



UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ



FACULTAD DE CIENCIAS

**“ESTUDIO DE PROPIEDADES
TOPOLÓGICAS DE REDES COMPLEJAS”**

TESIS

PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS APLICADAS

PRESENTA:

PEDRO CASTILLO CASTILLO

DIRECTOR TESIS:

DR. J. JESÚS ACOSTA ELÍAS

SAN LUIS POTOSÍ, S.L.P.

DICIEMBRE 2021

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS

Yo, Pedro Castillo Castillo, estudiante del Posgrado en Ciencias Aplicadas de la Facultad de Ciencias de la Universidad Autónoma de San Luis Potosí, como autor/(a) de la tesis “ESTUDIO DE PROPIEDADES TOPOLÓGICAS DE REDES COMPLEJAS”, declaro que la tesis es una obra original, inédita, auténtica, personal, que se han citado las fuentes correspondientes y que en su ejecución se respetaron las disposiciones legales vigentes que protegen los derechos de autor y de propiedad intelectual e industrial. Las ideas, doctrinas, resultados y conclusiones a los que he llegado son de mi absoluta responsabilidad.

Agradecimientos.

Gracias a todas las personas que estuvieron alrededor de esta investigación.

Gracias a mi asesor el Dr. J. Jesús Acosta Elías por ese ánimo que tuvo siempre para llevar adelante este proyecto de Redes Complejas.

Gracias a mis sinodales, Dr. Raúl Eduardo Balderas Navarro, Dr. Marcela Mejía Carlos, Dr. Jose Salomé Murguía Ibarra, Dr. Isaac Campos Cantón, Dr. Eric Campos Cantón, que han estado al pendiente de mi investigación.

Gracias al Dr. Pedro David Arjona Villicaña, que a pesar de no ser sinodal, me ha estado apoyando en la investigación, y a todos los investigadores del IICO y de Computación, y compañeros estudiantes con quienes he convivido en las buenas y en las malas para sacar adelante este proyecto.

Gracias a mi familia, papás e hijos que han esperado con paciencia a que yo pueda terminar mi Tesis y pueda obtener el grado.

Gracias a la Universidad que proporcionó el espacio y el acceso al Internet de donde se ha podido consultar bastante información que ayudó a realizar todo este trabajo.

Gracias al CONACYT, por el apoyo económico que me entregó y sin el cual no hubiera iniciado ni concluido el doctorado.

Dedicatoria.

Se lo dedico al esfuerzo de mi asesor y sinodales, a mi esposa, a mis hijos, a mis hermanos, a mi madre y a mi padre que se encuentra en el cielo, por esperar a que termine este posgrado.

Resumen

Un grafo o red en su definición más simple es un conjunto de puntos y líneas [57]. El cual tiene propiedades topológicas como son: El diámetro del grafo, el coeficiente de agrupamiento y la distribución de conectividad.

El diámetro del grafo, es la distancia más larga de las distancias más cortas de un vértice origen a un destino. Un triángulo está formado por tres vértices interconectados, entonces el coeficiente de agrupamiento, es la sumatoria de todos los triángulos que tiene un grafo entre todos los triángulos posibles que puede tener el grafo. La distribución de probabilidad se obtiene de la conectividad que tiene cada vértice del grafo.

La teoría de grafos aleatorios de Erdős & Rényi, fue desarrollada en 1959, donde la teoría mostraba herramientas para grafos aleatorios sin crecimiento y con crecimiento, las cuales fueron de gran utilidad para hacer estudios más profundos sobre grafos y retomar algunos estudios de los que se quedaron inconclusos, también aparecieron las computadoras que ayudaron a que se procesaran más datos y con ello hacer estudios con mayor volumen de información, como fueron los estudios que se hicieron sobre el Internet, la www y cita de artículos a finales de los 90's, entonces se demostró que las redes reales no se comportaban como redes aleatorias, ya que mostraban otras propiedades que no tienen las redes aleatorias como son un diámetro corto, una distribución de conectividad en forma ley de potencia, etc.

Motivado por estos resultados, el Dr. Barabási en 1999, propone el primer modelo de crecimiento de red compleja, el cual es un modelo de red sencillo con crecimiento por agregación y conectividad preferencial, del cual se pudo obtener una distribución de conectividad en forma de ley de potencia y que sirvió de estímulo para que otros investigadores que se dieron a la tarea de proponer nuevos modelos donde se pudieran mostrar otras propiedades

de las redes complejas, y se acercaran a las redes reales, tales como el recableado, aptitud, atractivo inicial, etc.

Lo cual los ayudó a entender mejor a las redes, ya que hay redes reales con la propiedad de nacimiento y muerte de enlaces, como es el Internet, www, etc. y también redes reales sin nacimiento ni muerte de enlaces, como son la cita de artículos y la cita de patentes de Estados Unidos, y que, sin embargo, de las redes reales se puede obtener una distribución de conectividad ley de potencia de enlaces entrantes y de enlaces salientes.

Para poder contribuir con un granito de arena en estos estudios de redes, se estudiaron muchos modelos ya propuestos hasta ahora y se descubrió que muestran una distribución ley potencia de los enlaces entrantes, y que los modelos que muestran una distribución de los enlaces salientes tienen nacimiento y muerte de enlaces.

Se puede explicar como obtener una distribución ley de potencia en los enlaces entrantes debido a la conectividad preferencial, pero no se puede explicar como generar una distribución de enlaces salientes en forma de ley potencia.

Por lo tanto, se revisó el modelo de copiado de Krapivsky [47], ya que el copiado podría servir para obtener ese resultado, sin embargo, el resultado que se obtuvo fue una distribución exponencial de los enlaces salientes, también se revisó el modelo propuesto por Ameerah [2], el cual se basa en krapivsky [45] y si se obtiene la distribución ley de potencia en el outdegree, pero artificialmente se obtiene el número de enlaces salientes del nodo nuevo, lo cual no es natural.

De manera que, no existe un modelo de red sin nacimiento y muerte de enlaces, y que se pueda obtener una distribución en forma de ley de potencia de enlaces salientes y entrantes.

Por eso mi principal objetivo es proponer un modelo de crecimiento de red compleja donde se pueda obtener una distribución ley de potencia de enlaces entrantes y salientes, y que la red no acepte nacimiento y muerte de enlaces en nodos existentes.

Para esto, se creó un modelo de crecimiento de red compleja con probabilidad p , donde con probabilidad p , se selecciona un nodo aleatorio y se copian sus enlaces, y con probabilidad $1 - p$, se selecciona un número aleatorio de enlaces, el número aleatorio de enlaces está entre $1 : N$ nodos, se obtiene como resultado una distribución de enlaces entrantes en

forma de ley de potencia cuando la probabilidad $p \rightarrow 1$ y en los demás valores de p , una distribución de enlaces entrantes en forma exponencial.

Se analizó el grado de los nodos en este modelo, y se observó que se incrementa conforme aumenta el número total de nodos en la red, a este fenómeno se le conoce como aceleración [24] y también produce un efecto en la distribución de enlaces entrantes, en el que de una red 1000 a 10,000 nodos varía el exponente, por ejemplo en este modelo tenemos con probabilidad $p = 0.99$, en una red de 1000 nodos se obtiene una distribución de enlaces entrantes con exponente $\gamma_{in} \approx 3.3$ y con 10,000 nodos se obtiene un exponente $\gamma_{in} \approx 4.3$. Para poder comparar los resultados del modelo con redes reales, se revisó los datos de la red cita de patentes de los Estados Unidos (Respaldo de cita de patentes desde el año 1975 a 1999), en la cual cuando se crea una patente, se cita a patentes ya existentes, pero a una patente que ya existe no se le puede agregar o eliminar las citas que haya hecho a otras patentes, por esta razón, esta red no tiene nacimiento ni muerte de enlaces, se obtuvo su distribución de enlaces salientes para 10^6 nodos y para $2 * 10^6$ nodos y se pudo observar que varía el exponente de la distribución de enlaces salientes en forma de ley de potencia.

