



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT
DE VALÈNCIA

REDES ALEATORIAS, DE PEQUEÑO MUNDO Y LIBRES DE ESCALA

TRABAJO FINAL DE MASTER
Máster en Investigación Matemática

Departamento de Matemática Aplicada
Universidad Politécnica de Valencia.

Presentado por:
Pedro A. Solares Hernández

Tutores:
J. Alberto Conejero Casares
Cristina Jordán Lluch

Curso 2016/2017

Índice general

1. Introducción	5
2. Notación y preliminares	11
3. Los Miserables	17
4. Redes aleatorias	27
El número de aristas en una red aleatoria	33
Distribución de Grado	35
5. Redes de pequeño mundo	37
Experimento de Stanley Milgram	37
Modelo de Watts & Strogatz	39
6. Redes libres de escala	43
Ajuste del exponente de la ley de potencias	46
Ajuste de la constante de la proporcionalidad de la ley de potencias	47
Tamaño de los hubs	49
Etimología del concepto <i>Libre de Escala</i>	50
7. Red de divisibilidad	53
8. Conclusión	65

Capítulo 1

Introducción

La *Teoría de Grafos* se considera que comenzó en el siglo XVIII con la resolución por *Euler* del conocido como problema de los puentes de Königsberg, habiendo sido desde entonces ampliamente desarrollada y siguiendo líneas muy distintas de estudio: desarrollo algorítmico, teoría algebraica de grafos, optimización... Es de destacar la *Teoría de Redes*, que surge a partir de la Teoría de Grafos, pero enfocándose al estudio de la estructura de las conexiones del grafo. Con la explosión de la cantidad de datos susceptibles de análisis, la Teoría de Redes ha cobrado una especial relevancia en la actualidad. En palabras de Watts, *las interconexiones que nos unen unos a otros, desde los enlaces en Internet hasta los principales centros intermodales de transporte y el mundo de las finanzas, pasando por las redes de relaciones sexuales por donde se transmite la pandemia del SIDA en todo el mundo, ofrecen un marco de referencia nuevo y dinámico para comprender nuestra sociedad global y sus cambios.*

Podemos identificar grafos con redes. El hecho de hablar de grafos o de redes viene dado por el tipo de problemas que se estén considerando, siendo la Teoría de Grafos mucho más teórica y formal y siendo la Teoría de Redes mucho más experimental. Debemos distinguir entre este tipo de redes y las redes asociadas a la obtención de un flujo máximo, las cuales constituyen otra de las partes de gran aplicabilidad de la Teoría de Grafos.

De esta teoría es un gran exponente *Barabási*, que junto con *Albert* des-

cubrieron las **redes libres de escala** a finales de los 90. La diferencia entre la *Teoría de Grafos* y la *Teoría de Redes* se explica en las siguientes Fig. 1.1 y Fig. 1.2:

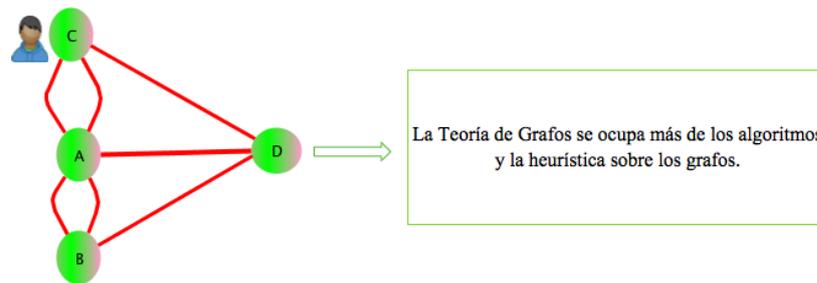


Figura 1.1: Grafo del problema de los puentes de Königsberg resuelto por Euler. Elaboración propia.

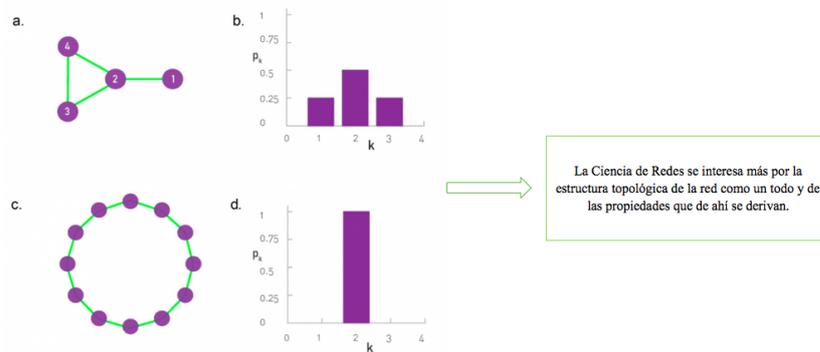


Figura 1.2: Fuente: [2]

Las redes, hablando en términos generales dentro de la Teoría de Grafos, modelizan contextos en los que se pueden distinguir dos tipos de elementos representados, respectivamente, por vértices y aristas. Podemos citar, como ejemplo de donde se aplican actualmente: la web, redes sociales, redes neuronales, redes eléctricas, ciencia computacional, interacciones entre genes. En la siguiente Fig. 1.3, se presenta una tendencia o evolución de las publicaciones que contienen temas relacionados a Networks

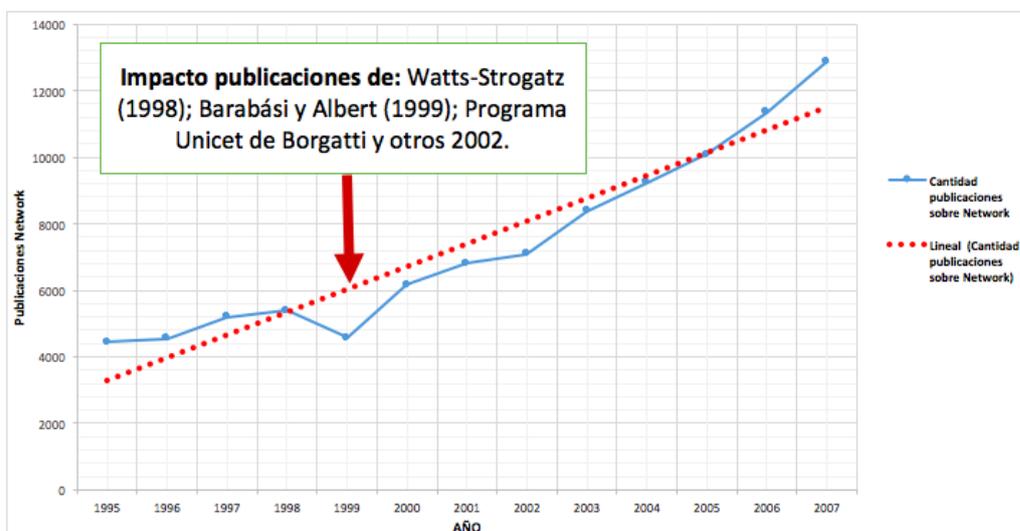


Figura 1.3: Evolución publicaciones de temas relacionados con *Networks*. Base de datos: PubMed-Social Networks. Elaboración propia.

Antes de la eclosión de la Teoría de Redes en las últimas dos décadas, la importancia del estudio sistemático de las relaciones entre individuos en el ámbito de la sociología ya fue vislumbrada por el sociólogo Ithiel de Sola Pool [3]. Este artículo, aunque publicado en los años 70, circuló durante casi 20 años en el mundo académico siendo un referente en cuanto a que planteaba muchas cuestiones sociológicas que sería interesante analizar y ver sus causas.

Un punto de inflexión fue la publicación del famoso artículo de Zachary “*An Information Flow Model for Conflict and Fission in Small Groups*”. Este antropólogo indagó sobre la manera de actuación de los diferentes miembros de un club universitario de karate para observar y estudiar las relaciones sociales entre los diferentes miembros. En dicho club había un desacuerdo entre el profesor de karate y el administrador sobre el precio de las cuotas: El profesor quería que se cobrara más para que le subieran el salario y el administrador no quería. Zachary se cuestionó, a medida que iba desarrollando su experimento, si era posible predecir cómo se iba a dividir el grupo. Para ello, modeló a los miembros del club como nodos. Las aristas conectan miembros que eran afines. Cada arista tenía un peso que representaba el grado de afinidad entre cada pareja de personas obtenido a partir de unas preguntas

sobre sus hábitos en común. Finalmente, al grafo ponderado resultante le aplicó el algoritmo de Ford Fulkerson para encontrar el flujo máximo en la red, tomando como fuente y sumidero, alternativamente, tanto al profesor como al administrador.

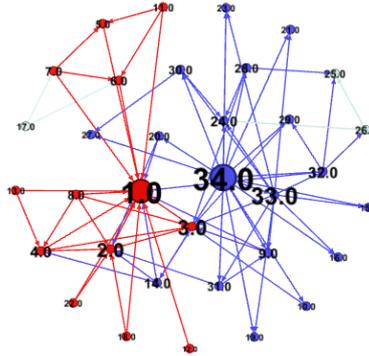


Figura 1.4: Red Club de Karate de Zachary. Elaboración propia en Gephi.

En la Fig. 1.4, se puede apreciar que los nodos 1 y 34 son los que tienen mayor conexiones y por tanto su grado es mayor a cualquier otro de la red, lo que implica que hay muchas relaciones entre los miembros del club en relación (excluyente) al director y el profesor.

La cortadura mínima del grafo permitió separar a los miembros en dos grupos, que fueron los que a la postre se formaron cuando el profesor decidió montar un nuevo club, con excepción de un alumno que aunque era más afín al círculo del administrador, estaba a punto de obtener el cinturón negro y por eso se fue con el profesor de karate. Desde entonces, los métodos de la Teoría de Grafos y la Teoría de Redes han permitido dar luz sobre la estructura de los grupos sociales. En primera instancia, se consideró que las redes sociales se podían modelar como redes aleatorias. Posteriormente, se comprobó que este modelo no permitía explicar fenómenos como la existencia de redes de mundo pequeño. Finalmente, el modelo de redes de libre de escala permite dar una descripción más adecuada de la estructura de redes sociales. En relación a la estructura del trabajo comentamos que distinguiremos tres partes:

1. En el Capítulo 2 se presentan los preliminares de Teoría de Grafos necesarios para poder estudiar las redes.
2. En el Capítulo 3 se utilizan dichos preliminares para mostrar el ejemplo de análisis de una red sencilla. En este caso se ha utilizado un ejemplo clásico de red propuesto por Knuth: la red de relaciones entre los personajes de la novela *Los Miserables* de Victor Hugo.
3. En el Capítulo 4 se estudian las **redes aleatorias**. Este modelo fue propuesto por *Erdős y Renyi*, modelo conocido como: *ER* en el que $G \sim G(N, p)$, donde G es un grafo aleatorio, N es el número de nodos de la red y p es la probabilidad de que dos nodos de G estén conectados. Se hace una demostración del por qué estos tipos de redes son aleatorios dentro de un entorno definido por un parámetro $P = \frac{1 \pm \epsilon}{n}$, para cualquier $\epsilon > 0$ suficientemente pequeño. Además de este modelo, también existe el modelo $G(N, L)$ en el que la red aleatoria solo depende del número de nodos y del cardinal de los enlaces, pero no se llega a definir ni a demostrar, dado que el modelo ER es el más usado para estos tipos de redes.
4. En el Capítulo 5 se estudian las **redes de pequeño mundo**, en la que se describe la teoría de *Los Seis Grados de Separación*, propuesta por *Milgram*. Esta teoría ha sido impulsada en los últimos años por *Watts y Strogatz*.
5. Posteriormente, en el Capítulo 6, se desarrolla la teoría de las **redes libres de escala** propuesta por *Barabási y Albert*. Esta teoría explica que la distribución de los grados de la red que obedece a una ley de potencia, debido a la presencia de ciertos tipos de nodos llamados **hubs** que se caracterizan por estar conectados a un gran número de otros nodos, es decir, que el grado de estos nodos es muy grande en relación a los demás nodos de toda la red. Este tipo de redes son libres de escalas, porque obedecen a una propiedad de la física-estadística denominada fase de transición [6], esto es, porque para que sea libre de escala ciertos comportamientos en algunos sistemas deben garantizar

dicha propiedad, debido al cambio drástico que pueden experimentar los componentes que conforman la red. El concepto de hub, se deduce también de dicha propiedad.

Las redes libres de escalas tienen características de fractalidad, por lo que se pueden emplear algunas nociones de la *Teoría del Caos* para poder estudiar algunos de sus comportamientos.

6. Finalmente, en el Capítulo 7, se estudia un ejemplo de red libre de escala que aparece al estudiar las relaciones de divisibilidad entre los números naturales [7].

Capítulo 2

Notación y preliminares

A continuación resumimos brevemente los conceptos de la teoría de grafos que nos serán de utilidad en el desarrollo del presente trabajo.

Definición 1 (Grafo). *Un grafo G se define como un par de conjuntos (V, E) donde:*

- V , llamado conjunto de vértices o nodos, es distinto de vacío .
- E es un conjunto de pares de elementos de V . En el caso de que los pares sean no ordenado a los elementos de E se les llama aristas, y en caso contrario arcos.

Un grafo se dice no dirigido en el caso de que los pares sean no ordenados y dirigido o digrafo en caso contrario.

Se ha acuñado el término teoría de redes para designar grafos que poseen un elevado número de vértices y aristas. Dado que las redes se representan de igual manera que los grafos, nos referiremos a ellas también mediante la notación $G = (V, E)$ y será de aplicación las nociones que introduciremos a lo largo de este capítulo.

