Сетевые характеристики социального капитала[1]

А.В. Леонидов

ФИАН, МФТИ

[1] M. Jackson, «A Typology of Social Capital and Associated Network Measures» [arXiv:1711.09504]

Разновидности социального капитала

- **Information Capital**: the ability to acquire valuable information and/or to spread it to other people who can use it through social connections.
- **Brokerage Capital**: being in a position to serve as an intermediary between others who wish to interact or transact.
- **Coordination and Leadership Capital**: being connected to others who do not interact with each other, and having the ability to coordinate others' behaviors.
- **Bridging Capital**: being an exclusive connector between otherwise disparate groups, with an ability to acquire, as well as control the flow of, valuable information.
- **Favor Capital**: the ability to exchange favors and safely transact with others through a combination of network position and repeated interaction and reciprocation.
- **Reputation Capital**: having others believe that a person or organization is reliable and/or provides consistently high quality advice, information, labor, goods, or services.
- **Community Capital**: the ability to sustain cooperative (aggregate social-welfare-maximizing) behavior in the running of institutions, the provision of public goods, the handling of commons, and/or collective action, within a community.

Definitions

- There are n individuals indexed by $i \in \{1, 2, ..n\}$.
- A *network* is a graph, represented by its adjacency matrix $\mathbf{g} \in [0,1]^{n \times n}$, where $g_{ij} > 0$ indicates the existence of an edge (a.k.a. link, tie, connection...) between i and j and $g_{ij} = 0$ indicates the absence of a edge.
- Let $G(n)$ denote the set of all admissible networks on n nodes.
- The *degree* of a node i in a network $\mathbf{g}$, denoted $d_i(\mathbf{g}) = \sum_j g_{ij}$, is the number of edges involving node i. In the case of a directed network, this is i's out-degree and indegree is $\sum_j g_{ji}$.

Definitions

- A *walk* between i and j is a succession of (not necessarily distinct) nodes $i = i^0, i^1, ..., i^M = j$ such that $g_{i^m i^{m+1}} = 1$ for all $m = 0, .., M-1$.
- A *path* in $\mathbf{g}$ between two nodes i and j is a succession of distinct nodes $i = i^0, i^1, ..., i^M = j$ such that $g_{i^m i^{m+1}} = 1$ for all $m = 0, .., M-1$.
- Two nodes i and j are connected (or path-connected) if there exists a path between them.
- A *geodesic* (shortest path) between nodes i and j is a path such that no other path between them involves a smaller number of edges.
- The number of geodesics between i and j is denoted $\nu_{\mathbf{g}}(i,j)$. Let $\nu_{\mathbf{g}}(k:i,j)$ denote the number of geodesics between i and j passing through k.
- The *distance* between nodes i and j, denoted $\ell_{\mathbf{g}}(i,j)$, is the number of edges involved in a geodesic between i and j. This is defined only for pairs of nodes that have a path between them and may be taken to be ∞ otherwise.

Definitions

- Let $N_i^\ell(\mathbf{g})$ be the set of individuals at distance ℓ from i in network g: $N_i^\ell(\mathbf{g}) = \{j : \ell(i,j) = \ell\}$.
- Let $N_i(\mathbf{g}) = N_i^1(\mathbf{g})$.
- Node i's degree is $d_i(\mathbf{g}) = |N_i(\mathbf{g})|$.
- Let $n_i^\ell(\mathbf{g}) = |N_i^\ell(\mathbf{g})|$ denote a higher order degree: the number of nodes at distance ℓ from node i.
- Let $clust_i(\mathbf{g})$ denote the clustering of node i: the fraction of pairs of i's neighbors who are connected to each other: $\sum_{kj \in N_i(\mathbf{g}), k<j} \frac{g_{kj}}{d_i(\mathbf{g})(d_i(\mathbf{g})-1)/2}$

- **Information capital** is related to an individual's ability to acquire valuable information as well as to spread it to other people who can use it through social connections. The associated measures of information capital based on networks account for how many people an individual can either send information to or receive it from.
- Let us assume that the chance that information is relayed decays with social distance. The decay of information with distance is captured via a parameter p, where we will generally presume that $0 < p < 1$.
- In addition, we will also include a parameter T that caps the number of times that information is relayed - we may think of this as the information's 'endurance'. This may reflect information becoming stale.

