Lección 3. Representación formal de las Redes Sociales

UCO ONLINE

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.

Estos lazos interpersonales entre miembros de un grupo fueron representados mediante los denominados sociogramas, que pueden definirse como gráficos en los que los individuos se representan como nodos y las relaciones entre ellos como líneas resultaron esenciales para descubrir las estructuras ocultas de los grupos mediante la identificación de protagonistas, alianzas y subgrupos, entre otras cosas

1. Introducción

Para analizar una red social es necesario realizar una descripción completa y rigurosa de sus relaciones resultando crucial una representación que nos permita obtener una representación de los datos de forma concisa y sistemática, poder automatizar el análisis de información y facilitar así el análisis.

Las representaciones empleadas en ARS utilizan grafos (figura 1) y matrices (figura 2).

	Luis	Caro	Pedro	María
Luis	--	1	1	0
Caro	0	--	1	0
Pedro	1	1	--	1
María	0	0	1	--

Figura 1. Representación de las relaciones de amistad mediante una matriz

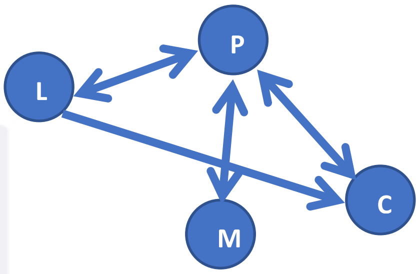


Figura 2. Representación de las relaciones de amistad mediante un grafo.

En esta sección se estudiarán los principales fundamentos de ambas representaciones.

2. Representación mediante Grafos

La utilización de grafos para representar las redes sociales es conocido como sociogramas [1]. Los analistas de redes utilizan esta representación gráfica que consiste en la representación de un círculo etiquetado para cada actor en la población y segmentos de línea entre pares de actores para representar el hecho que existe un vínculo entre ellos.

Entre las características de los sociogramas encontramos:

- Pueden representar un único tipo de relaciones entre los actores (simple), o más de un tipo de relación (múltiple).
- Cada vínculo o relación puede ser orientado o no. Los vínculos orientados pueden ser recíprocos (A nomina a B y viceversa) y suelen representarse con una flecha con doble punta.
- Los vínculos pueden representares con
    - Valores binarios: representan presencia o ausencia de vínculo
    - Con signos: representa un vínculo negativo, positivo o ningún vínculo
    - Ordinales: representan si el vínculo es fuerte, menos fuerte, etc.
    - Ponderado: miden un intervalo o nivel promedio.

2.1 Conceptos básicos de los Grafos

Un grafo G es un par de conjuntos (V,E) donde:

- $V=\{v_1,v_2,\ldots.v_n\}$ es el conjunto de vértices
- $E=\{(v_i,v_j),(v_{i'},v_{j'})\ldots\ldots\}$ es un conjunto de enlaces o ramas que representa los pares no ordenados de elementos de V.

Podemos encontrar diferentes clasificaciones de los grafos:

- En función de la orientación de los enlaces
  - **Grafos recíprocos:** los enlaces no llevan flechas (dirección) e indican que existe una comunicación recíproca entre los dos nodos.
  - **Grafos orientados:** los enlaces llevan flechas (dirección) que representan de donde a donde se produce el flujo de información.

- En función de los pesos de los enlaces
  - **Grafos binarios**: simplemente consideramos si existe el enlace o no.
  - **Grafos ponderados:** a cada rama del grafo se le asocia un número. El número asociado a cada rama puede indicar entre otras cosas una distancia, una capacidad, un valor temporal, etc…

Una vez definido el grafo, debemos tener presente los siguientes conceptos que nos dan información a nivel global de grafo:

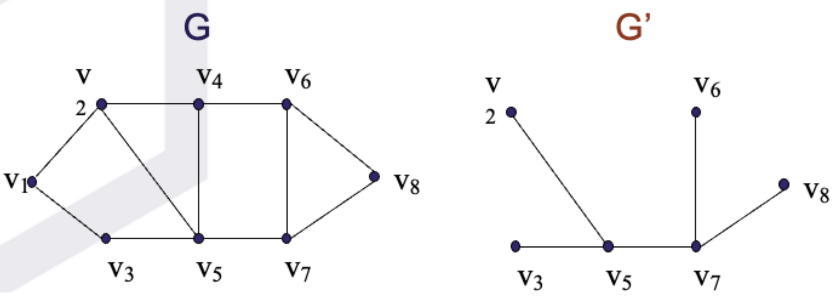
- **Tamaño** del grafo: número de nodos del grafo.
- La **distancia** entre dos nodos: es la longitud del camino entre dos nodos que es equivalente al número de ramas.
- El **diámetro** de un grafo: es la máxima distancia entre dos de sus nodos (considerando las distancias mínimas entre ellos)