Abstract

A graph or network in its simplest definition is a set of points and lines [57]. Which has topological properties such as: The diameter of the graph, the clustering coefficient, and the connectivity distribution.

The diameter of the graph is the longest of the shortest distances from a source vertex to a target. A triangle is formed by three interconnected vertices, so the clustering coefficient is the sum of all the triangles that a graph has among all the possible triangles that the graph can have. The probability distribution is obtained from the connectivity that each vertex of the graph has.

Erdős & Rényi's theory of random graphs was developed in 1959, where the theory showed tools for random graphs without growth and with growth, which were very useful to do more in-depth studies on graphs and retake some studies from which They were left unfinished, computers also appeared that helped to process more data and thus carry out studies with a greater volume of information, how were the studies that were done on the Internet, the www and citation of articles at the end of the '90s, Then it was shown that real networks did not behave like random networks, since they showed other properties that random networks do not have, such as a short diameter, a connectivity distribution in the form of a power-law, etc.

Motivated by these results, Dr. Barabási in 1999, proposed the first complex network growth model, which is a simple network model with growth by aggregation and preferential connectivity, from which it was possible to obtain a connectivity distribution in the form of power-law and that served as a stimulus for other researchers who undertook the task of proposing new models where other properties of complex networks could be shown and approached real networks, such as rewiring, aptitude, initial attractiveness, etc.

Which helped them to better understand networks, since there are real networks with the property of birth and death of links, like is the Internet, www, etc. and also real networks without the birth or death of links, how the citation of articles and the citation of United States patents are, and that, nevertheless, from real networks a power-law connectivity distribution of indegree and outdegree can be obtained.

To contribute with a grain of sand in these network studies, many models already proposed so far were studied and it was discovered that they show a power-law distribution of the indegree, and that the models that show a distribution of the outdegree have birth and death of links.

It can be explained how to obtain a power-law distribution on indegree due to preferential connectivity, but it cannot be explained how to generate a power-law distribution of the outdegree.

Therefore, the copying model of Krapivsky [47] was revised, since copying could serve to obtain this result, however, the result obtained was an exponential distribution of the outdegree, the model proposed by Ameerah [2], which is based on krapivsky [45] and if the power-law distribution is obtained in the outdegree, but artificially the number of outgoing links of the new node is obtained, which does not it's natural.

Thus, there is no network model without birth and death of links, and a power-law distribution of outdegree and indegree links can be obtained.

That is why my main objective is to propose a complex network growth model where a power-law distribution of indegree and outdegree can be obtained, and that the network does not accept the birth and death of links in existing nodes.

For this, a complex network growth model with probability p was created, where with probability p , a random node is selected and its links are copied, and with probability $1 - p$, a random number of links, the random number of links is between $1 : N$ nodes, the result is a distribution of indegree in the form of a power-law when the probability $p \rightarrow 1$ and in the other values of p , a distribution of indegree is exponential.

The degree of nodes in this model was analyzed, and it was observed that it increases as the total number of nodes in the network increases. This phenomenon is known as acceleration [24] and it also produces an effect on the distribution of indegree, in which the exponent varies from a network 1000 to 10,000 nodes, for example in this model we have with probability $p = 0.99$, in a network of 1000 nodes a distribution of indegree is obtained with exponent $\gamma_{in} \approx 3.3$ and with 10,000 nodes we get an exponent $\gamma_{in} \approx 4.3$.

To be able to compare the results of the model with real networks, the data of the United States patent citation network was reviewed (Support of patent citation from the year

1975 to 1999), in which when a patent is created, it is cited to existing patents, but to a patent that already exists you cannot add or delete the citations you have made to other patents, for this reason, this network has no birth or death of links, its distribution of outdegree was obtained for 10^6 nodes and $2 * 10^6$ nodes and it was observed that the exponent of the outdegree distribution varies in the form of a power law.

Índice general

1. Introducción	1
1.1. Solución de Euler	2
1.2. Grafos	4
1.2.1. Subgrafos	5
1.2.2. Árboles	5
1.3. Propiedades Topológicas	6
1.3.1. Diámetro del Grafo	7
1.3.2. Coeficiente de agrupamiento	8
1.3.3. Intermediación	8
1.3.4. Distribución de conectividad	9
1.4. Grafos ó redes aleatorias	10
1.4.1. Modelo de grafo aleatorio de Erdős-Rényi	10
1.5. Redes complejas reales	12
1.5.1. Redes Tecnológicas	13
1.5.2. Redes Sociales	15
1.5.3. Redes Biológicas	18
1.5.4. Redes de Información	20
2. Principales estudios de fenómenos y redes reales	26
2.1. El principio o ley de Pareto	26
2.2. Estudio de géneros y Especies del Dr. Willis	27
2.3. Estudio de Zipf	27
2.4. Estudio de Solla-Price	29

2.5. Estudio sobre el Internet.	29
2.6. Estudio del diámetro de la WWW	29
2.7. Estudio de las publicaciones de Physical Review D, de 1981 a 1997	30
2.8. Distribución de algunas Redes Reales	30
3. Principales modelos ya propuestos	32
3.1. Modelo de Yule	33
3.2. Modelo de red aleatoria Mundo Pequeño	33
3.3. Conectividad Preferencial	34
3.4. Modelo Barabási-Albert (BA)	34
3.5. Un modelo general de crecimiento geométrico para web libre de escalas pseudofractal	36
3.6. Modelo de Conectividad de redes aleatorias en crecimiento	38
3.7. Modelo de Organización de Crecimiento de redes aleatorias	38
3.8. Modelo de aptitud Bianconi-Barabási	40
3.9. Topología de redes en evolución: Eventos locales y universalidad	40
3.10. Modelo de estructura de crecimiento de red con atractivo inicial.	41
3.11. Modelo de edad de Dorogovtsev	42
3.12. Efecto del crecimiento acelerado en la estructura de las redes de comunica- ciones	42
3.13. Crecimiento de redes libres de escala en el mundo pequeño con coeficiente de variante ajustable	44
3.14. Modelo de crecimiento de red compleja con distribuciones de motivos triada ajustables	44
3.15. Modelo de grafos dirigidos libres de escala.	45
3.16. Modelo de Crecimiento de redes aleatorias por copiado de Krapivsky	45
3.17. Modelo de crecimiento de red con Superjoiners	46
3.18. Modelo de Esquivel	47
4. Problema y Objetivo	49

5. Modelo de crecimiento red propuesto.	51
5.1. Introducción	51
5.2. Modelo de crecimiento	52
5.3. Experimentos y resultados	53
5.4. Patentes de Estados Unidos de América y su γ_{out}	56
5.4.1. ¿Las redes complejas son libres de escala?	58
6. Discusión	60
7. Trabajo Futuro	62
8. Conclusiones.	63

Capítulo 1

Introducción

La migración colectiva de la mariposa Monarca, las Epidemias, el Internet, son algunos de los muchos fenómenos que se pueden observar en la naturaleza y que se les puede considerar como sistemas complejos. El estudio de los sistemas complejos ha generado un gran debate que revoluciona los cimientos científicos, principalmente por la presencia ubicua de propiedades “complejas”, en sistemas de naturaleza completamente distinta.

Una de las formas de ver un sistema complejo es a través de su red a las que se les llama redes complejas, las cuales son la columna vertebral de los sistemas complejos.

