Lección 4. Indicadores a nivel de red para analizar Redes Sociales

UCO ONLINE

Análisis de Redes Sociales

Indicadores a nivel de red para analizar Redes Sociales

UNIVERSIDAD D CÓRDOBA

El Análisis de Redes Sociales (ARS) estudia la interacción entre los actores en redes sociales. Es un área de investigación interdisciplinar en la que están implícitas la Sociología, la Psicología, la Antropología, la Física, las Matemáticas y la Informática, entre otras.

Los orígenes del ARS, para el desarrollo de conceptos sociológicos, puede situarse a comienzos de la década de 1930, donde el enfoque sociométrico se transforma como una forma de conceptualizar la estructura de las relaciones sociales establecidas entre pequeños grupos de individuos.

1. Introducción

El ARS implica tener en cuenta múltiples niveles de análisis (Figura 1). Las diferencias entre los actores se pueden interpretar en base a las limitaciones y oportunidades que surgen de la forma en que éstos están inmersos en las redes; la estructura y el comportamiento de las redes están basados en y activados por las interacciones locales entre los actores.

En esta lección examinaremos algunos de las propiedades esenciales de las redes sociales que nos permiten obtener información relevante de las redes sociales. Por ejemplo, el hecho de que existan muchas conexiones indica que la información se puede extender muy rápidamente, estas poblaciones son capaces de movilizar sus recursos y también de disponer de múltiples y diversas perspectivas para resolver los problemas.

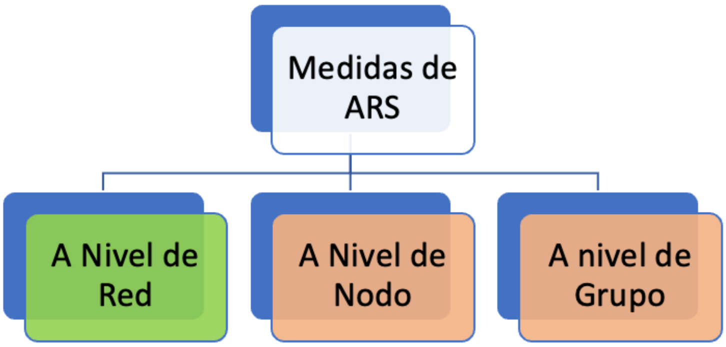


Figura 1. Diferentes niveles de análisis de redes sociales.

En un ARS podemos valorar diferentes medidas a distintos niveles que iremos viendo a lo largo del curso. En esta lección empezaremos con las medidas a nivel de red [1]:

- **Medidas globales (a nivel de red):** proporcionan información más compacta que permite evaluar la estructura global de la red, proporcionando información sobre propiedades importantes de los fenómenos sociales subyacentes.

- **Medidas locales (a nivel de nodos):** proporcionan información más específica evaluando la estructura y relaciones de cada nodo en la red. Están basadas en el concepto general de centralidad (redes no dirigidas) o prestigio (redes dirigidas), una medida general de la posición de un nodo en la estructura global de la red social. Se usan para identificar los nodos claves de la red. Muestran como las relaciones se concentran en unos pocos individuos, dando una idea de su poder social.

- **Medidas a nivel de grupo:** nos permiten analizar las diferentes agrupaciones que se pueden realizar entre los nodos y cómo de cohesionados están esos grupos con respecto al resto.

2. Indicadores a nivel de red

Los indicadores a nivel de red también son conocidos como medidas globales. Estos indicadores permiten evaluar la estructura global de la red para entender el comportamiento de los sistemas complejos (en este caso, sistemas sociales) que generan dichas redes [2].

Para el análisis de estas propiedades existe un conjunto de medidas que nos dan información sobre la estructura de una red y nos permite analizar las diferentes conexiones que existen en la red y cómo de cerca están unos actores de otros: tamaño, densidad, diámetro, distancia media, grado medio y coeficiente medio de agrupamiento.

Estos indicadores tienen la utilidad de servir sobre todo para el análisis comparativo de la cohesión relativa de diversas redes. No obstante, para redes pequeñas, a menudo resulta útil examinar directamente el grafo para tener una primera evaluación [3].

