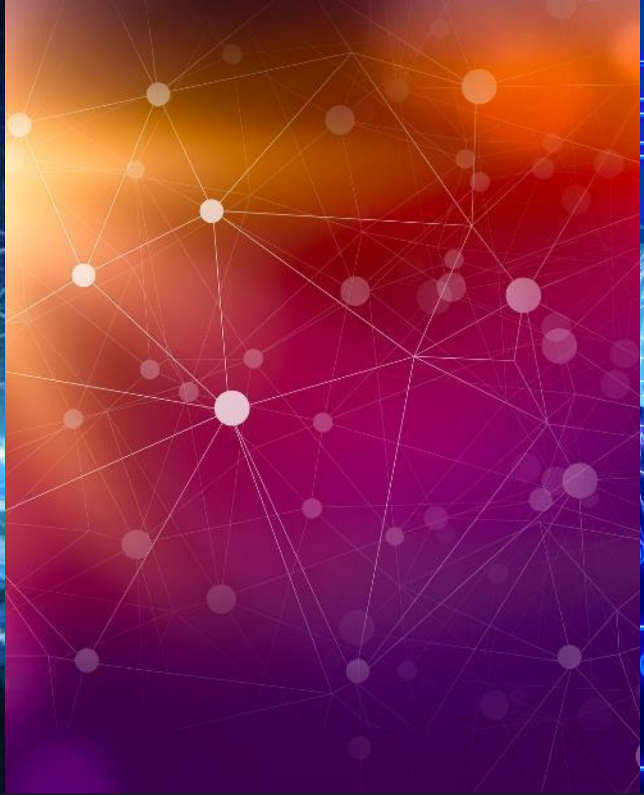
Redes complejas

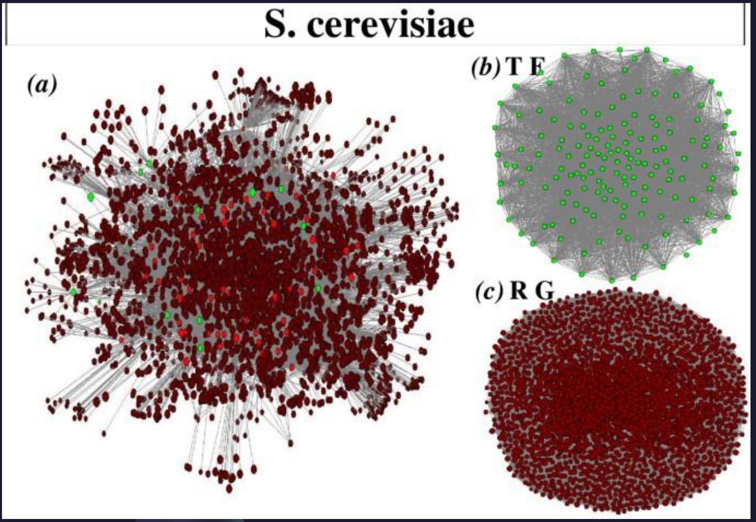
Moises Omar León Pineda

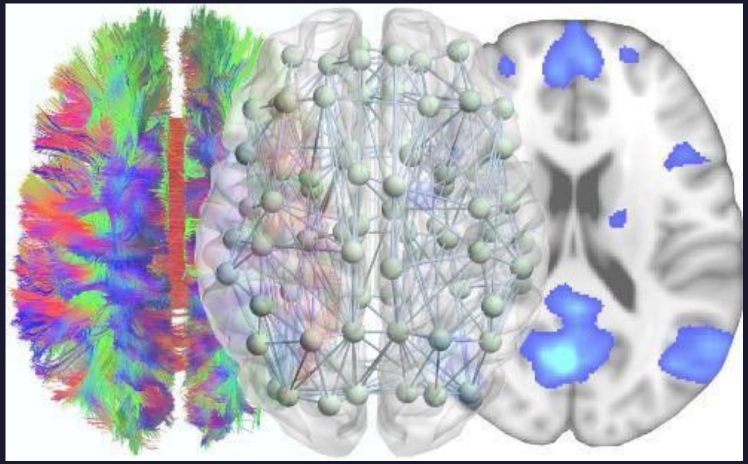



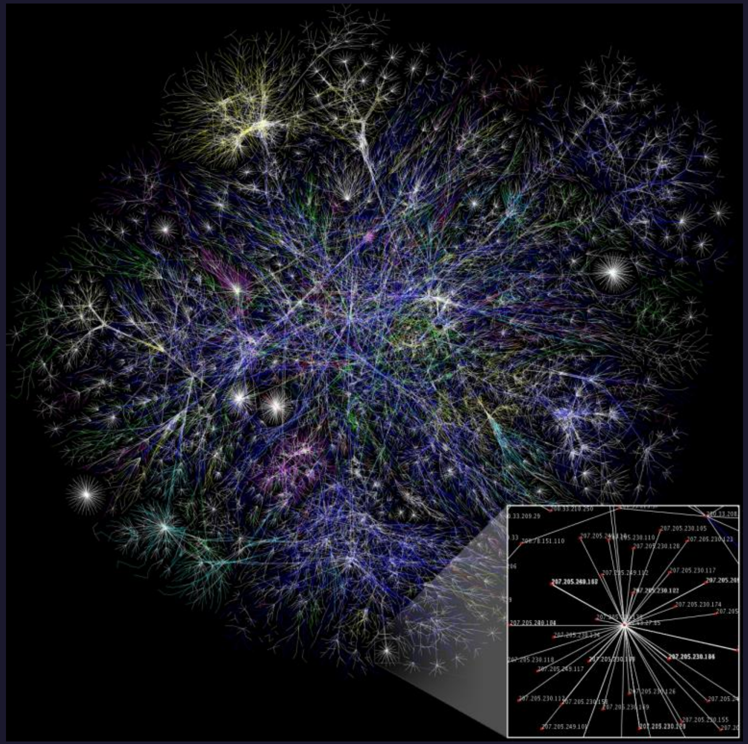
Introducción

- Las redes complejas son sistemas complejos formados por nodos interconectados entre sí. Estos nodos pueden representar personas, organizaciones, computadoras u otros elementos que estén relacionados entre sí.

Estas redes se pueden representar mediante grafos, con nodos interconectados por enlaces.

