

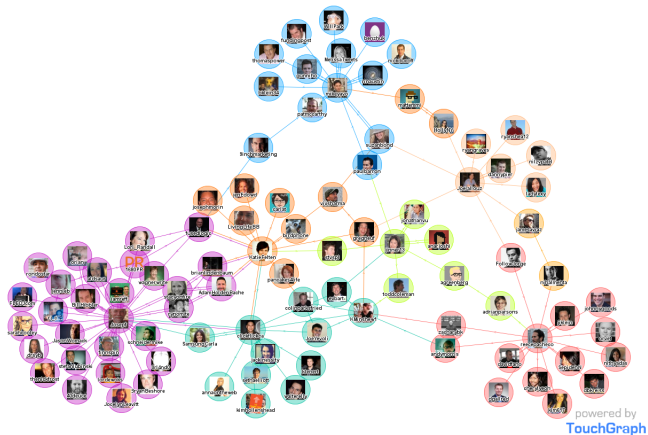
# UNIDAD 5: ANÁLISIS DE GRAFOS DE REDES SOCIALES

## REDES SOCIALES COMO GRAFOS

---

Blanca Vázquez y Gibran Fuentes-Pineda

Enero 2021



Las redes sociales contienen grandes cantidades de información.

¿Cómo identificar comunidades  
en las redes sociales?

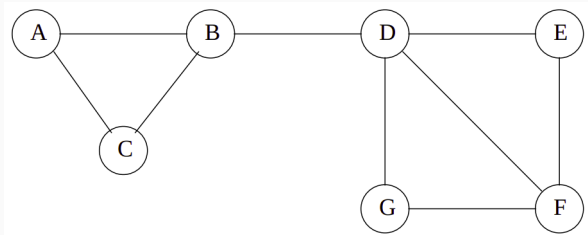
- Las comunidades las definimos como subconjuntos de nodos con conexiones entre los nodos.
- Los nodos pueden ser personas u otras entidades que forman la red.
- Las comunidades tienden a traslaparse.

- El algoritmo de *simrank* se usa para el descubrimiento de similitudes entre los nodos del grafo.
- El conteo de triángulos se usa medir las conexiones entre la comunidad.
- Diferentes algoritmos se usan para aproximar el tamaño de las vecindades en el grafo.

# ¿QUÉ ES UNA RED SOCIAL?

- Es una colección de entidades que participan en una red.
- Existe al menos una relación entre las entidades de la red.
- Se asume localidad (las relaciones tienden a formar grupos).
- Ejemplos: Facebook, twitter, instagram.

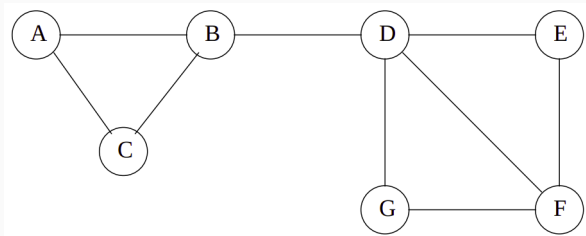
# REDES SOCIALES COMO GRAFOS



Las redes sociales usualmente son modeladas como grafos, debido a eso son llamadas: grafos sociales.

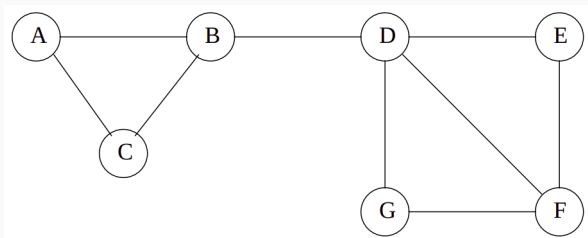
Imagen tomada de Aggarwal, 2016.

## REDES SOCIALES COMO GRAFOS

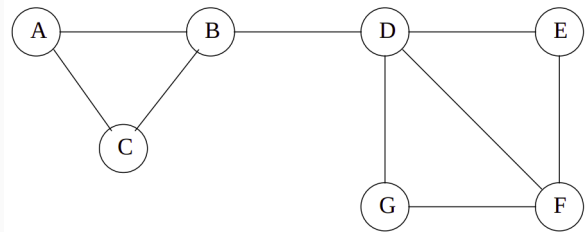


- Las entidades son los nodos A-G.
- La relación puede ser de **amigos**
- Por ejemplo: B tiene amistad con A,C y D.