En la figura 2 se muestra la red social de amistades entre 34 integrantes de un club de kárate en una universidad estadounidense en la década de 1970. A priori, se trata de una red pequeña, con solamente 34 nodos y 2 grupos diferenciados. Los dos grupos son representados por los dos colores de los nodos y enlaces. Además, el tamaño de los nodos está en función del grado de cada actor, esto es, del número de conexiones que tienen. Así, en cada grupo nos encontramos actores más representativos por tener un mayor número de conexiones.

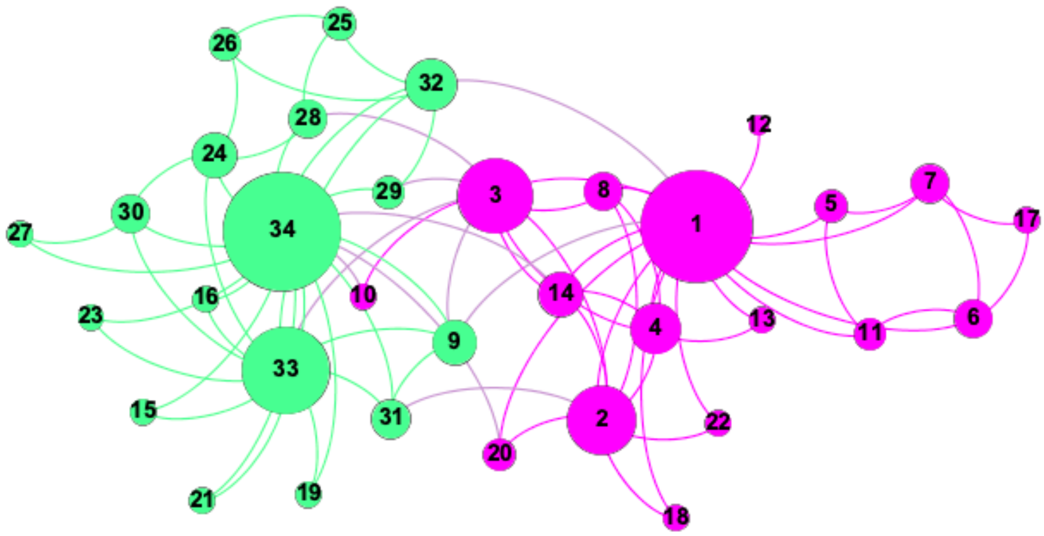


Figura 2. Red social de amistades de un club

La visualización de la matriz (Figura 3) también nos puede ayudar a ver la información, podemos ver que, al ser una relación de amistad, se ha considerado que es simétrica y por tanto los vínculos son recíprocos. Se puede ver que, al existir muchos valores con ceros, no se trata de un grafo muy denso. Se puede apreciar que determinados nodos cuentan con un gran número de unos, lo que es indicativo de que tienen un gran número de conexiones.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0
2	1	0	1	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0
3	1	1	0	1	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0
4	1	1	1	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1
10	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
11	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
17	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
20	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
22	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	1
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
28	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
29	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	1
31	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
32	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	1
33	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	1	0	1	0	1	1	0	0	0	0	0	1	1	1	0	1
34	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	0

Figura 3. Matriz que representa las amistades de un club

2.1 Tamaño de una red

El *tamaño de la red es* el número de nodos que tiene la red. Esta medida resulta muy relevante en el ARS. A medida que el grupo crece, la proporción de todos los lazos que pudiesen (lógicamente) estar presentes disminuirá, así por ejemplo no sería difícil que los 10 alumnos de un grupo de máster se conociesen entre sí bastante bien, pero sería difícil que los 100 alumnos de un Grado se conociesen entre sí.

De este modo, el tamaño es crítico para la estructura de las relaciones sociales a causa de los recursos limitados y las capacidades de que cada nodo dispone para construir y mantener lazos. En la Figura 4 y 5 podemos ver dos redes, una con 35 nodos y otra con 325, en las que se puede apreciar que las relaciones y comunicación se vuelve más compleja conforme aumenta el número de nodos.

