Lecture
Networks: an introduction

Computational Economics course
University of Florence, Italy

Contents

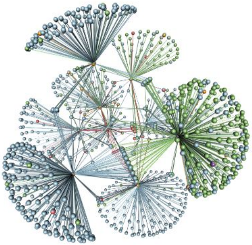
1 Introduction
- The basic idea
- Sources

2 Network
- Definition
- Kind of Networks
- Adjacency matrix
- Direct and Indirect Networks
- Direct and Indirect weighted Networks
- Paths

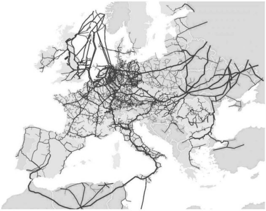
3 Network metrics

The basic idea

Network: a collection of points joined together in pairs by lines.

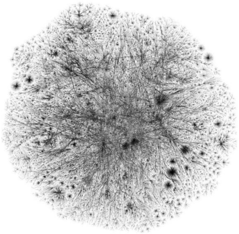


(a) HIV-1 protein

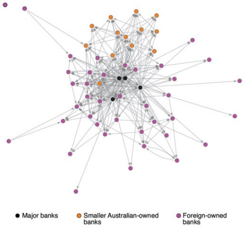


(b) Gas pipelines

The basic idea



(c) Internet



(d) Interbank

Why are networks so important?

Everything is a network.

Sources

- Main textbook: M.E.J. Newman, Networks: An Introduction. Oxford University Press, 2010
- Other sources: Barabasi, Vespignani, Dorogovtsev, Caldarelli...
- On Economic/Social networks: Jackson, Easley and Kleinberg, Vega-Redondo.
- Source for network data: http://math.nist.gov/RPozo/complex datasets.html (contains links to other repositories).
- Software packages: igraph (R), networkx (Python)
- Visualization software: Pajek, Gephi, Cytoscape...

Definition

A **network**, also called **graph** G, is a couple of sets V and E, and can be indicated as $G = (V, E)$, where:

- V is the set of VERTICES (or NODES)
- E is the set of EDGES (or LINKS)

Three elements:

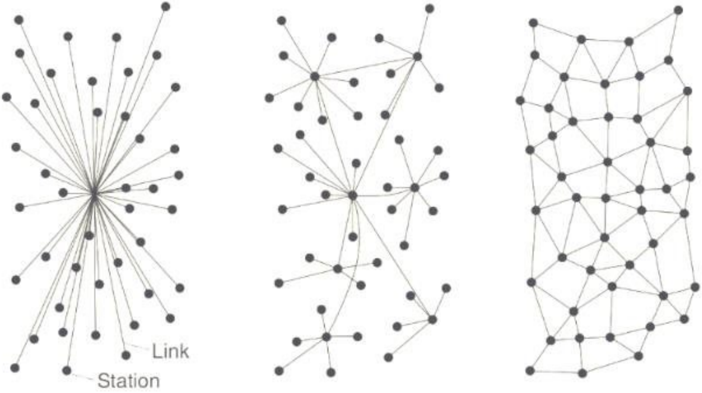
1. The graph is the **entire representation** of the network
2. Nodes are **objects that play a role** in the network
3. Edges are the **connections** between the nodes

The cardinality of V, $|V| = n$, is the number of vertices in the graph (order).

The cardinality of E, $|E| = m$, is the number of links in the graph (size).

Kind of Networks

1. **Centralized networks** built around a single, **centralized** "master" node.
2. **Decentralized** network connections across multiple masters. Each of these separate nodes is a central unit that interacts independently with other nodes.
3. **Distributed** networks are composed of equal, interconnected nodes.



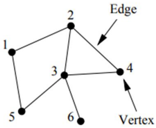
Centralised (A) Decentralised (B) Distributed (C)

Adjacency matrix

- The structure of a network can be described by a n x n matrix: the (weighted) adjacency matrix $A = \{a_{ij}\}$.
- If two nodes i and j are not joined by a link, $a_{ij} = 0$, otherwise, $a_{ij} = 1$.

$$A = \begin{cases} 1, & \text{if there is an edge between vertices } i \text{ and } j \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$



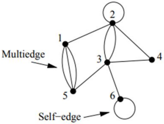
The degree of a node i is the number of its links: $k_i = \sum_j a_{ij}$.

Adjacency matrix

- There can be more than one edge between the same pair of vertices: **multiedges**.
- There can be edges that connect vertices to themselves: **self-edges**.

A network with neither self-edges nor multiedges is called a **simple network**.
A network with multiedges is called a **multi-graph**.

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 3 & 0 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix}.$$

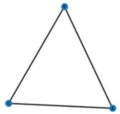


A multiedge is represented by setting the corresponding matrix element A_{ij} equal to the multiplicity of the edge.
A single self-edge from vertex i to itself is represented by setting the corresponding diagonal element A_{ii} of the matrix equal to 2.

