

Network Analysis and International Trade Flows

Introduction to Graphs and Networks

Luca De Benedictis¹

¹University of Macerata

Lezione 2

Part I

Graph Theory and Network Analysis

Graph

$$\mathcal{G} = (\mathcal{V}, \mathcal{L})$$

unit: **vertex**, $\text{node} \in \mathcal{V}$

link: $\text{line} \in \mathcal{L}$

directed (*arc*) $\in \mathcal{A}$;

undirected (*edge*) $\in \mathcal{E}$;

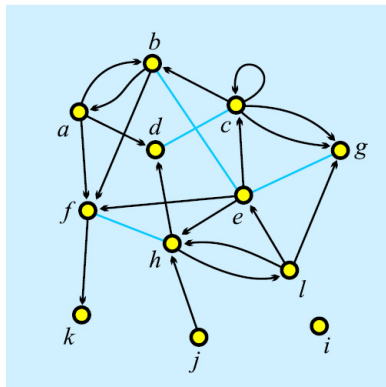
$$\mathcal{L} = \mathcal{A} \cup \mathcal{E}$$

$\mathcal{L}$ is a binary set $[0,1]$

dimensions:

$$n = |\mathcal{V}|, m = |\mathcal{L}|$$

example: $n = 12, m = 23$.



Note: We will analyze networks that do not contain *loops* ($c - c$) and *parallel arcs* ($c - g$).

Network: Graph + Data

$$\mathcal{N} = (\mathcal{V}, \mathcal{L}, \mathcal{P}, \mathcal{W})$$

- ▶ $\mathcal{P}$ is the *vertex value function*: (exogenous or endogenous) properties of vertices
- ▶ $\mathcal{W}$ is the *line value function*: (exogenous or endogenous) weights of lines

Three kind of networks:

- ▶ if $\mathcal{A} = 0 \longrightarrow \mathcal{N}$ is an **undirected** network;
- ▶ if $\mathcal{E} = 0 \longrightarrow \mathcal{N}$ is a **directed** network;
- ▶ if $\exists \mathcal{W} \longrightarrow \mathcal{N}$ is a **weighted** network.

Structural properties: Degrees, Neighborhood and Density

- ▶ **Indegree** and **outdegree**: is the number, **d**, of arcs received and sent by a vertex.
- ▶ **Neighborhood**: the set of vertices that are connected to any given $\mathcal{V}_i \in \mathcal{V}$ defines its *neighborhood* $\mathcal{V}_i^d \equiv \{j \in \mathcal{V} : ij \in \mathcal{L}\}$
 - ▶ A network is said to be *regular* if every vertex has the same number of links, $\mathcal{V}_i^d = \mathcal{V}^d$.
 - ▶ A *complete* network, $\mathcal{N}^c$, is a regular network in which $d = n - 1$.
- ▶ **Density**: is the number of lines in a network over the maximum possible number of lines
 - ▶ In a directed network $\gamma = \frac{m}{m_{\max}} = \frac{m}{n(n-1)}$

Structural properties: Centrality (1)

- Two definitions of centrality: **Closeness centrality** and **Betweenness centrality**.

The **Closeness centrality** of a vertex $\mathcal{V}_i$ is the number of other vertices divided by the sum of all distances between $\mathcal{V}_i$ and all others $\mathcal{V}_{j \neq i}$.

$$C_i^c = \frac{n-1}{\sum_{j \neq i} \delta_{ij}}. \quad (1)$$

For the whole network, if C_i^c* is the centrality of the vertex that attains the maximum closeness centrality score, the degree of **Closeness centralization** of a network is

$$C^c = \frac{\sum_{i=1}^n |C_i^c - C_i^c*|}{(n-2)(n-1)/(2n-3)}. \quad (2)$$

The closeness centralization is the variation in the closeness centrality of vertices divided by the maximum variation in closeness centrality scores possible in a network of the same size.

Structural properties: Centrality (2)

The **Betweenness centrality** of vertex $\mathcal{V}_i$ is the proportion of all geodesic distances between pairs of other vertices that include this vertex:

$$C_i^b = \sum_{j \neq k} \frac{\delta_{jk}^i}{\delta_{jk}} \quad (3)$$

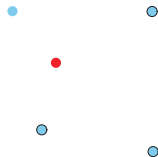
where δ_{jk} is the total number of shortest paths joining any two vertices $\mathcal{V}_k$ and $\mathcal{V}_j$, and δ_{jk}^i is the number of those paths that not only connect $\mathcal{V}_k$ and $\mathcal{V}_j$, but also go through $\mathcal{V}_i$.