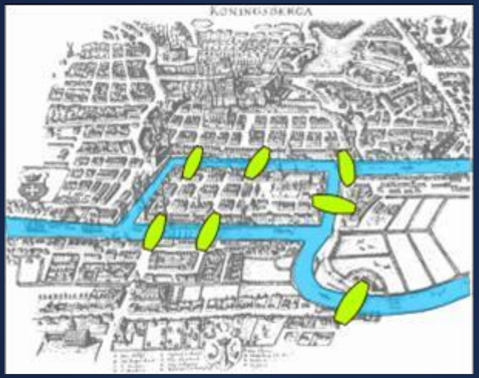






Red compleja	Nodos	Aristas/uniones
Red de actores	Actores	actuar en la misma película
Red de colaboraciones	Científicos	Coautores en una publicación
Redes de citas bibliográficas	Artículos científicos	Citación
La red de Facebook	Personas	Amistad en facebook
Redes metabólicas	Metabolitos	Aparecer en la misma reacción
Redes de interacción de proteínas	Proteínas	Interacción física
Redes de transcripción genética	Genes	Regulación
Red cerebral	Neuronas	Conexiones sinápticas
Internet	Routers	Líneas físicas
WWW	Páginas Web	Direcciones URL
Redes de Aeropuertos	Aeropuertos	Vuelos entre aeropuertos
Redes eléctricas	Centrales eléctricas	Conexiones eléctricas

