Networks and Graphs

Tom A.B. Snijders

University of Oxford

April 19, 2008

Outline

Graphs

A graph is a set of points (vertices, nodes) V
and a set of edges (lines) E;
the edges are unordered pairs of points.

If $\{i, j\}$ is an edge, i.e., $\{i, j\} \in E$,
then we say that i and j are tied, or related to each other.

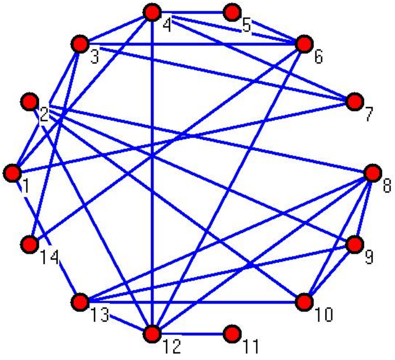
Graphs

A graph is a set of points (vertices, nodes) V
and a set of edges (lines) E;
the edges are unordered pairs of points.

If $\{i, j\}$ is an edge, i.e., $\{i, j\} \in E$,
then we say that i and j are tied, or related to each other.

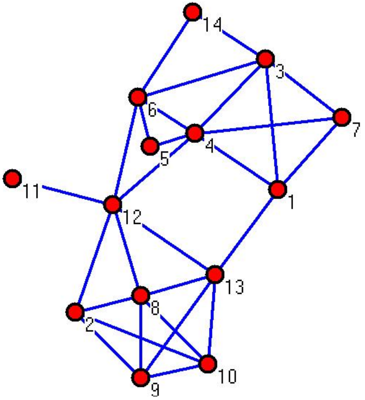
Usually in Social Network Analysis,
the points represent *social actors*;
and the graphs are *non-reflexive*:
$\{i, i\} \notin E$, points are not tied to themselves.

Doreian's network



An example of a graph
of 14 actors, from Doreian
(Social Networks 1988).
The actors are
the main political actors in a
Midwestern county in the US;
a tie means they can expect
public and private support
from each other
and they are also friends.

Doreian's network



This is the same graph!

Graphical representation

is a science and art of itself.

Digraphs

Often there is a directionality inherent in ties:

I like you, but you don't like me...

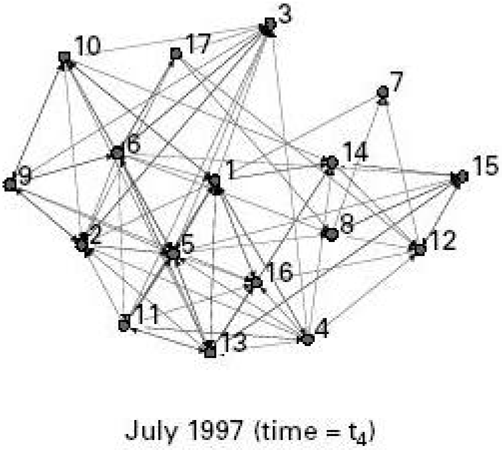
Digraphs

Often there is a directionality inherent in ties:
I like you, but you don't like me...

A directed graph (digraph) is a set of points V
and a set of arcs (directed edges) E;
the edges are ordered pairs of points.

If (i, j) is an arc, i.e., $(i, j) \in E$,
then we say that i is tied (or related) to j.
i is the *sender* of the tie,
j the *receiver*.

Paper Factory Graph



The trust network
of employees
in a paper factory
collected by
Rafael Wittek (1999)
(van de Bunt, Wittek,
de Klepper,
Int. Soc. 2005).

Adjacency matrix

A convenient representation of graphs and digraphs
(we often just say “graphs" when we also refer to digraphs)
is the adjacency matrix:

j is adjacent to *i* if there is a tie from *i* to *j*;
the adjacency matrix is the matrix (y_{ij}) with

$$y_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{if } (i,j) \notin E \end{cases}$$

Adjacency matrix

A convenient representation of graphs and digraphs
(we often just say “graphs" when we also refer to digraphs)
is the adjacency matrix:

j is adjacent to i if there is a tie from i to j;
the adjacency matrix is the matrix (y_{ij}) with

$$y_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E \\ 0 & \text{if } (i,j) \notin E \end{cases}$$

The diagonal of the adjacency matrix will be structurally zero
when there are no self-ties.

Adjacency matrix

The adjacency matrix for Doreian's network is

$$
\begin{pmatrix}
0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

On terminology

In mathematical terminology, a *relation* is the same as a graph.

