Equivalence Concepts for Social Networks

Tom A.B. Snijders

University of Oxford

May 2, 2012

Outline

Block modeling

The idea of block modeling is to bring out
some main features of the network
by dividing (partitioning) the nodes into categories
of ‘equivalent’ nodes.

The big question is what are meaningful types of equivalence.

Block modeling

The idea of block modeling is to bring out
some main features of the network
by dividing (partitioning) the nodes into categories
of 'equivalent' nodes.

The big question is what are meaningful types of equivalence.

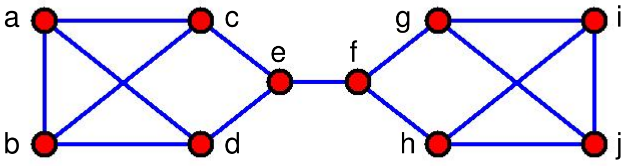
Lorrain and White (1971) defined that
nodes *a* and *b* are structurally equivalent,
if they relate to other nodes in the same way.

The Borgatti-Everett Network

The following is a network

proposed by Borgatti and Everett (1991).

Which nodes are structurally equivalent?

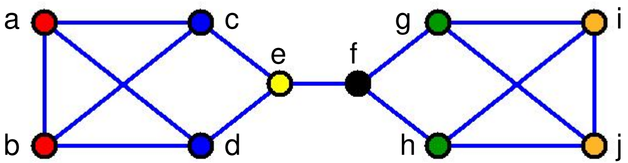


The Borgatti-Everett Network

The following is a network

proposed by Borgatti and Everett (1991).

Which nodes are structurally equivalent?



There are 6 colorings / equivalence classes

Image matrix

For a coloring / partition for a structural equivalence, the *image matrix* is the corresponding adjacency matrix.

The 0 and 1 diagonal entries are meaningful, unless the equivalence class has only one element.

Image matrix

For a coloring / partition for a structural equivalence, the *image matrix* is the corresponding adjacency matrix.

The 0 and 1 diagonal entries are meaningful, unless the equivalence class has only one element.

Usually, the vertices have to be rearranged so that each color indicates a set of successive nodes; then the adjacency matrix shows a *block structure*.

Image matrix

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

The adjacency matrix has a block structure: all blocks are either all – 0 or all – 1.

Adjacency matrix

$$\left(\begin{array}{cc|cc|c|c|cc|cc} . & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & . & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 1 & . & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & . & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & . & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & . & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & . & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & . & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & . & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & . \end{array}\right)$$

Approximate structural equivalence

For most empirically observed networks, hardly any nodes are structurally equivalent.
However, there may be groups of nodes that are *approximately structurally equivalent*.

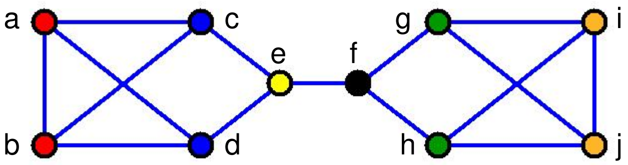
This is elaborated by defining the elements of the image matrix as the *proportion* of ties in the corresponding block of the adjacency matrix;
and striving for an image matrix with elements all of which are as close as possible to either 0 or 1.

As an exercise, you can run
"Operations – Blockmodeling" in Pajek
for Doreian's data set of 14 political actors,
and find approximate structural equivalence classes.

Uncheck the "short report" option,
and ask for 4 classes.
In the output, `com` means 'complete'.

Other equivalences

Here is the Borgatti and Everett (1991) network again:

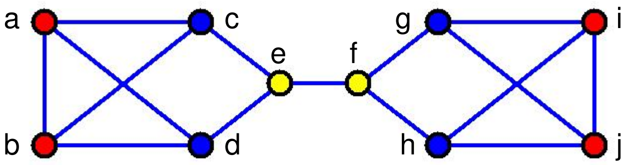


This is the structural equivalence coloring.

Do you see other possibilities of equivalence?

Other equivalences

Here is the Borgatti and Everett (1991) network again:



What about this coloring?

Doesn't is seem also a good representation of equivalence?

Regular equivalence

A coloring is a *regular equivalence*

(Sailer 1978; White and Reitz 1983)

if vertices with the same color

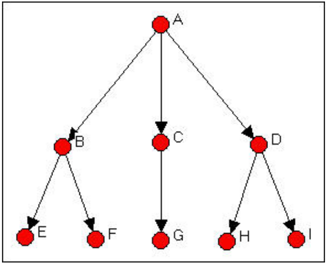
also have neighbors of the same color.