**Localidad:** es una propiedad de las redes sociales que indica que nodos y aristas de un grafo tienden a agruparse en comunidades.



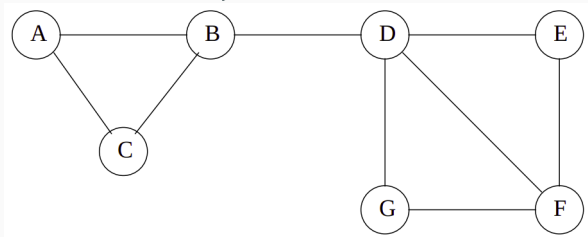
- 7 nodos, 9 aristas
- El número de combinaciones de los nodos y aristas son  $\binom{7}{2} = 21$  pares de nodos

Supongamos los nodos X, Y y Z, con aristas entre 'X y Y' y 'X y Z'.

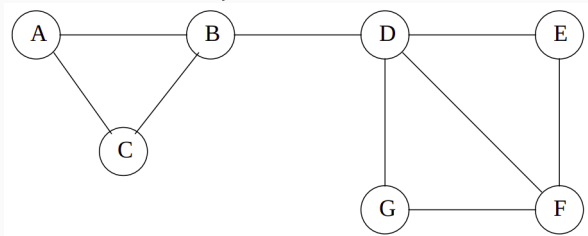
- ¿Cuál es la probabilidad que se espera de que exista una arista entre Y y Z?

La probabilidad de la arista (Y,Z) es:  $7/19 = .368$

Calcular la probabilidad de que exista la arista  $(Y,Z)$  dada las aristas  $(X,Y)$  y  $(X,Z)$ , donde  $X = A,C,E,G$

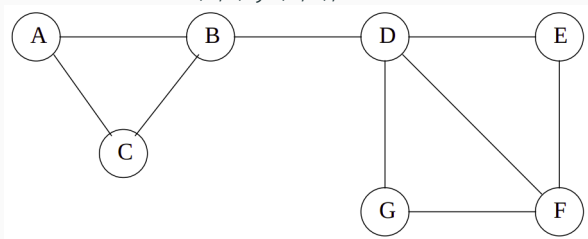


Calcular la probabilidad de que exista la arista  $(Y,Z)$  dada las aristas  $(X,Y)$  y  $(X,Z)$ , donde  $X = A,C,E,G$



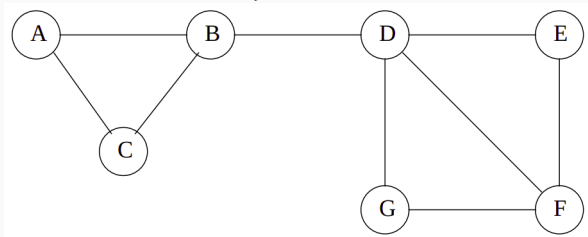
Existen 4 contribuciones positivas y 0 negativas.

Calcular la probabilidad de que exista la arista (Y,Z) dada las aristas (X,Y) y (X,Z), donde  $X = F$

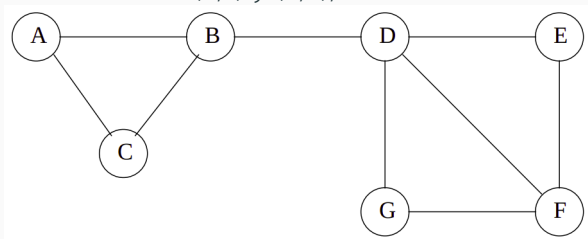


Existen 2 contribuciones positivas y 1 negativa.

Calcular la probabilidad de que exista la arista  $(Y,Z)$  dada las aristas  $(X,Y)$  y  $(X,Z)$ , donde  $X = B$



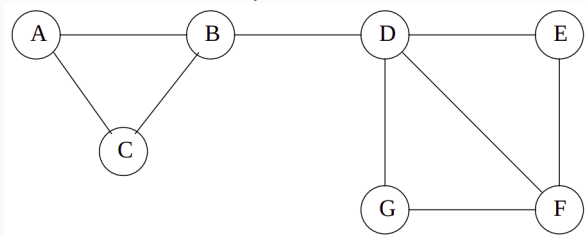
Calcular la probabilidad de que exista la arista  $(Y,Z)$  dada las aristas  $(X,Y)$  y  $(X,Z)$ , donde  $X = B$



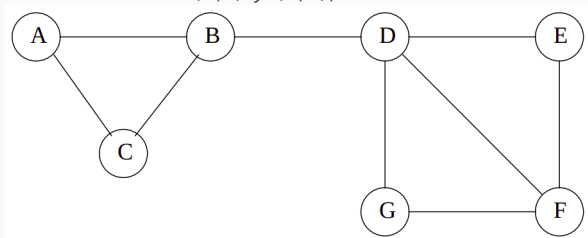
Existe 1 contribución positiva y 2 negativas.