Una red está estructurada en su forma más simple por nodos y enlaces [57], uno mismo, puede ser un nodo entrante de un salón de clases donde se convive con los compañeros de clase y maestros, como también se convive con otros grupos de alumnos y maestros de la facultad, formando así una red social.

Conforme se interactúa con las demás personas se va formando enlaces en los que posiblemente sea un vértice con pocos enlaces o si se tiene facilidad de palabra y sea sobresaliente en el estudio, entonces se podrá ser un nodo con muchos enlaces, en una sociedad compleja como la que se vive, se puede generar enlaces diariamente y también perder enlaces diariamente porque posiblemente se está cambiando de ambiente continuamente buscando tal vez vivir mejor o se está siendo afectado por el mundo que rodea, que podría ser el cambio de trabajo, el cambio de ser estudiante y pasar a ser un trabajador, posiblemente se piense que los estudiantes forman una sociedad y los trabajadores otra, pero en algunas circunstancias se interactúa como estudiante y trabajador, y en algunos casos como

jubilado, entonces estás tres sociedades funcionan como tres sistemas sociales complejos, pero forman de alguna manera un componente gigante que sería un sistema complejo de la sociedad.

Aquí trabajando en el escritorio, se es un nodo del sistema social de estudiantes, se está usando una computadora conectada a Internet, la cual es un nodo conectado a una red tecnológica como es el Internet, en el hogar y en el trabajo se hace uso de varios servicios como el Correo electrónico, la Energía eléctrica, el Internet, etc. En este tipo de comunicación con la sociedad se hace uso de herramientas como un idioma, el cual tiene un conjunto de palabras más usadas y otras menos usadas, también existen idiomas que son muy usados como el inglés y el español, e idiomas pocos usados como es el maya.

La gente se comunica en diferentes idiomas, formando enlaces de palabras de un mismo idioma o mezclando palabras de varios idiomas, eso dio como resultado el nacimiento de un nuevo idioma como es el español que se formó de varias lenguas en la época medieval. Ese mismo proceso se vive ahora donde el individuo usa palabras en inglés y en español para comunicarse con la sociedad y lo que puede dar en un futuro como resultado el nacimiento del idioma spanglish que se menciona muy a menudo en palabras coloquiales. La frecuencia de palabras que utilizamos en el idioma inglés no es una red, sin embargo, se puede obtener una distribución de la frecuencia de palabras [6].

En el mundo suceden fenómenos como los terremotos, muy pocos de alta intensidad y muchos de baja intensidad, si se hiciera un estudio sobre los terremotos se podría obtener una distribución de probabilidad que decaería como una ley de potencia, debido a que son muy pocos de alta intensidad y la mayoría de baja intensidad.

1.1. Solución de Euler

Hoy en día se puede observar redes complejas por dondequiera, pero antes del problema de Euler no se tenía idea de que era una red, puesto que ni siquiera se conocía el concepto de grafo.

Por lo tanto, se puede decir que el estudio de las redes empieza con la solución que dio Euler al problema de los puentes de Konigsberg en 1736, en la que desarrolla un grafo

donde indica que la cantidad de puentes usados para salir de los sitios, ha de ser la misma que la de ser usados para entrar, lo que quiere decir que la cantidad de puentes que inciden en el de cualquier sector de paso ha de ser par, y los puentes que tenía el río pregel en algunos casos no era par, de manera que, no hubo solución.

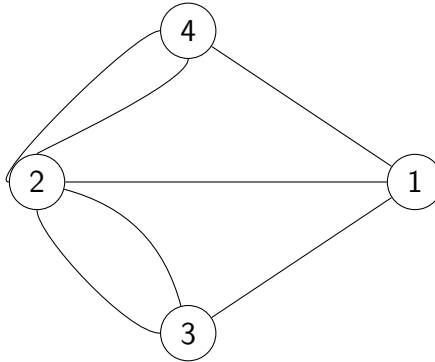


Figura 1.1: Grafo que representa los puentes y el río pregel en Konishberg, las líneas de color negro representan los puentes

El problema era, como visitar todos los sitios de Konishberg sin pasar más de una vez por el mismo puente y llegar al mismo punto de partida.

Este modelo creado por Euler que más tarde lo llamaron un grafo de rejilla, fue usado en los años siguientes para dar solución a problemas que se podían resolver con grafos de rejilla.

Como por ejemplo en 1847, Gustav Kirchhoff [43] utilizó la teoría de grafos para publicar sus leyes de los circuitos para calcular el voltaje y la corriente en los circuitos eléctricos.

La teoría de grafos fue mejorada y ampliada por los Dres. Erdős y Rényi. En la cual se definen varios conceptos de grafos como son los subgrafos, árboles, caminos entre vértices de los grafos, grado de conectividad de cada vértice, y un concepto muy importante para aquella época fue el descubrimiento de los grafos aleatorios, de los cuales en seguida se describen las definiciones de grafo y sus propiedades.

1.2. Grafos

Un grafo se define en su forma más simple como un conjunto de puntos unidos por líneas de acuerdo con Mark Newman [57] .

Una red se representa matemáticamente como un grafo, de acuerdo con Mark Newman, en un grafo en su forma más simple, un punto es un vértice y una línea es una arista.

Un grafo $G = (V, E)$ formalmente consiste en un conjunto de vértices V y en un conjunto de enlaces E entre ellos. Un enlace e_{ij} conecta dos vértices i y j .

La vecindad de vértices N para un vértice v_i se define como aquellos vértices inmediatamente conectados de tal modo que:

$$N_i = \{v_j\} : e_{ij} \in E \vee e_{ji} \in E.$$

El grado, que se representa como k_i de un vértice, k es definido como el número de vértices enlazados con uno dado i . De acuerdo con la figura 1.2 se puede decir que el vértice 2 tiene $k_i = 3$ y $k_5 = 1$, En esta expresión además se tiene que $|N_i|$.

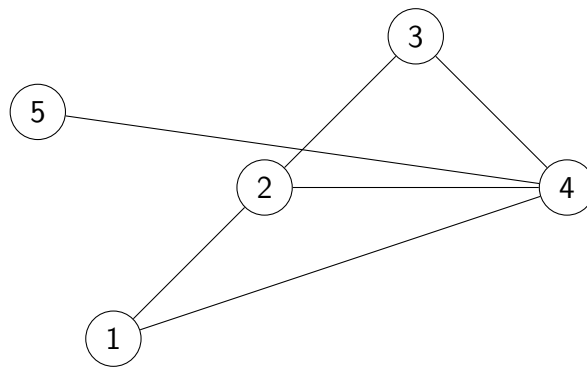


Figura 1.2: Grafo Ejemplo: Se muestra un grafo de 5 vértices y 6 aristas, los círculos representan los vértices y las líneas rectas representan las aristas.

1.2.1. Subgrafos

Un subgrafo es un subconjunto de vértices y aristas de un grafo, se puede considerar que el componente gigante de un grafo es un subgrafo, ver figura 1.3.

Un grafo G_2 que consiste en un conjunto P_2 de vértices y un conjunto E_2 de los enlaces es un subgrafo de un grafo $G_1 = \{P_1, E_1\}$ si todos los vértices P_2 también son vértices de P_1 y todos los enlaces en E_2 también son enlaces de E_1 . Los ejemplos más simples de subgrafos son ciclos, árboles y subgrafos completos (véase figuras. 1.3, 1.4. Un ciclo de orden k es un bucle cerrado de k enlaces tal que cada dos enlaces consecutivos, solo tienen un vértice en común. Una manera de imaginar gráficamente un ciclo es dibujando un polígono: un triángulo es un ciclo de orden 3, mientras que un rectángulo es un ciclo de orden 4. El grado promedio de un ciclo es igual a 2, puesto que cada vértice tiene dos enlaces.