Definición 2 (Subgrafos). *Dado el grafo $G = (V_G, E_G)$, se dice que el grafo $H = (V_H, E_H)$ es subgrafo de G si $V_H \subseteq V$ y $E_H \subseteq E$.*

Dado el grafo $G = (V_G, E_G)$ y el subconjunto de vértices V' de V , llamamos subgrafo inducido por V' al subgrafo $G[V']$ de G cuyo conjunto de vértices es V' y es maximal respecto al conjunto de aristas.

Dado un grafo dirigido G , llamamos grafo subyacente de G , al grafo no dirigido obtenido sustituyendo cada arco (u, v) por una arista (u, v) . En el caso de existir los arcos (u, v) y (v, u) , ambos se sustituyen por una única arista (u, v) .

El grafo subyacente del grafo dirigido G se denota \bar{G}

Los grafos admiten representación matricial, lo que facilita la programación de algoritmos para resolver problemas sobre ellos.

Definición 3 (Matriz de Adyacencia). Dado $G = (V, E)$, con $|V| = n$, llamamos matriz de adyacencia de G a la matriz $A = (a_{ij})$, de tamaño $n \times n$, donde:

$$a_{ij} = \begin{cases} 1 & \text{si } (v_i, v_j) \in E \\ 0 & \text{si } (v_i, v_j) \notin E \end{cases}$$

Como se deduce trivialmente de la definición, las matrices de adyacencia de grafos no dirigidos son simétricas.

El grado de un vértice se encuentra entre los conceptos básicos de la teoría de grafos pero que revisten gran importancia al trabajar en el campo de las redes.

Definición 4 (Grado de vértice en grafos no dirigidos). Dado $G = (V, E)$ un grafo no dirigido y v un vértice de G , se llama grado de v al número de aristas que inciden en él. Si la arista es un bucle en v , contribuye con dos unidades al valor del grado. Si un vértice tiene grado cero, entonces es un vértice aislado.

Podemos, por otra parte, obtener el valor del grado de un vértice, v_k , de un grafo no dirigido G a partir de su matriz de adyacencia, $A = (a_{ij})$, sin más que contar el número de unos en la fila o columna correspondiente de A .

$$\forall v_k \in V d(v_k) = \sum_{j=1}^n a_{kj} = \sum_{i=1}^n a_{ik}$$

Una propiedad fundamental que relaciona grados y número de aristas en grafos no dirigidos es la conocida como *Lema de darse la mano*.

Proposición 2.1 (Lema de darse la mano). *Sea $G = (V, E)$ un grafo no dirigido. Se verifica:*

- *La suma de los grados de los vértices de G coincide con el doble del número de sus aristas. Así:*

$$\sum_{v \in V} d(v) = 2|E|.$$

- *El número de vértices de grado impar de G es par.*

En el caso de grafos dirigidos las definiciones difieren a consecuencia de trabajar con pares ordenados. Hablamos en este caso de grado de entrada y grado de salida, entendiendo por grado la suma de ambos.

Definición 5 (Grado de vértice en grafos dirigidos). *Dado $G = (V, E)$ un grafo dirigido y v un vértice de G , se llama*

- *grado de entrada de v al número de arcos cuyo extremo final es v . Se denota $d_i(v)$.*
- *grado de salida de v al número de arcos cuyo extremo inicial es v . Se denota $d_o(v)$.*
- *grado de v a la suma del grado de entrada y grado de salida de v .*

Análogamente al caso no dirigido, dado un grafo dirigido G con matriz de adyacencia $A = (a_{ij})$ el grado de entrada y salida de un vértice v_k viene dado por las expresiones:

$$d_i(v_k) = \sum_{i=1}^n a_{ik}, \quad d_o(v_k) = \sum_{j=1}^n a_{kj}.$$

y encontramos una relación similar a la que proporciona el lema de darse la mano entre el número de arcos y la suma de grados de entrada y salida.

Proposición 2.2. *Sea $G = (V, E)$ un grafo no dirigido. Se verifica que: La suma de los grados de entrada de los vértices de G coincide con la suma de los grados de salida y a su vez con el número de sus arcos. Así:*

$$\sum_{v \in V} d_i(v) = \sum_{v \in V} d_o(v) = |E|.$$

Los conceptos de caminos y cadenas son de uso constante cuando se utiliza la teoría de grafos.

Definición 6. *Sea $G = (V, E)$. Una cadena es toda sucesión finita alterna de vértices y aristas (resp. arcos). Una cadena cerrada es toda cadena en la que los vértices inicial y final coinciden. Un camino es toda cadena en la que no se repiten ni vértices ni aristas (resp. arcos). Un ciclo es toda cadena cerrada en la que no se repite ninguna arista (resp. arco) ni vértice, a excepción del inicial y final. La longitud de una cadena es el número de aristas (resp. arcos) que la forman.*

Otra de las matrices asociadas a un grafo que utilizaremos en este trabajo es la matriz de acceso.

Definición 7 (Accesibilidad). *Dado un grafo $G = (V, E)$ se dice que el vértice v_i alcanza al vértice v_j en G si existe una cadena de longitud mayor o igual que cero de v_i a v_j . Por cadena de longitud cero entre v_i y v_j interpretamos que $i = j$, es decir, que todo vértice se alcanza a sí mismo.*

Se define la matriz de accesibilidad de G como la matriz $A = (a_{ij})$, de tamaño $n \times n$, donde:

$$a_{ij} = \begin{cases} 1, & \text{si } v_i \text{ alcanza a } v_j \\ 0, & \text{si } v_i \text{ no alcanza a } v_j \end{cases}$$

Los algoritmos de Búsqueda del Primero en Profundidad (Depth First Search -DFS) y de Búsqueda del Primero en Amplitud (Breadth First Search -BFS) permiten determinar los vértices alcanzados por un vértice dado en tiempo polinómico.

Relacionado con la accesibilidad encontramos el concepto de conexión, fundamental en la Teoría de Grafos.

Definición 8. *Sea un grafo no dirigido $G = (V, E)$. Se dice que un grafo G es conexo si todos sus vértices se alcanzan mutuamente. Llamamos componente conexa de G a todo subgrafo conexo H de G tal que no exista ningún otro subgrafo de G que lo contenga estrictamente.*

El asignar valores a los arcos o aristas permite modelar un mayor número de situaciones. Esta asignación se recoge en la llamada matriz de pesos del grafo.

Definición 9. *Se dice que un grafo $G = (V, E)$ es ponderado, si cada arista (resp. arco) (v_i, v_j) tiene un valor asociado, $p(v_i, v_j)$, al que se llama peso o coste .*

Dado un grafo ponderado G , $|V| = n$, llamamos matriz de pesos o matriz de costes de G a la matriz $P = (p_{i,j})$, de tamaño $n \times n$, cuyos elementos vienen definidos como sigue:

$$p_{i,j} = \begin{cases} 0, & \text{si } i = j \\ p(v_i, v_j), & \text{si } i \neq j \ \& \ (v_i, v_j) \in E \\ \infty, & \text{si } i \neq j \ \& \ (v_i, v_j) \notin E \end{cases}$$

Dado un camino $C = v_{i_1}, v_{i_2}, \dots, v_{i_k}$ en G , definimos el peso o coste de C como la suma de los pesos de las aristas (resp. arcos) que lo forman, es decir,

$$\sum_{i=1}^{k-1} p_{i,i+1}.$$

Uno de los problemas centrales de la Teoría de Grafos es el conocido como el *problema del camino más corto*, que persigue determinar, dado un grafo G ponderado y dos vértices u y v de G , el camino con menor peso de u a v , conocido habitualmente como camino más corto, o distancia, de u a v .

En grafos no dirigidos, a consecuencia de su simetría, el camino más corto de u a v coincide con el más corto de v a u , no ocurriendo así en grafos

dirigidos. En estos, no solo el peso de dichos caminos más cortos puede ser distinto sino que puede que incluso uno de ellos, o los dos, no existan. Hay que recordar que el camino más corto no hace en ningún momento referencia a la longitud del camino.

Entre los conceptos que constituyen uno de los pilares de la Teoría de Redes encontramos las medidas de centralidad que determinan la importancia relativa de un vértice dentro del grafo. Los conceptos básicos de estas medidas son los siguientes:

Definición 10. Sea $G = (V, E)$ un grafo conexo. Llamamos *excentricidad* $e(v)$ de un vértice v a la distancia de v al vértice más alejado de él, es decir, $e(v) = \max\{d(v, u) / u \in V(G)\}$

Se define el *radio* de G como la menor de sus excentricidades, es decir, $\text{rad}(G) = \min\{e(v) : v \in V(G)\}$, y como *diámetro* de G a la mayor de sus excentricidades, es decir, $\text{diam}(G) = \max\{e(v) : v \in V(G)\}$. Se define el *centro* de G , $c(G)$ como el subgrafo inducido por los vértices cuya excentricidad sea la mínima del grafo, es decir, $\text{rad}(G)$.

Definición 11. Sea un grafo ponderado $G = (V, E)$. Se define la *distancia* de un vértice v , $d(v)$, como la suma de sus distancias a todos los vértices del grafo. En el caso de que un grafo no fuera ponderado, se puede considerar que el peso de cada una de las aristas es 1.

Se define la *mediana* de G , $m(G)$, como el subgrafo inducido por los vértices que tienen mínima distancia.

Capítulo 3

Un ejemplo de análisis de red:

Los Miserables

Los conceptos de centralidad antes comentados, introducidos inicialmente por Bavelas a finales de los años 1940, constituyen uno de los conceptos más estudiados en el análisis de redes. Poder reconocer la centralidad de un nodo puede ayudar a determinar, por ejemplo, el impacto de una persona involucrada en una red social, la relevancia de una habitación en un edificio representado en sintaxis del espacio, la importancia de una carretera en una red urbana, o los componentes esenciales de una red de computadoras.

En este apartado vamos a considerar la base de datos¹ de la Novela *Los Miserables* de *Victor Hugo*, para explicar las interrelaciones entre los diferentes personajes, véase [4].

La red de *Los Miserables* está compuesta por 77 nodos, que representan a los personajes de la novela, y por 254 aristas que representan las conexiones entre estos. Concretamente, la existencia de la arista (v_i, v_j) indica que los personajes v_i y v_j se conocen.

Como se puede comprobar, la red es conexas, puesto que existen caminos entre dos personajes cualesquiera. En las Figuras 3.1 y 3.2 se muestra la mencionada red de amistades. Es de destacar el alto valor del grado del nodo

¹Para consultar la base de datos: <https://www.cise.ufl.edu/research/OptimaNetSci/tools/small-scale.html>

principal, *Jean Valjean*, el protagonista de la novela.

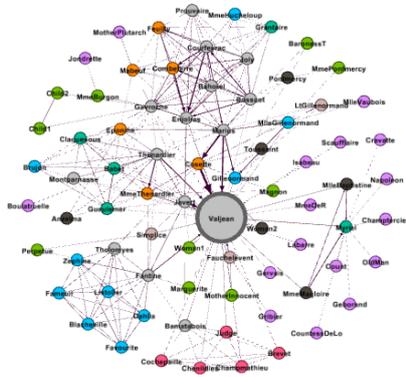


Figura 3.1: Red de los personajes de *Los Miserables*. Elaboración con Gephi

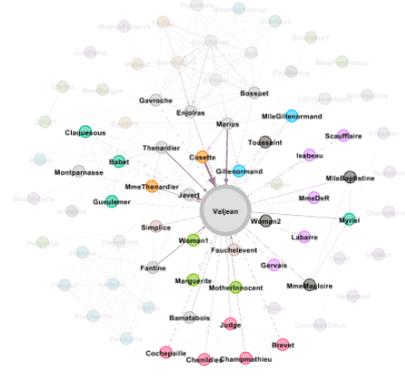


Figura 3.2: Jean Valjean y los personajes con los que coincide en la novela. Elaboración con Gephi

Introducimos a continuación un nuevo concepto relacionado con los grados de los nodos, el *grado medio* de un grafo, que indica cuál es la media de conexiones que tiene un nodo, de manera que se puede saber su popularidad.

Definición 12 (Grado medio). Dado un grafo no dirigido $G = (V, E)$, con $|V| = n$, llamamos **grado medio** de G al valor

$$\langle G \rangle = \frac{1}{n} \sum_{i=1}^n d_i = \frac{2|E|}{n} \quad (3.1)$$

donde d_i , $i = 1, 2, \dots, n$ representa los grados de los n vértices de G .

En nuestro caso, $\langle G \rangle = \frac{2 \cdot 254}{77} = 6,5974$.

En la Figura 3.3 podemos ver la distribución de grado correspondientes a la red. Por ejemplo, en el eje de x se representan los posibles valores de los grados, mientras que en el eje de la y , la frecuencia de los vértices.

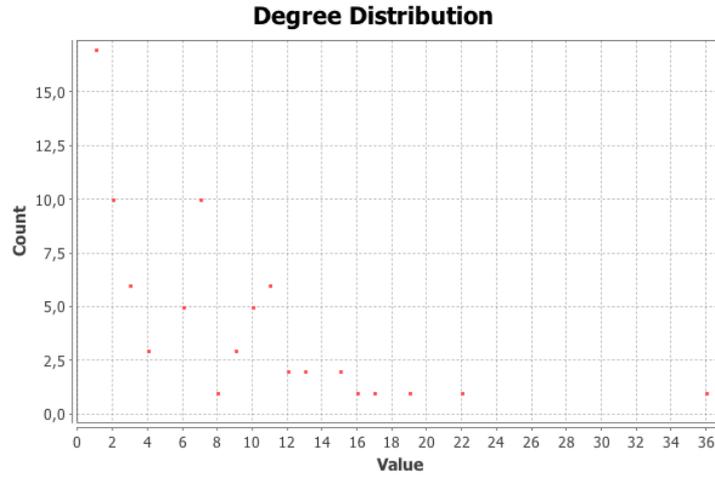


Figura 3.3: Distribución de frecuencia de los grados de los vértices en la red de *Los Miserables*. Elaboración propia con Gephi.