Information centrality

- Decay centrality

$$Dec_i(\mathbf{g}, p, T) = \sum_{\ell=1}^{T} p^{\ell} |N_i^{\ell}(\mathbf{g})|.$$

- Communication centrality. Let $\{p_{ij}\}$ be the matrix of probabilities of information transfer from i to j and $PInf(\mathbf{p}, T)_{ij}$ to be the probability that node j ends up hearing information that starts from node i if it is passed independently with probability $\mathbf{p}_{i'j'}$ from node i' to j' along each walk in the network, and running the whole process for T periods. Then communication centrality is

$$Com_i(\mathbf{p}, T) = \sum_j PInf(\mathbf{p}, T)_{ij}.$$

- Let $EInf(\mathbf{p}, T)_{ij}$ be the expected number of times that j will hear information that starts at node i and is passed according to the matrix $\mathbf{p}$ for T periods

$$EInf(\mathbf{p}, T)_{ij} = \sum_{\ell=1}^{T} \left[\mathbf{p}^{\ell}\right]_{ij}.$$

NB: this quantity includes walks, not paths!

- Then diffusion centrality is

$$Diff_i(\mathbf{p}, T) = \sum_j EInf(\mathbf{p}, T)_{ij} = \sum_j \sum_{\ell=1}^{T} [\mathbf{p}^{\ell}]_{ij}.$$

Valued relationships

- Let $\mathbf{v}$ be the $n \times n$ matrix in which entry v_{ij} is the value that j gets of information that comes from i.
- The corresponding generalisations of the above-introduced centralities read:

$$Dec_i(\mathbf{g}, p, T, \mathbf{v}) = \sum_{\ell=1}^{T} p^{\ell} \sum_{j : j \in N_i^{\ell}(\mathbf{g})} v_{ij},$$

$$Com_i(\mathbf{p}, T, \mathbf{v}) = \sum_{j} PInf(\mathbf{p}, T)_{ij} v_{ij},$$

$$Diff_i(\mathbf{p}, T, \mathbf{v}) = \sum_{\ell=1}^{T} \sum_{j} [\mathbf{p}^{\ell}]_{ij} v_{ij}.$$

Information capital: oriented graphs

- With directed networks, however, there is a clear distinction between sending and receiving. In that case, the definitions above are the appropriate ones for *sending* information. For *receiving* information, one needs to take care to account for paths from other nodes j to reach i, rather than the other way around.
- Decay centrality for oriented graphs:

$$Dec_i^{Rec}(\mathbf{g}, p, T, \mathbf{v}) = \sum_{\ell=1}^{T} \sum_{j: i \in N_j^{\ell}(\mathbf{g})} p^{\ell} v_{ji}$$

- Diffusion centrality for oriented graphs:

$$Diff_i^{Rec}(\mathbf{p}, T, \mathbf{v}) = \sum_{\ell=1}^{T} \sum_{j} [\mathbf{p}^{\ell}]_{ji} v_{ji}.$$

The Godfather Index

- **Information capital** is important as it embodies our ability to be well-informed and to inform others. What information capital misses is the uniqueness of our position in doing so. Can people get the same information via others?
- Two forms of social capital described by the same index:
  - **Brokerage capital** being in a position to serve as an intermediary between others who need to interact or transact.
  - **Coordination capital**, which might also be termed **leadership capital**: being situated as a 'friend-in-common' to others who cannot coordinate their actions directly, and thus being in a position to coordinate others' behaviors.
- Example: Medici were able to coordinate a number of families to provide armed men and act politically, while their key rivals the Albizzi and Strozzi, were not able to coordinate other families to counter the Medici. The Medici's position in the network was clearly different from other families, in particular in terms of serving as a connector of other families.