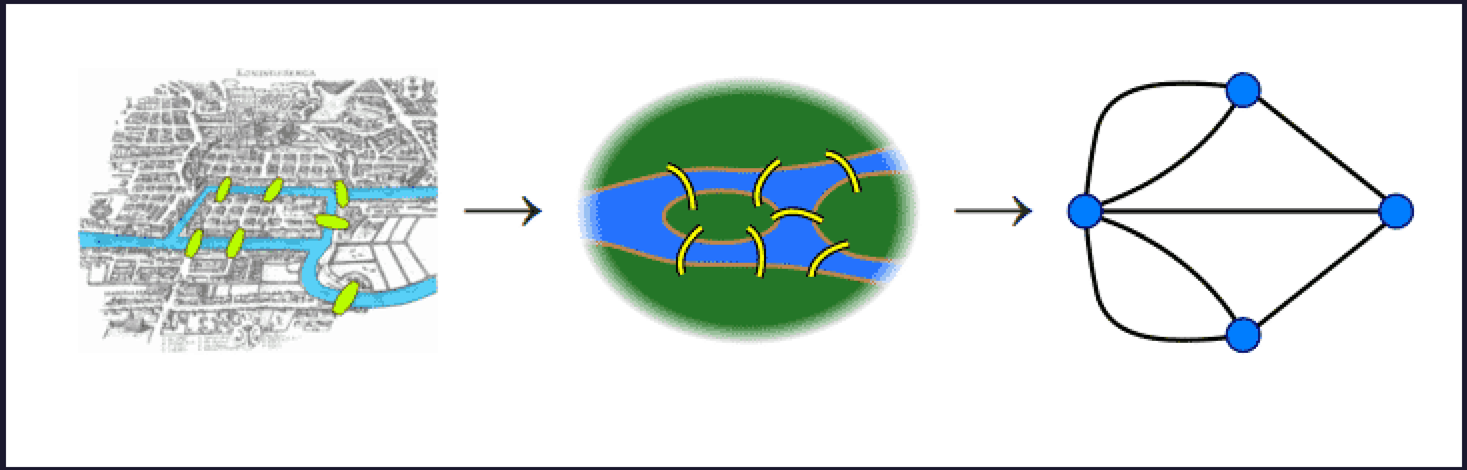
Teoría de grafos



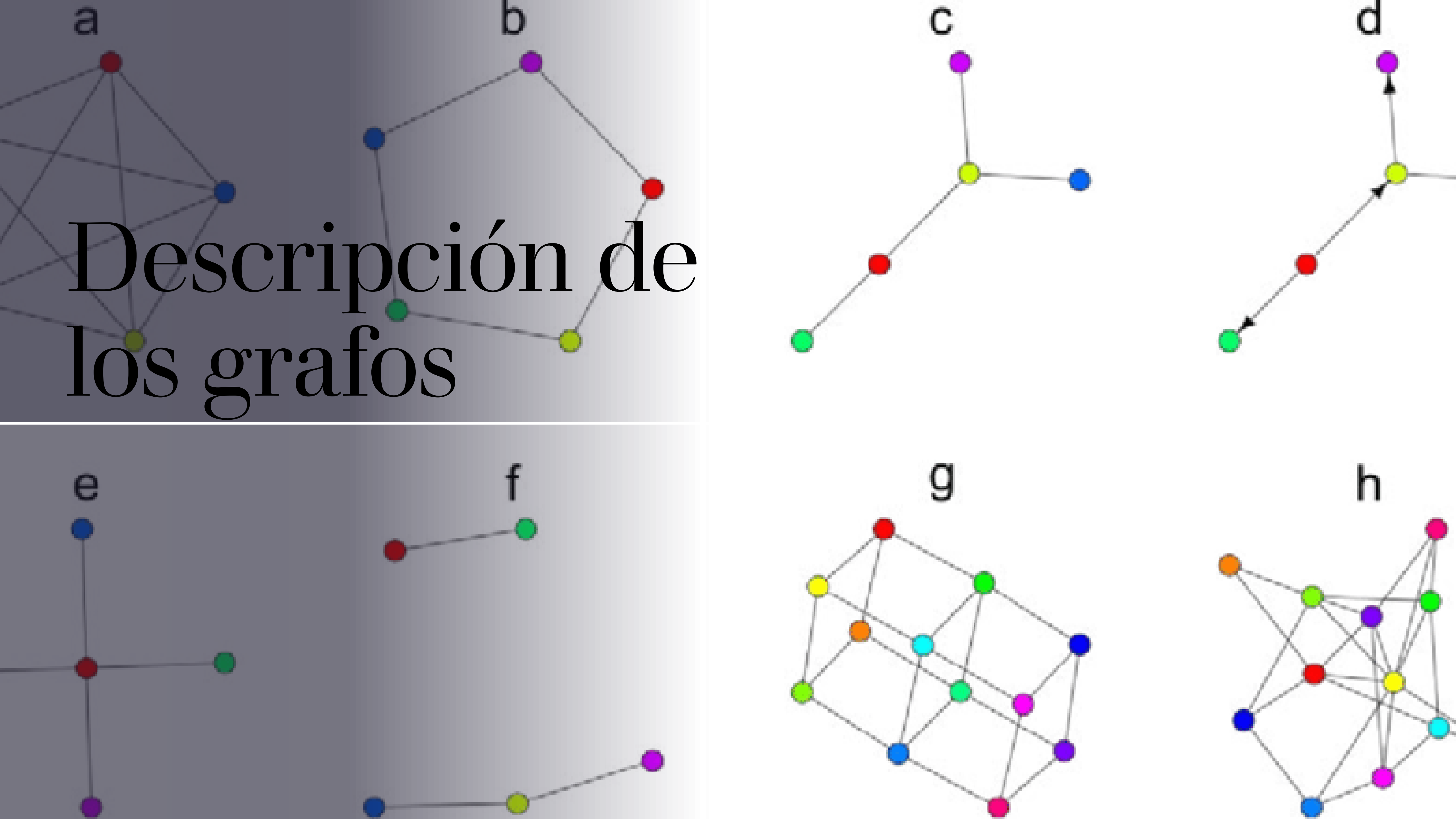
Leonhard Euler, 1736

El problema de los puentes de Königsberg

Para poder recorrer **los puentes de Königsberg**, **los vértices “intermedios” deben tener un número par de aristas**. Es decir, deben tener **una vía para entrar y una vía para salir**. Sólo los puntos de inicio y salida pueden tener un número impar de aristas, porque, evidentemente, **nunca “entramos” al punto de inicio y nunca “salimos” del punto de llegada.**



Descripción de los grafos



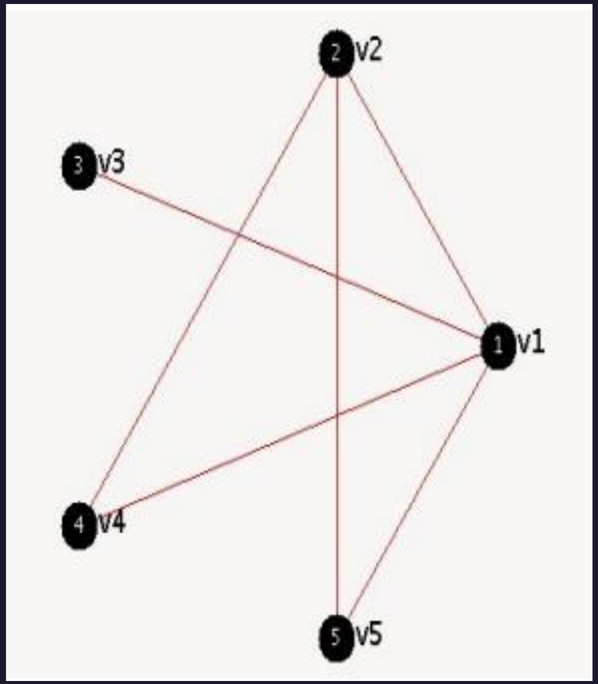
Grafo

Un conjunto de vértices (o nodos) y un conjunto de parejas ordenadas, llamadas aristas entre dos nodos

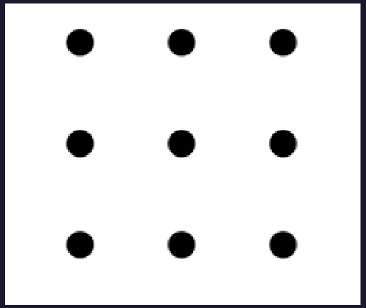
$$G = (V, E)$$

$$V = \{v_1, v_2, \ldots, v_n\} \quad E \subseteq \mathrm{V} \times \mathrm{V}$$

$$E = \{e_1, e_2, \ldots, e_m \,|\, e_i = (v_i, v_j)\} \quad v_i, v_j \in V$$

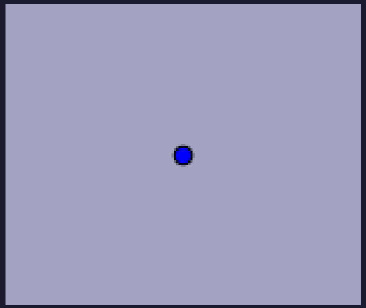


Algunos tipos de grafos



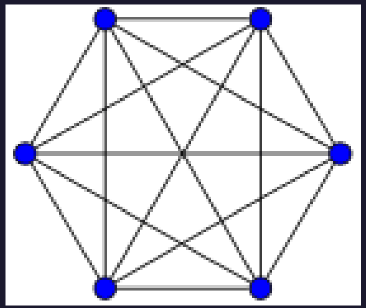
Grafo vacío

grafo sin aristas, sólo con nodos



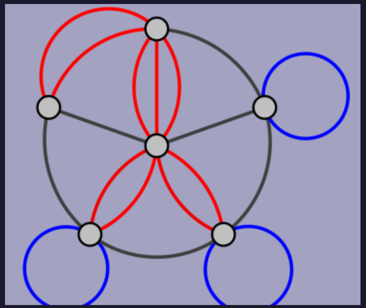
Grafo nulo

el que no tiene vértices (y por lo tanto no tiene aristas)



N -grafo completo

Todos los vértices están conectados



Multigrafos

Existen aristas múltiples

Representación gráfica de grafos

LISTA DE ARISTAS:

- la lista de aristas es una lista de L pares de nodos (nj , ni) , o sus etiquetas (j, i) indicando que el nodo j está unido al i por una arista.