Por otro lado es interesante también estudiar el histograma de frecuencias relativas de los valores de los grados en una grafo, tomando como 1 el doble de la suma de todos los grados del grafo.

También, se puede estudiar la distribución relativa de los grados de las redes, a los fines de ver cuál es su representación o importancia en dicha red.

Definición 13. Sea $G = (V, E)$ una red. La distribución de la frecuencia relativa de los grados se considera a partir de los valores $\{p_k\}_{k=1}^{\infty}$, donde p_k es la probabilidad de que un vértice seleccionado aleatoriamente en la red tenga grado k en G .

Los valores p_k se pueden estimar como $p_k = \frac{n_k}{n}$, siendo $n = |V|$. La suma de estas probabilidades cumple:

$$\sum_{k=1}^{\infty} p_k = 1 \quad (3.2)$$

Ello implica que la suma de las probabilidades es acumulativa, de modo que podemos representar su distribución, como se indica en la Figura 3.4:

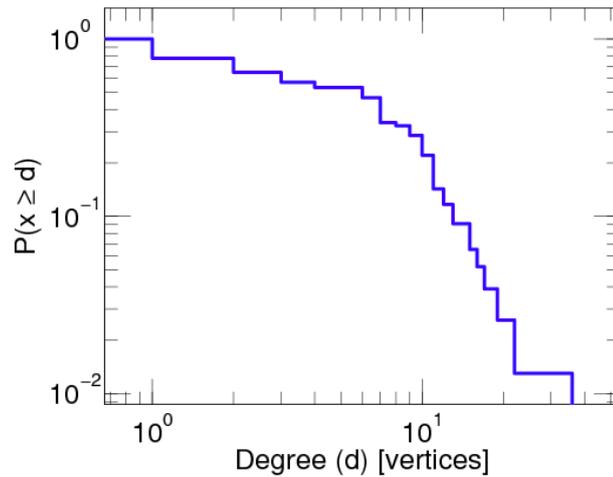


Figura 3.4: Distribución acumulada de los grados de los vértices de la red de *Los Miserables*. Fuente: <http://konect.uni-koblenz.de/img/bidd.ax.moreno>

Utilizando la expresión (3,3) se puede expresar una estimación la cantidad de nodos de grado k en términos de la probabilidad p_k . La determinación de la probabilidad p_k es de suma importancia en el estudio de las distintas propiedades de los diferentes componentes de la red.

Se puede estimar la esperanza de la variable aleatoria *grado de un vértice* a partir de las probabilidades p_k . Denotaremos a este valor por $\langle k \rangle$. Esta esperanza se puede calcular a partir de la siguiente expresión:

$$\langle k \rangle = \sum_{k=0}^{\infty} kp_k \quad (3.3)$$

En esta línea, *La Ley de Metcalfe* sobre el valor de una red afirma que el valor de la misma es proporcional al cuadrado del número de sus nodos, es decir, n^2 .

Aunque en una red pequeña es fácil determinar si un nodo es accesible desde otro (y la distancia entre esa pareja de nodos), no lo es así para redes con un número grande de vértices. Para hablar de distancias entre nodos recurriremos a la matriz de adyacencia A_{ij} de la red.

Consideremos una red no ponderada $G = (V, E)$ con n vértices enumerados como v_1, \dots, v_n . Denotaremos por d_{ij} a la distancia entre los vértices v_i

y v_j con $v_i, v_j \in V$.

- $d_{ij} = 1$ siempre que haya un enlace entre los vértices v_i y v_j , en cuyo caso $A_{ij} = 1$. Para el resto de casos se tiene $A_{ij} = 0$.
- $d_{ij} = 2$. De existir un camino de longitud 2 entre i y j , entonces $A_{ik}A_{kj} = 1$ ($A_{ik}A_{kj} = 1$ será igual a 0, en cualquier otro caso). El número de caminos de longitud 2 sobre la red puede estimarse como $n_{ij}^{(2)} = \sum_{k=1}^n A_{ik}A_{jk} = A_{ij}^2$.
- $d_{ij} = D$. Si hay un trayecto cuya longitud es D , entre i y j , entonces $A_{i_1} \dots A_{i_2} \dots A_{i_D} = 1$ ($A_{i_1} \dots A_{i_D} = 0$ en cualquier otro caso). Se puede probar inductivamente que el número de trayectos de longitud D entre v_i y v_j se puede estimar a partir del elemento ij de la potencia D -ésima de A . $n_{ij}^{(D)} = A_{ij}^D$

En grafos ponderados, las distancias se pueden obtener por un método computacional más eficiente mediante una aplicación reiterada del algoritmo de Dijkstra, o bien de otros algoritmos de caminos mínimos como Bellman-Ford o Floyd-Warshall.

Volviendo al ejemplo de *Los Miserables*, a partir del cálculo de distancias en la red se puede obtener su diámetro, que en este caso es $d_{max} = 5$. En la Figura 3.5 se representan las excentricidades de cada uno de los vértices mediante tres colores:

1. Verde: que representa el 12,99 % de los nodos, tiene una excentricidad de 3.
2. Azul: que representa el 61,04 % de los nodos, tiene una excentricidad de 4.
3. Rojo: que representa el 25,97 % de los nodos, tiene una excentricidad de 5.

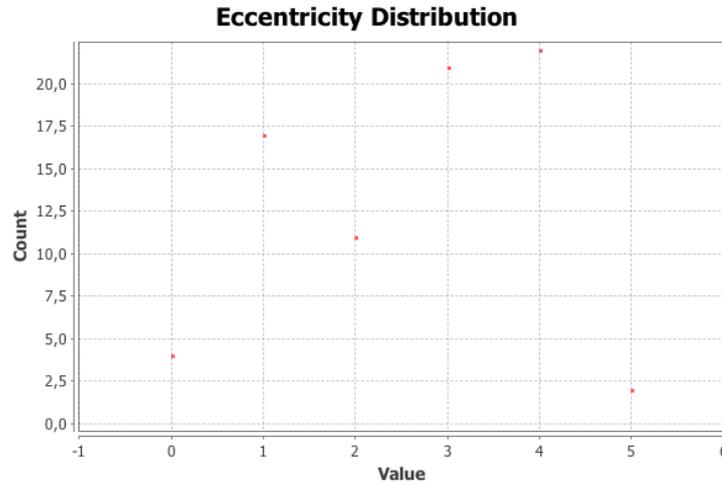


Figura 3.5: Excentricidades de los vértices de la red de *Los Miserables*. Elaboración propia en Gephi.

A partir del cálculo de los caminos más cortos entre dos vértices, se puede determinar la *intermediación* de un vértice, como la frecuencia con la que aparece ese vértice en el camino más corto entre otros dos vértices cualesquiera, o la *cercanía* del mismo, como la distancia media de ese vértice a cualquiera de la red.

En general, se puede estimar la **longitud promedio entre dos vértices** cualesquiera de la red como el promedio de la distancia entre todos los pares de vértices que pertenecen a la red. En el caso de que una red sea dirigida, $\langle d \rangle$ es expresado de la siguiente manera:

$$\langle d \rangle = \frac{1}{n(n-1)} \sum_{i,j=1;i \neq j}^n d_{i,j} \quad (3.4)$$

y si la red no es dirigida, basta tomar $1 \leq j < i \leq n$ en este sumatorio.

Para el caso de Los Miserables, la distancia media es: $\langle d \rangle = 2,4$.

La *densidad* ρ de una red $G(V, E)$ se estima mediante:

$$\rho = \frac{|E|}{CP} \quad (3.5)$$

donde en (CP) es la Conexión Potencial que cuantifica la cantidad de cone-

xiones posibles en toda la red y se calcula como:

$$CP = \frac{n(n-1)}{2} \quad (3.6)$$

que en nuestro caso da $CP = \frac{77 \cdot 76}{2} = 2,926$, luego, $\rho = \frac{254}{2,926} \approx 0,087$.

La densidad indica cómo de cerca está la red de ser completa. Se dice que una red es *completa* si dos vértices cualesquiera están conectados por una arista. En ese caso la red tendría una densidad igual a 1.

Para el caso que nos compete de Los Miserables, la densidad es $\rho = 0,087$, lo que quiere decir que la mayoría de sus vértices no están conectados entre sí y, por tanto, la red no es completa. La completitud, depende mucho de las características de la red. Dado que la novela de Los Miserables gira en torno a la vida del protagonista Jean Valjean, la densidad de la red es relativamente baja.

A continuación, se presentan la Figura 3.6 que representa la cantidad en cada una de las 6 comunidades determinadas por Gephi y en la Figura 3.7 se representa cada una de ellas con un color distinto.

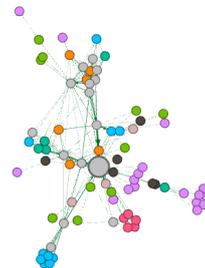
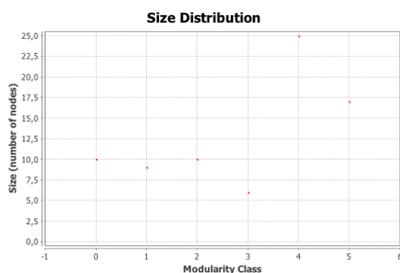


Figura 3.6: Distribución de la modularidad en Los Miserables. Elaboración con Gephi.

Figura 3.7: Comunidades en Los Miserables. Elaboración en Gephi.

Una forma de medir la importancia de un vértice en una red es mediante el cálculo del Pagerank. Esta medida no sólo tiene en cuenta si un vértice está muy conectado (tiene un grado alto), sino que tiene en cuenta si otros vértices muy conectados están conectados directamente con él

Definición 14 (PageRank- $PR(A)$). Sea $G(V, E)$ una red, el PageRank de un vértice $v_j \in V$, $PR(v)$, se determina mediante:

$$PR(v_j) = (1 - \lambda) + \lambda \sum_{j=1, (v_i, v_j) \in E}^n \frac{PR(v_i)}{d_{out}(i)} \quad (3.7)$$

donde:

- $PR(v)$ es el PageRank del vértice v .
- λ es un factor de amortiguación que tiene un valor entre 0 y 1, e indica qué importancia se le da a las conexiones existentes respecto a posibles conexiones aleatorias con otros vértices.
- $PR(v_i)$ son los valores de PageRank que tienen cada uno de los vértices enlazados con v_j .
- $d_{out}(i)$ es el grado (de salida) del vértice i .

En caso del ejemplo Los Miserables, la distribución de los $PR(v_j)$, $j = 1, \dots, n$ se aprecia en la Figura 3.8



Figura 3.8: Distribución de valores de PageRank para los vértices de la red de *Los Miserables*. Elaboración propia en Gephi.

El *coeficiente de agrupamiento* o de *clustering* de un vértice da una medida del agrupamiento de los vértices adyacentes a uno dado.

En el caso de redes no dirigidas, el coeficiente de clustering de un vértice $v_i \in V$ con grado k_i se calcula como

$$c_i = \frac{L_i}{k_i(k_i - 1)} \quad (3.8)$$

siendo L_i el número de conexiones entre los vértices adyacentes a i . Para redes no dirigidas se estima como

$$c_i = \frac{2|L|}{k_i(k_i - 1)} \quad (3.9)$$

De esta manera, el *coeficiente medio de clustering* de la red G se obtiene como

$$\bar{C} = \frac{1}{N} \sum_{i=1}^n c_i \quad (3.10)$$

El coeficiente medio de clustering permite ver si una red es de mundo pequeño al ponerse en comparación con respecto al valor de una red aleatoria de similares características.

Para el ejemplo de Los Miserables, el coeficientes medio de clustering es de $\bar{C} = 0,736$. Su representación en forma de red y de distribución es la que se muestra en la siguiente Figura 3.9:

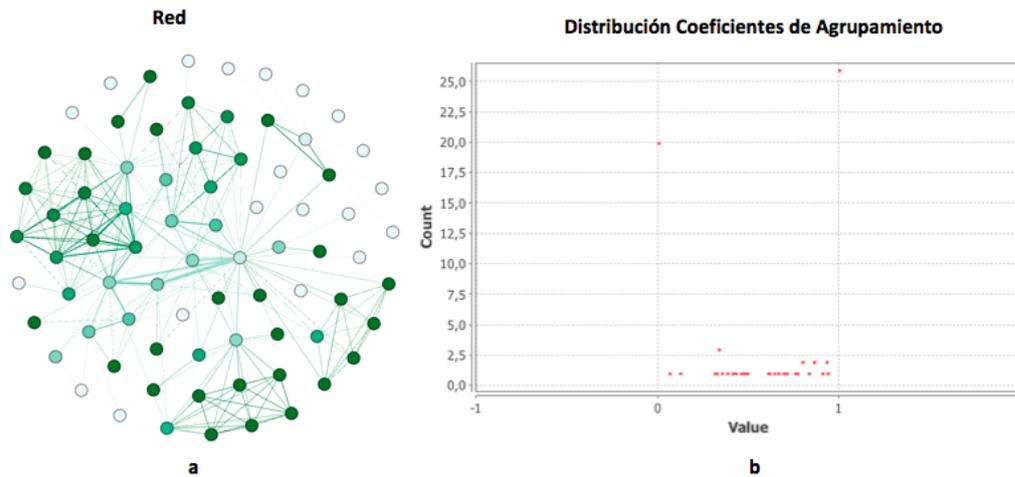


Figura 3.9: Coeficiente Medio de Agrupamiento. Elaboración propia en Gephi.

