

Análisis de Comunidades

JULIO 2021



Que es un Grafo?

La red es un conjunto de relaciones entre dos o mas entidades. Las relaciones se representan gráficamente por líneas y los elementos o entidades por nodos.

Los grafos son realmente una red, es necesario estudiar algunos conceptos básicos y formales y notaciones de teoría de grafos, junto con algunas propiedades fundamentales que caracterizan una red.

Se componen de tres partes, el conjunto de vértices, nodos o puntos, el conjunto de aristas, enlaces o líneas que pueden ser dirigidas o no dirigidas y los pesos que pueden contener o no las aristas

El Análisis de las Redes Sociales (ARS) es fruto del aporte de la matemática (análisis cuantitativo por medio de la teoría de grafos), informática (software de análisis de redes) y las ciencias humanísticas (análisis cualitativo por medio de sociogramas); y consiste en estudiar la relación entre actores: datos, objetos, grupos de personas, organizaciones, etc.

Tipos de Grafos

Un grafo puede ser dirigido o no dirigido, bipartito o ponderado, en donde las aristas pueden contener pesos o no.

Grafo simple. o simplemente *grafo* es aquel que acepta una sola una arista uniendo dos vértices cualesquiera. Esto es equivalente a decir que una arista cualquiera es la única que une dos vértices específicos. Es la definición estándar de un grafo.

No Dirigido. son grafos en los cuales sus aristas son no dirigidas. La relación entre los nodos es reciproca.

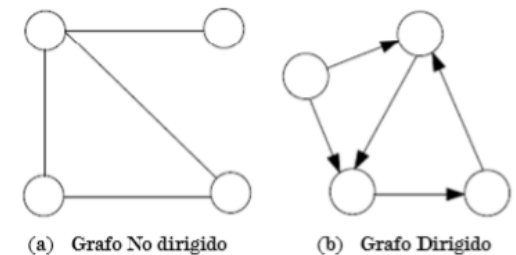
Grafo dirigido. Son grafos en los cuales se ha añadido una *orientación* a las aristas, representada gráficamente por una flecha. La dirección no es bidireccional.

Grafo etiquetado. Grafos en los cuales se ha añadido un *peso* a las aristas (número entero generalmente) o un *etiquetado* a los vértices.

Grafo aleatorio. Grafo cuyas aristas están asociadas a una *probabilidad*.

Hipergrafo. Grafos en los cuales las aristas tienen más de dos extremos, es decir, las aristas son incidentes a 3 o más vértices.

Grafo infinito. Grafos con conjunto de vértices y aristas de *cardinal infinito*.



Redes Sociales

A partir del año 1995 surgieron las redes sociales en Internet, sitios web dedicados a brindar servicios de comunicación. En la década del año 2000 se popularizaron con el surgimiento de Facebook y Twitter, diariamente se genera y comparte una inmensidad de contenido que puede ser analizado para descubrir información relevante respecto a las comunidades virtuales. Esta información puede usarse para conocer sus intereses, establecer redes, así como fácilmente una empresa puede obtener el grado de aceptación de sus clientes hacia un producto y/o servicio.

Objetivos de un análisis de redes sociales:

Detectar Nodos Influyentes

Detectar Comunidades por tipos de asociaciones

Medidas en un Grafo

La medida Grado Medio permite conocer el grado de entrada y salida de un nodo, cuantas veces está conectado un nodo dentro de toda la red. El grado de entrada y el grado de salida definen el rol de influyente y el rol de comunicador, respectivamente dentro de la red.

La medida “Diámetro de red” brinda el número de nodos necesarios para transportarse en toda la red de un extremo a otro.

Betweenness (Intermediación): Es una medida que brinda el número de rutas más cortas entre dos nodos que pasan a través de un nodo en particular. Una alta intermediación puede sugerir que el nodo está conectando varias partes de la red, es decir, estos nodos que mantienen la red unida.

Closeness (Cercanía): Es una medida que indica qué tan cerca está un nodo de todos los demás nodos de una red. Una alta cercanía significa que hay una gran distancia promedio a otros nodos en la red.

Eccentricity (Excentricidad): Excentricidad es una medida que brinda la distancia entre un nodo y el nodo que está más alejado de él. Una excentricidad alta significa que el nodo más alejado de la red está muy lejos, y una excentricidad baja significa que el nodo más alejado en realidad está bastante cerca.

Modularidad. Se refiere al agrupamiento natural de las redes sociales. Dentro de un grupo, siempre existen sub-grupos de personas. Estos sub-grupos determinan las dinámicas de todo el grupo.

