometric transformation. Rotate, reflect, translate, scale, Procrustes analysis.

Cheat sheet: which visualization should I take?

What for?

Layout the network to give a clear picture of its structure.

- Usually: use the quick layout button.
- For comparison with other software, visone offers other algorithms (MDS, spring embedder, spectral, circular, . . .).
- Link routing bends edges to reduce crossings.
- Label placement reduces overlap.

Highlight node and edge properties.

- Attributes (external or network analytic) can be mapped to graphical variables (size, color, label, coordinates, . . .).
- For displaying importance: centrality and status layout.

Geometric transformation. Rotate, reflect, translate, scale, Procrustes analysis.

Cheat sheet: which visualization should I take?

What for?

Layout the network to give a clear picture of its structure.

- Usually: use the quick layout button.
- For comparison with other software, visone offers other algorithms (MDS, spring embedder, spectral, circular, . . .).
- Link routing bends edges to reduce crossings.
- Label placement reduces overlap.

Highlight node and edge properties.

- Attributes (external or network analytic) can be mapped to graphical variables (size, color, label, coordinates, . . .).
- For displaying importance: centrality and status layout.

Geometric transformation. Rotate, reflect, translate, scale, Procrustes analysis.

Cheat sheet: which visualization should I take?

What for?

Layout the network to give a clear picture of its structure.

- Usually: use the quick layout button.
- For comparison with other software, visone offers other algorithms (MDS, spring embedder, spectral, circular, . . .).
- Link routing bends edges to reduce crossings.
- Label placement reduces overlap.

Highlight node and edge properties.

- Attributes (external or network analytic) can be mapped to graphical variables (size, color, label, coordinates, . . .).
- For displaying importance: centrality and status layout.

Geometric transformation. Rotate, reflect, translate, scale, Procrustes analysis.

Cheat sheet: importing networks from `csv` tables.

Import in visone via `open (Adjacency Matrix Files)`.

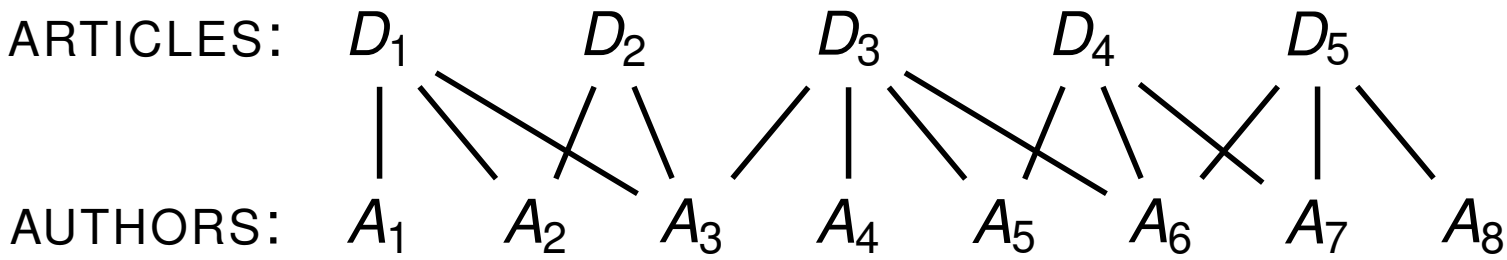
Take care of setting the correct options.

- `network type`: one-mode vs. two-mode network;
- `link attribute type`: numbers or labels;
- `row/column labels`: present or not (in latter case, actors are numbered);
- `cell delimiter`: which character indicates a new entry in the adjacency matrix? (must match export format!);
- `textframe` character to indicate beginning/end of labels (labels might contain cell delimiter);
- `merge empty cells` ignore repeated cell delimiters (e. g., two spaces);

Use the `file view` tab to find out the right options.

Side note: two-mode networks.

Edges connecting actors (objects) of different type.
E. g., membership in organizations, written articles, . . .