Los subgrafos completos del orden k , contienen k vértices y todos los posibles $k(k-1)/2$ enlaces, en otras palabras, están completamente conectados.

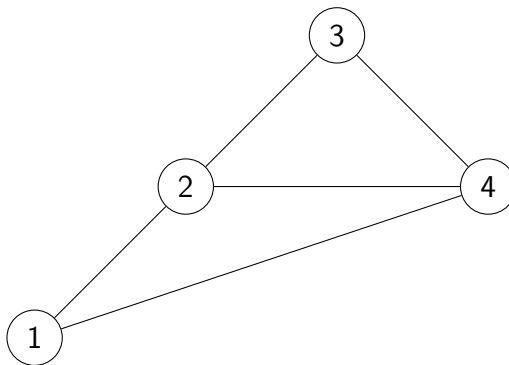


Figura 1.3: Subgrafo Ejemplo: Se muestra un grafo de 4 vértices y 5 aristas, se puede decir que es un subgrafo del grafo de la figura 1.2.

1.2.2. Árboles

Un árbol es un grafo conectado y no dirigido que no contiene bucles cerrados [57]. Por “conectado” se quiere decir que todos los vértices del grafo son accesibles desde todos los demás a través de algún camino del grafo. Un grafo también puede constar de dos o

más partes, desconectados entre sí, y si una parte individual no tiene bucles, también se denomina árbol. Si todas las partes del grafo son árboles, el grafo completo se denomina bosque.

Se puede considerar que una de las propiedades más importantes de los árboles es que no tienen ciclos.

Un árbol de n vértices tiene $n - 1$ enlaces, lo cual muestra que todos los n vértices están conectados.

El grado promedio de un árbol de orden k es $\langle k \rangle = 2 - 2/k$, acercándose a 2, para un árbol muy grande.

Un grafo con enlaces dirigidos y sin ciclos se considera que es un árbol dirigido, el cual puede tener más de $n - 1$ aristas.

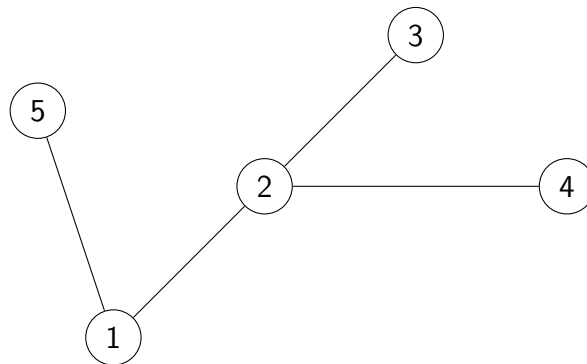


Figura 1.4: Árbol Ejemplo: Se muestra un grafo de 5 vértices y 4 aristas, se puede decir que es un árbol debido a que no tiene ciclos. 1.2.

1.3. Propiedades Topológicas

Los grafos tienen propiedades topológicas como son, la distribución de conectividad, el diámetro de red, el coeficiente de agrupamiento, y otras propiedades. Una de las propiedades topológicas más importantes en la que se hará especial énfasis tanto en el trabajo relacionado como en los resultados es **la distribución de conectividad**, debido a que tanto en los estudios de las redes reales como de los modelos que se hablará más adelante, lo primero que se observa, como resultado es su pendiente de la distribución de

conectividad.

A través de las propiedades se puede observar los procesos microscópicos por ejemplo cuando un usuario conecta su computadora a Internet a través de un proveedor, la propiedad macroscópica puede ser la distribución de enlaces entrantes.

1.3.1. Diámetro del Grafo

Diámetro del grafo (D), se define como el máximo de las distancias (d) más cortas entre cualesquiera de dos vértices del grafo es decir:

$$D = \max_{i,j \neq i} d(n_i, n_j) \tag{1.1}$$

Donde D es el diámetro del grafo, $d(n_i, n_j)$ es la distancia del vértice n_i al vértice n_j del grafo de la Figura 1.5.

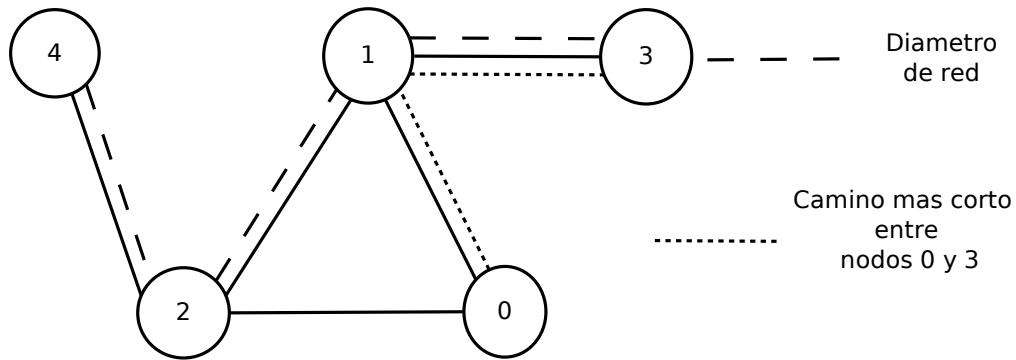


Figura 1.5: La ruta más corta es la ruta con el menor número de enlaces (saltos) entre dos vértices. Por ejemplo, la ruta más corta entre los nodos 0 y 3 es 2. El diámetro en un grafo o red es la ruta más corta de la distancia más larga entre los dos vértices más lejanos. El diámetro de este grafo es 3, porque la ruta más corta entre los dos nodos más lejanos es 3.

De acuerdo con la tabla 2.2, se puede observar que el diámetro de la red de Cita de

Patentes es de 22, y que el diámetro más pequeño que se puede observar en esa tabla es el de la red wiki-vote con 7.

Como aplicación de este concepto Milgram [55] realizó un experimento de envío de cartas. Cuando Milgram preguntó a un amigo cuántos intercambios de cartas creía que serían necesarios para completar el trayecto a través de la red social, este respondió que tal vez “unos cien” sería una buena estimación. Sin embargo, cuando el experimento se hubo completado (con algunas cartas perdidas de por medio), el resultado fue sorprendente: las cartas sólo habían necesitado unos cinco intercambios para alcanzar al desconocido de Boston. Cinco individuos intermedios, lo que equivalía a seis conexiones en la red o “seis grados de separación”[38]. El experimento había sido lanzado desde la Universidad de Harvard en Cambridge Masachusets. Es decir, el diámetro de red entre la población mundial es sólo seis conexiones de la red.

1.3.2. Coeficiente de agrupamiento

Se define el coeficiente de agrupamiento como la fracción promedio de pares de vecinos de un vértice que también son vecinos entre sí.

En un clique en el que la conectividad es máxima. Para un grafo dirigido, e_{ij} es distinto de e_{ji} , y por lo tanto, para cada vecino N_i hay $k_i(k_i - 1)$ enlaces que podrían existir entre los vértices de la vecindad (k_i es el grado del vértice i para el total (entrantes + salientes)).

De esta forma el grado de agrupamiento en los grafos dirigidos está dado por:

$$C_i = \frac{|e_{ij}|}{k_i(k_i-1)} : v_j, v_k \in N_i, e_{jk} \in E.$$

1.3.3. Intermediación

Se define grado de intermediación como la cantidad de enlaces que pasan en un vértice intermedio del grafo, de un vértice origen a un vértice destino, por ejemplo la computadora está en casa, conectada a través del Internet a una página web en Alemania, el vértice intermedio sería a un switch de alta velocidad en Houston que reproduce la conexión hacia

Alemania, existen muchos enlaces que atraviesan un vértice de intermediación, así como hay pocos enlaces que atraviesan un vértice intermedio.

