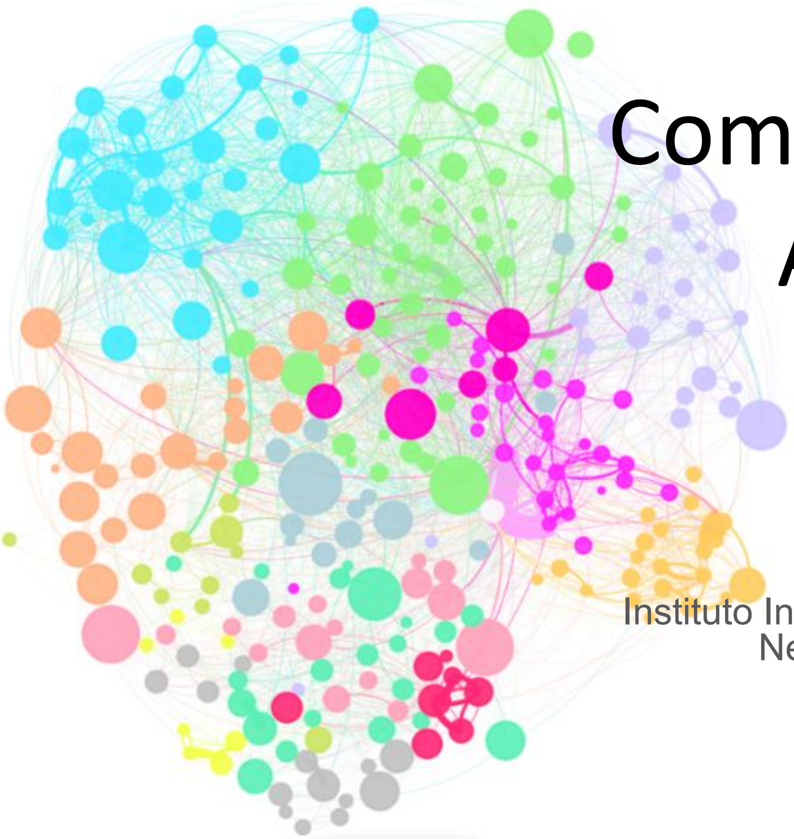


Complex Economic Networks: Analysis, Applications and Data

Viktoriya Semeshenko

Instituto Interdisciplinario de Economía Política (IIEP-UBA-CONICET)
Networks and Complex Systems Laboratory (**Netlab**)

vika.semeshenko@gmail.com
@vsemesh @Netlab_IIEP

Class 2. Economic Networks: Labor Mobility

Sources used in the Source:

- Barabási: Albert-László Barabási. Network Science
- Menczer et al: Filippo Menczer, Santo Fortunato, Clayton A. Davis, A First Course in Network Science. Cambridge University Press 2020

Our Story: Labor Mobility and Network Science

- Labor mobility: job switchings between firms of different industries
  - underlying dynamic and structural aspects of the labor market
- Individuals change jobs
  - capabilities and knowledge exchange between industries
  - local productive structure
- Where does this information comes from?
  - Data: administrative records of private formal employment AFIP
- **What do we want to know and How?**

Our Story: Labor Mobility and Network Science

- ➢ Extract structure of labor interactions and characterize the network
- ➢ What metrics do we need for this?
- ➢ How informative are the derived networks?