The **Betweenness centralization** is the variation in the betweenness centrality of vertices divided by the maximum variation in betweenness centrality scores possible in a network of the same size.

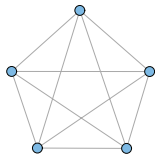
$$C^b = \sum_{i=1}^n |C_i^b - C_i^{b*}|. \quad (4)$$

Some families of Graphs

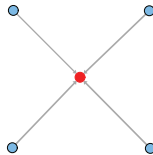
Graphs:



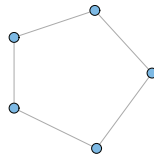
Empty Graph₁



Complete Graph₁



Star Graph₁



Circle Graph₁

Properties:

Density=0
Closeness=0
Betweenness=0

Density=1
Closeness=0
Betweenness=0

Density=0.4
Closeness=1
Betweenness=1

Density=0.5
Closeness=0
Betweenness=0

The properties of vertices are not necessarily the ones of the entire network. In this complete network all vertices have closeness centrality equal to 1, and betweenness centrality equal to 0.

Part II

The World Trade Network

Our bilateral trade flows dataset

Our dataset:

- ▶ From **Subramanian and Wei (2007)**, “The WTO promotes trade, strongly but unevenly”, *JIE*, 72, 151-175.
 - ▶ **Aggregate bilateral imports** as reported by the importing country and measured in U.S. dollars, reported in the *IMF Direction of Trade Statistics*;
 - ▶ in the first part of the analysis we transformed the data in a **dichotomous variable**.
- ▶ **Time span** from 1950 to 2000.
- ▶ Observations for 60 to 157 **countries**.
- ▶ **Geographical partition**: macro continental regions (WB classification).
- ▶ **Institutional partition**: WTO membership

Properties of World Trade Data

Trade flows intensities

	1950	1960	1990	2000
Countries reporting trade flows	60	113	145	157
Total number of flows	1649	3655	10289	11938
Value of total imports (mln US dollar at const. prices)	1585.04	3205.92	22217.38	34100.35
Countries making up 50% of trade	23	26	23	24
Flows making up 50% of trade	46	70	68	78
Countries making up 90% of trade	57	88	82	82
Flows making up 90% of trade	340	634	748	855
share of world imports accounted for by the top 10% flows (number of flows)	77.52 (165)	81.50 (365)	93.16 (1029)	93.45 (1194)
share of world imports accounted for by the top 1% flows (number of flows)	29.25 (17)	37.34 (36)	58.64 (103)	58.17 (119)
Average of export markets	27.95	32.65	70.96	76.04
Median of export markets	24	25.5	60	67
Average of import markets	29.45	38.88	74.02	78.54
Median of import markets	27	33	66	71.50
Correlation between tot.import value and flows by country	0.68	0.69	0.54	0.53
Correlation between tot.export value and flows by country	0.58	0.60	0.56	0.55

Source: our elaboration on S-W data.

The World Trade Network properties

The **evolution** of the World Trade Network (1950-2000)

	1950	1960	1970	1980	1990	2000
No. Countries	60	113	130	143	145	157
No. Arcs	1649	3655	6593	8180	10289	11938
Density	0.4658	0.2888	0.3931	0.4028	0.4928	0.4874
In-Degree Closeness Centralization	0.5260	0.6005	0.5647	0.5801	0.5108	0.5159
Out-Degree Closeness Centralization	0.4743	0.5464	0.5100	0.4383	0.4688	0.4836
In-Degree St.Dev.	14.132	24.024	30.7904	37.0516	37.486	39.0725
Out-Degree St.Dev.	15.550	26.307	31.9825	32.8688	35.8637	41.4161
Betweenness Centralization	0.0417	0.0627	0.0355	0.0317	0.0161	0.0158

Source: our elaboration on S-W data.

The World Trade Network properties

The **evolution** of the World Trade Network (1960-2000)
within a constant subset of countries

	1960	1970	1980	1990	2000
No. Countries	113	113	113	113	113
No. Arcs	3655	5807	6522	7355	6964
Density	0.289 [*]	0.459 [*]	0.515 [*]	0.581 [*]	0.550 [*]
In-Degree Closeness Centralization	0.6005	0.5190	0.4800	0.3866	0.3547
Out-Degree Closeness Centralization	0.5464	0.4920	0.3809	0.3776	0.3547
In-Degree St.Dev.	24.02	26.16	30.01	28.04	28.54
Out-Degree St.Dev.	26.31	28.78	25.91	27.84	30.72
Betweenness Centralization	0.0627	0.0308	0.0155	0.0097	0.0065

Note: Here the network and its indices are computed including only the group of countries for which data are available over the entire time span 1960-2000.