The Godfather Index

- First candidate: betweenness centrality

$$Bet_i(\mathbf{g}) = \frac{2}{(n-1)(n-2)} \sum_{(j,k):j\neq i\neq k\neq j} \frac{\nu_{\mathbf{g}}(i:j,k)}{\nu_{\mathbf{g}}(j,k)}.$$

- Drawbacks:
  - gives equal credit to all intermediaries on a shortest path, regardless of how many intermediaries there are on the path;
  - weights all paths equally, even though nodes at great distances are much less likely to interact via some long chain of intermediaries than nodes who are fairly close to each other
- Possible solution: weighting with a decreasing function $f(l)$:

$$\sum_{(j,k):j\neq i\neq k\neq j} f(\ell_{\mathbf{g}}(j,k)) \frac{\nu_{\mathbf{g}}(i:j,k)}{\nu_{\mathbf{g}}(j,k)}.$$

The Godfather Index

- The basic idea is to measure brokerage and coordination capital by the *number* of pairs of a person's friends who are not friends with each other

$$GF_i(\mathbf{g}) = \sum_{k>j} g_{ik} g_{ij} (1 - g_{kj}) = |\{k \neq j : g_{ik} = g_{ij} = 1, g_{kj} = 0\}|.$$

- The godfather index has an inverse relationship to clustering, but weighted by the number of pairs of a node's neighbors.

$$GF_i(\mathbf{g}) = (1 - clust_i(\mathbf{g})) d_i(\mathbf{g})(d_i(\mathbf{g}) - 1)/2.$$

- In a directed and/or weighted case, the measure becomes:

$$GF_i(\mathbf{g}) = \sum_{k>j : g_{kj} = g_{jk} = 0} g_{ki} g_{ji}$$

The Godfather Index

- Godfather Index does not account for the fact that a pair of i's neighbors j and k might also have some other friend in common who can serve as an intermediary, therefore uniqueness is important.
- One possibility: divide by the number of intermediaries

$$|\{k,j \in N_i(\mathbf{g}) : \ell(k,j) = 2, \nu_{\mathbf{g}}(k,j) = 1\}|,$$

- Another possibility: give no value to connectors of j and k if there is more than connector ("ego betweenness centrality"):

$$\sum_{k,j \in N_i(\mathbf{g}) : j \neq k, \ell(k,j)=2} \frac{\nu_{\mathbf{g}}(i:j,k)}{\nu_{\mathbf{g}}(j,k)}.$$

Bridging capital

- **Bridging capital**: having a position as a unique or vital connector between otherwise disparate groups, with an ability to acquire and control the flow of valuable knowledge.
- Let $\mathbf{p} - ij$ indicate the matrix $\mathbf{p}$ with entry ij set to 0. The criticality or importance as a connector of link ij is

$$Crit_{ij}(\mathbf{p}, T, \mathbf{v}) = \sum_{i'j'} v_{i'j'} (EInf(\mathbf{p}, T) - EInf(\mathbf{p} - ij, T))_{i'j'}.$$

- The bridging capital of a node i is

$$Brid_i(\mathbf{p}, T) = \sum_j Crit_{ij}(\mathbf{p}, T) = \sum_j \sum_{i'j'} v_{i'j'} \left(\sum_{t=1}^{T} \mathbf{p}^t - (\mathbf{p} - ij)^t \right)_{i'j'}.$$

Favor, reputation and community capital

- **Favor capital**: the ability of an individual to exchange favors and safely transact with others through a combination of network position and repeated interaction. Can be measured via support, counting the number of friends that a person has who are supported by a friend in common:

$$Supp_i = |\{j \in N_i(\mathbf{g}) : [\mathbf{g}^2]_{ij} > 0\}|.$$

- **Reputation capital** , in the form of the beliefs of others for how a given individual or organization will perform in producing a good or service, can be a very valuable form of capital
- **Community capital** to be the ability to sustain cooperative (aggregate social-welfare-maximizing) behavior in transacting, the running of institutions, the provision of public goods, the handling of commons and externalities, and/or collective action, within a community.