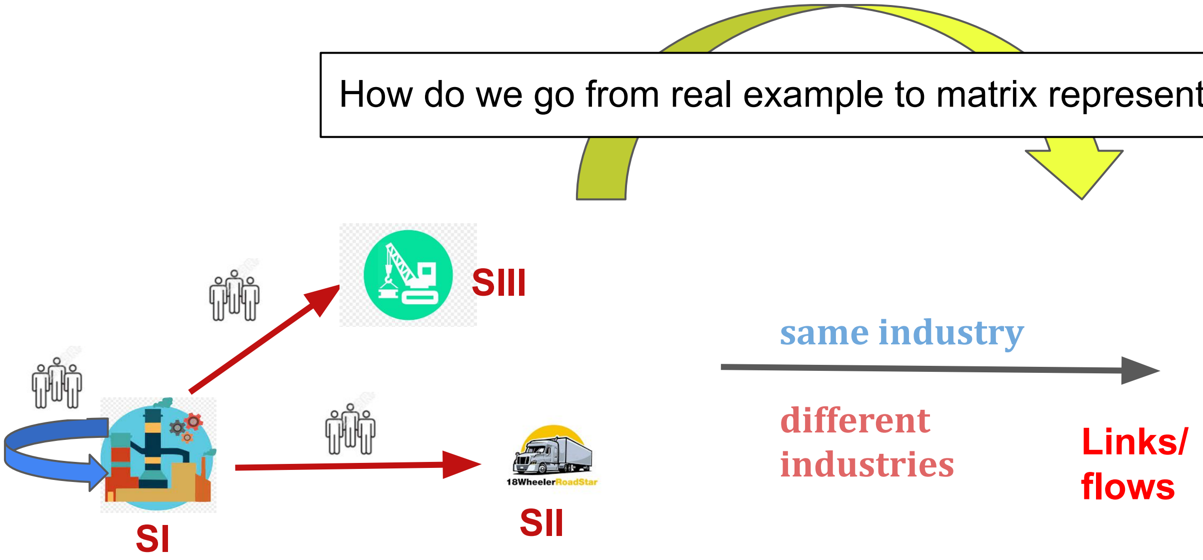
THE ROLE OF NETWORKS

- ➢ Behind each system studied in complexity there is an intricate wiring diagram, or a network, that defines the interactions between the component

- ➢ We will never understand complex system unless we map out and understand the networks behind them

Our Story: Labor Mobility transitions

How do we go from real example to matrix representation?



Nodes/
Sectors

t \ t+1	I	II	III
I			
II			
III			

employment transition
/ adjacency matrix

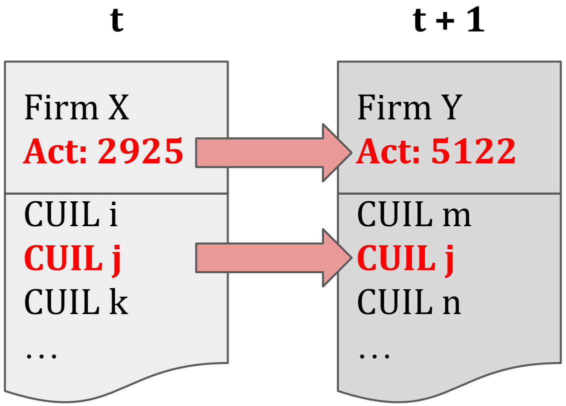
underlies the need to obtain accurate maps of the network

Understanding Data

- **Number** of private formal employment **transitions** between companies of different economic activities (CIIU Rev 3) at 4 digits
    - The International Standard Industrial Classification of All Economic Activities (ISIC)
  - Flows from panels of individuals
  - Transition **Series** 1996-2020 (large)
    - year-to-year (observed, **T=24**)

DDJJ firms
(monthly records)