MATRIZ DE ADYACENCIA:

- La matriz de adyacencia A de un grafo simple de N nodos es una matriz cuadrada N × N cuyos elementos son:

$$A_{il}=\begin{cases} 1 \ \ \text{si el nodo j está unido al i por } una\ arista \\ 0 \text{ cualquier otra situación} \end{cases}$$

MATRIZ DE INCIDENCIA:

- es una matriz B N × L donde N es el número de vértices y L el número de uniones o aristas l = (i, j) y tal que
-

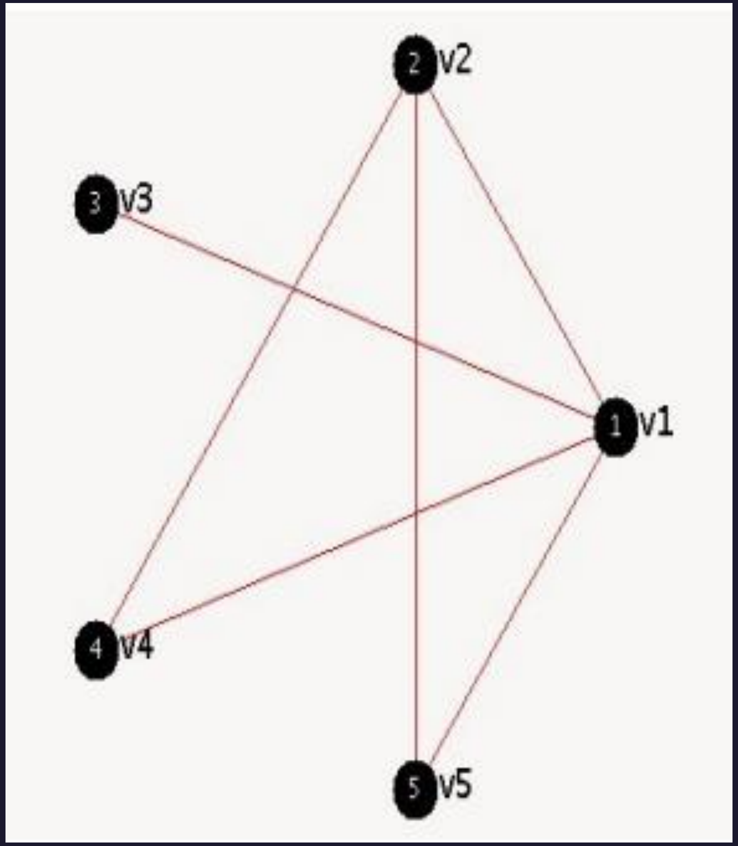
$$B_{il}=\begin{cases} 1 \text{ si la unión l contiene al vértice i} \\ 0 \text{ cualquier otra situación} \end{cases}$$

Grafos no dirigidos

a_{ij}= 1 si $(v_i, v_j) \in V$

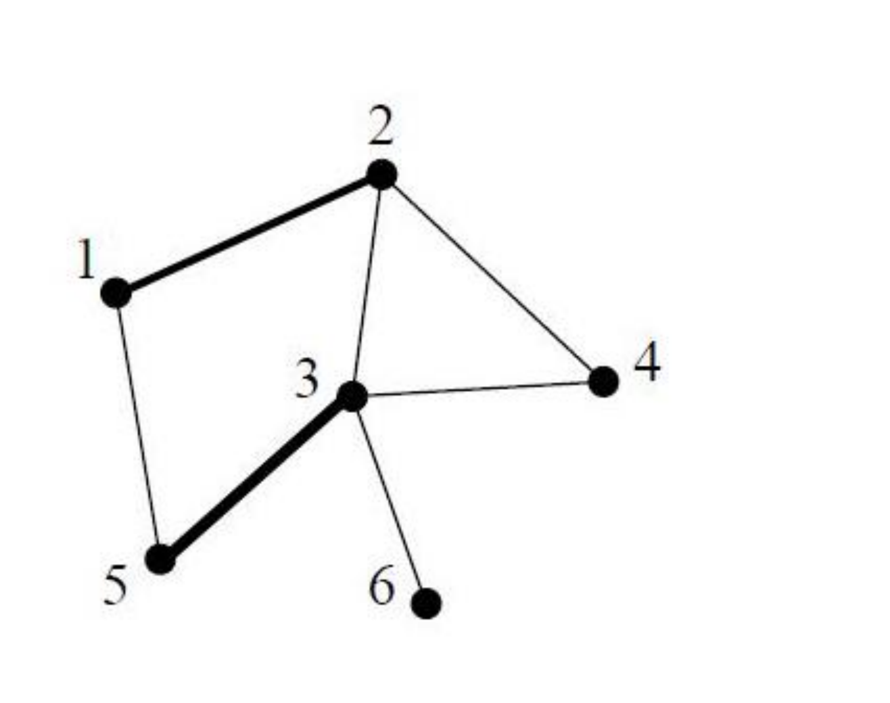
a_{ij}= 0 si $(v_i, v_j) \notin V$

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



Gráfos dirigidos y pesaos

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



a_{ij}= w_{ij},si $(v_i, v_j) \in V$

a_{ij}= 0 si $(v_i, v_j) \notin V$

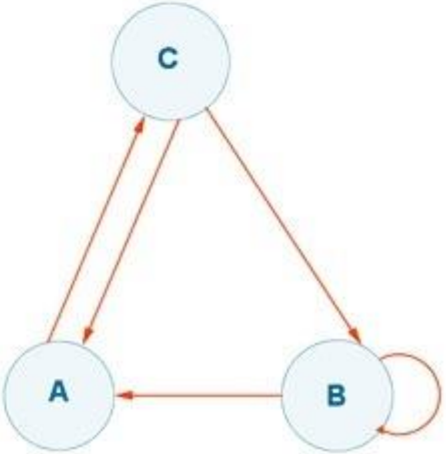
A es una matriz simétrica.

A=A^t

$A_{ij} = A_{ji}$

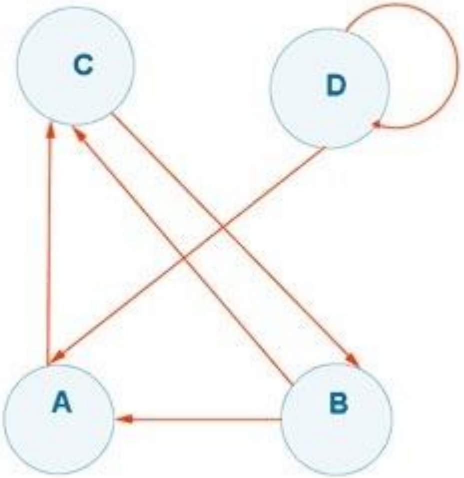
1.

$$M_1 = \begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \end{array}$$



2.

$$M_2 = \begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \end{array}$$



3.