La medida de intermediación de un vértice se obtiene al contar las veces que este aparece en los caminos cortos que conectan a todos los pares de vértices de la Red, a estos vértices se les llama vértices puente.

1.3.4. Distribución de conectividad

El grado de un vértice i en un grafo es el número de conexiones del vértice i con los otros vértices, y también se le conoce como conectividad.

El grado de los vértices en redes se distribuyen estadísticamente, por lo tanto para cada vértice, se introduce la probabilidad $p(k, i)$ de que un vértice i en el grafo de tamaño N tiene k conexiones, de manera que la distribución de grado $P(k)$ es la sumatoria de $p(k, i)$ en una red aleatoria entre el número total de vértices N .

$$P(k) = \frac{1}{N} \sum_{i=1}^N p(k, i) \quad (1.2)$$

Si todos los vértices de un grafo aleatorio son estadísticamente equivalentes, cada uno de ellos tendría el mismo grado de distribución de probabilidad $p(k)$.

Entonces para encontrar el grado medio de cada vértice (\bar{k}) es $\bar{k} = \sum_{i=1}^N k_i P(k)$.

De los grafos sin crecimiento, que son los modelados por Erdős y Rényi (ER), se obtiene una distribución en forma de Poisson, Erdős y Rényi mostraban un conjunto de vértices de los cuales cada vértice tenía la capacidad de enlazarse en forma aleatoria a otro vértice, de lo cual se podía obtener una distribución de red aleatoria, la cual se muestra de la forma:

$$P(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!} \quad (1.3)$$

Que es una distribución de probabilidad tipo poisson.

Cuando se genera el crecimiento de un grafo aleatorio en la que el vértice nuevo se enlaza en forma aleatoria a vértices existentes de la red, se obtiene una distribución en forma

exponencial, la cual se representa matemáticamente:

$$P(k) \propto e^{-k/\bar{k}} \quad (1.4)$$

1.4. Grafos ó redes aleatorias

El estudio sistemático de los grafos aleatorios fue iniciado en 1959 por dos matemáticos Erdős y Rényi (ER) [29], con el propósito original de estudiar mediante métodos probabilísticos, las propiedades de los grafos en función de un incremento aleatorio del número de conexiones en el grafo, logrando con esto un gran avance en la teoría matemática clásica de grafos que revolucionaría la forma en que se pueden modelar estos problemas, describiendo una red con topología aleatoria por medio de un grafo aleatorio, estableciendo los fundamentos de la teoría de redes aleatorias.

De acuerdo con Erdős y Rényi [29], Un grafo aleatorio es aquel en que el grado de cada vértice es una variable aleatoria $k_i \in [0, N]$.

El incremento de vértices en cada paso del tiempo en un grafo aleatorio, es una evolución del grafo [29].

A continuación se hace una breve explicación de como se obtiene la distribución de un grafo aleatorio

1.4.1. Modelo de grafo aleatorio de Erdős-Rényi

El modelo de grafo aleatorio fue introducido por Erdős-rényi en el año 1959 [29], en el cual se toman n vértices aislados, sin agregación de más vértices, se asigna una probabilidad p , para cada vértice origen, y con probabilidad p , el vértice origen se enlaza a $n - 1$ vértices destino existentes en la grafo, se generan enlaces de todos los vértices como vértices origen. Si la probabilidad $p = 1$, entonces al final del proceso de enlazamiento, se puede observar que cada uno de los vértices tiene $n - 1$ enlaces ver Figura 1.6.

Para un segundo ejemplo, se crea un algoritmo con las reglas de los grafos aleatorios descritas al principio, dado 10,000 vértices aislados y cada vértice con probabilidad $p = 1$,

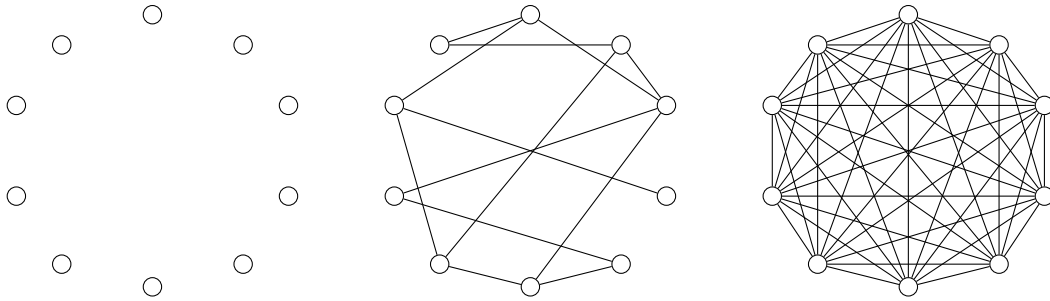


Figura 1.6: En la figura de la izquierda se tienen 10 vértices aislados, y se enlazan entre ellos, con probabilidad $p = 1$, en la figura de la derecha se tienen $n(n - 1)$ enlaces =90 enlaces.

se enlaza con un vértice al azar de los 10,000-1 vértices destino existentes, y se obtiene la distribución de un grafo aleatorio en forma poisson, ver Figura 1.7,

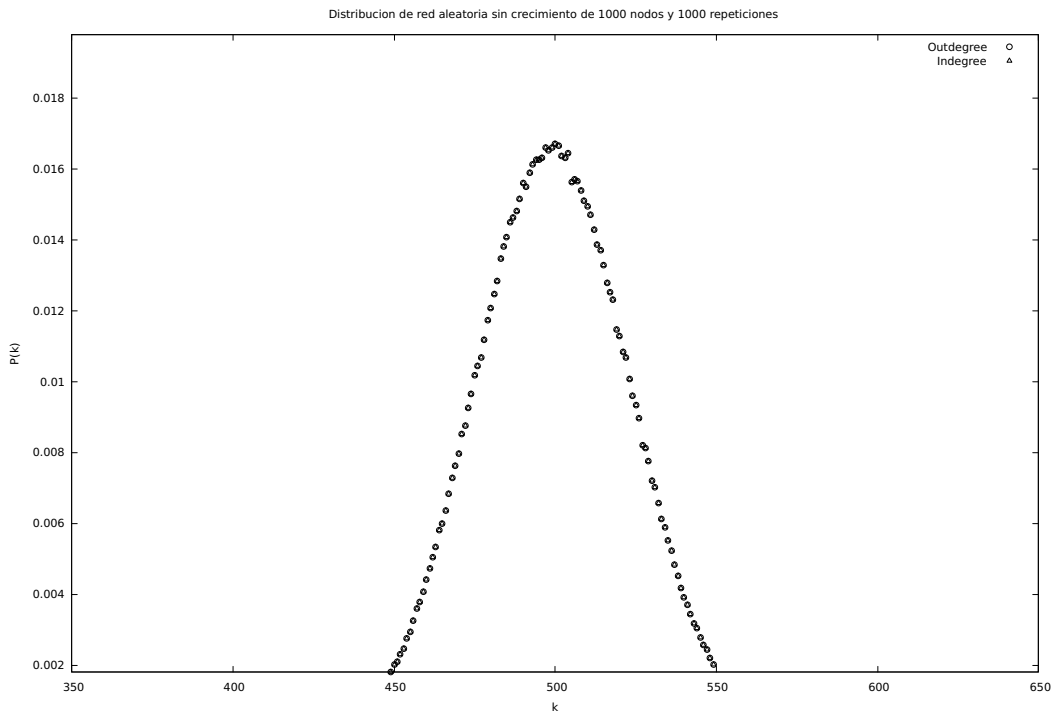


Figura 1.7: Distribución de un grafo aleatorio con 1000 vértices, sin crecimiento y ejes lineales.