Capítulo 4

Redes aleatorias

Definición 15. Sea una red $G = (V, E)$ con n vértices. Diremos que G es aleatoria si las aristas que conectan dos vértices cualesquiera se generan aleatoriamente.

El inicio de la teoría sobre las redes aleatorias se debe a los matemáticos húngaros *Erdős y Rényi*, y siendo conocidas como modelos ER, debido a las iniciales de sus autores.

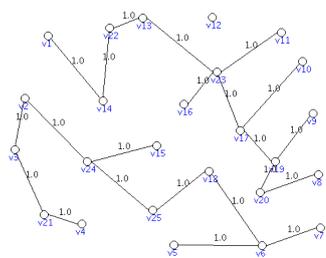


Figura 4.1: Red aleatoria no conexa. Probabilidad $0 < p < 1$. Elaboración propia en SWGraph.

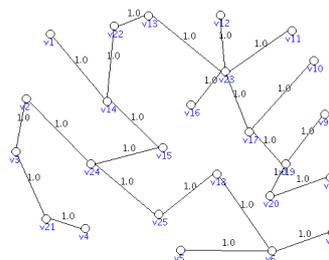


Figura 4.2: Red aleatoria conexa con probabilidad 1. Elaboración propia en SWGraph.

Se puede hablar de dos modelos de redes aleatorias:

1. $G(N, L)$. Redes aleatorias con N vértices y L aristas donde los nodos se conectan entre sí, aleatoriamente, utilizando esas L aristas.

2. $G(N, p)$. Redes aleatorias con N vértices donde las aristas se generan siguiendo una distribución de probabilidad binomial que nos indica si dos vértices están conectados o no.

A lo largo de este capítulo nos centraremos en este segundo modelo. A continuación, introducimos algunas nociones sobre espacios de probabilidad, que serán necesarias para enunciar algunos resultados sobre grafos aleatorios:

Definición 16. Diremos que (Ω, \mathcal{F}, P) es un espacio de probabilidad, donde:

- Ω es el conjunto (llamado *espacio muestral*) de los posibles sucesos elementales.
- \mathcal{F} es el conjunto de todos los sucesos que ocurren aleatoriamente (no solo los elementales). En caso de que todos los sucesos sean elementales, \mathcal{F} se define como un σ -álgebra sobre el conjunto Ω . De hecho, al par (Ω, \mathcal{F}) , es lo que se conoce como espacio de medida.
- P es la función de probabilidad, que asigna una probabilidad a todo suceso y que verifica los llamados axiomas de Kolmogórov.

Decimos que un suceso ε_n ocurre con alta probabilidad, **cap** por brevedad, en el espacio de probabilidad $G(n, p)$ si

$$\lim_{n \rightarrow \infty} P(G \sim G(N, p) \text{ cumple } \varepsilon_n) = 1 \quad (4.1)$$

[5].

Enunciamos a continuación los axiomas de Kolmogórov:

- $P(\Omega) = 1$, que indica que la ocurrencia estocástica de Ω se dará **cap**.
- La probabilidad de un suceso ε_n es un número comprendido entre 0 y 1: $0 \leq P(\varepsilon_n) \leq 1$.
- La probabilidad de la la unión de sucesos incompatibles dos a dos es la suma de las probabilidades de cada uno de los sucesos:

$$P(\cup_{i=1}^n E_i) = P(E_1) + P(E_2) + \dots + P(E_n) \quad (4.2)$$

siendo E_1, E_2, \dots, E_n sucesos incompatibles

De todo ello se deduce que:

$$P(\emptyset) = 1 - P(\Omega) = 0, \quad (4.3)$$

lo que significa que la probabilidad de que aparezca el conjunto vacío es cero.

Se puede entender un grafo aleatorio $G \sim G(N, p)$ como una variable aleatoria, cuya probabilidad de ocurrencia viene dada por P .

La conexión de la red va a depender del valor asignado al parámetro ϵ que va a definir el entorno en el que se desarrolla el proceso estocástico de transición del modelo *Erdős-Renyi*. El clásico resultado de *Erdős-Renyi* establece que el gráfico aleatorio $G(N, p)$ experimenta una transición de fase aguda alrededor de $\frac{1 \pm \epsilon}{n}$, para cualquier $\epsilon > 0$ [6].

A continuación, vamos a incluir las pruebas de los resultados de **cap y su dependencia del valor de ϵ** . La demostración que se presenta, es una prueba de los autores *Michael Krivelevich y Benny Sudakov* en el año 2012 en su artículo *The phase transition in random graphs – a simple proof* y se basan en el conocido algoritmo **Depth First Search (DFS)** de búsqueda de datos. Describimos el algoritmo brevemente aplicado a un grafo dado $G = (V, E)$ para facilitar la comprensión de la demostración siguiente. En la ejecución del algoritmo se consideran tres conjuntos:

1. S es el conjunto de vértices cuya exploración se completa en cada etapa del algoritmo.
2. T el conjunto de vértices no visitados.
3. $U = V \setminus (S \cup T)$, guardándose estos vértices en una pila LIFO

También se supone que hemos establecido algún tipo de orden σ en los vértices de G , de modo que el algoritmo prioriza los vértices en función de σ .

El algoritmo se inicializa con $S = U = \emptyset$ y $T = V$ siendo el test de parada $U \cup T = \emptyset$. Mientras el conjunto U sea distinto de vacío, el algoritmo comprueba, siguiendo el orden indicado por σ si el último vértice considerado u en U tiene algún vecino v en T , en cuyo caso añade v a U borrándolo de T . En

caso de que u no tenga ningún vecino, se elimina de U y se mueve a S . Si U es vacío, el algoritmo escoge el primer vértice de T de acuerdo con la ordenación σ , lo borra de T y lo añade a U . A fin de terminar la exploración, cuando tanto el conjunto S como el T son vacíos, momento en que la estructura en componentes conexas del grafo ha quedado ya determinada, hacemos que el algoritmo consulte los vértices en $S = V$, no consultados todavía. Obsérvese que el algoritmo DFS comienza a determinar una componente conexa C de G cuando el primer vértice de C entra en el conjunto, hasta ahora vacío, U , terminando de completarla cuando el conjunto U vuelve a estar vacío. Llamaremos etapa (en inglés “epoch”) el periodo de tiempo entre dos vaciados consecutivos de U , por lo que cada etapa corresponde a una componente conexa.

Destacamos las siguientes propiedades del algoritmo que se deducen de su funcionamiento:

1. En cada vuelta del algoritmo, siempre hay un vértice que se mueve de T a U , o de U a S (esto nos garantiza el recorrido en todos los vértices de la red) y, por tanto,
2. en cualquier etapa del algoritmo, se observa que el grafo G no tiene aristas entre los actuales S y T
3. El conjunto U siempre se alarga por un camino (de hecho, cuando un vértice u se añade a U , es porque u es un vecino del último vértice v en U , por lo que u aumenta el camino generado por U , del cual v es el último vértice).

En nuestro caso, hacemos correr el DFS sobre el grafo aleatorio $G \sim G(n, p)$, considerando como σ el conjunto de vértices $V = \{1, 2, \dots, n\}$ a fin de tener la permutación identidad. Si le pasamos al algoritmo la sucesión de variables aleatorias idénticamente distribuidas siguiendo una distribución de Bernoulli de parámetro p $\bar{X} = (X_i)_{i=1}^N$, de manera que que la i -ésima consulta positiva se corresponde con $X_i = 1$ y la negativa con $X_i = 0$, obtendremos un grafo que se distribuye claramente como $G(n, p)$. Por tanto

estudiar las componentes conexas de G se reduce al estudio de la propiedades de la sucesión aleatoria X_i .

Obsérvese que si $T \neq \emptyset$, cada respuesta positiva a una consulta da como resultado, como se comentó anteriormente, que un vértice se mueva de T a U y, por lo tanto, después de t consultas y asumiendo todavía que $T \neq \emptyset$, tenemos $|S \cup U| \geq \sum_{i=1}^t X_i$. (En realidad la última desigualdad es estricta porque el primer vértice de cada componente conexa se mueve "libremente" de T a U , es decir, sin necesidad de obtener una respuesta positiva

La parte probabilística de nuestra argumentación queda reflejada en el siguiente lema.

Lema 1. *Sea $\epsilon > 0$ una constante suficientemente pequeña. Sea $\bar{X} = (X_i)_{i=1}^N$ una sucesión de variables aleatorias i.i.d. que sigue una distribución de Bernoulli, con parámetro p .*

1. *Sea $p = \frac{1-\epsilon}{n}$. Sea $k = \frac{7}{\epsilon^2}$, entonces **cap** no hay ningún intervalo de longitud kn en $[N]$, en el que al menos k de las variables aleatorias X_i toman el valor 1.*
2. *Sea $p = \frac{1+\epsilon}{n}$. Sea $N_0 = \frac{\epsilon n^2}{2}$, entonces **cap**, $|\sum_{i=1}^{N_0} X_i - \frac{\epsilon(1+\epsilon)n}{2}| \leq n^{2/3}$*

Demostración. 1. Para un intervalo dado I de longitud kn en $[N]$, la suma $\sum_{i \in I} X_i$ se distribuye binomialmente, con parámetros kn y p . Aplicando una cota estándar de tipo de *Chernoff* [1], que da límites exponencialmente decrecientes en las distribuciones de cola de sumas de variables aleatorias independientes, a la parte superior de la cola $B(kn, p)$ y aplicando la cota de la unión, vemos que la probabilidad de la existencia de un intervalo, violando la aserción del lema, es a lo sumo:

$$(N - k + 1)P[B(kn, p) \geq k] < n^2 \cdot e^{-\frac{2}{3}(1-\epsilon)k} < n^2 \cdot e^{-\frac{\epsilon^2(1-\epsilon)}{3} \frac{7}{\epsilon^2} \log n} = \mathcal{O}(1),$$

para un $\epsilon > 0$ suficientemente pequeño.

2. La suma $\sum_{i=1}^{N_0} X_i$ se distribuye binomialmente con parámetros N_0 y p . Por tanto, su esperanza $N_0 P = \frac{\epsilon n^2 p}{2} = \frac{\epsilon(1+\epsilon)n}{2}$, y su desviación estándar

es de orden n . Aplicando la desigualdad de *Chebyshev*, obtenemos la estimación requerida. □

EL anterior lema nos permite formular y probar el siguiente resultado.

Teorema 1. *Sea $\epsilon > 0$ una constante suficientemente pequeña. Sea $G \sim G(N, p)$ un grafo aleatorio.*

1. *Sea $p = \frac{1-\epsilon}{n}$. Entonces, **cap** todas las componentes de G son de tamaño, como máximo, $\frac{7}{\epsilon^2} \log n$.*
2. *Sea $p = \frac{1+\epsilon}{n}$,. Entonces, **cap** G contiene una trayectoria de longitud al menos $\frac{\epsilon^2 n}{5}$.*

En ambos casos, hacemos uso del algoritmo DFS sobre el grafo aleatorio $G \sim G(N, p)$ y asumimos que la sucesión $\bar{X} = (X_i)_{i=1}^N$ de variables aleatorias, que definen el grafo aleatorio $G \sim G(N, p)$ y son guía para el algoritmo DFS, satisface la parte correspondiente al Lema 2.

Demostración. 1. Procedemos por reducción al absurdo, es decir, suponemos que G contiene una componente conexa C con más de $k = \frac{7}{\epsilon^2} \log n$ vértices. Consideremos la fase del DFS en que se generó C y el momento dentro de ella en que el algoritmo ha encontrado el vértice $(k+1)$ -ésimo de C y está a punto de moverlo a U . Denotamos $\Delta S = S \cap C$ ese momento. Entonces $|\Delta S \cup U| = k$, y por lo tanto el algoritmo proporciona exactamente k respuestas positivas a sus consultas a las variables aleatorias X_i durante la mencionada fase, siendo cada una de las respuestas positivas responsable de encontrarse con un nuevo vértice de C , después de que el primer vértice de C se pusiera en U al principio de la fase. En ese momento, durante dicha fase, sólo se han consultado pares de enlaces con extremos en $\Delta S \cup U$, y el número de tales pares es por tanto, como máximo, $\binom{k}{2} + k(n-k) < kn$. De aquí se deduce que la sucesión \bar{X} contiene un intervalo de longitud, a lo sumo, kn con al menos k 1's dentro, lo que está en contradicción con la propiedad 1 del Lema 2.

2. Supongamos que la sucesión \bar{X} satisface la propiedad 2 del Lema 2. Afirmamos que después de las primeras $N_0 = \frac{\epsilon n^2}{2}$ iteraciones del algoritmo DFS, el conjunto U contiene al menos $\frac{\epsilon^2 n}{5}$ vértices (con los elementos de U formando un camino de la longitud deseada en ese momento).

Obsérvese primero que $|S| < \frac{n}{3}$ en el instante N_0 . De hecho, si fuera $|S| \geq n$, si miramos el instante t donde $|S| = \frac{n}{3}$ (la existencia de tal instante está asegurada ya que los vértices fluyen hacia S uno por uno). En ese momento $|U| \leq 1 + \sum_{i=1}^t X_i < \frac{n}{3}$ por la propiedad 2 de Lemma 2. Luego, $|T| = n - |S| - |U| \geq \frac{n}{3}$ y el algoritmo ha examinado todos los $|S| \cdot |T| \geq \frac{n^2}{9} > N_0$ pares entre S y T (y encontrado que no son aristas del grafo), lo cual es una contradicción.