labor switch, individual j



Network Representation and information

A. Transition

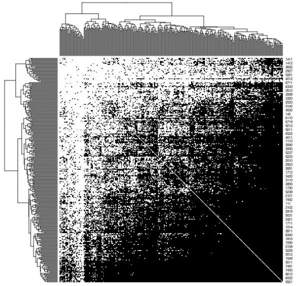
t \ t+1	I	II	III
I			
II			
III			

B. List of edges

from	to	weight
I	I	100
I	II	80
II	I	70
II	III	150

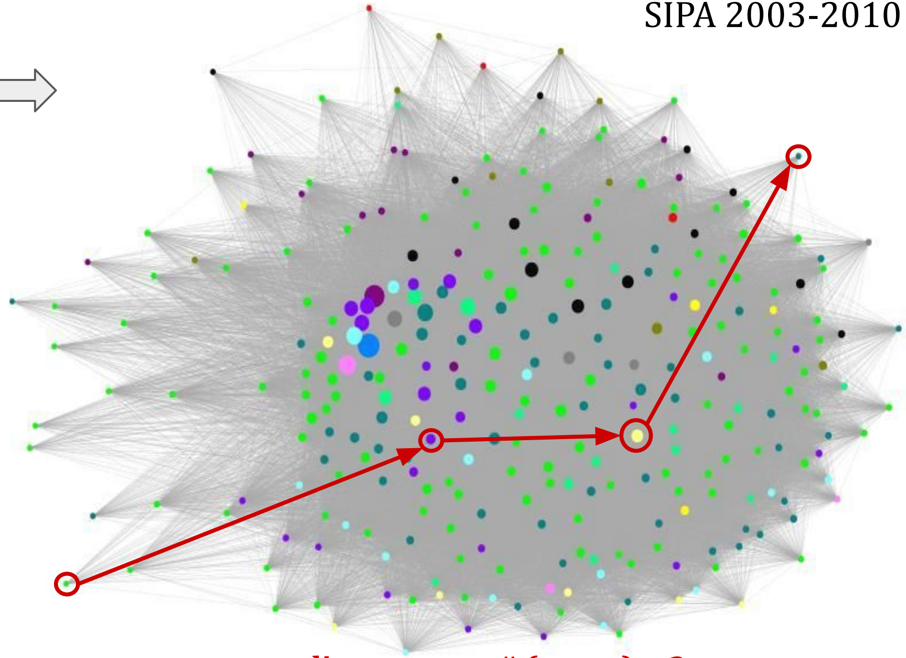
SIPA

C. Heatmaps



D. Grafos

Interindustry
Labor Flows
SIPA 2003-2010



We need a bit more of knowledge: going back to social networks

Many types of networks display a few fundamental features.

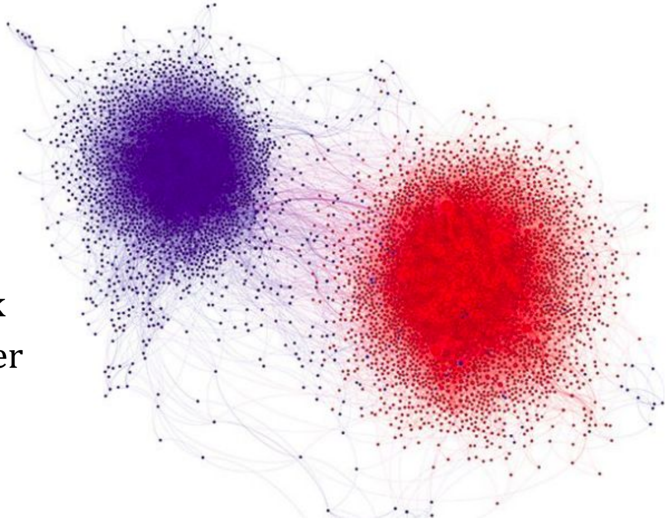
We need **three** characteristics: similarity between neighbors, short paths connecting nodes, and clustering formed by common neighbors

Social networks provide us with familiar cases to illustrate these features, because they are the most extensively studied category of networks.

Assortativity: “birds of a feather flock together”

- In a social network, nodes may have many properties, such as age, gender identity, ethnicity, sexual preference, location, topics of interests, and so on.
- Often, nodes that are connected to each other in a social network tend to be similar in their features: **assortativity**

Image: A retweet network on Twitter, among people sharing posts about US politics. Links represent retweets of posts that used hashtags such as #tcot and #p2, associated with conservative (red) and progressive (blue) messages, respectively, around the 2010 US midterm election. When Bob retweets Alice, we draw a directed link from Alice to Bob to indicate that a message has propagated from her to him. The direction of the links is not shown.



Network Assortativity

- Assortativity based on degree is called *degree assortativity* or *degree correlation*: high-degree nodes tend to be connected to other high-degree nodes, while low-degree nodes tend to have other low-degree nodes as neighbors. The corresponding networks are called assortative.
- Assortative networks have a *core–periphery* structure
- Networks where high-degree nodes tend to be connected to low-degree nodes and vice versa are called *disassortative*
- Pearson's correlation coefficient is a

common way to measure correlation;

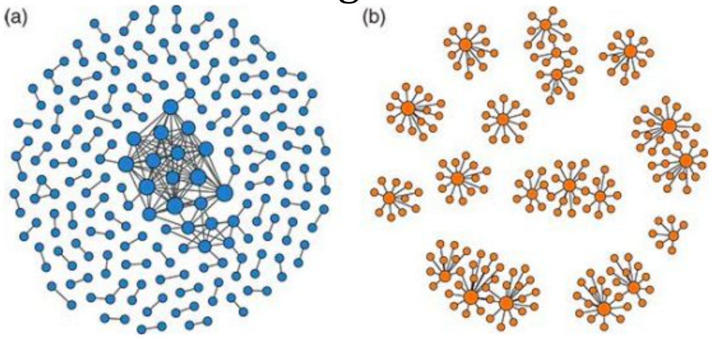


Fig. 2.2 Network degree assortativity illustrated by (a) an assortative network and (b) a disassortative network.