A nivel de nodo dentro de un grafo tenemos la medida de grado que es ampliamente utilizada en el ARS y además es la base de otros indicadores que se verán durante este curso:

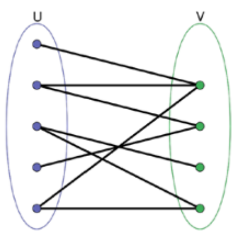
- **Grado de un nodo**: es el número vecinos que tiene dicho nodo. Dos nodos de un grafo son vecinos o adyacentes si existe una rama que los conecta.
    - En los grafos dirigidos se calcula el grado de entrada y el grado de salida.
    - En los grafos ponderados, el grado se puede promediar por el valor asociado a cada rama.
    - Un grafo se dice que es **regular** si todos los nodos tienen el mismo grado.

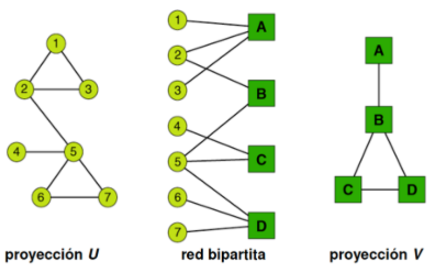
Medidas relacionadas con la cohesión y conectividad de un grafo que nos permiten determinar influencias y situaciones de poder frente a grupos que puedan quedarse inconexas:

- **Subgrafo:** un grafo G'=(V',E') es un subgrafo de un grafo G=(V,E) si V' es un subconjunto de V y E' es un subconjunto de E

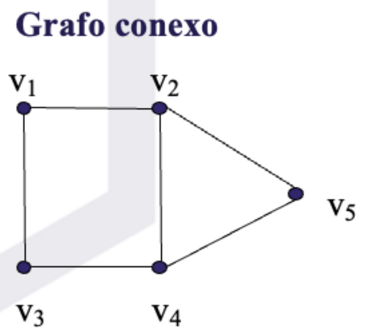


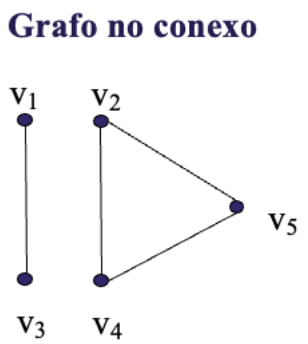
- **Grafo bipartito:** un grafo es ***bipartito*** si sus elementos pueden dividirse en dos conjuntos disjuntos. Se pueden obtener dos proyecciones.



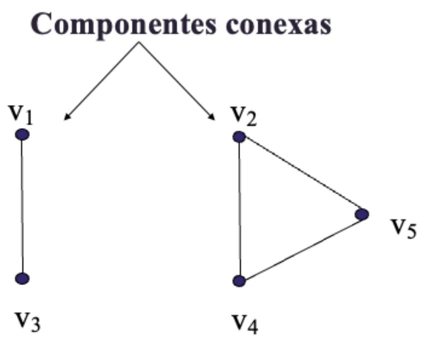


- **Grafo conexo:** un grafo es conexo si para cada par de nodos del grafo existe al menos un camino que los une.

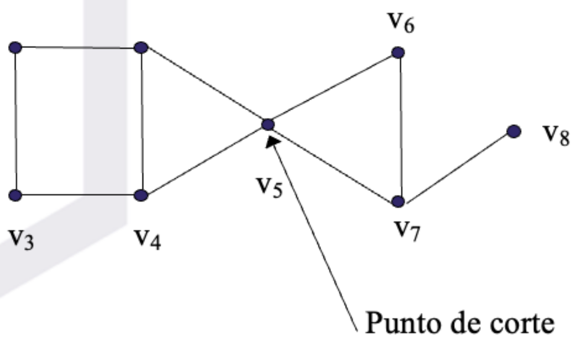




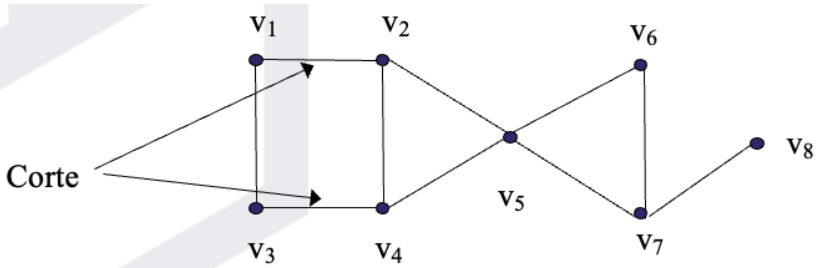
- **Componente conexa:** es cada uno de los subgrafos máximos que son conexos. Si la mayor componente conexa de la red incluye una fracción significativa de la misma, se denomina componente gigante.