Volvamos al instante N_0 . Si $|S| < \frac{n}{3}$ y $|U| < \frac{\epsilon^2 n}{5}$ entonces tenemos que $T \neq \emptyset$. Esto significa que el algoritmo sigue obteniendo las componentes conexas de G , y cada respuesta positiva se tradujo en mover un vértice desde T a U (algunos de estos vértices ya pueden haberse movido de U a S). Por la propiedad 2 del Lema 2, el número de respuestas positivas en este momento es al menos $\frac{\epsilon(1+\epsilon)n}{2} - n^{2/3}$. Por lo tanto, tenemos $|S \cup U| \geq \frac{\epsilon(1+\epsilon)n}{2} - n^{2/3}$. Si $|U| \leq \frac{\epsilon^2 n}{5}$, entonces $|S| \geq \frac{\epsilon n}{2} + \frac{3\epsilon^2 n}{10} - n^{2/3}$. Todos los pares, $|S||T| \geq |S|(n - |S| - \frac{\epsilon^2 n}{5})$, entre S y T han sido probados por el algoritmo (y dado respuesta negativa). Por tanto obtenemos, utilizando la hipótesis de que $|S| < \frac{n}{3}$, que :

$$\begin{aligned} \frac{\epsilon n^2}{2} = N_0 &\geq |S|(n - |S| - \frac{\epsilon^2 n}{5}) \geq (\frac{\epsilon n}{2} + \frac{3\epsilon^2 n}{10} - n^{2/3})(n - \frac{\epsilon n}{2} - \frac{\epsilon^2 n}{2} + n^{2/3}) \\ &= \frac{\epsilon n^2}{2} + \frac{\epsilon^2 n^2}{20} - \mathcal{O}(\epsilon^3)n^2 > \frac{\epsilon n^2}{2} \end{aligned}$$

Lo que es obviamente una contradicción. \square

El número de aristas en una red aleatoria

Muchas veces el interés del investigador se centra en determinar el cardinal de enlaces que se necesita para diseñar una determinada red, de la que se conoce la cantidad de nodos N y la probabilidad p de sus conexiones. Esto

es debido a que para valores similares de nodos, e incluso, para la misma probabilidad p , se pueden tener configuraciones ligeramente distintas entre el diseño de una red y otra, dado que sus enlaces pueden variar.

La probabilidad de que el vértice v de una red aleatoria tenga exactamente grado L enlaces teniendo en cuenta:

1. La probabilidad de que el vértice v esté conectado a L vértices concretos y no a los restantes es $p^L(1-p)^{\frac{N(N-1)}{2}-L}$

2. Existen

$$\binom{\frac{N(N-1)}{2}}{L}$$

posibilidades de combinar los N vértices de L en L

es:

$$p_L = \binom{\frac{N(N-1)}{2}}{L} p^L (1-p)^{\frac{N(N-1)}{2}-L} \quad (4.4)$$

Dado que se trata de una distribución binomial, el número esperado de enlaces en una red aleatoria es:

$$\langle L \rangle = \sum_{L=0}^{\frac{N(N-1)}{2}} L p_L = p \frac{N(N-1)}{2} \quad (4.5)$$

A partir de la definición de grado medio $\langle k \rangle$ obtenemos

$$= \frac{2\langle L \rangle}{N} = p(N-1), \quad (4.6)$$

y

$$\sigma_k = p(1-p)\sqrt{N-1} \quad (4.7)$$

Se deduce por tanto que conforme aumenta el tamaño de la red, la distribución se va volviendo más estrecha, es decir, tenemos una seguridad creciente de que el grado de un nodo está cerca del valor de $\langle k \rangle$.

Distribución de Grado

La mayoría de las redes reales son dispersas $\langle k \rangle \ll N$. En ese caso y cuando N es grande, la distribución binomial se puede aproximar por una de Poisson, más cómoda de manejar:

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \quad (4.8)$$

con esperanza $\langle k \rangle$ y

$$\sigma_k = \langle k \rangle^{1/2} \quad (4.9)$$

Como podemos apreciar en las relaciones anteriores que, para estimar la probabilidad de que el nodo elegido al azar tenga determinado grado, en la distribución Binomial discreta se necesitan los parámetros N y p_k , mientras que en la Poisson discreta sólo es necesario el parámetro k . Esta diferencia nos induce a considerar una u otra, dependiendo de la información de que se disponga, pero, normalmente, se emplea la Poisson, dada su simplicidad, ya que sólo se necesita un parámetro para las estimaciones de los estadísticos requeridos prediciendo que la distribución del grado es la misma en redes de distintos tamaños en tanto en cuanto tengan el mismo grado medio.

Capítulo 5

Redes de pequeño mundo

Uno de los resultados más importantes y, que al día de hoy, fascina a la comunidad matemática, tiene que ver con la teoría del fenómeno de *Pequeño Mundo*, también conocido como *Teoría de los Seis Grados de Separación*. Esta tesis afirma que si consideramos dos personas aleatoriamente en dos lugares distintos en cualquier parte del mundo o en cualquier entorno, seguramente existen seis personas que se conocen y mediante los cuáles se puede establecer una cadena que conecte con esa otra persona.

Las Redes de Pequeño Mundo revela que la sociedad humana es una red social con forma de mundo pequeño y que están interconectados entre sí, con estructura de red, cuyos nódulos son personas y los enlaces, las interrelaciones entre ellos.

Experimento de Stanley Milgram

Stanley Milgram, (*Nueva York: 15 de agosto de 1933/ 20 de diciembre de 1984*), psicólogo graduado en la Universidad de Yale y egresado de la Universidad de Harvard en 1960. Famoso por sus experimentos del **Mundo Pequeño**, apoyo de la teoría de **Los Seis Grados de Separación**, también condujo el famoso: *Experimento de Milgram*, cuyo fin era determinar el nivel de obediencia de los seres humanos en relación a la autoridad.

Las redes de Mundo Pequeño revela que la sociedad humana es una red social con forma de mundo pequeño y que están interconectados entre sí, con estructura de red, cuyos nódulos son personas y los enlaces, los tipos de conocimientos/relaciones entre ellos.

El experimento de Milgram consistió en:

- Elección al azar de los participantes (en pares de personas conocidas), en las ciudades norteamericanas **Omaha, Wichita y Boston**, de modo que estas ciudades representan en origen y el destino de una cadena de correspondencia que se enviaría a dos conocidos por medio de otros que los conectarían.
- Directamente, a los participantes que residían en Omaha y Wichita "origen", se les enviaba cartas con informaciones que detallaban el objetivo del estudio, además de información básica acerca de la persona "destino" que debía ser contactado en Boston. Contenían además una lista en la que las personas que participaban debían inscribir sus nombres, y tarjetas de respuesta pre-dirigidas a Harvard (cabe resaltar que este aspecto era muy importante, porque en caso de que el proceso se interrumpiera por cualquier razón, se sabría perfectamente, en cual nódulo se daría dicha ruptura del eslabón).
- En caso de que el participante (primer eslabón de la cadena de correspondencias) conociera a la persona destino (último eslabón), a esta conexión se le llamaba "relación de tú a tú".
- En el caso de que la persona no conociera directamente al destinatario, entonces debía considerar a la otra persona más cercana que permitiera conectarlo con la de origen.
- Por último, cuando el destinatario recibía la comunicación, los investigadores (en Harvard) comenzaban a contar la cantidad de veces que la comunicación había pasado por personas previas.

Como resultado del estudio, de las 296 cartas enviadas, solo llegaron a su destino 64 con distancia promedio de relaciones de entre 5,5 y 6. Esto dio lugar a la teoría de los Seis grados de Separación.

Lo más curioso es que se han llevado a cabo estudios de similar naturaleza, obteniéndose resultados muy parecidos.

Modelo de Watts & Strogatz

El experimento de Milgram es, según la crítica, rudimentario y no favorece la calidad de la información levantada, puesto que no garantiza que las relaciones sean las de distancia mínima (menor diámetro entre dos nodos en la cadena de correspondencias) propiciando factores de clasterizaciones no óptimas, es decir, muchos agrupamientos con pequeños número de vértices.

Como respuesta a ello, en 1998, los matemáticos *Watts y Strogatz* publicaron un famoso artículo en la revista *Nature* titulado *Collective dynamics of small-world networks*, en la que presentan un nuevo algoritmo dinámico de agrupamiento de pequeños mundos, en base a dos criterios esencialmente fundamentales: a) distancias medias pequeñas entre dos nodos seleccionados al azar y un b) coeficiente de clasterización alto [8].

El algoritmo de Watts-Strogatz

A continuación describimos los pasos del algoritmo que propusieron:

- Se establece una red inicial unidimensional con n nodos, estos nodos se pueden disponer en forma de *anillo k -Regular* de tal forma que cada uno de los vértices se una con $2k$ vecinos. Esto hace que el número total de aristas sea $k \cdot n$.
- La probabilidad de conectar un nodo con otro cualquiera es p . Para cada nodo v_i . se toma v_j con $1 \leq i < j \leq n$, y se reconectan v_i con v_j con probabilidad p . Esto se hace eliminando la arista (v_i, v_j) y añadiendo

una arista (v_i, v_k) , donde $1 \leq k \leq n$, siendo $k \neq i$ y k no pudiendo ser ninguno de los vértices que se encuentren conectados con v_i .

Para un grafo con $p = 0$ se puede ver que la conectividad es la misma y de valor $2k$ y su $C(0) \sim \frac{3}{4}$. Por otro lado, un valor no nulo de p introduce desorden en la red de tal forma que la conectividad no es uniforme, manteniendo todavía de media un valor de $2k$. Esta construcción nos permite sintonizar el gráfico entre la regularidad ($p = 0$) y el desorden ($p = 1$), y por lo tanto sondear la región intermedia $0 < p < 1$, sobre la cual poco se sabe.[8]

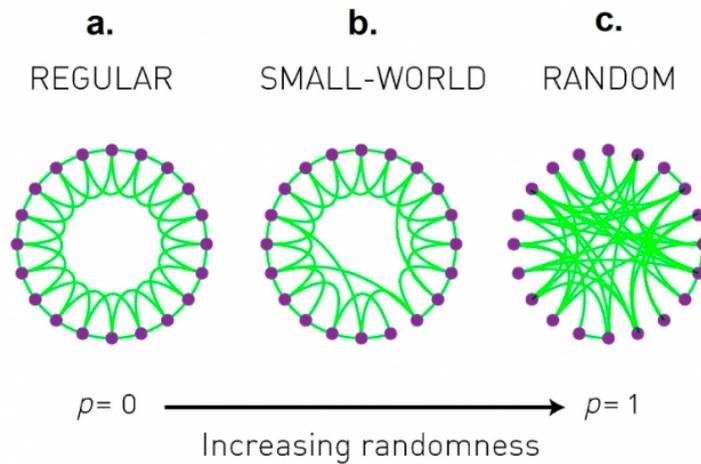


Figura 5.1: Fuente: [2].

Intepretación de las redes de pequeño mundo

Una interpretación de las redes de pequeño mundo es que la distancia entre dos nodos elegidos aleatoriamente en una red es corta. Sin embargo, a ¿qué llamamos *distancia corta*? Veamos, matemáticamente, cómo podemos interpretar este hecho.

Consideremos una red aleatoria donde estimamos el grado medio de la red como $\langle k \rangle$. Consideremos la siguiente notación para determinar la estimación del número de vértices que se encuentran a distancia d de uno dado:

$$\langle k \rangle^d \tag{5.1}$$

Esto implica que el número esperado de vértices a distancia d de uno dado se puede estimar de manera recursiva como $\langle k \rangle^d = \langle k \rangle \times \langle k \rangle^{d-1}$. De esta manera el número de vértices a distancia d de uno dado se puede estimar como

$$n(d) \approx 1 + \langle k \rangle + \langle k \rangle^2 + \langle k \rangle^3 + \dots + \langle k \rangle^d = \frac{\langle k \rangle^{d+1}}{\langle k \rangle - 1} \quad (5.2)$$

de modo que, a lo sumo,

$$n(d_{max}) \approx n \quad (5.3)$$

Para valores de $\langle k \rangle \gg 1$, es fácil apreciar que la expresión anterior nos queda de la siguiente manera

$$\langle k \rangle^{d_{max}} \approx n \quad (5.4)$$

de ahí que el diámetro de una red aleatoria sea:

$$d_{max} \approx \frac{\log(n)}{\log(\langle k \rangle)} \quad (5.5)$$

En [2] se encuentra la siguiente interpretación de una red de pequeño mundo: “... En general, $\log(n) \ll n$, por lo tanto la dependencia de $\langle d \rangle$ en $\ln(N)$ implica que las distancias en una red aleatoria son órdenes de magnitud menores que el tamaño de la red. Consecuentemente, por pequeño en el “fenómeno del mundo pequeño” queremos decir que la longitud promedio de la trayectoria o el diámetro depende logarítmicamente del tamaño del sistema. Por lo tanto, “pequeño” significa que $\langle d \rangle$ es proporcional a $\log(n)$, en lugar de n o alguna potencia de n ”.

Capítulo 6

Redes libres de escala

Durante muchos años (entre finales del siglo XIX y principios del XX) se consideraban las redes como si tuvieran un elemento regulador, e incluso, se les estudiaba bajo nociones euclidianas, en busca de una simplicidad. Posteriormente, en más de 40 de los últimos casi 60 años, las redes se trataban siguiendo la teoría del modelo Erdős-Renyi, en el que la cantidad de aristas que conectan vértices siguen un promedio y, por tanto, la frecuencia con el que un vértice tiene un determinado grado sigue una distribución de Poisson (Figura 6.1).