Para un tercer ejemplo se crea un algoritmo con las reglas de grafos aleatorios descritos

al principio, se hace el crecimiento de un grafo hasta 10,000 vértices y en cada paso del tiempo, un vértice nuevo se enlaza con un vértice al azar de los $n - 1$ vértices destino existentes con probabilidad $p = 1$, se procesaron los datos y se obtuvo la distribución de un grafo aleatorio en forma exponencial, ver Figura 1.8,

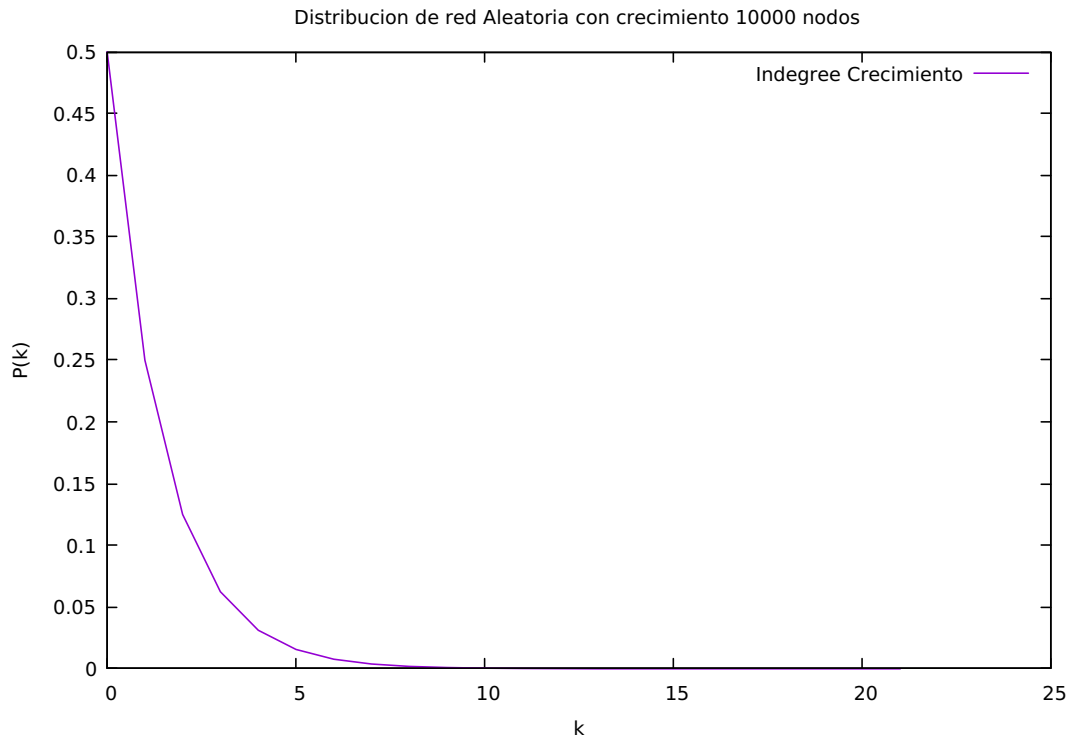


Figura 1.8: Distribución de enlaces entrantes de un grafo aleatorio con 10,000 vértices, grafo con crecimiento y ejes normales.

1.5. Redes complejas reales

Las redes son representadas matemáticamente a través de grafos, y se toman en cuenta las definiciones de grafo como definiciones de red, como es el caso del diámetro del grafo que será diámetro de red. En matemáticas, los grafos tienen vértices y aristas, en el caso de redes se tiene nodos y enlaces, por ejemplo el **enlace** terrestre de una ciudad a otra es una carretera.

En esta sección se verán algunas estructuras de redes reales de diferentes tipos. Existen redes reales naturales como las redes biológicas, y también redes reales fabricadas por el hombre como es el caso del Internet, Cita de Artículos, WWW, etc. Enseguida se describen algunas redes complejas de acuerdo a su clasificación.

1.5.1. Redes Tecnológicas

La red Internet

Generalmente la topología de Internet se estudia a escala de sistemas autónomos (AS por sus siglas en inglés).

Un AS es gestionado por una autoridad administrativa única, al interior, el encaminamiento se lleva a cabo usando algoritmos de enrutamiento como OSPF o RIP. Los AS's se comunican entre sí usando BGP (Border Gateway Protocol). Los AS's pueden ser de diferentes tamaños, por ejemplo la Universidad Autónoma de San Luis Potosí, es un sistema autónomo relativamente pequeño comparado con el AS Uninet S.A. de CV; una empresa que se originó en Teléfonos de México para proveer el servicio de Internet a usuarios finales en nuestro País, Ver Figura 1.9.

El primer estudio que se tiene acerca del Internet es el hecho por [34].

Red de energía eléctrica

La red de energía eléctrica está formada por las líneas de transmisión de alta tensión que proporcionan transporte de energía eléctrica a larga distancia, hay enlaces entrantes y salientes de los países debido a que hay países que producen su propia energía y pueden exportar la sobrante, y también hay países que necesitan más energía de la que producen, entonces ellos la compran.

Las líneas de suministro de energía locales de baja tensión son normalmente excluidas, ver Figura 1.10.

Los nodos en una red de energía corresponden a las estaciones de generación y conmutación de subestaciones y los enlaces corresponden a las líneas de alta tensión. La topología de las redes eléctricas no es difícil de determinar.

Las redes suelen ser supervisadas por una autoridad única y los mapas completos de las

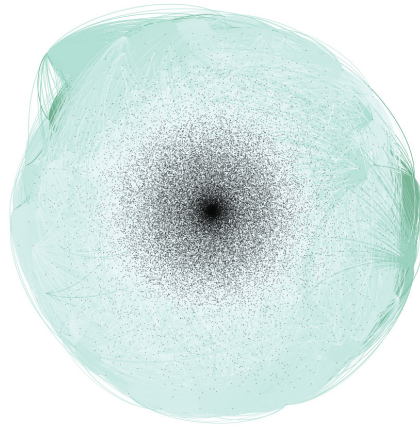


Figura 1.9: Estructura de Internet a nivel de sistemas autónomos. Datos proporcionados por CAIDA (The Cooperative Association for Internet Data Analysis), reconstruido a partir de las tablas BGP.

redes están disponibles, de hecho, son datos muy completos sobre las redes de energía (así como otras redes relacionadas con la energía, tales como oleoductos y gasoductos) que se encuentran disponibles en editoriales especializadas, ya sea en papel o en formato electrónico.

El estudio sobre las red eléctrica fue hecho por Watts-strogatz [73].

Red de telefonía Fija

La red de telefonía fija está formada por dispositivos de teléfono que hay en casa o empresa, y los conmutadores que reciben las llamadas y las retransmiten a los teléfonos destino, donde la llamada y el alambre sirven de conexión de un teléfono origen a un destino, esta tecnología fue creada antes del siglo XX, y se fue extendiendo a todo el mundo, que ahora el mundo está conectado vía telefónica.