Paths and Distances

- In networks **distance** is a challenging concept. A physical distance is replaced by *path length*. A path is a route that runs along the links of the network. A *path's length* represents the number of links the path contains
- The shortest path between nodes *i* and *j* is the path with the fewest number of links
- The diameter of a network, denoted by d_{max}, is the maximum shortest path in the network (the largest distance recorded between any pair of nodes)
- The average path length, denoted by $\langle d \rangle$, is the average distance between all pairs of nodes in the network

Social Distance

- The average path length, characterizes how close or far we expect nodes to be in a network
- Intuitively, in a grid-like network like road networks and power grids, paths can be long
- Is this typical of many real-world networks?
- →Consider a few social networks, in which this question has been explored extensively

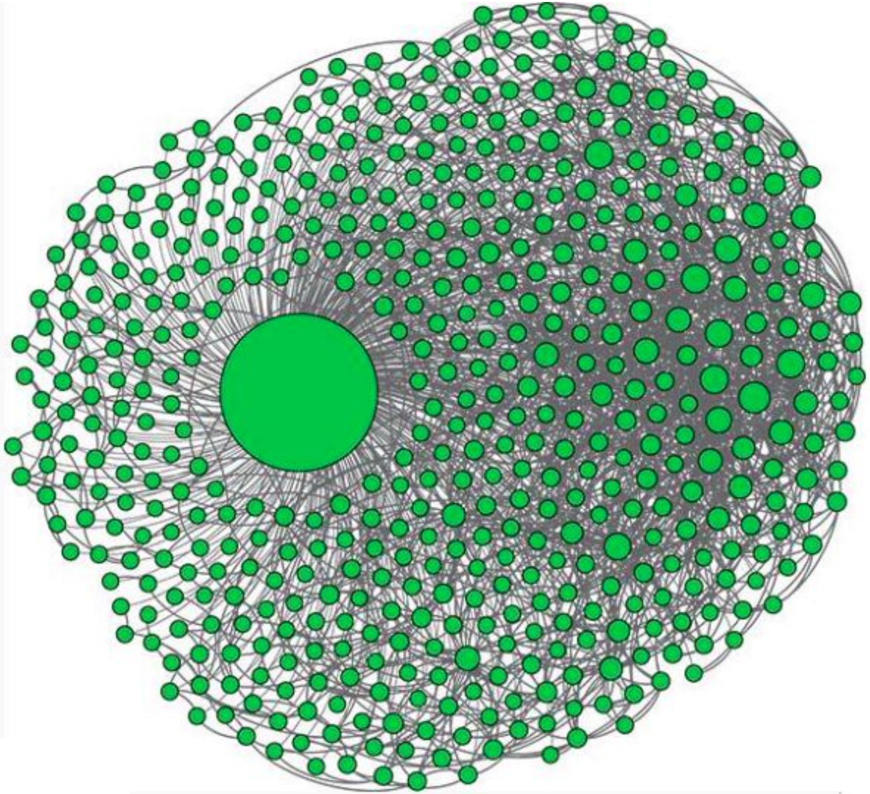
Coauthorship networks

- Nodes are scholars, when a publication is co-authored by two or more scholars, we can infer links between them in the network.
- Paul Erdős, a mathematician, made critical contributions to network science.
- Mathematicians are fond of studying their distance in the coauthorship network from the particular node corresponding to Erdős. They call this distance their *Erdős number.*
- Many mathematicians have a very small Erdős number.
- There is even an online tool to compute the Erdős number for mathematicians (www.ams.org/mathscinet/collaborationDistance.html)

Erdős collaboration network

Image: illustrates the network of collaborations involving Erdős and his over 500 coauthors. In reality, scholars are not just close to Erdős; they are close to everyone. This is typical of collaboration networks: there are short paths among all pairs of nodes. Pick any two scholars and they will not be very far from each other.