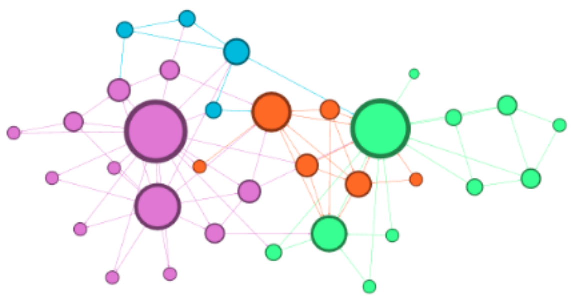


Figura 4. Red social de tamaño 34

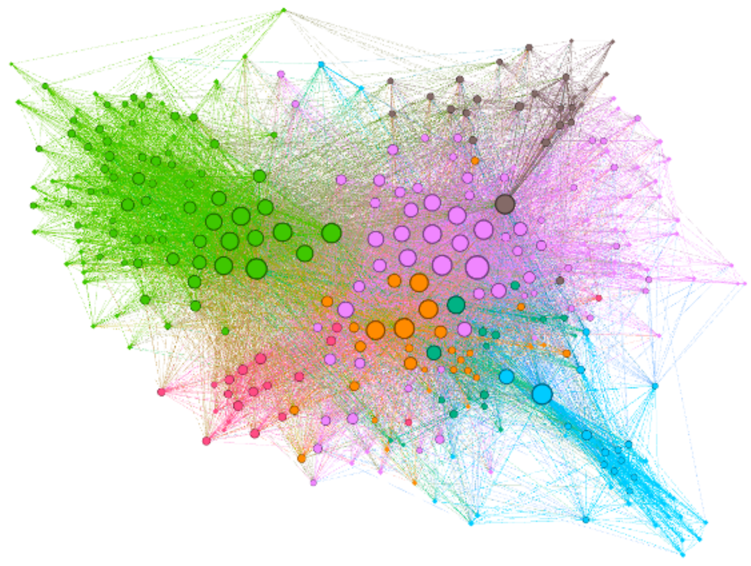


Figura 5. Red social de tamaño 325

2.2 Densidad de una red

Representa el número de vínculos que se establecen entre los nodos con respecto al número máximo que pudiera establecerse si todos los actores estuvieran conectados directamente por una línea con todos los demás.

$$D = \frac{L}{L_{max}}$$

- **L**: número de enlaces
- **Lmax**: el número de enlaces máximos si todos los nodos estuviesen conectados
    - Grafo dirigido: (n x (n-1)) siendo n el número de nodos
    - Grafo no dirigido: (n x (n-1))/2 siendo n el número de nodos

Ejemplo con datos dirigidos y una red con 10 actores → Lmax = 90 relaciones con lazos.

Ejemplo con datos no dirigidos y una red con 10 actores → con lazos recíprocos o simétricos el número Lmax = 45 (ya que la relación AB, sería la misma que BA).

En figura 6 se muestra una red con una densidad de 0.9412 lo que indica que es una red cerca a se completa, tienen la mayoría de los enlaces. En el caso de la figura 7 vemos una red poco densa con un valor de 0.1765, donde muchas de las conexiones no están presentes.