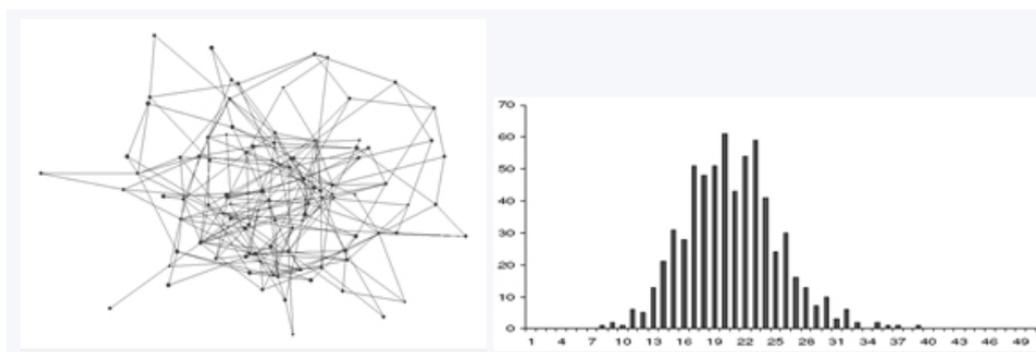


Figura 6.1: Red aleatoria. Fuente: [?]

En el año 1998, Barabási junto a otros investigadores, se centraron en un proyecto para mapear la World Wide Web (WWW). En principio pensaban

que se encontrarían con un modelo más del tipo Erdős-Renyi. Sin embargo, al igual que pasó a Lorenz con el efecto mariposa y el caos, estos investigadores se encontraron con una sorpresa: el modelo seguía un patrón, en el que pocas páginas estaban altamente conectadas. Ello implicaba que la totalidad de la Web estaba parcializada en muy pocas páginas y de ahí el nacimiento del término hubs. Por su parte, la mayoría de las demás páginas (más del 80 %) estaban enlazadas, como máximo a otras cuatro.

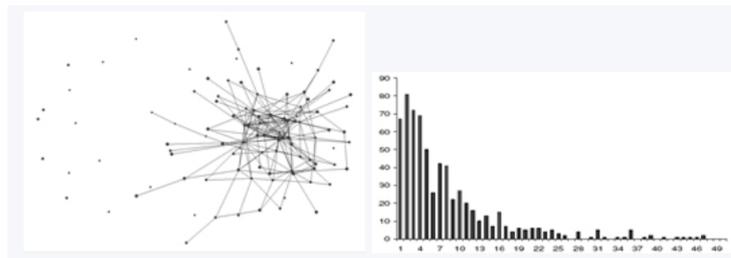


Figura 6.2: Red Libre de Escala. Fuente: o[?]

En la Figura 6.3 que se presentan a continuación, se observa la distribución de cableados de una red eléctrica, en la que se forman hubs (2 exactamente: los nodos de mayor radio en verde y azul), en los que se concentran muchas conexiones. Esto induce a que se puede estar perfilando que la distribución de la frecuencia de los grados siga una ley de potencia.

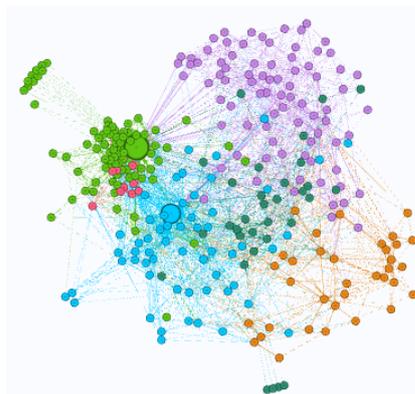


Figura 6.3: Distribución red eléctrica Oeste EE.UU. Elaboración propia en Gephi.

Otros ejemplos pragmáticos, entre los que se da esta ley de potencias están relacionados con apariciones de casos frecuentes de enfermedades de transmisión sexual, agendas telefónicas, nexos sintácticos entre palabras en un texto o discurso, citas bibliográficas entre miembros de la comunidad académica, colaboraciones en reportes de investigación, alianzas tecnológicas, relaciones entre actores de cine, sinapsis neuronales, contactos entre personas en una organización, cadenas alimentarias, conexiones entre organismos vinculados al metabolismo o proteínas reguladoras, propagación de enfermedades y virus informáticos.

Estas redes siguen un fenómeno inesperado: si se desconectaran el 80 % de sus nodos con menos conexiones, el resto seguiría conectado. De esta propiedad se deduce otra, de similar importancia, *si se elimina entre el 5 y el 10 % de de los “hubs”, sería suficiente para que el sistema se “desmoronase”, permitiendo, incluso que la red rompiera su conexión.* Además de estas propiedades, las redes obtenidas a partir de datos reales suelen cumplir esta otras propiedades

Además de esta propiedad, las RLE muestra otras más:

- Jerarquía de conglomeraciones o *clusters*, permitiendo la aparición de picos, lo que es lo mismo, presenta la ley de potencia.
- El exponente $-\gamma$ en la ley de potencias $k^{-\gamma}$ suele estar entre 2 y 3.
- Debido a la característica de su estructura, cualquier nodo está conectado con cualquier otro con muy pocos grados de separación.
- Existencia de nodos cuyo valor de conectividad supera varias veces el número promedio, lo que no es propio de las distribuciones aleatorias (modelo ER), como ejemplo, la que rige la tabla de estaturas de una población, donde nunca se encontrará una persona que esté conectada mil o un millón de veces más que otra.

En virtud de la distribución singular de las redes libres de escala, muchos de los métodos estadísticos, como son: *muestreo, análisis de varianza, generalización, coeficientes de correlación* no son pertinentes para tratarlas, dado

que ello supone articular *distribuciones normales, linealidad y regímenes estables*, lo que no se da en este tipo de redes, dado que su distribución sigue una ley de potencia, de transición de fase y libres de escala. Sobre esto último, vamos a intentar dar algunas nociones, que se esbozan a continuación.

Definición 17 (Red Libre de Escala). *Una red libre de escala es una red cuya distribución de grados siguen una ley de potencias de exponente negativo, $p_k \sim k^{-\gamma}$ con $\gamma > 0$.*

En general, en los ejemplos de la naturaleza y de la vida real se tiene $2 < \gamma < 3$. De esta manera, estas redes contienen una cantidad relativamente pequeña de nodos muy conectados, de grado alto, llamados “*hubs*”, y una gran mayoría de vértices que tengan pocas conexiones.

Ajuste del exponente de la ley de potencias

Otra característica es que la distribución de la red no es aleatoria, dado que la distribución de Poisson presenta malos ajustes de la distribución potencial de los grados y esto se debe, en gran medida, a los hubs. Sin embargo, en una escala logarítmica de los ejes ($\log - \log$), los puntos de datos forman una línea recta aproximada.

Aplicando logaritmo a ambos lados de la expresión $p_k \sim k^{-\gamma}$, obtenemos:

$$\log p_k \sim -\gamma \log k \quad (6.1)$$

De modo, que si la relación se mantiene, entonces la esperanza del $\log p_k$ depende linealmente del $\log k$, cuyo coeficiente (pendiente) es el valor del exponente γ . En el caso de la WWW, se sabe que es una red dirigida, dado que los archivos consultados tienen conexiones entrantes (in) y de salida (out), de ahí que tengamos las siguientes relaciones:

$$p_{k_{in}} \sim k^{-\gamma_{in}} \quad (6.2)$$

$$p_{k_{out}} \sim k^{-\gamma_{out}} \quad (6.3)$$

En la Figura 6.4, se puede apreciar como se realizan estos ajustes:

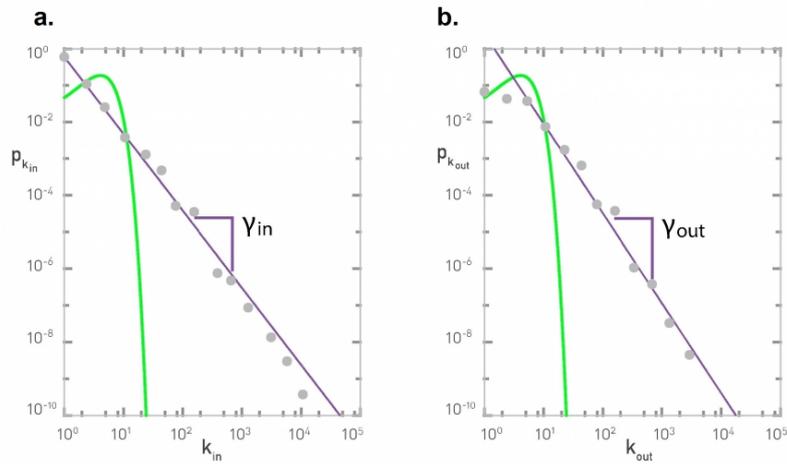


Figura 6.4: Fuente: [2]

Ajuste de la constante de la proporcionalidad de la ley de potencias

Una vez estimada γ , dado que la probabilidad de que un vértice de grado k viene dada por

$$p_k = Ck^{-\gamma}, \quad (6.4)$$

queda por estimar la constante de proporcionalidad C . Como $\sum_{k=1}^{\infty} p_k = 1$, combinando ambas expresiones se tiene

$$C \sum_{k=1}^{\infty} k^{-\gamma} = 1 \quad (6.5)$$

Esto implica que:

$$C = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}} = \frac{1}{\zeta(\gamma)} \quad (6.6)$$

donde $\zeta(\gamma)$ es la función zeta de Riemann y, por tanto, para $k > 0$, la distribución discreta de las probabilidades de la ley de potencias tiene la

forma:

$$p_k = \frac{k^{-\gamma}}{\zeta(\gamma)} \quad (6.7)$$

Hay que tener en cuenta que la expresión anterior diverge para el valor de $k = 0$. En general se suele considerar las estimaciones de los grados para las componentes conexas de las redes con lo que esta situación no se da.

Mediante un razonamiento análogo pero considerando la versión continua de las estimaciones anteriores se puede obtener

$$p_k = (\gamma - 1)k_{min}^{\gamma-1}k^{-\gamma} \quad (6.8)$$

donde k_{min} es el grado mínimo de cualquier vértice de la red.

Hubs

Como hemos indicado anteriormente, las distribuciones de grados de las redes aleatorias y de las redes de escala no son comparables, debido en buena medida a los hubs. En la Figura 6.5 se presenta la diferencia entre la distribución de una red aleatoria versus a una red libre de escala.

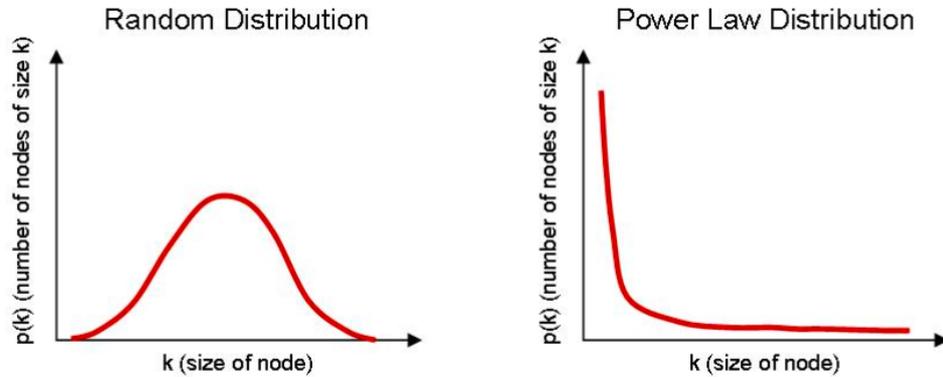


Figura 6.5: Fuente: Wikipedia.

Las principales diferencias entre ambas distribuciones son las siguientes:

- Para valores de k pequeños, la ley de potencia está por encima de la función de Poisson, indicando que una red libre de escala tiene un

gran número de vértices de grado bajo, la mayoría de los cuales están ausentes en una red aleatoria.

- Para k entorno a $\langle k \rangle$ la distribución de Poisson está por encima de la ley de potencia, lo que indica que en una red aleatoria hay un exceso de vértices de grado próximo a $\langle k \rangle$, en comparación a las redes libres de escala.
- Para k grande, la ley de potencia está otra vez por encima de la curva de Poisson, lo que indica que la probabilidad de observar un nodo de alto grado, o hub, es varios órdenes de magnitud más alta en una red libre de escala que en una aleatoria.

Tamaño de los hubs

Hay redes cuyo número de vértices es súmamente grande. Por ejemplo, se estima el tamaño de la WWW en $n \approx 10^{12}$, la red neuronal de un cerebro humano en $n \approx 10^{11}$ y la población de seres humanos de nuestro planeta $n \approx 7 \times 10^9$. Podemos encontrar redes que sean de un tamaño inmenso, sin embargo, ello nos dice que, en todas las redes que el ser humano ha descubierto, existe *finitud*. Esto quiere decir, que aún somos capaces de contar la cantidad de nodos y conexiones en las redes.

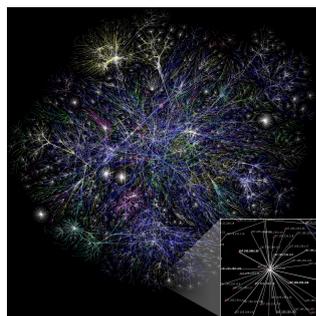


Figura 6.6: Internet Backbone. Fuente: https://en.wikipedia.org/wiki/Internet_backbone

En este contexto, sabemos que somos capaces de contar todos los vér-

tices y sus aristas en una red libre de escala. Veamos como se puede estimar a priori el grado de los hubs.

Supongamos que hay un valor k_{max} a partir del cual se espera que como máximo haya un vértice con grado mayor que k_{max} . A partir de la expresión (6.8) y teniendo en cuenta que $\int_{k_{max}}^{\infty} = \frac{1}{N}$, se obtiene que

$$k_{max} = k_{min} N^{\frac{1}{\gamma-1}} \quad (6.9)$$

La expresión anterior nos lleva a la siguiente conclusión: El tamaño del hub en una red, es directamente proporcional al número de nodos y de conexiones, es decir, cuanto mayor es una red, mayor es el grado de su mayor hub. En la Figura 6.7 se puede ver la relación existente entre los k_{max} de las redes libres de escala y de las redes aleatorias.