La red de telefonía celular

La red de telefonía celular es una red de teléfonos inalámbrica que está formada por cada teléfono celular, cada estación base o cada antena de retransmisión, las cuales se consideran que son los nodos y enlaces son las llamadas, hoy en día, un teléfono celular

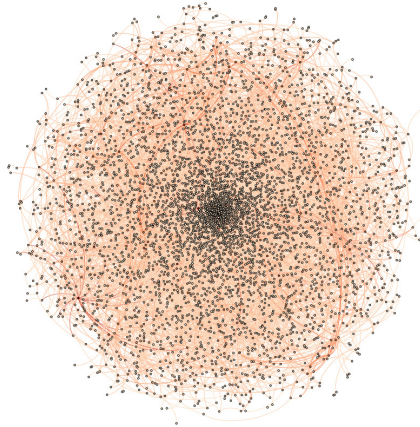


Figura 1.10: Red no dirigida y no ponderada que representa la topología de la red eléctrica de los Estados del Oeste de los Estados Unidos. Los datos fueron recopilados por D. Watts y S. Strogatz y puesto a disposición en la web.[73]

también puede ser un nodo de Internet, debido a que se pueden enviar o recibir mensajes. La telefonía celular está conectada con la fija, ya que se puede hacer una llamada desde su dispositivo móvil a cualquier teléfono fijo, desde esta manera se podría decir la red de telefonía fija y celular están enlazadas formando una red de telefonía.

1.5.2. Redes Sociales

Una red social es un conjunto de personas o grupos de personas con algún patrón de contactos o interacciones entre ellos [68], entre las redes sociales, se puede mencionar las redes de amigos, la red de estudiantes, la red de trabajadores, la red de personajes que aparecen en una novela etc.

A continuación se hará una breve descripción de algunas de ellas, para una mayor claridad.

Red de Amigos

Un grupo de amigos (nodos) están en contacto por su amistad (enlaces) y por diferentes medios tienen comunicación, generan puntos de reunión ya sea para compartir conoci-

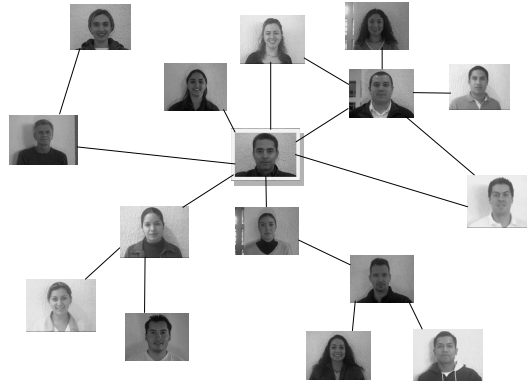


Figura 1.11: Las redes de amistades es una red social no dirigida debido a que A es amigo de B, entonces B es amigo de A.

mientos, alegrías, tristezas etc. Lo cual hace que se mantenga su amistad (enlace).

Se puede considerar que es una red no dirigida debido a que si A es amigo de B, entonces B es amigo de A. Ver figura 1.11

Red de trabajadores

Existe una relación de trabajo (enlace) entre el obrero y el supervisor (nodos).



Figura 1.12: Red de trabajadores.

O la relación que existe trabajador a trabajador ya sea que comparte las mismas herramientas o forma parte de una línea de producción en la que el producto en proceso pase

por un trabajador y luego por el siguiente trabajador ver Figura 1.12.

Red de la novela *Les Misérables*.

En la figura 1.13 se puede ver una red de la co-aparición de los personajes de la novela *Los Miserables*, donde los nodos corresponden a los personajes y los enlaces a la aparición de los personajes en la novela.

Después de realizar un análisis visual sólo se puede concluir que la gráfica tiene nueve grupos. Esto es probablemente un poco de la información que el creador de la novela quería transmitir.

La gráfica no tiene ninguna información relacionada con el número de nodos explorados, o lo que representan los grupos y el número de elementos que tiene cada grupo, etc.

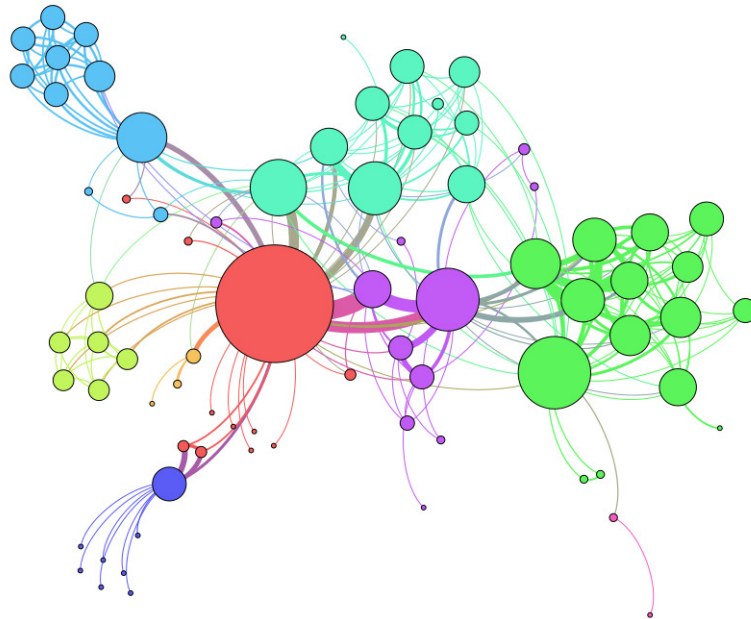


Figura 1.13: Red de la aparición de los personajes en la novela *Les Misérables*. [44]

Correo Electrónico de Enron.

El conjunto de datos de correo electrónico de Enron fue recogido y preparado por el Proyecto CALO (Un asistente cognitivo que aprende y organiza). Contiene datos de cerca

de 150 usuarios, en su mayoría altos directivos de Enron, organizados en carpetas. Contiene aproximadamente medio millón de mensajes. Estos datos fueron originalmente hechos públicos en la web de la Comisión Federal Reguladora de Energía durante su investigación. Esta información es valiosa, es la única colección importante de correo electrónico real que es público. La razón de que otros conjuntos de datos no son públicos se debe a preocupaciones sobre la privacidad.

Los cerca de 150 usuarios del correo electrónico son los nodos y los mensajes son los enlaces, Ver Figura 1.14.

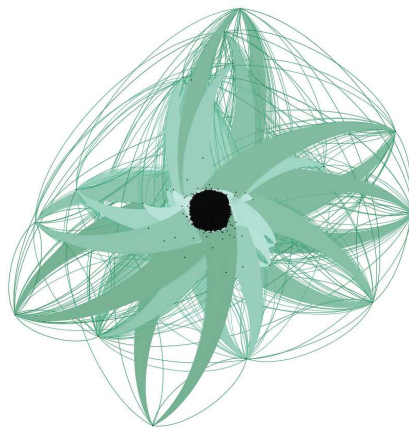


Figura 1.14: Red del correo electrónico de Enron. Esta formada por los mensajes que se enviaban entre sí, los ejecutivos de la empresa en los estados unidos.[53]

1.5.3. Redes Biológicas

Algunos sistemas biológicos se pueden representar como redes, como son las redes ecológicas, redes de iteración proteína-proteína, etc. Enseguida se describen algunas de ellas.

Redes Ecológicas

En las redes ecológicas, tales como depredador-presa, la presa (nodo) en defensa activa puede tener algunas de las reacciones que utilizan los organismos ante la presencia de un depredador (nodo). Como ejemplos se pueden citar el escapar, amenazar, contra-atacar,



Figura 1.15: Red depredador-presa, En la primer caja está una serpiente(depredador) y un ratón (presa), La serpiente se come al ratón.

moverse a escondidas, o construir y habitar una madriguera con túneles de salida de emergencia [7] [67]. El depredador se come (enlace) a la presa, ver Figura 1.15.

Red de Interacción proteína-proteína

La red de interacción proteína-proteína es una red metabólica que actúa produciendo reacciones celulares, donde las proteínas son los nodos y la interacción proteína-proteína son los enlaces.