Calcular la probabilidad de que exista la arista  $(Y,Z)$  dada las aristas  $(X,Y)$  y  $(X,Z)$ , donde  $X = D$

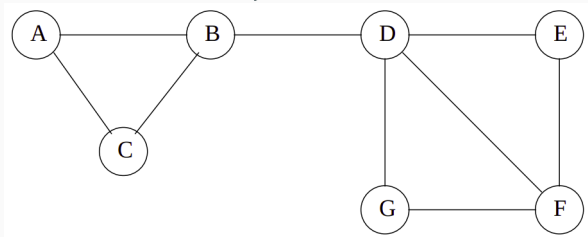


Calcular la probabilidad de que exista la arista  $(Y,Z)$  dada las aristas  $(X,Y)$  y  $(X,Z)$ , donde  $X = D$



Existe 2 contribuciones positivas y 4 negativas.

Calcular la probabilidad de que exista la arista  $(Y,Z)$  dada las aristas  $(X,Y)$  y  $(X,Z)$ , donde  $X = D$



En total, el grafo tiene 9 contribuciones positivas y 7 negativas.  
La probabilidad de que exista la arista  $(Y,Z) = 7/15 = 0.46$ .

Además de las redes de amistad, existen diferentes tipos de redes:

- Redes telefónicas
- Redes de correo electrónico
- Redes colaborativas

# REDES TELEFÓNICAS

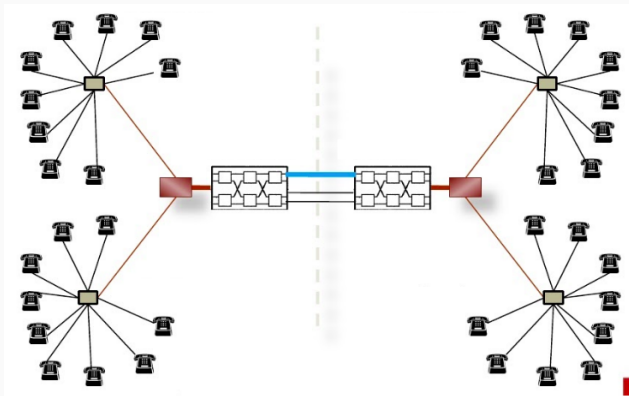


Imagen tomada de Villegas, 2008.

Los nodos son representados por los números telefónicos y las aristas son las llamadas entre los nodos. Las comunidades pueden ser: miembros de un club, compañía, amigos.

# REDES DE CORREO ELECTRÓNICO

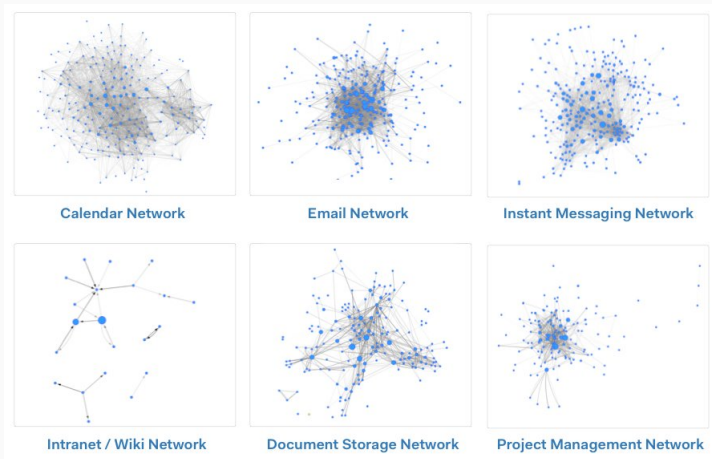


Imagen tomada de Worklytics , 2017.

Los nodos son representados por las direcciones electrónicas y las aristas son al menos un correo entre direcciones.

- Grupos de investigación
  - Publicación de artículos
  - Líneas de investigación
  - Wikipedia (artículos, autores, editores)

- Biológicas
- Infraestructura
- Marketing