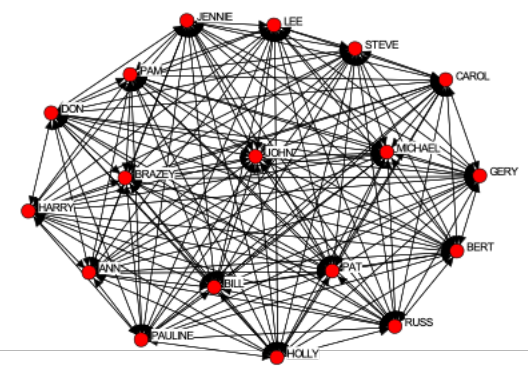


Figura 6. Red social con densidad = 0.9412

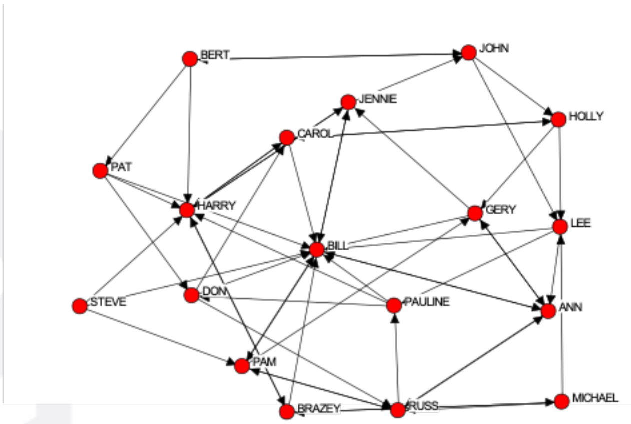


Figura 7. Red social con densidad = 0.1765

2.3 Diámetro de una red

La definición de la distancia entre actores en una red es usada por muchos algoritmos para definir propiedades más complejas de las posiciones de los individuos y de la estructura de la red en su conjunto. Una medida de distancia especialmente utilizada es el camino mínimo o distancia geodésica que se define como el número de relaciones en el camino más corto posible de un nodo a otro. Si dos nodos son adyacentes, la distancia entre ellos es una (un mensaje necesitaría un paso para llegar desde el emisor hasta el receptor). Si el nodo A tiene un enlace a B y B tienen un enlace a C, entonces los nodos A y C están a una distancia de dos. El estudio de las diferentes distancias entre cada par de nodos puede ser importante para entender las diferencias entre nodos en las limitaciones y oportunidades que tienen como resultado de su posición.

El diámetro de una red se define como la longitud del camino mínimo más largo de la red. Esta medida nos permite conocer la proximidad entre pares de actores en la red, indicando cómo de lejos están en el peor de los casos.

Las redes más dispersas suelen tener un mayor diámetro que las más densas al existir menos caminos entre cada par de nodos. La figura 8 muestra el camino más largo entre todos los caminos mínimos entre cada par de actores.

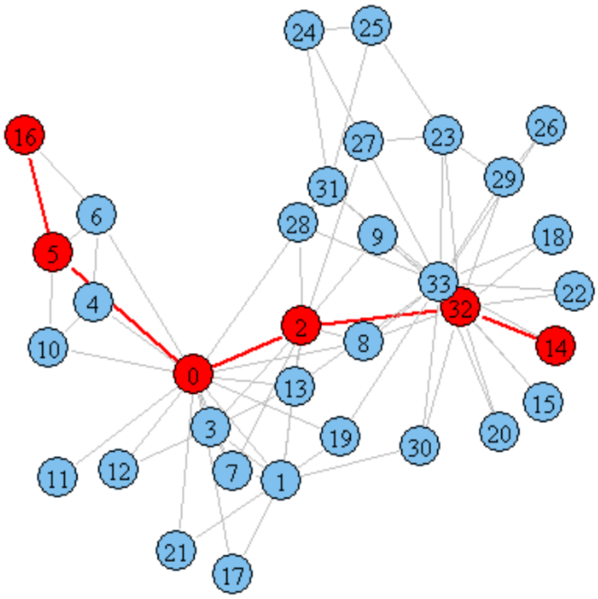


Figura 8. Camino mínimo más largo (diámetro = 6)

2.4 Distancia media

Aunque la distancia geodésica es importante, el resto de las distancias también debe de tenerse en cuenta. Cualquier relación entre dos nodos puede proveer a cada otro con oportunidades y limitaciones. A veces es interesante conocer de cuántas maneras se pueden conectar dos nodos con una distancia dada. El hecho de tener múltiples conexiones puede indicar una conexión más fuerte entre dos nodos que una sola relación.

En este contexto la distancia media resulta un valor muy apropiado para determinar cómo de lejos están los distintos actores en media. Si las distancias medias son largas, indica que es necesario mucho tiempo para que una información pueda difundirse.