$$M_3 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Teoría de redes

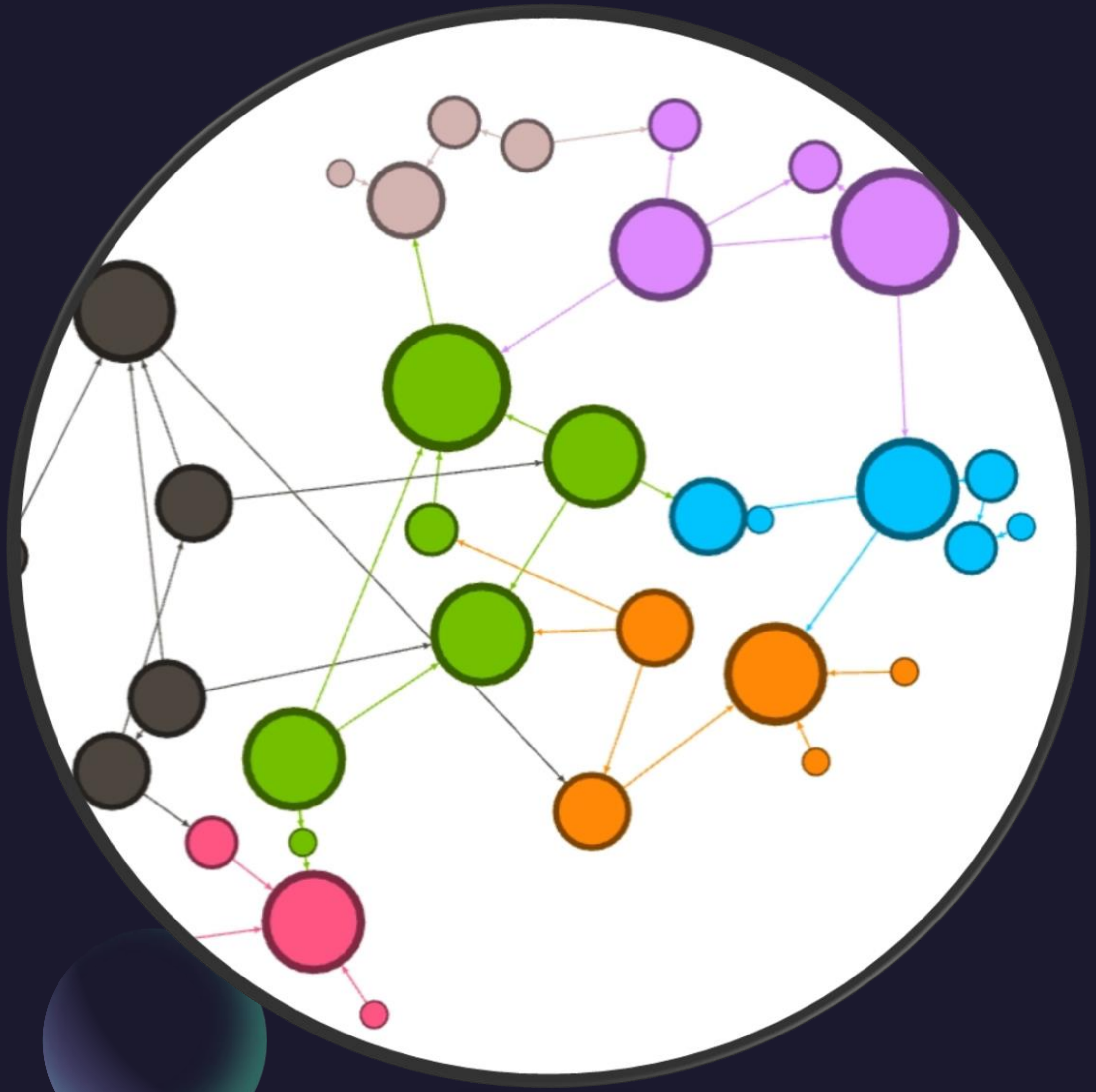
ASPECTOS ESTRUCTURALES:

- Centralidad de nodos
- Efecto de Mundo pequeño
- Libertad de Escala
- Resilencia y Robustez
- Modularidad
- Anidamiento y Estructura Jerarquica

ASPECTOS DINÁMICOS:

- Transiciones de fase y fenómenos emergentes
- Percolación
- Propagación de epidemias
- Difusión y caminatas aleatorias
- Sincronización
- Network control

Aspectos Estructurales Básicos



Grado de un nodo

- En una red no dirigida se define el grado de un nodo i como el número total de aristas incidentes en dicho nodo y se denota por k_i
- En el caso de redes no pesadas el grado de un nodo se puede calcular directamente a partir de la matriz de adyacencia:

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$k_i = \sum_{j=1}^{N} A_{ij} = \sum_{j=1}^{N} A_{ji}$$

- En una red dirigida se distingue entre grado entrante (in-degree) de un nodo i como el número total de nodos que apuntan a dicho nodo y se representa por k_i^{in}, y grado saliente (out degree) de un nodo i como el número total de nodos a los que apunta el nodo i y se denota por k_i^{out}

$$k_i^{in} = \sum_{j=1}^{N} A_{ij}$$

$$k_i^{out} = \sum_{j=1}^{N} A_{ij}$$

Secuencia de grados:

En una red no dirigida es la secuencia ordenada de los grados de todos los nodos de la red

$$\{K_i\} = \{K_1, \ldots, K_N\}.$$

Esta definición se puede generalizar para un grado dirigido a las secuencias de grados entrantes

$$\{k_i^{in}\} = \{k_1^{in}, \ldots, k_N^{in}\}$$

y a la secuencia de grados salientes

$$\{k_i^{out}\} = \{k_1^{out}, \ldots, k_N^{out}\}.$$

GRADO PROMEDIO DE LA RED:

Dada la secuencia de grados de un grafo no dirigido se define el grado medio ⟨k⟩ en la forma:

$$\langle k \rangle = \frac{1}{N} \sum_i k_i$$

Se tiene trivialmente que para una red simple

$$L = \frac{1}{2} \langle k \rangle N.$$

El grado máximo en una red simple no dirigida se define como

$$K = \max_i k_i.$$

Para redes dirigidas:

Cada enlace tiene un extremo in y otro out

$$M = \sum_{i=1}^{N} k_i^{in} = \sum_{j=1}^{N} k_i^{out} = \sum_{i,j} A_{ij}$$

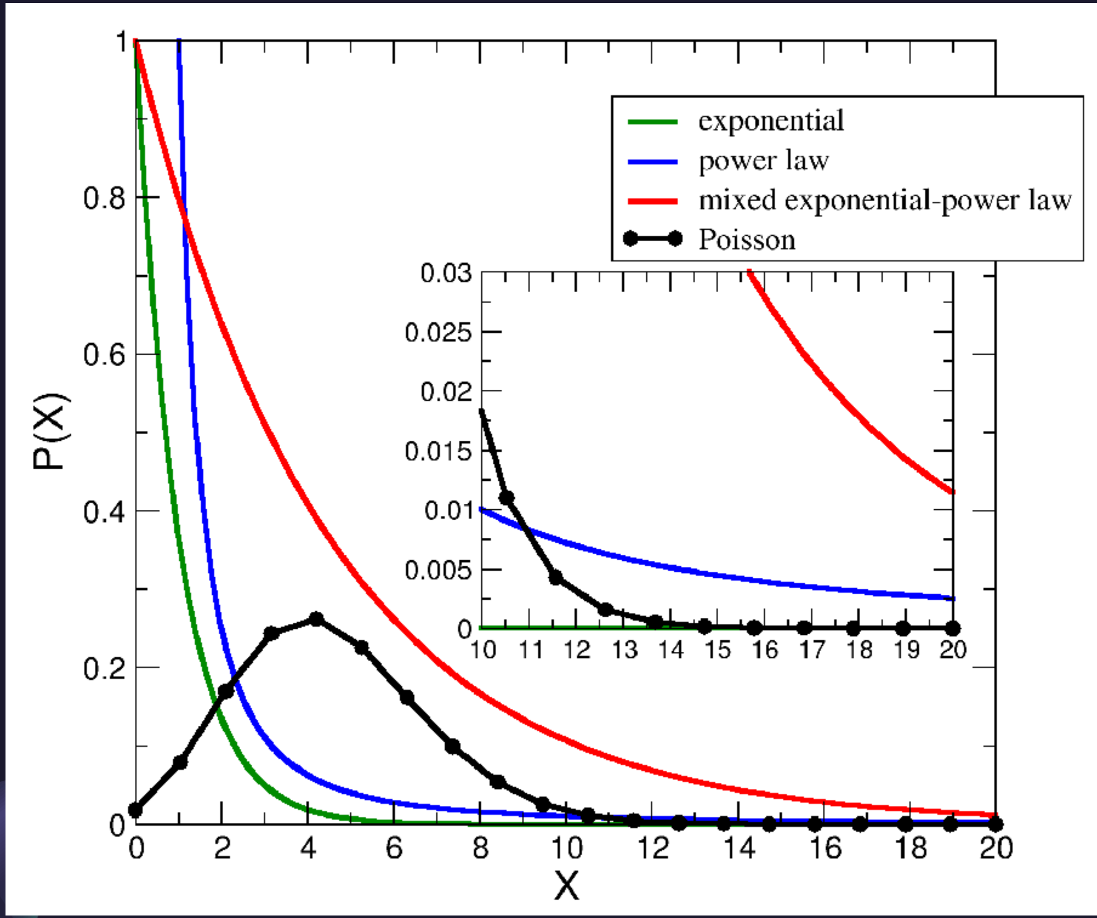
- Grado medio

$$\left.\begin{array}{l} \langle k^{in} \rangle = \frac{1}{N} \sum_i K_i^{in} \\ \langle k^{out} \rangle = \frac{1}{N} \sum_i K_i^{out} \end{array}\right] = \frac{1}{N} \sum_{i,j} A_{ij}$$

- Además se pueden definir el grado máximo saliente y entrante en la forma:
  - $K^{in} = \max_i k_i^{in}$
  - $k^{out} = \max_i k_i^{out}$

Distribución de probabilidad de grados de un grafo:

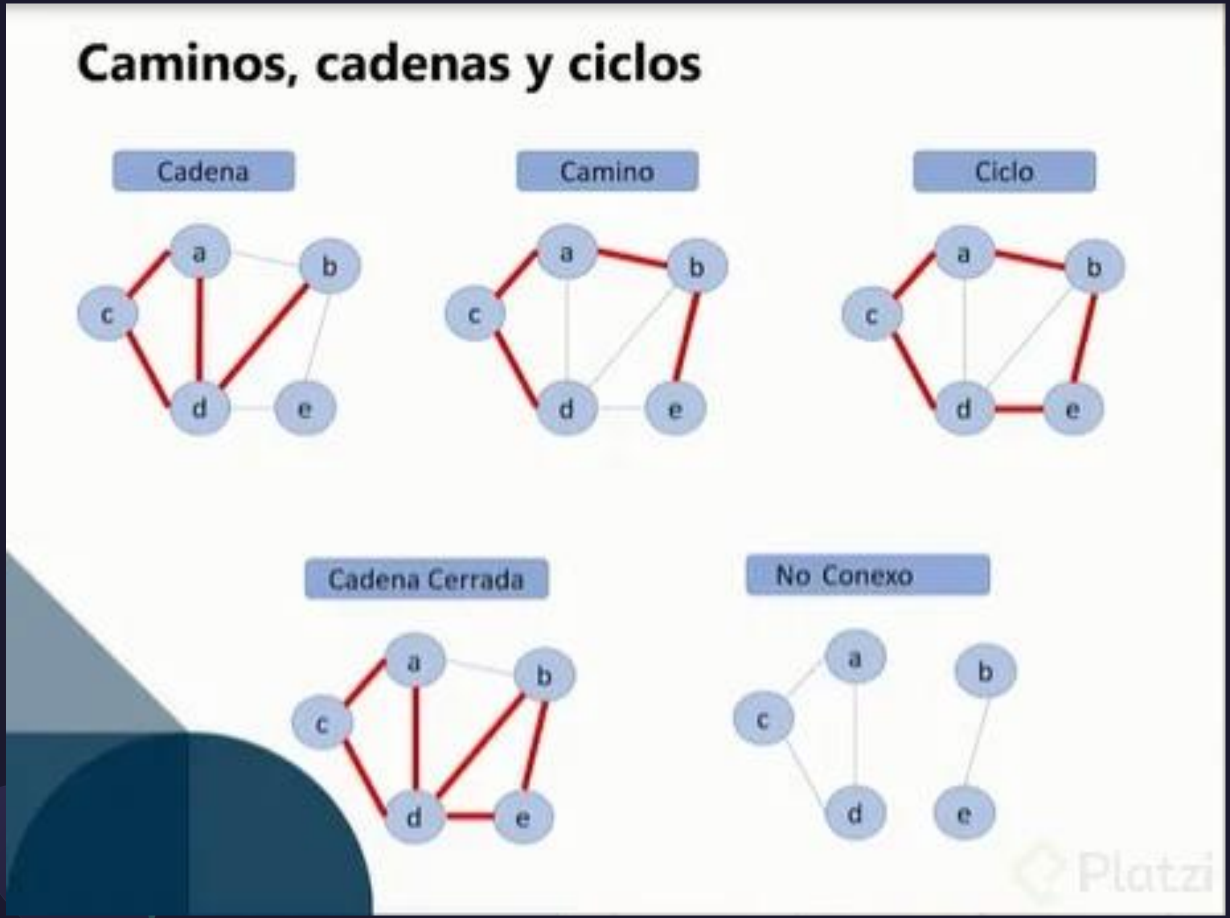
El grado de un nodo es una propiedad local de la red pero considerando la secuencia de grados podemos determinar algunas propiedades globales de la red. La organización y estructura global de una red inducida por su secuencia de grados está caracterizada por la distribución de probabilidad de grados de la red P (k)