Source: Menczer et al: Filippo Menczer, Santo Fortunato, Clayton A. Davis, [A First Course in Network Science. Cambridge University Press 2020]



Six degrees of separation: Small World

- Not only collaboration networks, but pretty much all social networks have very short paths among nodes
- It is likely to know someone who knows someone who knows someone… and in a few steps get to anyone on the planet!
- *Six Degrees of Kevin Bacon* is a fun game that originates from such a network
  - Play this game online at The Oracle of Bacon (oracleofbacon.org)
  - The website pulls data to build the network from the Internet Movie Database (IMDB.com)
  - The Wiki game. You will be amazed at how quickly you can reach any target with a bit of practice. Wikipedia has short paths.
- Milgram experiment: small world

Random Networks and Small-World Networks

The Random Network (RN) Model

- From a Cocktail Party to Random Networks
- A random network consists of N nodes where each node pair is connected with probability p
- There are two definitions of a random network:
  - G(N, L) Model: N labeled nodes are connected with L randomly placed links (Erdős and Rényi, papers on random networks). Model fixes the total number of links L.
  - G(N, p) Model: Each pair of N labeled nodes is connected with probability p, a model introduced by Gilbert. The model fixes the probability p that two nodes are connected.

Rapoport (1951) demonstrated that if we increase the average degree of a network, we observe an abrupt transition from disconnected nodes to a graph with a giant component.

Random Networks: Number of links expected

- Each random network generated with the same parameters N, p looks slightly different
- How many links we expect for a particular realization of a random network with fixed N and p?

The average degree of a random network is

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1)$$

Average Distance:

$$\langle d \rangle \propto \frac{\ln N}{\ln \langle k \rangle}$$

Clustering Coefficient:

$$\langle C \rangle = \frac{\langle k \rangle}{N}$$

in a random network the chance of observing a hub decreases faster than exponentially.

The probability that a random network has exactly L links is the product of three terms:

- The probability that L of the attempts to connect the $N(N-1)/2$ pairs of nodes have resulted in a link, which is p^L.
- The probability that the remaining $N(N-1)/2 - L$ attempts have not resulted in a link, which is $(1-p)^{N(N-1)/2-L}$.
- A combinational factor,

$$\binom{\frac{N(N-1)}{2}}{L} \qquad (3.0)$$

counting the number of different ways we can place L links among $N(N-1)/2$ node pairs.

We can therefore write the probability that a particular realization of a random network has exactly L links as

$$p_L = \binom{\frac{N(N-1)}{2}}{L} p^L (1-p)^{\frac{N(N-1)}{2}-L} \qquad (3.1)$$

As (3.1) is a binomial distribution (BOX 3.3), the expected number of links in a random graph is

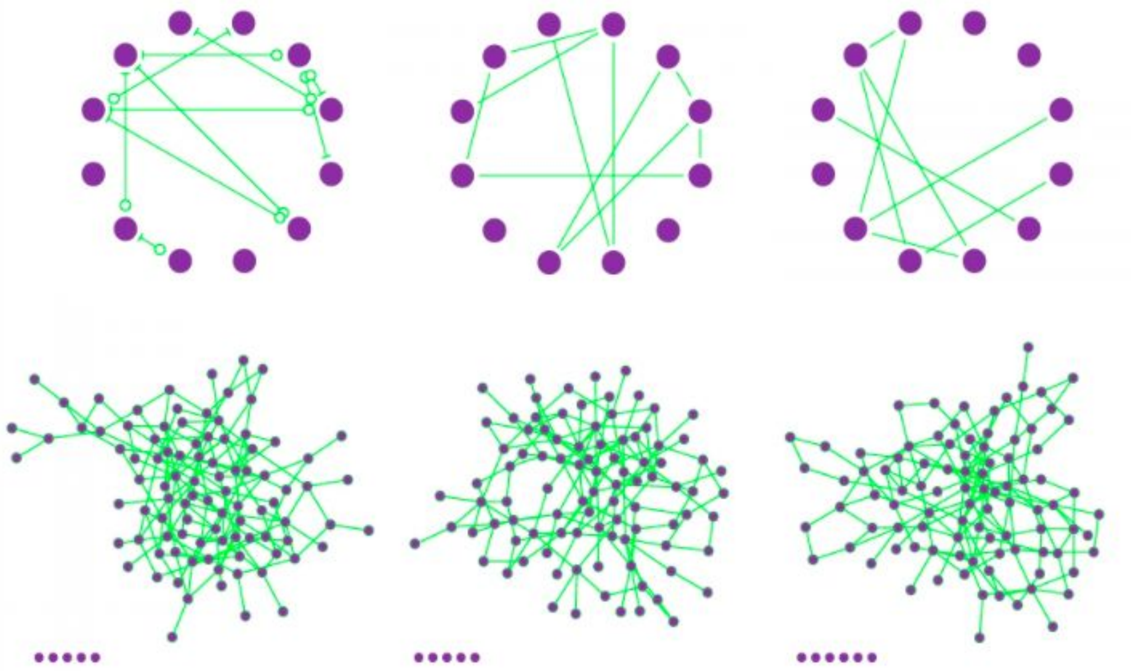
\

$$\langle L \rangle = \sum_{L=0}^{\frac{N(N-1)}{2}} L p_L = p\frac{N(N-1)}{2} \qquad (3.2)$$