[*] indicates that the density is significantly different from the null hypothesis of $\gamma=1$ with $p=0.0002$.

Source: Our elaboration on S-W data.

The World Trade Network in 1950

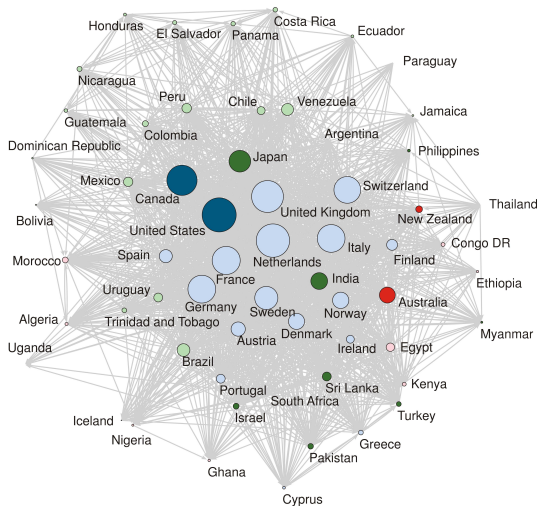


Figure: The size of each vertex is given by its betweenness index.

The World Trade Network in 2000

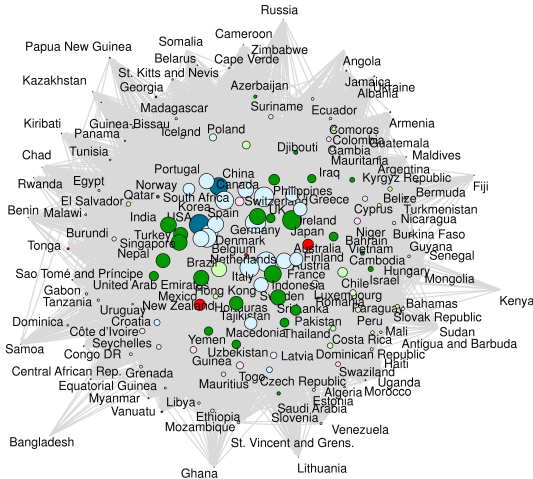
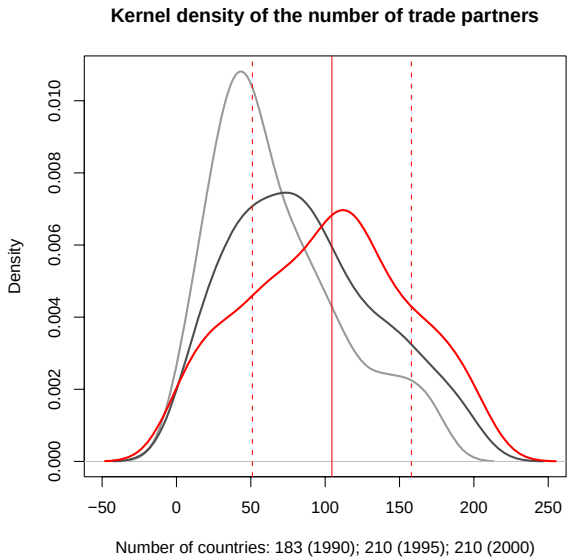


Figure: The size of each vertex is given by its betweenness index.

Correlation between GDP per capita and Betweenness centrality = 0.59; the one between GDP and Betweenness centrality = 0.75

Trade partners in 1990 - 1995 - 2000



Changes in the network structure

From the network indices we observe that

- ▶ **Density increases over time** and the increase is statistically significant $\rightsquigarrow$ countries are more tightly interconnected
- ▶ **Centralization decreases over time** $\rightsquigarrow$ less polarized trading system (more 'globalization' in the proper sense of the world)
- ▶ **Average degree increases (slowly) over time** $\rightsquigarrow$ more trading partners but still selection
- ▶ Differences in distribution of **export (outdegree) and import (indegree)** $\rightsquigarrow$ changes in the two flows due to different causes.
- ▶ **Structural (geodesic) distance** often not related to geographical distance.

Other findings

Apart from the increase in density, different pattern between exports and imports, and decrease in centrality, applications of NA to trade flows found that the WTN has the following **properties**:

- ▶ **Scale-free degree distribution** (high heterogeneity)
- ▶ **Small-world property** (average path length between any pair grows logarithmically with system size)
- ▶ **High clustering** (neighbors are likely to be connected)
- ▶ **Disassortativity** (central countries are more related with peripheral countries)
- ▶ **Degree-degree correlation** (hierarchical organization)

All this evidence on the topological properties of the WTN has implications for trade theory and trade empirics.