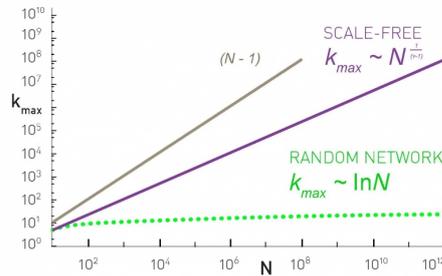


Figura 6.7: Mayor grado k_{max} según tipo de distribución. Fuente: [2]

Etimología del concepto *Libre de Escala*

La física estadística, es el apoyo conceptual y técnico desde donde se concibe los modelos de las redes libres de escala. En particular estos se apoyan en la teoría de las transiciones de fase.

Para entender el concepto *libre de escala*, primero debemos conocer lo que se entiende por *momento de la distribución de grados*. Estos se

definen de la siguiente manera. Para una red en el que el grado mínimo de un vértice sea k_{min} tenemos:

$$\langle k^n \rangle = \sum_{k_{min}}^{\infty} k^n p_k \approx \int_{k_{min}}^{\infty} k^n p_k dk \quad (6.10)$$

A continuación, detallamos la importancia de los momentos inferiores, es decir, cuando $n = 1, 2, 3$:

- $n = 1$: El primer momento se corresponde con el valor esperado del grado de un vértice, $\langle k \rangle$.
- $n = 2$: El segundo momento, $\langle k^2 \rangle$, nos proporciona el cálculo de la varianza $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$, midiendo la propagación en los grados. De manera directa, podemos obtener la desviación estándar, aplicando la raíz cuadrada a la varianza: $\sigma = \sqrt{\sigma^2}$
- $n = 3$: El tercer momento, $\langle k^3 \rangle$, determina la asimetría de una distribución, cuán simétrica es la probabilidad p_k alrededor del promedio $\langle k \rangle$.

Más concretamente, en una red libre de escala el n -ésimo momento de la distribución de grados viene dado por:

$$\langle k^n \rangle = \int_{k_{min}}^{k_{max}} k^n p_k dk = C \frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n - \gamma + 1} \quad (6.11)$$

Hasta ahora conocíamos que k_{min} , al ser el grado menor, se entiende muchas veces como un valor fijo y que k_{max} es proporcional al tamaño de la red.

Como ya se conoce la ley de potencias asociada a la distribución de los grados en una red libre de escala, solo hace falta entender al valor máximo al que puede llegar k_{max} . Esto significa, saber cuál es su límite asintótico. Este límite nos permite conocer el valor del $\langle k^n \rangle$, mediante la interacción entre n y γ , lo que quiere decir que:

- Si $n - \gamma + 1 \leq 0$ entonces el primer término en la ecuación anterior, $(k_{max}^{n-\gamma+1})$ tiende a cero a medida que k_{max} aumenta. Por lo tanto, todos los momentos que satisfacen $n \leq \gamma - 1$ son finitos.
- Si $n - \gamma + 1 > 0$ entonces $\langle k^n \rangle$ va al infinito si $k_{max} \rightarrow \infty$. Por lo tanto, todos los momentos mayores que $\gamma - 1$ divergen.

En el caso de que las redes tengan un valor $2 < \gamma < 3$, se tendría que en este último caso, para $n \geq 3$. Es por ello, que seguramente podemos alcanzar a cualquier persona con una de conocidos de longitud 3, y bajar incluso de la estimación que los datos de Facebook daban de 4.

Capítulo 7

La red de divisibilidad de los números naturales

A continuación, vamos a presentar una descripción de los resultados de un artículo reciente, que parte de un conjunto numéricamente rico, como es el de los número naturales, y en el que a partir de la relación de divisibilidad entre sus elementos se haya una red libre de escala con interesantes propiedades. El trabajo al que nos referimos al trabajo de Shekatkar, Bhagwat y Ambika *Divisibility patterns of natural numbers on a complex network* publicado en Scientific Reports. [7]

Se considera el conjunto de los número naturales $\mathbb{N} = 1, 2, 3, \dots$ como los vértices de la red. Si un número divide a otro entonces sus vértices están conectados por una arista. En la siguiente figura que se presenta a continuación, hay dos redes con $n = 16$ y $n = 32$ nodos.

En la Figura 7.1, se puede apreciar que en función del grado de cada vértice, éstos se colorean de manera diferente. También se aprecia que van apareciendo, de forma impredecible, los número primos. Estos van a ser los que propicien que esta red sea libre de escala.

La red se irá construyendo de manera iterativa como se indica a continuación:

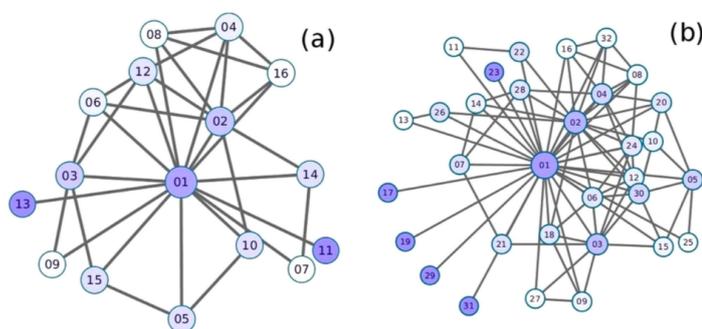


Figura 7.1: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4571582/pdf/srep14280.pdf>

1. En el tiempo $t = 1$ la red comienza con un solo vértice $v_1 = 1$ y en cada tiempo t , se agrega a la red un vértice con el número $n = t$.
2. En el paso t , el vértice v_t nodo se conecta a todos los vértices existentes cuyos números dividen a t .

Atendiendo a esto, en el estudio se hace crecer la red hasta $n = 2^{25}$, lo que ha supuesto un coste computacional alto ya que el número de nodos asciende a $N = 3355432$. Estos resultados, se pueden ver muy bien utilizando un tipo de gráfico llamado (log – log), dado que la distribución del conjunto numérico sigue una ley de potencia, como se aprecia a continuación:

Esta gráfica representa una distribución que se rige por la ley de potencia, en el que pocos nodos (hubs) tienen alto grado de conexiones. Dichos nodos son los números naturales más bajos.

La distribución de las probabilidades de que un vértice tenga un determinado grado $\{p_k\}$ sigue una ley de potencia $p_k \sim k^{-\gamma}$ con $\gamma = -2$ que describe la pendiente negativa de la distribución.

Al calcular el coeficiente de clusterización local para cada vértice $v_i \in V$, se obtiene la siguiente figura:

En la Figura 7.3, se puede ver que el comportamiento asintótico es

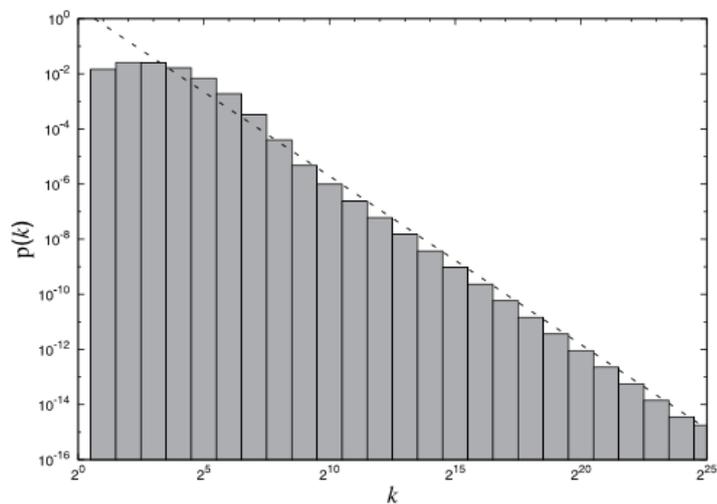


Figura 7.2: Distribución de los grados en la red de divisibilidad de los números naturales: Gráfico tipo log – log. Fuente: [7]

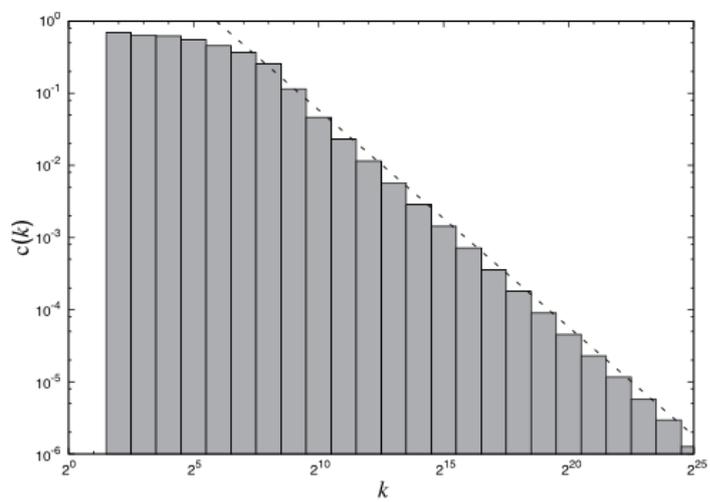


Figura 7.3: Dependencia del coeficiente de agrupamiento local en grado. Fuente: Artículo

compatible con una ley de potencia y es un comportamiento similar al que suele observarse en redes reales. En este trabajo, los autores introducen una nueva medida del coeficiente de agrupamiento que han denominado: *semejanza de estiramiento*. Esta medida describe el suave desplazamiento de los coeficientes de clusterización al añadir de uno en uno los vértices de la red. Esto permite hacer un análisis no estacionario de la red. Ver Figura 7.4:

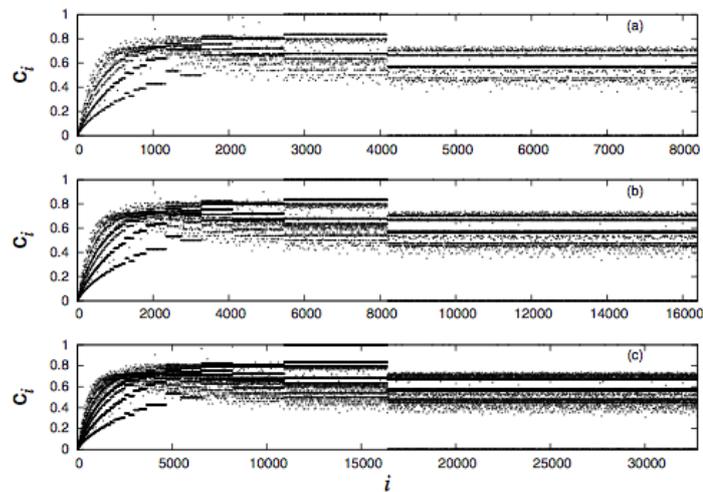


Figura 7.4: Coeficiente de agrupamiento local en función del índice de nodos para tres tamaños distintos de la red $n = 8192, 16384$ y 32768 . Fuente: [7].

Se observa un estiramiento global del perfil del modelo a medida que n se hace muy grande. Este estiramiento se va suavizando y parece alcanzar una cota, aunque si ubicamos cualquier región dentro de la gráfica nos encontramos con valores c_i dispersos y, hasta ahora, es el único conjunto de datos que exhibe tal comportamiento, de los estudiados hasta ahora. Se puede apreciar que los saltos más abruptos se dan al pasar por las n que son potencias de 2.

De igual manera, los números primos también provocan ciertos saltos que aparecen de manera irregular. Para explicar este fenómeno, consideremos el siguiente argumento: tomemos cualquier número primo P_r

que esté en el intervalo $(n/3, n/2)$. Se deduce que para números inferiores, sólo se conecta al vértice del 1, mientras que para números mayores se conecta con sus múltiplos. Teniendo a lo sumo entre 0 y n a lo sumo un múltiplo, $2P_r$.

De ahí que la subred generada por los tres primos (se considera el 1 como primo) $1, P_r, 2P_r$, constituya un triángulo haciendo que el coeficiente de agrupamiento converja a 1 para cualquiera que sea el valor de P_r . Hay otra banda de números con coeficiente de agrupamiento exactamente 0 entre $n/2$ y n que son también números primos. Esto es, porque todos los primos en este rango están conectados sólo a 1 haciendo su agrupación 0.

Otro perfil de semejanza de estiramiento que se deriva de este conjunto de datos, es el que se logra mediante la iteración sucesiva y el análisis entre los coeficientes de clusterización, $\Delta c_i = c_{i+1} - c_i$:

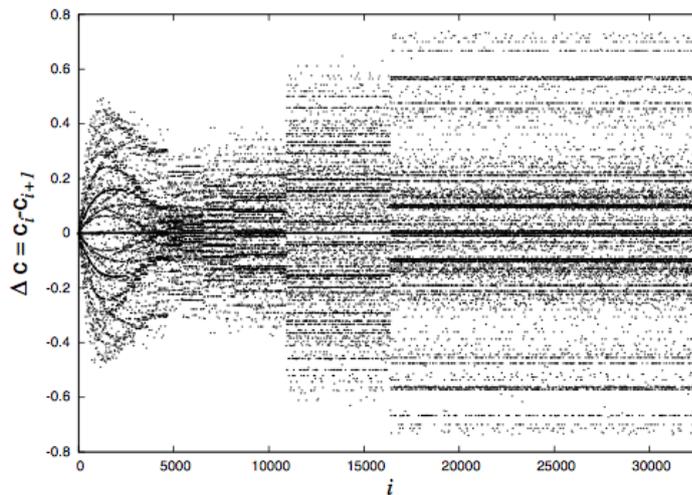


Figura 7.5: Diferencia entre coeficientes de agrupamientos de sucesivos de los vértices v_i e v_{i+1} en función del índice i . Fuente: Artículo. [7]

Veamos con detenimiento, el coeficiente de clusterización para los números compuestos que están entre $(n/2, n)$. Para un vértice i , los únicos vecinos son los divisores propios de i , es decir m tal que $1 \ll m < i$ e i

es divisible por m . Veamos esto:

Sea

$$m = \prod_{j=1}^k P_j^{i_j}$$

la factorización de n como el producto de potencias primas (distintas). El teorema fundamental de la aritmética afirma que tal factorización es única hasta un reordenamiento de los primos p'_j . Se puede observar que cada divisor m de n es de la forma:

$$m = \prod_{j=1}^k P_j^{l_j}$$

donde $0 \leq l_j \leq i_j$.