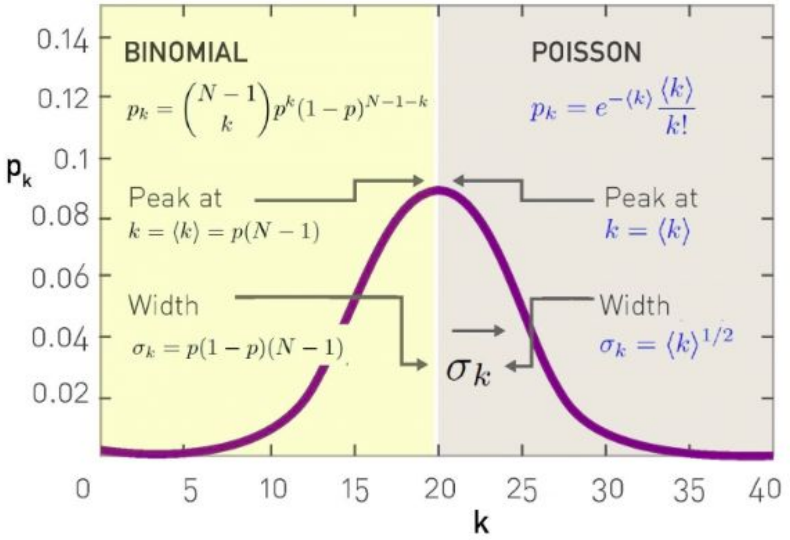
Random Networks (RN)

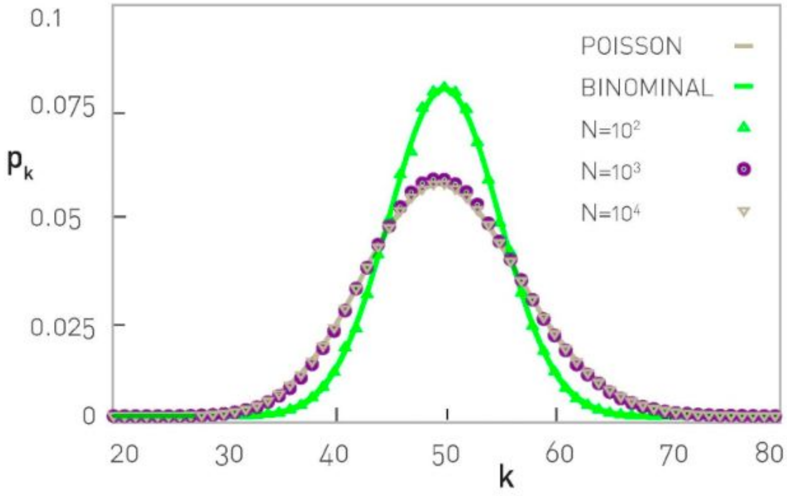
3 realizations of a RNs generated with the same parameters: p=1/6 and N=12. Despite the identical parameters, the networks not only look different, but they have a different number of links:
(L=10, 10, 8).

3 realizations of a RN with p=0.03 and N=100. Several nodes have degree k=0, shown as isolated nodes at the bottom