Medidas en un Grafo

OutDegree, que mide la cantidad de conexiones que salen de cada Nodo

Weighted OutDegree. Esta estadística pondera el número de interacciones de salida tomando en cuenta su importancia.

InDegree, que mide la cantidad de conexiones de entrada a cada Nodo

Weighted InDegree, que incluye no solo las interacciones de entrada, pero también la importancia que se le da a dichas entradas.

El **algoritmo HITS** (acrónimo del inglés *Hypertext Induced Topic Selection*, también conocido como *hubs y autoridades*) es un algoritmo de análisis de enlaces

Medidas en un Grafo

Coeficiente de agrupamiento: nivel de agrupamiento de los nodos, para saber cómo de cohesionados o integrados están los agentes/actores.

Camino característico: mide el grado de separación de los nodos, para determinar lo contrario al punto anterior: cómo de separados o alejados están, y poder buscar así medidas para juntar más la relación entre agentes/actores.

Densidad: un grafo puede ser denso (cuando tiene muchas aristas) o disperso (muy pocas aristas). En este sentido, se puede interpretar como que hay mucha o poca conexión.

Coeficiente de clustering: Esta métrica permite calcular la cohesión de la red de forma sencilla tanto a nivel local como global, dando una idea de cómo de unida esta la comunidad. Dicho de otra forma, cuanto más cohesionada este la red mayor será la interacción de los usuarios entre sí y, por tanto, habrá una distribución del trabajo más equilibrada.

Agoritmos

los algoritmos de distribución siguen un principio básico basado en la fuerza, los nodos conectados se atraen entre sí y los nodos no enlazados se separan.

Fruchterman Reingold : Es un tipo de distribución que ordena todos los nodos creando una circunferencia, es funcional para tener una visualización de cada uno de los nodos que conforma la red.

Noverlap: Permite que no exista superposición entre los nodos. Puede darse el caso que nodos de un tamaño grande oculten a nodos pequeños.

Expansión: Esta distribución permite, como su nombre lo indica, expandir la red sin cambiar el diseño que ya posee.

Visualización de Redes

Principios estéticos que permitan facilitar la visualización de la red:

- Minimizar los cruces entre enlaces
- No permitir que los nodos se superpongan a enlaces
- En redes ponderadas, hacer que la longitud del enlace sea proporcional a su peso

Principales layouts:

Layout aleatorio Su funcionamiento es simple y persigue la idea más ingenua de la visualización de los grafos ya que consiste en ubicar los nodos de manera aleatoria dentro del plano y unirlos mediante los enlaces. Si bien tiene como ventaja la rapidez de cálculo, tiene como desventajas que no muestra la estructura real de la red.


Layout de árbol Una de las distribuciones más clásicas de un grafo consistente en ubicar a los nodos debajo su antecesor común, dando una representación fiable para redes con estructuras jerárquicas.

Layout radial Esta distribución, derivada del layout de árbol, consiste en ubicar al nodo central en el plano y distribuir el resto de los nodos en capas concéntricas alrededor de este en base al peso de sus enlaces.

Layouts guiados por fuerzas Si bien los layouts guiados por fuerzas no son en sí mismos un tipo de distribución, sí engloban una serie de algoritmos que basan la representación de los grafos en un modelo de objetos unidos por fuerzas. L

Ejercicio: Generación de Matriz de Datos

1. Enumerar 20 personas que puedan responder afirmativamente al menos una de las siguientes preguntas:
 1. ¿Si me voy de vacaciones esta persona vendría a cuidar a mi mascota?
 2. ¿Me prestaría dinero para una emergencia?
 3. ¿Escucharía mis problemas?
2. Escribir los nombres en un archivo en Excel (Estos nombres serán los nodos de la red)
 1. El Excel tendrá 2 columnas: id y label, id corresponde a un numero que puede ser incremental.
 2. Guardar el archivo como formato CSV
3. Elaborar una nueva matriz que corresponde a los enlaces:
 1. En un nuevo archivo colocar 3 columnas: Source, Target y Weight
 2. Establecer quienes se conocen entre si, por ejemplo ¿Juan (1) conoce a Francisco (2)? Si la respuesta es sí, Anotamos la relación (source – target)
 3. ¿Francisco (2) conoce a Adriana (3)? (No es necesario repetir que 2 conoce a 1 porque esa relación ya se estableció, así que seguimos escalonadamente)



The screenshot shows the Excel ribbon with the 'Inicio' tab selected. The 'Pegar' (Paste) dropdown menu is open, showing options like 'Cortar' (Cut), 'Copiar' (Copy), 'Copiar formato' (Copy format), and 'Portapapeles' (Clipboard). The active cell is B22. Below the ribbon, a table with 2 columns (A and B) and 21 rows is visible. Column A contains IDs from 1 to 21, and Column B contains names. The names are: Juan, Francisco, Adriana, Joan, Cristina, Felipe, Maria, Pablo, Ana, Montserrat, Carmen, Eduard, Laia, Pedro, Marina, John, Pam, Julia, Andrew, and Guadalupe.

