Today we talk about Borgatti-Everett network model, also called "core-periphery"

In BA we can have nodes with high degree but also low degree.
Nodes with high degree are at the "periphery" of the network

The original paper was published in 2000 (BA in 1999, ER in 1960+)

It all starts from the adjacency matrix, we can write it as 4 blocks

$$M = \begin{pmatrix} \boxed{CC} & \boxed{CP} \\ \boxed{PC} & \boxed{PP} \end{pmatrix}$$

there are some 0s and some 1s

note: the diagonal is full of 0s because no self-links

basically full of 1s

basically full of 0s

In the original paper, CC is full of ones, PP is full of zeroes. PC and CP can be full of one for examples.

Once you select what nodes are in the core, and what nodes are in the periphery, the network is determined, there's no randomness.

If in CP and PC we put some zeroes and some ones, we can introduce randomness.

BA and ER were network models defined bu equations.
BE is a model defined by rules.

How do we know if a network is a core-periphery type of network?

Craig and Von-Peter approach:

$$M = \begin{pmatrix} CC & CP \\ PC & PP \end{pmatrix} = \begin{pmatrix} 1 & RR \\ CR & 0 \end{pmatrix}$$

Row regular: at least one link in each row must be present in this block

column regular: at least one link in each column must be present in this block

Note: in the core, for us, there are not self loops, so it's full of ones but not in the diagonal.

If the adjacency matrix can be schematized like this, we have a core-perypery-like network

Ex

$N = 5$ $\qquad N_c = 2$ $\qquad N_p = 3$

$$M = \left(\begin{array}{cc|ccc} \times & 1 & 1 & 0 & 1 \\ 1 & \times & 0 & 1 & 0 \\ \hline 1 & 0 & \times & 0 & 0 \\ 0 & 1 & 0 & \times & 0 \\ 1 & 0 & 0 & 0 & \times \end{array}\right)$$

this is a core-periphery type of network

We introduce the following "error score"

$$\ell = \frac{\ell_{cc} + \ell_{pp} + \ell_{cp} + \ell_{pc}}{\underbrace{\sum_i \sum_j E_{ij}}}$$

it's the number of ones in the adjacency matrix

We calculate l_{cc}, l_{pp}, l_{cp}, l_{pc}

We start by writing the base adjacency matrix M of the network.
We write a candidate adjacency matrix with the $\begin{pmatrix} 1 & RR \\ CR & 0 \end{pmatrix}$ structure

l_{cc} is the number of differences between the base matrix and the candidate in that region.

The other definitions follow easily.

We try different candidates, and the candidate matrix that produces the lowest error score is the "best fit"

Between candidates, we can change the number of nodes we think are in the core, and the number of nodes we think are in the periphery, and the costruction of CR or RR.

NOTE: The original matrix is reordered, we put the nodes with higher degree in the first positions, before doing any comparing

There's also another metric, made from Everett and Borgatti themselfes to estimate if a network is core-periphery-like, very simple.

There's also another way, much more difficult, that we will not show in detail, made by Holme.

The metric is: a value measured in a certain way for our graph, MINUS the average of that same value measured in n random generated graphs.
This idea is pretty neat, we are not interested in the absolute "value", but in a comparison with an average of that same metric found in other graphs.
We will use a similar idea many times in the future.

The generated networks are generated with having in mind a certain core-periphery structure.

$$\underbrace{V(G)}_{\text{This is what we are really interested in!}} = \underbrace{v(G)}_{\substack{\text{a certain value calculated for the graph G of interest.} \\ \text{too much complicated in this case}}} - \underbrace{avg\left(v(G')\right)}_{\text{the average of the same value but averaged over n random generated G' graphs}}$$

BA gives us a way to generate a model with nodes with high degree and nodes with low degree.
BE can obtain the same results as well!

Mr Ching and Lu were asking themselfes:
BE is a deterministic model or stochastic?

whenever we have a degree distribution that has a tail that goes like a $\underbrace{\text{power-law}}_{k^{-\beta}}$, we have something like this:

$k^{-\beta}$ with $\beta = [2, 3]$

$$d \propto \ln \ln (N)$$

and not like $d \propto \ln(N)$

$N \equiv$ number of nodes

$d \equiv$ the average distance between two nodes

They also showed that, in those conditions, we can find a "core", so a "subgraph", where:

The average shortest path lenght ℓ, in a BA model with N nodes goes like:

$$\ell \sim \frac{\ln(N)}{\ln(\ln(N))}$$

So a BA (since has exponent 3), while it's not core-periphery (since it is "fully" random), it has a core-like structure!

it would be improper way of talking