Los vértices $m = \prod_{j=1}^k P_j^{l_j}$ y $m' = \prod_{j=1}^k P_j^{l'_j}$ tal que $m < m'$ están conectados si, y sólo si, $l_j \leq l'_j$ para todo j . De este modo, el coeficiente de clusterización de i en la red que tiene una cantidad de nodos n viene dado por:

$$c_i = \binom{s-1}{2}^{-1} \left[\left(\sum_{j_1=0}^{i_1} \sum_{j_2=0}^{i_2} \cdots \sum_{j_k=0}^{i_k} (l_1+1)(l_2+1) \cdots (l_k+1) - 1 \right) - [s-1] \right]. \quad (7.1)$$

donde $s = (i_1+1)(i_2+1) \cdots (i_k+1)$, y por tanto

$$c_i = \binom{s-1}{2}^{-1} \left(\prod_{j=1}^k \binom{i_j+2}{2} - 2s + 1 \right) \quad (7.2)$$

Así, el valor de c_i depende del número de factores primos que son diferentes de i y de las potencias i_j derivados de la factorización de los primos de P_j , pero no de los primos que aparecen allí. De este modo, para cualesquiera valores dados de $i_1, i_2, i_3, \dots, i_k$, el valor del coeficientes de agrupamiento c_i permanece constante para todos los valores de i en el intervalo en cuestión: $n/2 < i \leq n$ tal que $i = \prod_{j=1}^k P_j^{l_j}$ para algún

conjunto de k primos distintos de los que aparecen en la factorización de i . Esto es lo que explica que el comportamiento de la gráfica de coeficientes de agrupamiento local mantenga una cota horizontal constante a partir de n grandes.

Este conjunto de datos, también exhibe una distribución de sus grados medios, que lógicamente, va a depender del tamaño de n y se puede calcular a partir de la expresión:

$$\frac{2|E|}{n} \quad (7.3)$$

donde $|E|$ es el número de aristas y n el número de vértices de la red. El valor de $|E|$ es igual a la suma de los elementos en la parte triangular inferior (o superior) de la matriz de adyacencia, siempre que $A_{ij} = 1$. En este caso, diremos que el elemento j es el divisor del primer elemento i si, y sólo si, $i > j$ y j divide a i . De esta idea se parte para llegar a la siguiente expresión que aproxima el valor de los grados medios para valores de n grandes, véase [7]:

$$2 \ln(n) - 1,6912 \quad (7.4)$$

Esto significa que el grado medio de la red crece logarítmicamente según el valor de n , según se muestra en la Figura 7.6:

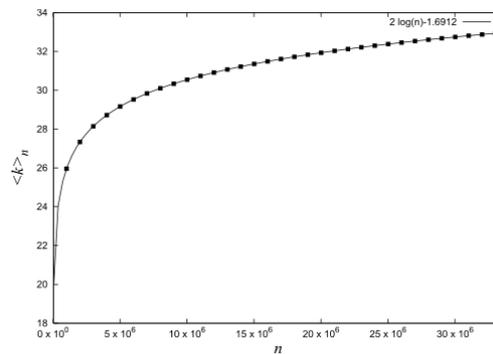


Figura 7.6: Distribución del grado medio en función del tamaño de la red n . Fuente [7].

Una medida alternativa para ver el nivel de clusterización de una red es el *coeficiente de agrupamiento global*. Esta medida cuantifica la densidad de tripletes cerrados en la red. Un triplete conectado (T_p) en la red es el conjunto de 3 nodos conectados entre sí con al menos 2 enlaces. Un triplete cerrado es el conjunto de 3 nodos conectados entre sí con exactamente 3 enlaces. Un triángulo (T_n) en la red cuenta como tres tripletes cerrados (uno centrado en cada nodo del triángulo). Dicho esto, el coeficiente de agrupación global de la red se define entonces como:

$$C = \frac{3 \times |T_n|}{|T_p| \text{ conectados}} \quad (7.5)$$

Si fijamos un vértice i , podemos estimar la cantidad de triángulos en la red en los que i es el vértice más pequeño:

$$T_n = \sum_{i=1}^n \sum_{k=2}^{\lfloor \frac{n}{i} \rfloor} \lfloor \frac{n}{ki} \rfloor \quad (7.6)$$

aproximando mediante integración la expresión anterior, se obtiene:

$$T_n \sim n \sum_{i=1}^n \int_{x=2}^{n/i} \frac{1}{ix} dx \sim n \sum_{i=1}^n \frac{1}{i} \left(\ln \frac{n}{i} - A \right) \quad (7.7)$$

La expresión anterior está acotada por:

$$n \sum_{i=1}^n \frac{1}{i} \left(\sqrt{\frac{n}{i}} - A \right) \sim n\sqrt{n} \int_{x=1}^n \frac{dx}{x^{3/2}} - An \int_{x=1}^n \frac{dx}{x} \sim Bn - An \ln(n) \quad (7.8)$$

donde A y B son constantes. Esto permite ver que la cantidad de triángulos en una red de este tipo de tamaño n viene estimada por el siguiente orden de magnitud

$$T_n \leq O(n) + O(n \ln(n)) + o(n^2) \quad (7.9)$$

y en particular se optiene que $T_n = o(n^2)$.

Sea $U(n)$ el número de tripletes conectados en la red después de del paso en tiempo $t = n$. Entonces, $U(n)$ viene dado por:

$$U(n) = \sum_{i=1}^n (d_i^2 - d_i). \quad (7.10)$$

donde d_i es el grado del vértice i . De este modo se puede expresar $U(n)$ en términos de los valores de la esperanza de las variables k^2 y k , siendo k el valor esperado como grado de un vértice en dicha red.

$$n\langle k^2 \rangle - n\langle k \rangle \sim O(n\langle k^2 \rangle) \quad (7.11)$$

En la Figura 7.3 hemos visto que la red sigue una distribución de potencia con $k^{-\gamma}$, con $\gamma = -2$. Así, la esperanza de k^2 se puede estimar como:

$$\langle k^2 \rangle = \sum_{i=1}^n k^2 p(k) \sim n, \quad (7.12)$$

y por tanto

$$U_n \sim n\langle k^2 \rangle = O(n^2) \quad (7.13)$$

Tomando las expresiones (7.5), (7.9) y (7.14) es fácil ver que el coeficiente global de agrupamiento decae a cero cuando $n \rightarrow \infty$. La comprobación numérica de estos resultados se puede hallar en la Figura 7.7:

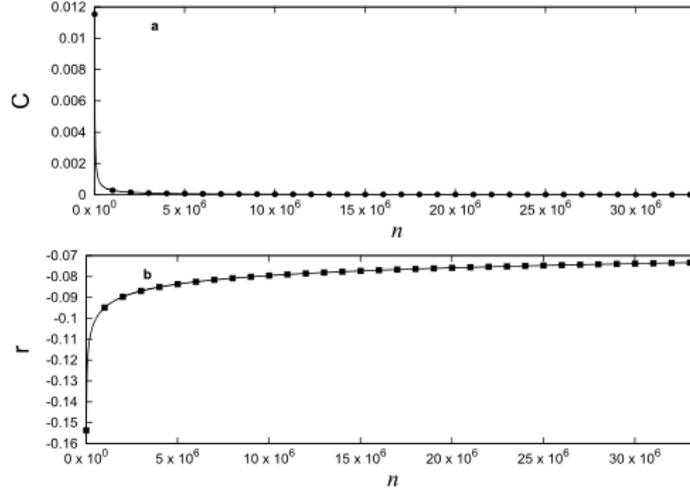


Figura 7.7: Coeficiente de clustering global y coeficiente de asortatividad en función del tamaño de la red. Fuente: [7].

En la Figura 7.7 **a**. El coeficiente de agrupamiento global disminuye a 0 a medida que $n \rightarrow \infty$.

El *coeficiente de asortividad* indica la *correlación de grados* en la red. Es un estadístico importante que tiende a cuantificar la estructura de la red. Por ejemplo, si en una red los vértices de grado alto tienden a conectarse a vértices de bajo grado (correlaciones de grado negativo), quiere decir que la red es *disortiva* en estructura, mientras que si los vértices de grado similar tienden a conectarse entre sí, la red es *asortativa*.

La mayoría de las redes reales, excepto las redes sociales, son disortivas y esto se ha explicado con el hecho de que el estado disortivo es el estado más probable de las redes libres de escala. El carácter *asortivo/disortivo* de las redes puede cuantificarse utilizando el coeficiente de asortatividad, que se calcula empleando la siguiente fórmula:

$$r = \frac{\sum_{1 \leq i < j \leq n} (A_{ji} - d_i d_j / 2|E|) k_i k_j}{\sum_{ij} (d_i \delta_{ij} - d_i d_j / 2|E|) d_i d_j} \quad (7.14)$$

donde d_i es el grado del vértice v_i , A_{ij} es el elemento (i, j) de la matriz

de adyacencia, $|E|$ es el total del número de aristas en la red y δ_{ij} es el delta de Kronecker. El resultado numérico y gráfico se representa en la Figura 7.7. El coeficiente de asortividad r , también parece llegar a 0 asintóticamente aunque siempre permanece negativo.

Con el propósito de tener una buena estimación del exponente que describe la caída que exhibe la distribución de los grados que siguen una ley de potencia, la pendiente γ se puede estimar utilizando el siguiente modelo:

$$\alpha = 1 + n \left[\sum_{i=1}^n \ln \frac{d_i}{d_i - \frac{1}{2}} \right]^{-1} \quad (7.15)$$

donde d_i son los grados de los vértices v_i tales que $i \geq i_{min}$, siendo i_{min} un valor a partir del cual la distribución de los grados sigue una ley de potencias. Dicho i_{min} se puede estimar a partir del estimador de Kolmogorov-Smirnov. En este ejemplo, para $n = 2^{15}$, el valor de α es aproximadamente -2 .

Capítulo 8

Conclusión

Nadie ha sugerido que la tarea sea simple, o que el punto de inicio más estratégico se identifique fácilmente, o que el analista será necesariamente capaz de comenzar en este punto aún si logra identificarlo, o que re-tendrá su carácter estratégico una vez que él haya arrancado en forma promisorio. Las situaciones cambian, los grupos se forman y se disuelven, las interrelaciones mutan; las redes permanecen. Lo que importa es comenzar.

Whitten y Wolfe (1973: 740).

En la presente memoria se han abordado básicamente los tres grandes tipos que aparecen en la Teoría de Redes, específicamente, las aleatorias, las de pequeño mundo y las libres de escala. Hemos visto cómo han surgido estas teorías, desde *Euler hasta Barabási y Albert*, y se han descritos sus elementos más importantes.

Se ha estudiado la importancia de las redes libres de escala en cuanto a la aplicación a la vida real: Aplicaciones en la biología, en la neurociencia, la informática, redes eléctricas, redes procedentes de la ingeniería de procesos dinámicos complejos, la WWW, las redes sociales, el cine, el estudio de la evolución de virus, que permiten la vida. Ello muestra que la Teoría de Redes está presente en todo sistema y subsistemas que podamos imaginar. Dentro de estas redes cabe destacar la importancia

de los *hubs*, que son unos tipos de nodos que tienen un alto grado de conexión, y que exhiben propiedades muy robustas, como es el caso de que si desconectamos el 80 % de los nodos de una red, sin tocar a éstos, la red quedaría aun conectada.

Por último, hemos estudiado la divisibilidad de los números naturales desde el punto de vista de la Teoría de Redes, viendo que incluso, los hubs no son estacionarios, exhibiendo otra propiedad interesante denominada, *semejanza de estiramiento* [7]. Dicha propiedad puede permitir el estudio de la evolución de ciertos microorganismos que presentan un comportamiento no estacionario.

Sin lugar a dudas, estamos en presencia de una de las ciencias más ricas y, que su producción científica crece de un año a otro, prácticamente siguiendo una ley de potencia.

Bibliografía

- [1] Noga Alon and Joel H Spencer. *The Probabilistic Method*. John Wiley & Sons, 2004.
- [2] Albert-Laszlo Barabasi. *Network Science*. Cambridge University Press, 2007.
- [3] Ithiel de Sola Pool and Manfred Kochen. Contacts and influence. *Social networks*, 1(1):5–51, 1978.
- [4] Donald Ervin Knuth. *The Stanford GraphBase: a platform for combinatorial computing*, volume 37. Addison-Wesley Reading, 1993.
- [5] Michael Krivelevich and Benny Sudakov. The phase transition in random graphs: A simple proof. *Random Structures & Algorithms*, 43(2):131–138, 2013.
- [6] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Física estadística*, volume 3. Reverté, 1988.
- [7] Snehal M Shekatkar, Chandrasheel Bhagwat, and G Ambika. Divisibility patterns of natural numbers on a complex network. *Scientific Reports*, 5, 2015.
- [8] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, 1998.