Camino en un grafo

De manera formal se define un camino en un grafo como una secuencia de nodos tal que dos nodos consecutivos en la secuencia están conectados por una arista.

. Cada camino tiene su longitud definida como el número de aristas que se pasan al seguirlo incluyendo posibles repeticiones en caminos que se intersectan asimismo.



CAMINO MÁS CORTO

- Definimos el camino más corto entre dos nodos i y j en una red como el camino (camino dirigido en el caso de una red dirigida) de mínima longitud (medida en número de aristas que hay en el camino para llegar de un nodo al otro).
- Definimos la distancia más pequeña entre uno nodo i y j y la denotamos como d(ni, nj) = d(i, j) como la longitud del camino más corto entre dichos nodos

LONGITUD PROMEDIO DE CAMINOS

- Promedio de las distancias más pequeñas entre dos nodos cualesquiera de la red (estamos suponiendo que existen un camino conectando cualquiera dos nodos de la red):

$$l = \frac{1}{N(N-1)} \sum_{j \neq i}^{N} d(n_i, n_j)$$

El diámetro D de una red conectada se define como el máximo de la distancias más cortas entre cualquiera dos nodos de la red es decir:

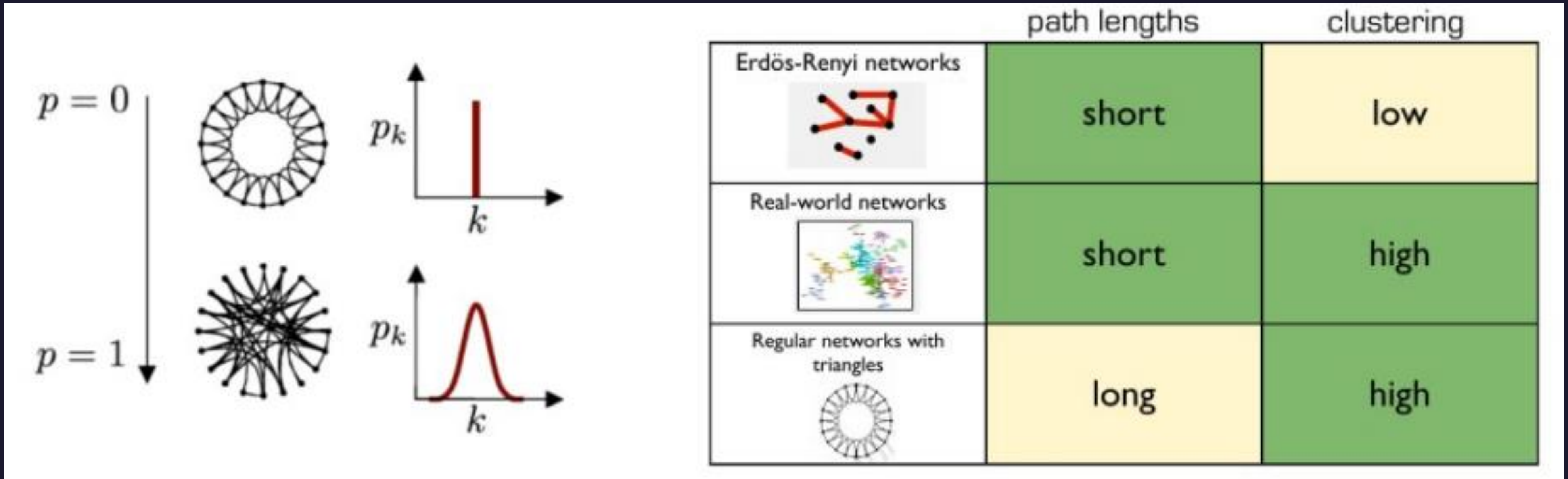
$$D = \max_{i,j \neq i} d(n_i, n_j)$$

Propiedad de pequeño mundo

La propiedad de pequeño mundo en redes complejas suele estar asociada con la existencia de un alto grado de agrupamiento de nodos (clustering) en la red.