	A	B
1	Id	Label
2		1 Juan
3		2 Francisco
4		3 Adriana
5		4 Joan
6		5 Cristina
7		6 Felipe
8		7 Maria
9		8 Pablo
10		9 Ana
11		10 Montserrat
12		11 Carmen
13		12 Eduard
14		13 Laia
15		14 Pedro
16		15 Marina
17		16 John
18		17 Pam
19		18 Julia
20		19 Andrew
21		20 Guadalupe

	A	B	C	D
1	Source	Target	Weight	Type
2	1	2	3	Undirected
3	1	3	3	Undirected
4	1	4	3	Undirected
5	1	5	5	Undirected
6	1	7	2	Undirected
7	1	8	2	Undirected
8	1	9	2	Undirected
9	1	10	2	Undirected
10	1	13	1	Undirected
11	1	14	1	Undirected
12	1	19	1	Undirected
13	1	20	1	Undirected
14	2	3	3	Undirected
15	2	4	3	Undirected
16	2	5	3	Undirected
17	2	7	2	Undirected
18	2	19	1	Undirected
19	2	20	1	Undirected
20	3	4	3	Undirected
21	3	5	3	Undirected
22	3	6	2	Undirected
23	3	7	2	Undirected

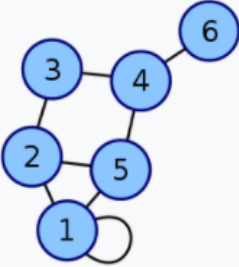
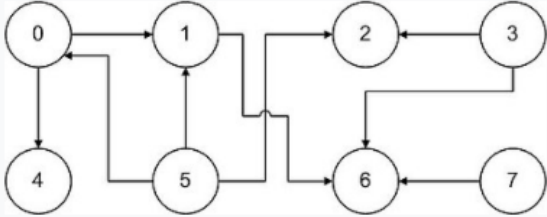
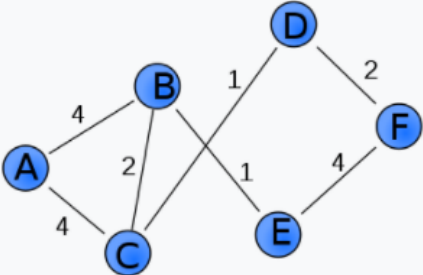
Ejercicio: Generación de Matriz de Datos