(*a* and *b* are neighbors if they are tied to each other.)

Regular equivalence

A coloring is a *regular equivalence*
(Sailer 1978; White and Reitz 1983)
if vertices with the same color
also have neighbors of the same color.
(*a* and *b* are neighbors if they are tied to each other.)

This definition is a nice mathematical representation
of the sociological concept of *role*:
the color / role determines
to which other colors / roles you should be tied;
what is required is being tied to *some* actors in this role,
not to *all* actors of this role.



Wasserman and Faust (1994)
give the following example.

You can look in Hanneman's
text (Section 15)
for further examples.

One graph can have many different colorings
that all are regular equivalences!

Stochastic equivalence

The classical concepts of equivalence in networks
can be applied to cases of approximate equivalence
by maximizing some measure of adequacy,
that measures how well the observed block structure
corresponds to what would be predicted
in the case of exact equivalence.
All nodes are classified in one of the classes.

Probability models provide another way
to express the deviations between observations and
the idealized concept of ("exact") equivalence.

For a probability distribution of the ties in a graph,
a coloring is a *stochastic equivalence*
(Fienberg and Wasserman, 1981)
if nodes with the same color have
the same *probability distribution* of ties with other nodes.

More formally:
the probability distribution of the graph must
remain the same when equivalent nodes are exchanged.
Such a distribution is called a *stochastic block model*.

The stochastic block model is a kind of
Latent Structure Analysis (LSA).
The basic idea of LSA, proposed by Lazarsfeld & Henry (1968),
is that there exist latent (i.e. unobserved) variables such that
the observations are *conditionally independent*
given the latent structure (= latent variables).

The stochastic block model is a kind of
Latent Structure Analysis (LSA).
The basic idea of LSA, proposed by Lazarsfeld & Henry (1968), is that there exist latent (i.e. unobserved) variables such that the observations are *conditionally independent* given the latent structure (= latent variables).
The *structural model* then specifies the latent variables and the *measurement model* specifies how the observations depend on the latent variables.

The stochastic block model is a kind of
Latent Structure Analysis (LSA).
The basic idea of LSA, proposed by Lazarsfeld & Henry (1968), is that there exist latent (i.e. unobserved) variables such that the observations are *conditionally independent* given the latent structure (= latent variables).
The *structural model* then specifies the latent variables and the *measurement model* specifies how the observations depend on the latent variables.
LSA has been extended to measurement models that specify not conditional independence, but more generally also allow simple, restricted, types of dependence.

The stochastic block model is a latent structure model where the latent structure is the node coloring,
which has to be recovered from the observed network;

for each pair of nodes *i* and *j*, the colors of these nodes determine the probability of a tie
or (for valued / multivariate networks)
of a certain tie configuration between *i* and *j*;
conditional on the coloring, the tie variables are independent.

The stochastic block model is a latent structure model where
the latent structure is the node coloring,
which has to be recovered from the observed network;

for each pair of nodes *i* and *j*, the colors of these nodes
determine the probability of a tie
or (for valued / multivariate networks)
of a certain tie configuration between *i* and *j*;
conditional on the coloring, the tie variables are independent.

This is a ‘rough’ type of network model,
which is useful for bringing out the global structure.

Often we are interested in *cohesive blocks*:

1-blocks on diagonal, 0-blocks off-diagonal:

high within-group density,

low between-group density.

Often we are interested in *cohesive blocks*:
1-blocks on diagonal, 0-blocks off-diagonal:
high within-group density,
low between-group density.

Blockmodeling is, however, much more general:
any difference between probabilities of ties
within and between groups is permitted.

Example of blockmodeling using Pajek:

From

`http://vlado.fmf.uni-lj.si/pub/networks/course/blockmodels.pdf`

pages 8–16.

Literature

- Steve Borgatti and Martin Everett,
  Notions of positions in social network analysis. In P. V. Marsden (Ed.): Sociological Methodology, 1992, pp. 1–35.
- Patrick Doreian, Vladimir Batagelj, and Anuška Ferligoj,
  Positional Analyses of Sociometric Data. In Peter Carrington, John Scott, Stanley Wasserman (eds.),
  Models and Methods in Social Network Analysis. Cambridge University Press, 2005, pp. 77-97.
- R package blockmodeling
- Krzysztof Nowicki and Tom Snijders,
  Estimation and prediction for stochastic blockstructures, *Journal of the American Statistical Association*, 96 (2001), 1077–1087.
- Tom Snijders and Krzysztof Nowicki ,
  Manual for BLOCKS version 1.6. Groningen: ICS, University of Groningen. Available from the StOCNET website.