Random Networks: Degree Distribution



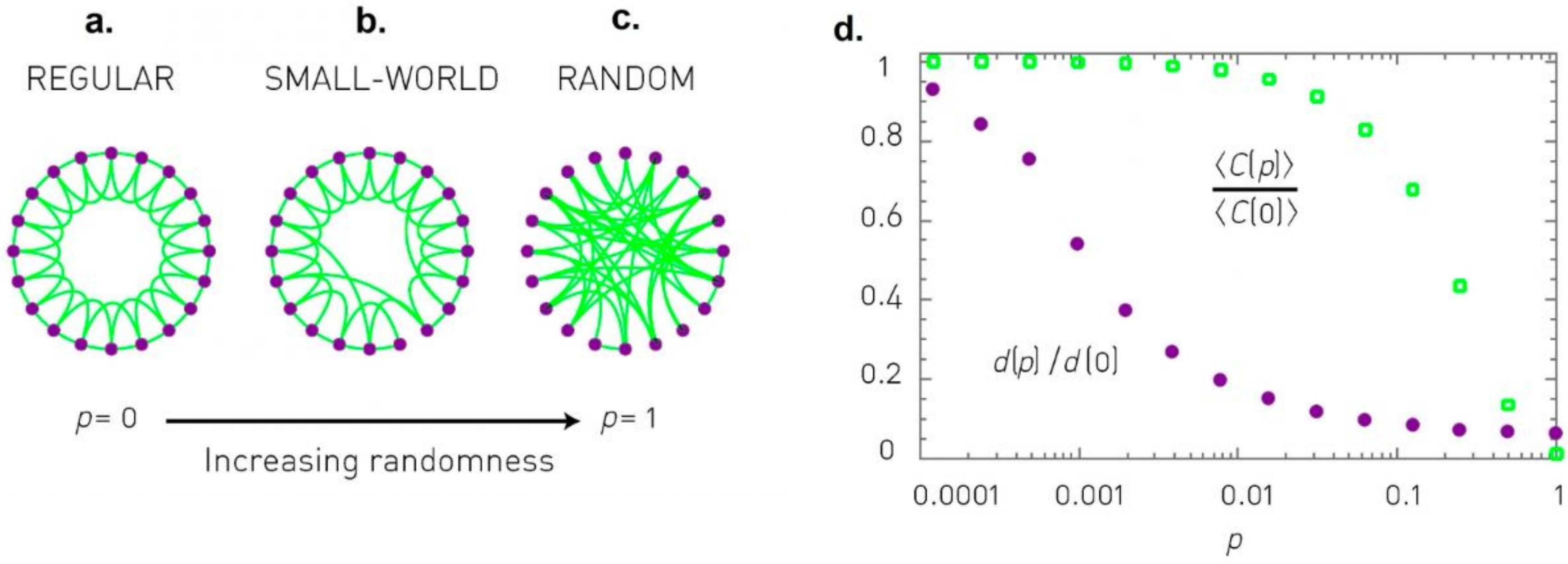


<u>Binomial vs. Poisson Degree Distribution</u>

The exact form of the degree distribution of a random network is the binomial distribution. For N ›› ‹k› the binomial is well approximated by a Poisson distribution. Both describe the same distribution, they have the identical properties, but they are expressed in terms of different parameters: The binomial distribution depends on p and N, while the Poisson distribution depends only on ‹k›.

Small-Worlds: Watts-Strogatz Model

- proposed an extension of the random network model



Small-World

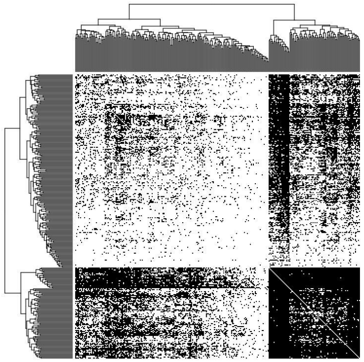
- The small world property is defined by $\langle d \rangle \approx \frac{\ln N}{\ln \langle k \rangle}$ describing the dependence of the average distance in a network on N and ‹k›
- In general lnN << N, hence the dependence of <d> on lnN implies that the distances in a random network are orders of magnitude smaller than the size of the network
- By small in the "small world phenomenon" means that the average path length or the diameter depends logarithmically on the system size. Hence, “small” means that <d> is proportional to lnN, rather than N or some power of N
- The 1/ln ‹k› term implies that the denser the network, the smaller is the distance between the nodes

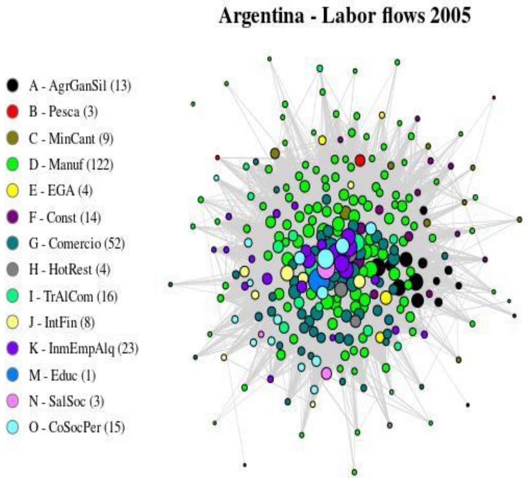
Going back to labor flow networks…

How do we characterize them?

