

Análisis de Redes Sociales

Guillermo Jiménez Díaz (gjimenez@ucm.es)
Alberto Díaz (albertodiaz@fdi.ucm.es)

2 de octubre de 2014

Prefacio

Estos son los apuntes de la asignatura Análisis de Redes Sociales, impartida en la Facultad de Informática de la Universidad Complutense de Madrid por los profesores Guillermo Jiménez Díaz y Alberto Díaz, del Departamento de Ingeniería del Software e Inteligencia Artificial.

Este material ha sido desarrollado a partir de distintas fuentes, destacando como referencia principal el libro *Network Science* de Laszlo Barabasi y el material de la asignatura *Social Network Analysis* impartido por Lada Adamic a través de Coursera.

Tema 2: Modelos de redes: El modelo de red aleatoria

La visualización de una red permite observar y calcular algunas de las propiedades vistas hasta ahora. Sin embargo, muchas de las redes reales son complejas de visualizar.

Un **modelo** es una representación simple de un sistema complejo del que podemos extraer y derivar propiedades matemáticamente. Por tanto, mediante un modelo que se aproxime a una red real podremos predecir propiedades de la misma. Además, los modelos nos ayudan a entender cómo y por qué se han formado las redes tal y como son ahora.

En este y en sucesivos capítulos estudiaremos algunos de los modelos más conocidos de redes sociales. El primero de ellos es el **modelo de red aleatoria**.

Modelo de Red Aleatoria: algoritmo de construcción

Una red aleatoria es una red en la que cada uno de los enlaces entre dos nodos se ha creado siguiendo un proceso completamente aleatorio. De manera más matemática podemos decir que una red aleatoria es una red que tiene N nodos donde cada nodo puede estar conectado con otro con una probabilidad p . Se representa como $G(N, p)$ y se construye siguiendo el siguiente algoritmo:

1. Crear N nodos aislados
2. Seleccionar un par de nodos y generar un número aleatorio entre 0 y 1. Si es mayor que p entonces añadimos un enlace entre ellos. En otro caso, los dejamos desconectados.
3. Repetir el paso 2 para los $\frac{N(N-1)}{2}$ pares de nodos de la red.

El modelo de red aleatoria también se conoce como el modelo de Erdős-Renyi en honor de los dos matemáticos húngaros que las estudiaron inicialmente y que proporcionaron gran cantidad de información sobre las propiedades contenidas en este tipo de redes.

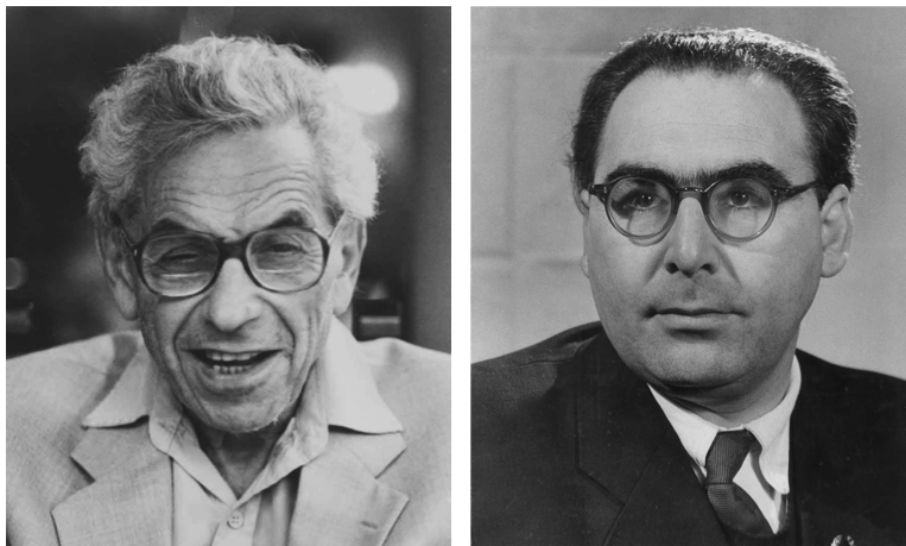


Figure 1: Pal Erdős y Alfred Renyi, matemáticos húngaros principales investigadores del modelo de red aleatoria.

El modelo de red aleatoria también puede caracterizarse como $G(N, L)$, donde L es el número de enlaces de la red. En este caso, la red se forma seleccionando aleatoriamente 2 nodos y, si no existe un enlace entre ellos, se añade. Este proceso se repite hasta conseguir un total de L enlaces. Esta forma de caracterizar la red se usa menos.

Número de enlaces

Las redes aleatorias $G(N, p)$ tienen un número variable de enlaces (L). Sin embargo, usando probabilidades básicas podemos calcular la probabilidad de que la red tenga exactamente L enlaces.

Para aproximar este cálculo se utiliza la *distribución binomial* que describe el número de éxitos (x) que se pueden conseguir en la realización de n experimentos independientes donde la probabilidad de acierto es p y la de fracaso, $1 - p$.

$$B(n, x, p) \rightarrow p_x = \binom{n}{x} p^x (1 - p)^{n-x}$$

De esta función de distribución nos interesan las siguientes métricas:

- Media de la distribución: $\langle x \rangle = \sum_{x=1}^N x \cdot p_x = n \cdot p$
- Desviación estándar de la distribución: $\sigma_x = [p \cdot (1 - p) \cdot n]^{\frac{1}{2}}$

De acuerdo a esto podemos representar la probabilidad de una red de tener exactamente L enlaces como:

$$p_L = B(L_{max}, L, p) \rightarrow p_L = \binom{L_{max}}{L} p^L (1 - p)^{L_{max}-L}$$

De acuerdo a esto podemos calcular el número medio de enlaces esperados en $G(N, p)$ como:

$$\langle L \rangle = p \cdot L_{max} = p \cdot \binom{N}{2} = p \cdot \frac{N(N-1)}{2}$$

Por tanto, el grado medio de esta red¹ es:

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p \cdot (N - 1)$$

¹Si la red se representa como $G(L, p)$ entonces ya sabemos el número exacto de enlaces por lo que $\langle k \rangle = \frac{2L}{N}$.