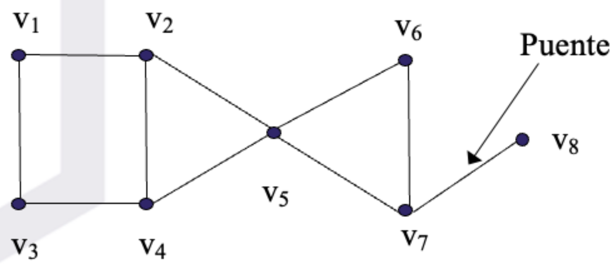
- **Punto de corte:** es un nodo que desconecta un grafo conexo.



- Un **corte**: es un conjunto de ramas que desconecta un grafo conexo



- Un **puente:** es un corte compuesto por una única rama.



Finalmente se estudian los enlaces que existen entre nodos. El estudio nos permiten determinar las diferentes relaciones entre cada dos nodos de la red:

- Un **paseo** de un nodo **u** a un nodo **v** es una secuencia de vértices y ramas que unen dichos vértices. El número de ramas del paseo es su **longitud.**
- Un **sendero** es un paseo en el cual no se repiten ramas**.** El número de ramas del paseo es su **longitud.**
- Un **camino** es un paseo en el cual todos los vértices $\{v_0,v_1,....v_k\}$ son distintos. El número de ramas del paseo es su **longitud.**
- Un **camino mínimo** entre dos nodos es aquel de menor longitud de entre todos los posibles caminos entre ambos nodos. El número de ramas del paseo es su **longitud.** En ARS suele denominarse **distancia geodésica**.
- Un **paseo cerrado** es un paseo $\{v_0,v_1,....v_k\}$ tal que $v_0 = v_k$**.** El número de ramas del paseo es su **longitud.**

- Un **circuito** es un paseo cerrado en el que no se repiten ramas. El número de ramas del paseo es su **longitud.**
- Un **ciclo** es un circuito en el que no se repiten vértices. El número de ramas del paseo es su **longitud.**

3. Representación mediante Matrices

Las matrices permiten llevar a cabo varias operaciones matemáticas que resultan muy eficientes e intuitivas en el ARS. Los analistas de redes utilizan esta representación gráfica que consiste en la representación de una matriz donde los actores son identificados en las filas y columnas y en cada celda se representa el vínculo que existen entre ellos.

	Luis	Caro	Pedro	María
Luis	--	1	1	0
Caro	0	--	1	0
Pedro	1	1	--	1
María	0	0	1	--

El valor de la diagonal dependiendo del estudio puede ser irrelevante si no interesa la relación con uno mismo y muy relevante si interesa la relación con uno mismo.

Entre las características de la representación en matrices encontramos:

- Pueden representar un único tipo de relaciones entre los actores (simple), o más de un tipo de relación (múltiple).
- Cada vínculo o relación puede ser orientado o no. Los vínculos orientados pueden ser recíprocos (A nomina a B y viceversa).
- Los vínculos pueden representares con
    - Valores binarios: representan presencia o ausencia de vínculo. Se representan con una matriz con unos y ceros, indicando la presencia o ausencia de cada relación posible entre pares de actores

- Con signos: representa un vínculo negativo, positivo o ningún vínculo. Se representan con una matriz con -1, 0 y +1.
- Ordinales: representan si el vínculo es fuerte, menos fuerte, etc. Se representan con una matriz donde la magnitud numérica del vínculo medido se coloca como un elemento de la matriz.
- Ponderado: miden un intervalo o nivel promedio. Se representan con una matriz donde la magnitud numérica del vínculo medido se coloca como un elemento de la matriz

La relación entre la matriz y un grafo se muestra en la figura 3. Las filas representan el origen de los vínculos dirigidos y las columnas los destinos.

	Luis	Caro	Pedro	María
Luis	--	1	1	0
Caro	0	--	1	0
Pedro	1	1	--	1
María	0	0	1	--

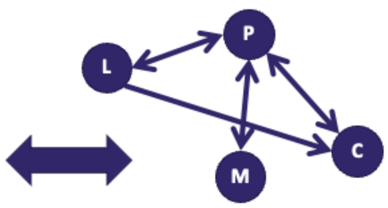


Figura 3. Relación entre la representación basada en matrices y grafos.