Therefore we do not like to talk about the relation from i to j, but rather about the *tie* from i to j.

Element y_{ij} of the adjacency matrix is called a *tie variable*.

On terminology

In mathematical terminology, a *relation* is the same as a graph.

Therefore we do not like to talk about the relation from i to j, but rather about the *tie* from i to j.

Element y_{ij} of the adjacency matrix is called a *tie variable*.

The *density* of a graph is
the fraction of pairs of nodes that are tied.

For Doreian's network, there are 28 edges and

$$\frac{14 \times 13}{2} = 91 \text{ unordered pairs, so the density is } \frac{28}{91} = 0.31.$$

Graphs with attributes

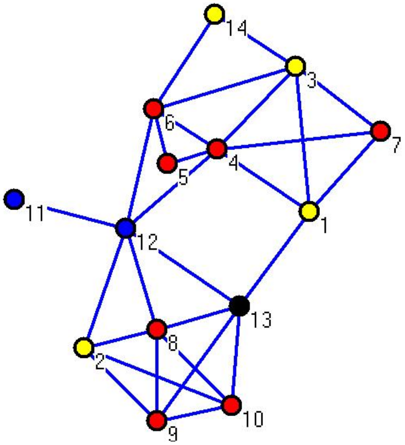
Practically always, there is more of interest in a network than merely the structure represented by the graph.

Often the nodes have attributes, or values;
also the arcs may have attributes or values.

A *valued (di)graph* is a (di)graph where edges have values.
An important instance is the *signed (di)graph*,
where edge (arc) values are – and + .

Categorical attributes often are referred to as *colors*.
The colored network may tell a more extensive story...

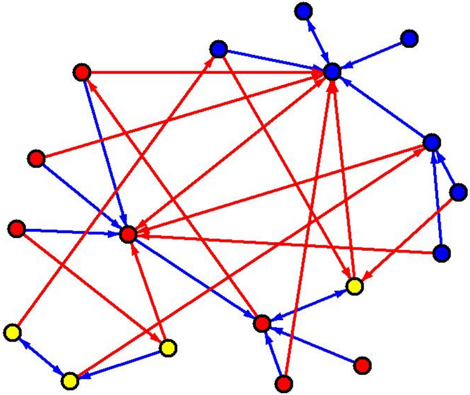
Example: Doreian's network



Colors:

- council members
- officers
- mayor
- former council members

Example: Signed digraph



Sampson's data (1969)
'crisis in a cloister'
at time T4,
"+" blue arcs: best liked
"-" red arcs: least liked.

- young turks
- loyal opposition
- outcasts

Graphs with attributes

Another extension is *multivariate* or *multiplex* networks, where several relations are given on the same set of points; e.g., trust, liking, and cooperation in a work department.

Complete and incomplete networks

Networks where for each pair of actors it is known
if they are tied, are called *complete* or *entire* networks.

The term *complete* is more common,
but invites confusion with the *complete graph*
which is the graph with density 1.

Complete and incomplete networks

Networks where for each pair of actors it is known
if they are tied, are called *complete* or *entire* networks.

The term *complete* is more common,
but invites confusion with the *complete graph*
which is the graph with density 1.

Networks data where for a set of actors,
the outgoing ties are known for each actor,
are called *personal* or *egocentric* networks.
These can be collected through surveys of a large population.

Complete and incomplete networks

Intermediate forms are less common but very important;

e.g., survey members are asked to indicate ties
between all or some of their network members;

or a *snowball sample* is carried out,
where some network members themselves are interviewed.

Network sampling is an interesting area,
in which only little research has been done.

Introductory books on social network analysis

- Stanley Wasserman and Katherine Faust ,
  Social Network Analysis: Methods and Applications.
  Cambridge University Press, 1994.
- Peter Carrington, John Scott, Stanley Wasserman (eds.),
  Models and Methods in Social Network Analysis.
  Cambridge University Press, 2005.
- Alain Degenne and Michel Forsé,
  Introducing Social Networks. Sage, 1999.
- John Scott,
  Social Network Analysis: A Handbook. 2nd edition.
  Sage, 2000.

Books from complementary viewpoints

- Duncan Watts,
  Six Degrees. The Science of a Connected Age.
  W.W. Norton, 2003.
  (A popular account linking work about networks
  by physicists and computer scientists
  with social science work about network analysis.)
- Gabrielle Demange and Myrna Wooders (eds.),
  Group formation in economics. CUP, 2005.
  (A reader of papers on networks
  published in the economics literature.)