Direct and Indirect Networks

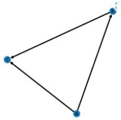
1. If E is a set of non-ordered pairs of distinct elements in V, we have a **binary undirected network**. For undirected networks the adjacency matrix is symmetric, $A = A^T$.

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$



2. If E is a set of ordered pairs of distinct elements in V, we have a **binary directed network**. This means that a link connecting node i to j is different from a link connecting j to i.

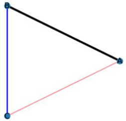
$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$



Direct and Indirect weighted Networks

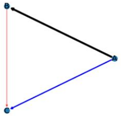
1. If a binary undirected network is associated with a function $w : V \to \mathbb{R}^+$ we have a **weighted undirected network**: w_{ij} is the weight of the link $i - j$. If $w : V \to \mathbb{N}$, the weight can be seen as:
   - numbers quantifying the relationship between nodes.
   - a measure of how many single links are present between any two nodes.

$$A = \begin{bmatrix} 0 & 10 & 0.5 \\ 10 & 0 & 5 \\ 0.5 & 5 & 0 \end{bmatrix}$$



2. If weights are associated with a directed network we have a **weighted directed network**.

$$A = \begin{bmatrix} 0 & 10 & 5 \\ 0 & 0 & 1 \\ 0 & 5 & 0 \end{bmatrix}$$



Paths

- A **path** in a network is any sequence of vertices such that every consecutive pair of vertices in the sequence is connected by an edge in the network.
- The **length** of a path is the number of edges traversed along the path.
- The total number of paths of length 2 from node i to node j is:

$$N_{ij}^{(2)} = \sum_{k=1}^{n} a_{ik} a_{kj} = (A^2)_{ij}. \tag{2}$$

and generalizing the number of paths $i \to j$ of length r is $(A^r)_{ij}$.

- The total number of **cycles** of length r is thus:

$$L_r = \sum_{i=1}^{n} (A^r)_{ii} = Tr[A^r]. \tag{3}$$

What if the network is indirect? Proof: $L_r = \sum_i^n \phi_i$, where ϕ_i are the eigenvalues of A.

Network metrics

- The **distance** between two vertices i, j is the shortest number of edges to go from i to j.
- The **neighbors** of a vertex i are all vertices j which are connected to that vertex by a single edge $(d_{ij} = 1)$:

$$d_{ij} = \min\{\sum_{k,i \in P_{ij}} a_{ki}\}, \tag{4}$$

  where P_{ij} is a path connecting vertex i to vertex j.
- The **diameter** of a graph is given by the maximum of all distances between pairs.

Network metrics

- The **size**: $L = \sum_{i<j} a_{ij}$ (undirected) and $L = \sum_{i,j} a_{ij}$ (directed).
- The **density** is a measure of the density of links in the network:
  - Undirected network: $c^{(und)} = \frac{2L}{n(n-1)}$
  - Directed networks: $c^{(dir)} = \frac{L}{n(n-1)}$
- The **degree** of a node is the number of neighbors:
  - Undirected network: $k_i = \sum_j a_{ij}$.
  - Directed networks: indegree $k_i^{in} = \sum_j a_{ji}$, outdegree $k_i^{out} = \sum_j a_{ij}$, total $=$ $k_i^{tot} = k_i^{in} + k_i^{out}$.
- For a weighted networks the **strength** of a node is:

$$s_i = \sum_j w_{ij}. \tag{5}$$

- Similarly one defines the in-strength and the out-strength

Network metrics

- The **eccentricity** of a node is the inverse of the maximum distance of that node from any other possible node of the network:

$$e_i = \frac{1}{\max_{\forall j \in N} d_{ij}}. \tag{6}$$

- The **center** of a graph is the set of all vertices of minimum eccentricity, that is, the set of all vertices u where the greatest distance $d(u, v)$ to other vertices v is minimal.
- The **diameter** of graph is the maximum distance between the pair of vertices.
- The minimum among all the maximum distances between a vertex to all other vertices is considered as the **radius** of the graph. It is denoted as $r(G)$.

Centrality Measures

What can we do with networks?

- The centrality of a node (or of an edge) of a network is a measure of the importance of the node:
  - how influential is a person in a social network
  - how critical is an element in an infrastructure network,
  - what is the disease-spreading capacity of an individual
  - what is the most systemically important financial institution
- Loose definition: many metrics!
  - Degree centrality → **POPULARITY**.
  - Betweenness centrality → **BRIDGE**.
  - Closeness centrality → **CENTRALNESS**.
  - Eigenvector centrality → **INFLUENCE**.
  - Others