En grafos dirigidos la distancia media se calcula:

$$d_{media} = \frac{1}{L_{max}} \sum_{i,j \neq i} d_{ij},$$

Donde,

- d_{ij} es la longitud del camino mínimo entre los nodos i y j.
- L_{max} número de enlaces máximos en grafos dirigidos.

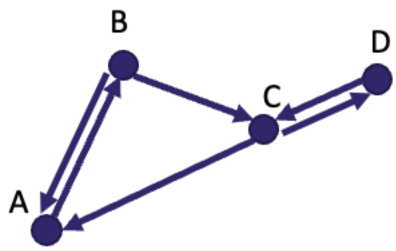
En grafos no dirigidos:

$$d_{media} = \frac{1}{L_{max}} \sum_{i,j>i} d_{ij,}$$

Donde,

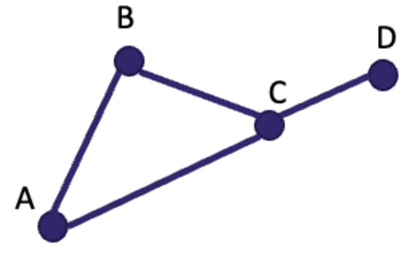
- $d_{ij} = d_{ji}$
- d_{ij} es la longitud del camino mínimo entre los nodos i y j
- L_{max} número de enlaces máximos en grafos no dirigidos

Ejemplo del cálculo de la distancia media en grafos dirigidos:



- $d_{media} = \frac{1}{L_{max}} \sum_{i,j\neq i} d_{ij,}$
- $d_{media} = \frac{1}{12} (d_{AB} + d_{AC} + d_{AD} + d_{BA} + d_{BC} + d_{BD}$
- $+ d_{CA} + d_{CB} + d_{CD} + d_{DA} + d_{DB} + d_{DC})$
- $d_{media} = \frac{1}{12} (1+2+3+1+1+2+1+2+1+2+3+1)$
- $d_{media} = \frac{20}{12} = 1{,}667$

Ejemplo del cálculo de la distancia media en grafos no dirigidos:



- $d_{media} = \frac{1}{L_{max}} \sum_{i,j>i} d_{ij,}$
- $d_{media} = \frac{1}{6}(d_{AB} + d_{AC} + d_{AD} + d_{BC} + d_{BD} + d_{CD})$
- $d_{media} = \frac{1}{6}(1 + 1 + 2 + 1 + 2 + 1)$
- $d_{media} = \frac{8}{6} = 1{,}333$

2.5 Grado medio

El grado de un nodo es el número de conexiones/enlaces que tiene y nos determina cómo de relacionados están los nodos en media. Si es un grafo dirigido se deben mirar los enlaces de entrada y salida.

Con respecto a los enlaces de salida, nos da información de un nodo considerándolo como la fuente de la información. Tanto los nodos con pocos lazos de salida (nodos periféricos de la red) como los nodos con muchos lazos de salida (nodos en el centro de la red) suelen tener patrones de conducta más predecibles que los nodos con algunos lazos.

Con respecto a los enlaces de entrada, nos da información de un nodo considerándolo como el receptor de la información. Los nodos que reciben muchos enlaces son considerados como prestigiosos y también pueden ser más poderosos. No obstante, el hecho de tener demasiados enlaces puede producir una sobrecarga de información, debido a recibir mensajes contradictorios de varias fuentes.

El grado medio nos da información del número medio de conexiones en la red, la varianza de este valor nos puede determinar si existen muchas diferencias entre los diferentes nodos. En grafos no dirigidos el grado medio se calcula:

$$\mathrm{k_{media}} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{k_i}$$

Donde,

- k_i: $grado\ del\ nodo\ i,$ $número\ de\ enlaces\ que\ los\ conecta\ con\ otros$
- N: número de nodos de la red

En grafos dirigidos:

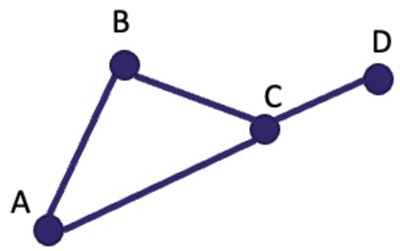
$$\mathrm{k_{media}^{in}} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{k_i^{in}}$$

$$\mathrm{k_{media}^{out}} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{k_i^{out}}$$

Donde,

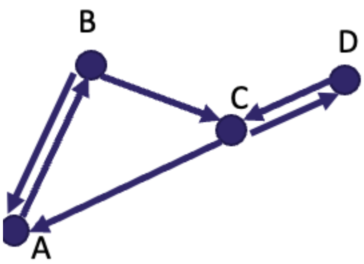
- $k_i^{in} grado\ de\ entrada\ del\ nodo\ i,$ $número\ de\ enlaces\ que\ recibe$
- $k_i^{out} grado\ de\ salida\ del\ nodo\ i$, número de enlaces que salen de él
- N número de nodos de la red

Ejemplo del cálculo del grado medio en grafos no dirigidos:



- $k_{media} = \frac{1}{4}(k_A + k_B + k_C + k_D)$
- $k_{media} = \frac{1}{4}(2 + 2 + 3 + 1)$
- $k_{media} = \frac{7}{5} = 1{,}4$

Ejemplo del cálculo del grado medio en grafos dirigidos:



- $k_{media}^{in} = \frac{1}{N}\sum_{i=1}^{N} k_i^{in} = \frac{1}{4}\left(k_A^{in} + k_B^{in} + k_C^{in} + k_D^{in}\right)$
- $k_{media}^{in} = \frac{1}{4}(2 + 1 + 2 + 1) = 6/4 = 1{,}5$
- $k_{media}^{out} = \frac{1}{N}\sum_{i=1}^{N} k_i^{out} = \frac{1}{4}(k_A^{out} + k_B^{out} + k_C^{out} + k_D^{out})$
- $k_{media}^{out} = \frac{1}{4}(1 + 2 + 2 + 1) = 6/4 = 1{,}5$

2.6 Coeficiente medio de agrupamiento

Indica la probabilidad de que dos vecinos de un nodo de la red escogidos aleatoriamente estén conectados entre sí.

El coeficiente de agrupamiento de cada nodo se calcula:

$$C_i = \frac{2L_i}{k_i(k_i - 1)}$$

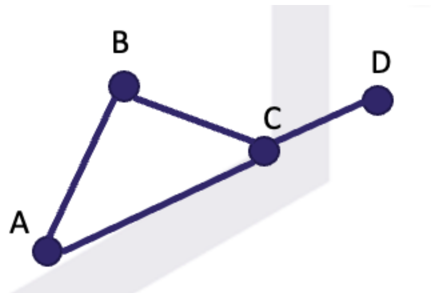
Donde,

- k_i: $grado\ del\ nodo\ i,$ $número\ de\ enlaces\ que\ los\ conecta\ con\ otros$
- N: número de nodos de la red
- L_i: $número\ de\ enlaces\ entre\ los\ vecinos\ del\ nodo\ i$
- C_i: $coeficiente\ de\ agrupamiento\ del\ nodo\ i.$
  - $C_i = 0\ indica\ que\ ninguno\ de\ los\ vecinos\ de\ i\ están\ conectados\ entre\ sí$
  - $C_i = 1\ indica\ que\ todos\ están\ conectados$

El Coeficiente medio de agrupamiento se calcula, por tanto, como:

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

Ejemplo del cálculo del coeficiente medio de agrupamiento:



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

$$\langle C \rangle = \frac{1}{4}(C_A + C_B + C_C + C_D) = \frac{1}{4}\left(\frac{2}{2} + \frac{2}{2} + \frac{2}{6} + 0\right) = \frac{7}{12} = 0{,}583$$

Referencias

[1] Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2018). *Analyzing social networks*. Sage.

[2] Dey, N., Borah, S., Babo, R., & Ashour, A. S. (2018). *Social network analytics: computational research methods and techniques*. Academic Press.

[3] Scott, J. (2000). Social Network Analysis. A handbook. 2º edition. Sage.