3.1 Conceptos básicos de las matrices

Tipos de matrices:

- **Matriz asimétrica**: representa vínculos dirigidos (vínculos que van de un origen a un receptor). Es decir, el elemento i,j no necesariamente es igual que el elemento j,i.
- ***Matriz simétrica:*** representa vínculos no dirigidos (vínculos recíprocos o concurrencia), es decir, el elemento i,j sería igual al elemento j,i.
- **Matriz de adyacencia**: representa con un 1 que existe un vínculo entre i,j y con un 0 que no existe dicho vínculo.

Operaciones con matrices que nos permiten obtener información relevante en un estudio de ARS:

- **Permutaciones en matrices**: es el cambio del orden de las líneas y columnas. En el ejemplo permutamos a Pedro por Caro.

	Luis	Caro	Pedro	María
Luis	--	1	1	0
Caro	0	--	1	0
Pedro	1	1	--	1
María	0	0	1	--

	Luis	Pedro	Caro	María
Luis	--	1	1	0
Pedro	1	--	1	1
Caro	0	1	--	0
María	0	1	0	--

- **Bloques de la matriz**: los bloques se forman trazando líneas de división a través de las filas y columnas de la matriz. En el ejemplo una vez permutado Pedro por Caro podemos hacer la división entre Pedro y Caro y se divide en función del sexo de los actores. Permite entender cómo algunos conjuntos de actores se encuentran "inmersos" en roles sociales o en entidades grandes.

	Luis	Pedro	Caro	María
Luis	--	1	1	0
Pedro	1	--	1	1
Caro	0	1	--	0
María	0	1	0	--

- **Transposición de una matriz:** es el intercambio de filas y columnas de forma tal que *i se* convierte en *j y viceversa.* Si se toma la matriz traspuesta de una matriz de adyacencia directa y se examinan sus vectores fila, se estará observando los orígenes de los vínculos dirigidos a un actor. La correlación entre una matriz de adyacencia y la traspuesta de esa matriz, mide el grado de reciprocidad de los vínculos.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \Rightarrow A^t = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix}$$

- **Inversa de una matriz**: origina una matriz que, al multiplicarla por la matriz original, produce una nueva matriz con números 1 en la diagonal principal y cero en las demás posiciones. Las matrices inversas suelen utilizarse para hacer otros cálculos en el análisis de redes sociales, pueden verse como un “opuesto” de la matriz original. Hay distintas formas de calcular la inversa de una matriz: método de Gauss-Jordan, usando determinantes, ….

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 2 & 4 \end{pmatrix},\ A^{-1} = \begin{pmatrix} 0 & -2 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{pmatrix}$$

- **Suma/Resta de matrices**: se suma o resta cada elemento correspondiente i,j de dos (o más) matrices. Las matrices deben tener el mismo número de elementos i y j. Los valores de i y j deben estar en el mismo orden en cada matriz. Se emplean en análisis de redes cuando se intenta simplificar o reducir la complejidad de múltiples datos a formas simples.

$$A + B = \begin{pmatrix} 1 & -3 & -5 \\ -2 & 4 & 6 \end{pmatrix} + \begin{pmatrix} 4 & 5 & -3 \\ 1 & 8 & -5 \end{pmatrix} = \begin{pmatrix} 5 & 2 & -8 \\ -1 & 12 & 1 \end{pmatrix}$$

- **Multiplicación de matrices**: para multiplicar dos matrices, éstas deben ser compatibles con la multiplicación. El número de filas en la primera matriz debe ser igual al número de columnas en la segunda
    - Para multiplicar dos matrices, se comienza en la esquina superior izquierda de la primera matriz, y se multiplica cada celda en la primera fila de la primera matriz por el valor en cada celda de la primera columna de la segunda matriz, y luego se suman los resultados.
        - Se procede de igual forma para cada celda en cada fila de la primera matriz, multiplicada por la columna en la segunda.

$$A \cdot B = \begin{pmatrix} -2 & -3 & 0 & 1 \\ 3 & -4 & 2 & -1 \\ 0 & 1 & 6 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -3 \\ 5 & 0 \\ 6 & -2 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -18 & 8 \\ -4 & -15 \\ 39 & -8 \end{pmatrix}$$

- **Multiplicación de matrices aplicada a la matriz de adyacencia:** la matriz de adyacencia (A), nos dice si hay un camino entre dos actores de longitud 1. Un 1 representa la presencia de un camino, un 0 representa la ausencia de éste. Si multiplicamos la matriz de adyacencia por sí misma (A^2) muestra el número de caminos entre dos nodos que tienen longitud dos, y si se vuelve a multiplicar obtendríamos el número de caminos de longitud 3, y así se puede ir obteniendo los caminos de diferente longitud.

A =

0	1	1	0
0	0	1	0
1	1	0	1
0	0	1	0

A^2=

1	1	1	1
1	1	0	1
0	1	3	0
1	1	0	1

Referencias

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