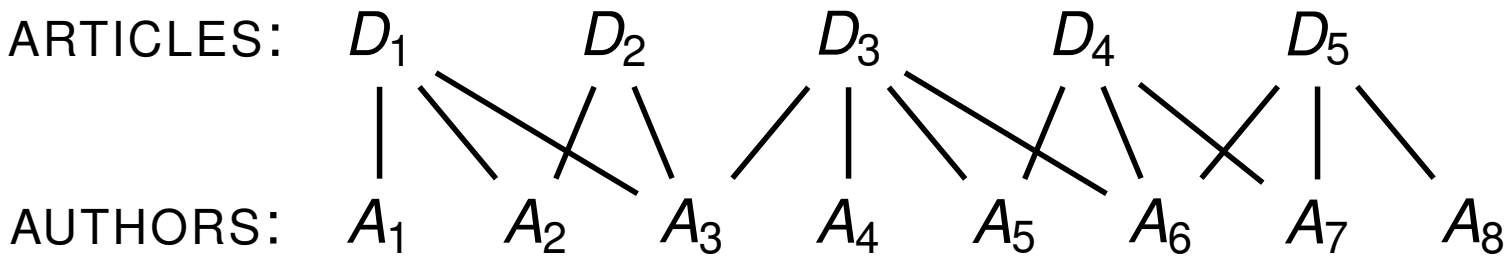


Representation in rectangular matrices.

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
D_1	1	1	1	0	0	0	0	0
D_2	0	1	1	0	0	0	0	0
D_3	0	0	1	1	1	1	0	0
D_4	0	0	0	0	1	1	1	0
D_5	0	0	0	0	0	1	1	1

Side note: two-mode networks.

Edges connecting actors (objects) of different type.
E. g., membership in organizations, written articles, . . .



Representation in rectangular matrices.

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
D_1	1	1	1	0	0	0	0	0
D_2	0	1	1	0	0	0	0	0
D_3	0	0	1	1	1	1	0	0
D_4	0	0	0	0	1	1	1	0
D_5	0	0	0	0	0	1	1	1

Cheat sheet: transformation tab.

Network transformations operate on different levels.

links

- `merge` putting together parallel links;
- `split` links, e. g., by multiplicity;
- `simplify`, modify `direction`, `...`

nodes

- `group by attribute` collapsing actors with common properties;

network

- `two-mode` transformation to one-mode (e. g., co-author, co-citation);
- `link graph`: ties become nodes;

Exercise: analysis and visualization of networks.

Import of adjacency matrices and attribute tables.

- Open the file `egoredes/know_each_other.csv`
- Import node attributes from the file `attributes.csv`

Explore the social environment of this migrant.

- Display the network structure and simultaneously visualize different attributes, e. g.,
  - country of origin (`Afrm`);
  - city in which the actor lives (`Acit`);
  - type of relationship to ego (`Arel`).
- Describe the different clusters. Which characteristics explain ties in this network?

Who is important?

- Compute, visualize, and interpret different centralities.

Outline.

Advanced attribute management.

In this trail you'll learn many things about managing attributes

- further possibilities to select nodes and links;
- types of attributes and their implications;
- operations on attributes;
- use of non-uniform link-strength and link-length in analysis;
- grouping actors by attributes and collapsing groups.

Introduction to the dataset.

NEWCOMB FRATERNITY DATA 15 matrices recording weekly sociometric **preference rankings** from 17 men attending the University of Michigan in the fall of 1956.

(1) best friend
(2) 2nd best friend
...
(16) least friend

Available, e. g., as a Ucinet dataset, or on the Pajek Website.

Coding of the Newcomb fraternity data.

Data represented in 15 adjacency matrices of order 17×17 in files `newfrat01.csv` to `newfrat15.csv`

			j	
	·	12	2	7
i	1	·	**5**	10
	3	5	·	8
	6	1	3	·

Entry in row i and colum j gives the **rank** of j in the list of i's friends (ordered from the best friend to the least).

Row labels and column labels (actor names) are not included in the files—but the ordering is the same in all files.

Exercise: advanced attribute management.

Data input and preparation.

- open `newfrat01.csv`;
- create a new attribute encoding tie strength;
- encode the tie-strength in the thickness of lines.

Popularity of actors.

- Compute in-degree weighted by tie strength dependent on all 16 nominations . . .
- and once again dependent on the top-three nominations.
- Visualize these two popularity scores in one image.
- Does the ordering change? How do you interpret this? What does it tell you about the actors?

Outline.

Task assignment.

Data input and preparation.

- Open the file `egoredes/egonet_original.graphml`
- Look at the edge attributes. Why is this network so dense?
- Make an appropriate choice to delete some of the ties.

Explore the social environment of this migrant.

- Display the network structure and simultaneously visualize different attributes (similar to the second exercise).
- This time additionally look at attributes coding which actors are "important" for this migrant (e. g., frequency of contact, "*how close do you feel to . . .* ?", etc.)

Save one or more network images with which you can describe the social environment, the acculturation level, etc. of this migrant. Use these for your presentation on Friday.

Outline.

Dynamic networks (recently added features of visone).

Here you'll learn about

- visualization of dynamic networks;
- an interface to `RSiena` for modeling network dynamics.

That’s it!