What metrics shall we use?

Flow Network: Predominant year-to-year structure





Source: Elaboración propia SIPA, año 2005

- Single large connected component
- Relevant link density (non-sparse matrices)
- High Degree (*k*)
- High reciprocity of ties
- High clustering coefficient
- Short paths (reduced diameter: 3 steps)

Metrics	Average
Nodes (Order)	**287**
Edges (size)	26.961
Weights	278.117
CComp	1,10
Densidad	0,33
⟨k⟩	188
⟨kin⟩	94
⟨kout⟩	94
⟨knn⟩	**285**
Assortativity	-0,26
Reciprocidad	0,69
CCoef	0,65
LCP	1,68
Diámetro	3,40

What does it mean in terms of labor flows?

Sectors

Links &/industries

… of labor

… of links

Average ties by sec

… de lazos

Triángles

Caminos e/sec

→high reciprocity (0.69) in the connections and a high transitivity or global clustering (0.65), meaning that 2 out of 3 of the possible connections between three nodes (triples) constituted closed triangles.

Core-Periphery decomposition

We can decompose the network to reveal its core–periphery structure. This is accomplished by iteratively filtering out shells of low-degree nodes and focusing on the remaining, denser and denser cores.

The degree of each node can be used to separate a network into distinct portions, called shells, based on their position in the core–periphery structure of the network.

Low-degree outer shells correspond to the periphery.

As they are removed, or peeled away, what remains is a denser and denser inner subnetwork, the core.

Core-Periphery decomposition

Formally, the k-core decomposition algorithm starts by setting $k=0$. Then it proceeds iteratively. Each iteration corresponds to a value of k and consists of a few simple steps:

1. Recursively remove all nodes of degree k, until there are no more left.

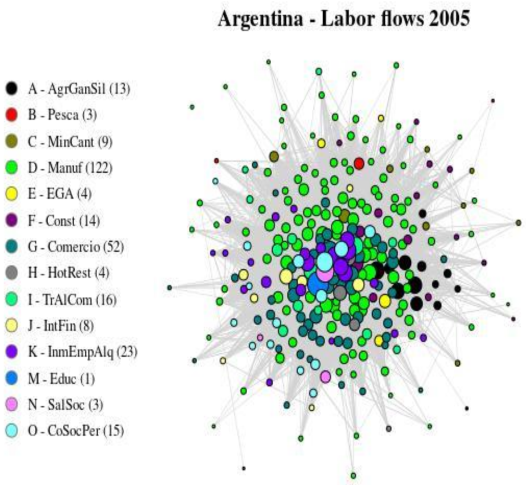
2. The removed nodes make up the k-shell and the remaining nodes make up the ($k+1$)-core, because they all have degree $k+1$ or larger.

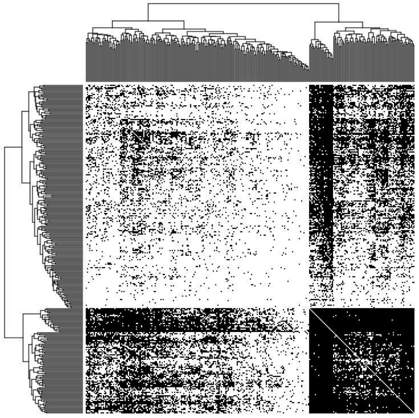
3. If there are no nodes left in the core, terminate; else, increment k for the next iteration.

Flow Network (cont.)

- **Properties**: Small world
  - High clustering coefficient (transitivity)
  - Short average paths

- Core-periphery structure
  - Core: subset of high degree nodes, highly interconnected
  - Periphery: subset of low degree nodes, connected (almost only) with core nodes





HANDS-ON

1. Implement the proposed Standing Ovation (model 2) assuming spatial structure. You can try to implement some synthetic networks (random, small-worlds), o directly try some real world data networks from the repository sources offered in the course.
2. Implement PD game on the spatial structure. Change the rule of updating strategy: instead of imitation mechanism suggest another and see what happens with the global level of cooperation.
3. References on financial networks:
    a. Network models and financial stability
    b. Systemic risk in banking ecosystems

see you next class!