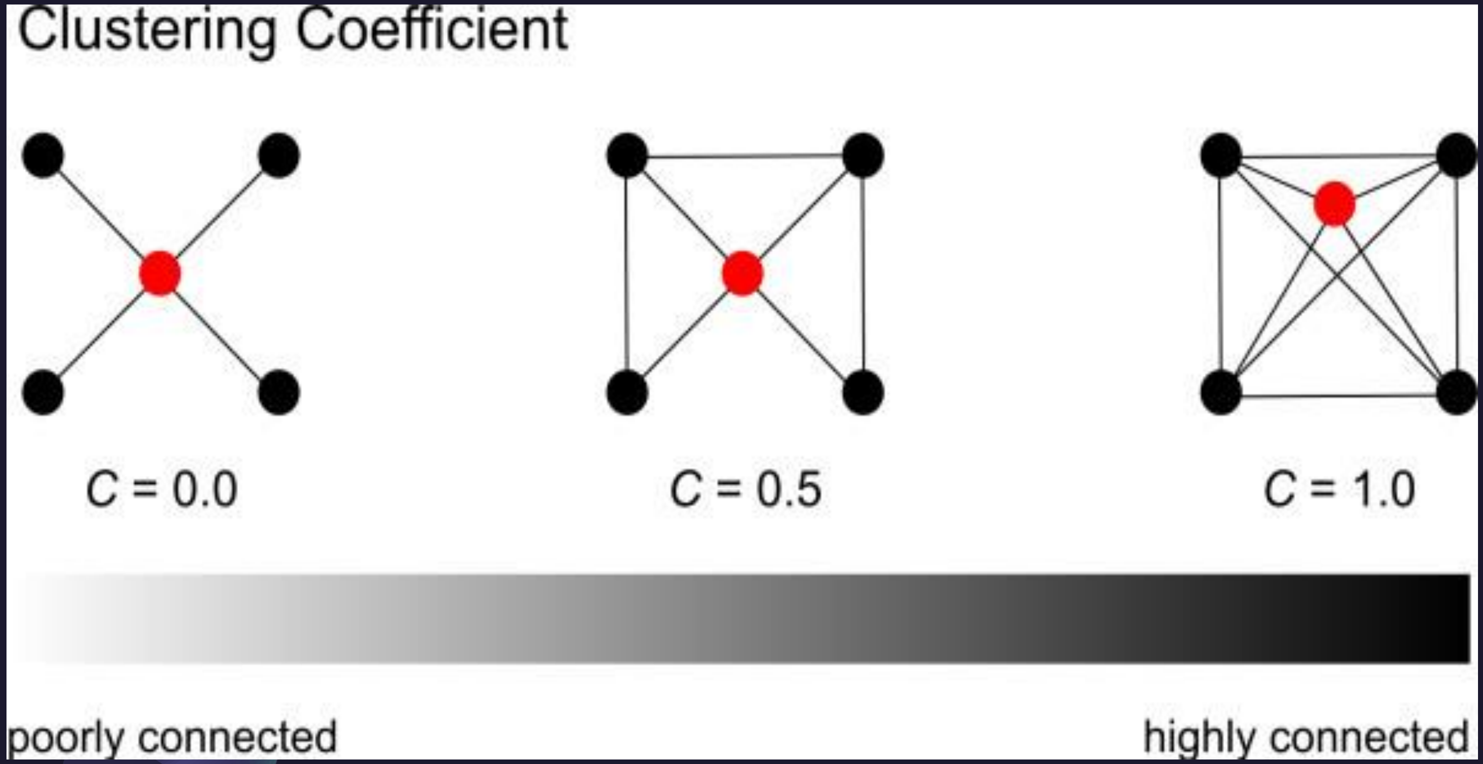
Por esta razón Watts y Strogatz definieron las redes tipo pequeño mundo como aquellas que tienen ℓ pequeño y C grande

$$l \sim \log N$$



Coeficiente de agrupamiento (clustering coeficient)

Es una medida del grado en el que los nodos de una red tienen a agruparse entre ellos.



Dado un grafo G(N, L) no dirigido se define el coeficiente de agrupamiento de un nodo i de dicho grafo y se denota como C_i al cociente:

$$C_i = \frac{\text{número de pares (l, m) de nodos vecinos de i conectados por aristas}}{\text{número total de pares que puedrían existir de nodos vecinos de } i}$$

Dado que el número de vecinos de un nodo i es k_i el denominador en la expresión anterior es simplemente

$$\frac{k_i(k_i - 1)}{2}$$

- En el caso de un grafo dirigido se tiene la definición:

$$C_i = \frac{\text{número de pares (l, m) de nodos vecinos de i conectados por aristas dirigidas}}{\text{número total de pares dirigidos que podrían existir de nodos vecinos de i}}$$

A Ci se le llama también coeficiente de agrupamiento local. Una definición alternativa para un grafo no dirigido es la siguiente:

$$C_i = \frac{\text{número de triángulos que contienen el nodo i}}{\text{número de tripletes que contienen al nodo i incidente a las dos aristas del triplete}}$$

donde un triplete es un grafo con 3 nodos y dos aristas.

Por otra parte se define el coeficiente de clustering de la red simplemente como

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i$$

- Para los gráficos ponderados, hay varias formas de definir la agrupación. El que se usa en Networkx se define como el promedio geométrico de los pesos de aristas del subgrafo

$$c_i = \frac{1}{k_i(k_i-1)}\sum_{jk}(\widehat{w_{ij}}\widehat{w_{ik}}\widehat{w_{jk}})^{\frac{1}{3}}$$

Aquí los pesos de los bordes están normalizados por el máximo peso en la red, $\hat{w}_{ij}$=w_{ij} /max(w) y la contribución de cada triángulo depende de todos los pesos de sus aristas.

Así, los triángulos en los que el peso de cada arista es igual a max(w) contribuyen con la unidad a la suma, mientras que los triángulos que tienen un eslabón con un peso insignificante tendrán una contribución insignificante. Esta definición se puede reescribir como

$$\tilde{C}_i = C_i \bar{I}_i \qquad \text{con} \qquad \bar{I}_i = \frac{1}{2t_i} \sum_{j,k} \left(\widehat{w_{ij}} \widehat{w_{ik}} \widehat{w_{jk}} \right)^{\frac{1}{3}}$$

donde Ci es el coeficiente de agrupamiento no ponderado e Ii denota la intensidad promedio "normalizada" de los triángulos en los que vértice i participa

Average clustering

- El coeficiente de agrupamiento para el gráfico es el promedio,

$$C = \frac{1}{n} \sum_{v \in G} C_v$$

- Dónde n es el número de nodos en G

NetworkX
Network Analysis in Python

- NetworkX es un paquete de Python para la creación, manipulación y estudio de la estructura, dinámica y funciones de redes complejas.

- Tutorial de networkx:

- Networkx/tutorial_full (1).ipynb at main · MLeon28/Networkx (github.com)

- Analizando la estructura de la red:

- Networkx/Untitled (1).ipynb at main · MLeon28/Networkx (github.com)

Gracias

Moises Omar León Pineda

moises.leon@cinvestav.mx

MLeon8 – GitHub :
https://github.com/MLeon8