4. Establecer los pesos:

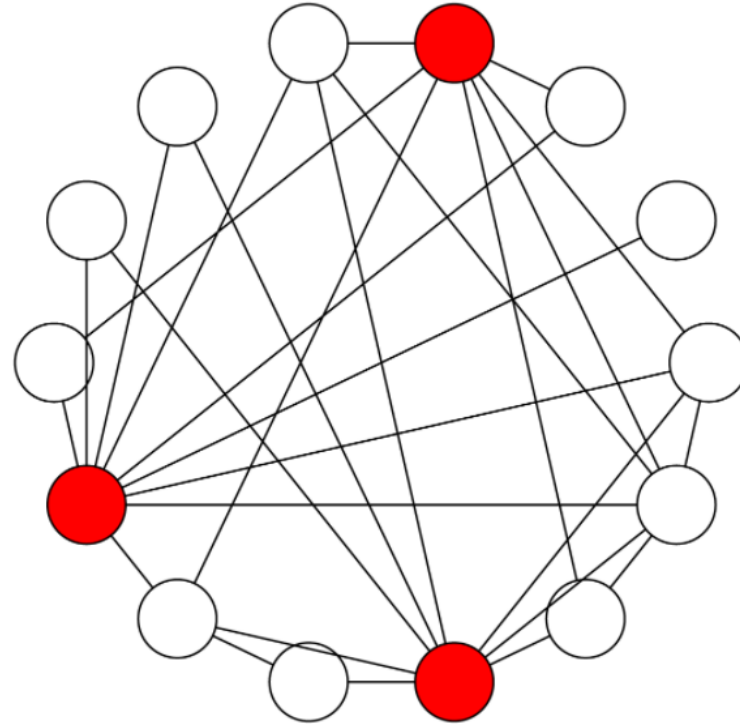
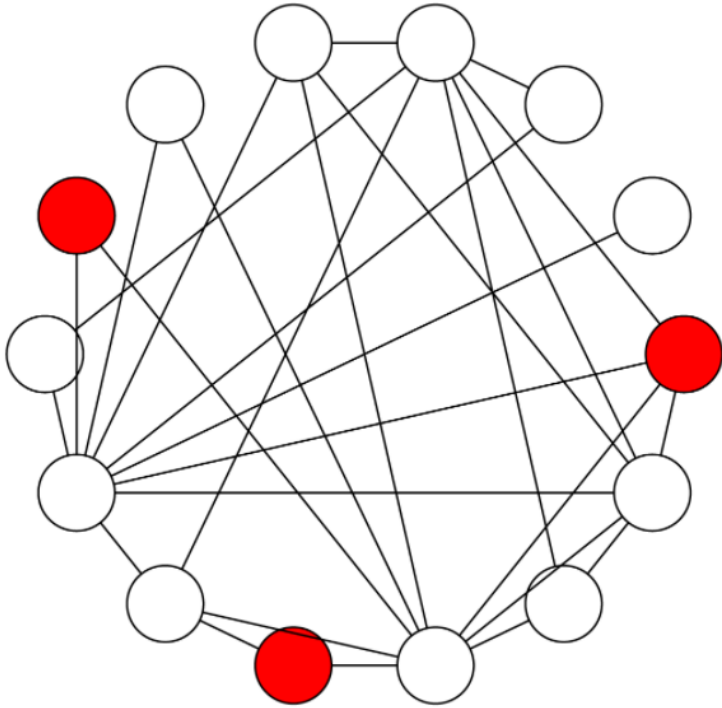
4. Para poner el peso (weight) de los enlaces haremos la siguiente pregunta:
 - Si no fuera por mí ¿ Juan (1) y Francisco (2) se conocerían? Y asignaremos los siguientes valores:
 - 1 improbable
 - 2 Probable
 - 3 Seguramente
 - En type lo pondremos como undirected ya que la red no es dirigida pues se permite el flujo de ambos lados

Ejercicio: Construir Matriz de Adyacencia

La matriz de adyacencia es una matriz cuadrada que se utiliza como una forma de representar relaciones binarias.

Grafo no dirigido	Matriz de adyacencia
	$\begin{bmatrix} 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$
Grafo dirigido	Matriz de adyacencia
	$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$
Grafo ponderado	Matriz de adyacencia
	$\begin{bmatrix} 0 & 4 & 4 & 0 & 0 & 0 \\ 4 & 0 & 2 & 0 & 1 & 0 \\ 4 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 2 & 4 & 0 \end{bmatrix}$

Espejismo de la Mayoría



Espejismo de la Mayoría

El espejismo de la mayoría es un grafo donde hay catorce personas, que once piensan blanco y tres piensan rojo. Da igual lo que sea blanco y rojo. Te preguntan, ¿qué piensa la mayoría de esta sociedad de catorce? Y entonces te dan los puntitos y las rayitas que dicen quién es amigo de quién. Y te enseñan el grafo, y yo te lo enseño y digo: «¿Qué piensa la mayoría?». Y bueno, no hace falta ser doctor en Matemáticas, hay once puntos blancos y tres rojos. Y tú dices: «La mayoría piensa blanco». Y entonces, cuando ahora empiezas a preguntarle a los puntos blancos, ellos ven que la mayoría piensa rojo. O sea que desde fuera la mayoría piensa blanco, pero dentro de la burbuja, de tu burbuja social, es muy claro. O sea, porque como los rojos son los conectores, casi todos están conectados a uno o dos conectores y la mayoría, todos los que están dentro de la comunidad, perciben que es la mayoría de su entorno, de su pueblo, de su «timeline» piensan que es rojo.

Espejismo de la Mayoría

¿puede una **minoría** acabar imponiendo sus convicciones a una **mayoría**? Un estudio de la Universidad de Pennsylvania concluye que sí, siempre y cuando esa minoría alcanza una masa crítica de al menos el **25% de la población**. Es el único ingrediente que se necesita, aseguran, para conseguir provocar un **cambio social** a gran escala.



“Una vez alcanzas el número mágico del 25% se produce un cambio dramático en el comportamiento de la gente: de golpe y de forma muy rápida, la mayoría de la población comienza a adoptar ese nuevo comportamiento.”

Damon Centola

“Es el lado oscuro. Esos mismos puntos de inflexión en las dinámicas sociales se pueden usar por parte de gobiernos y organizaciones para intentar controlar a la población. Es el caso del gobierno de China, que crea falsas cuentas en las redes sociales del país, como Weibo, para influenciar las discusiones online”, alerta Centola.