Es importante desarrollar enfoques confiables de proteoma (totalidad de proteínas en una célula), para una mejor comprensión de funciones de la proteína. Enfoques genómicos se han utilizado para predecir las funciones de un gran número de genes sobre la base de sus secuencias. Sin embargo, como se sabe, las proteínas rara vez actúan solas a nivel bioquímico, sino que interactúan con otras proteínas como un conjunto para llevar a cabo determinadas tareas celulares. Sus funciones son sistémicas; juntas son más que la suma de sus partes [28].

Tradicionalmente, las interacciones de proteínas se estudiaron individualmente por técnicas genéticas, bioquímicas y biofísicas centradas en unas pocas proteínas a la vez. Cada

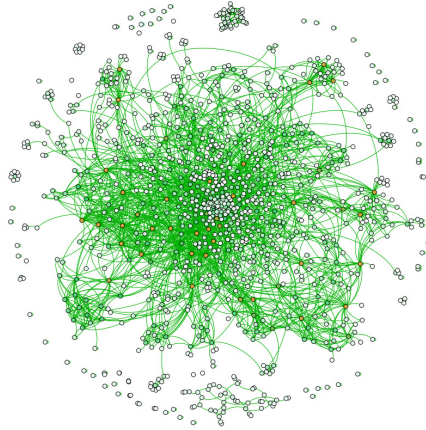


Figura 1.16: Red de proteínas. Las proteínas son macromoléculas biológicas formadas por unidades denominadas aminoácidos y juegan un papel muy importante en el funcionamiento de cualquier organismo. Los métodos de detección de interacciones han llevado al descubrimiento de miles de interacciones entre proteínas. La relevancia que poseen los conjuntos de datos a gran escala es importante para la biología actual.[65]

vez más se da uno cuenta de que la disección de la circuitería genética y bioquímica de una célula nos impide la comprensión de los procesos biológicos como un todo.

Las interacciones proteína-proteína son determinantes clave de la función de las proteínas. Se cree que todos los procesos biológicos son esenciales y se llevan a cabo con precisión a través de interacciones proteína-proteína [28], ver Figura 1.16.

1.5.4. Redes de Información

Las redes de información son redes formadas por el cúmulo de datos que va formando el hombre, como es el caso de la cita de patentes, cita de artículos, la www, etc. Y están organizadas de alguna manera, como es el caso de la cita de patentes que cuando nace una patente se hace referencia a patentes anteriores que pudieron servir de base de conocimiento de esta nueva patente.

Al crear un nuevo artículo se hace referencia a artículos anteriores que ayudan a sostener el conocimiento que se escribe en este nuevo artículo, como también la www. Que hace referencia a otras páginas para sostener el contenido de la página, entonces tanto

el artículo, la patente y la página web son los nuevos nodos, las referencias son los enlaces que en el caso de patentes y de citas de artículos forman redes acíclicas debido a que no se puede hacer que un artículo ya escrito, se le agreguen referencias, pero si se puede referenciar, y en páginas web si podría suceder que una página existente se enlace a otra ya existente, puesto que cada página puede estar siendo modificada constantemente.

Red world wide web (WWW)

La red de la WWW está formada por nodos que son las páginas HTML, y los enlaces que son los enlaces entre páginas. Una página puede hacer referencia a otras para darle más credibilidad a sus datos o para traer información de alguna base de datos y presentarla en la página que se está observando.

La red www. Se encuentra alrededor del mundo, y se puede consultar vía dispositivos como una PC, las tabletas, etc.

Una página web se puede dividir en una o varias páginas que son como subpáginas llamadas hiperenlaces de una principal para que el usuario pueda visualizar más información de la misma, enseguida se muestra un ejemplo una página web, que es la página wikipedia

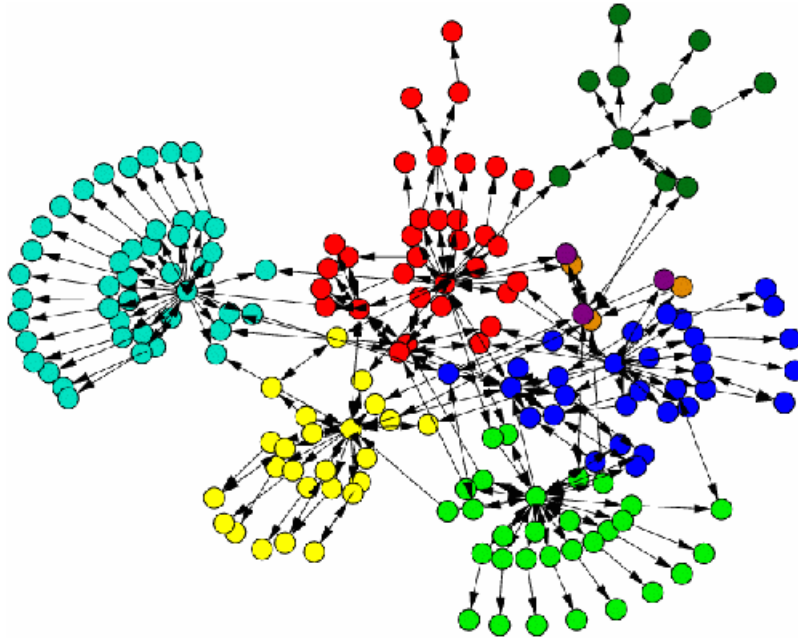


Figura 1.17: Red de www.

La Wikipedia es una enciclopedia libre y políglota de la Fundación Wikimedia (una organización sin ánimo de lucro). Sus más de 20 millones de artículos en 282 idiomas y dialectos han sido redactados conjuntamente por voluntarios de todo el mundo, y prácticamente cualquier persona con acceso al proyecto puede editarlos. Iniciada en enero de 2001 por Jimmy Wales y Larry Sanger, es actualmente la mayor y más popular obra de consulta en Internet. Al crear un artículo se puede libremente modificar, pero si se quiere proteger, el autor se da de alta en wikimedia y el artículo lo puede dejar protegido por el nombre del usuario o por el nombre del grupo a que pertenece el autor. En los artículos que no están protegidos, quedan grabados los datos personales del equipo desde donde fue editado el artículo.

La palabra Wikipedia, nombre propio creado originalmente por Larry Sanger a principios de 2001, es la contracción de wiki, una tecnología para crear sitios web colaborativos, procedente a su vez de wikiwiki, 'rápido' en hawaiano, y enciclopedia, 'enciclopedia' en inglés, Este nombre es pronunciado en inglés como wikipidia.

Lista de hiper-enlaces en una página de wikipedia

href="../../../../articles/b/r/a/Brasil.html"

href="../../../../articles/i/d/i/Idioma_portugu%C3%A9s.html"

href="../../../../articles/i/d/i/Idioma_franc%C3%A9s.html"

href="../../../../articles/r/e/g/Reggae.html"

href="../../../../articles/m/%C3%BA/s/M%C3%BAsica_salsa.html"

href="../../../../articles/1/9/8/1980.html"

href="../../../../articles/l/o/s/Los_Kjarkas_9584.html"

href="../../../../articles/a/n/t/Antillas.html"

href="../../../../articles/z/o/u/Zouk.html"

href="../../../../articles/c/a/t/Especial%7ECategories_701a.html"

href="../../../../articles/m/%C3%BA/s/Categor%C3%ADa%7EM%C3%BAsica_de_Brasil_4fbd.html"

href="../../../../articles/m/%C3%BA/s/Categor%C3%ADa%7EM%C3%BAsica_de_Bolivia_4646.html"

href="../../../../articles/l/a/m/Lambada.html"

href="../../../../articles/l/a/m/Discusi%C3%B3n%7ELambada_d67d.html"

href="http://es.wikipedia.org"

href="../../../../index.html"

href="../../../../index.html"

href="../../../../articles/c/o/m/Portal%7EComunidad_753e.html"

href="../../../../articles/a/c/t/Portal%7EActualidad_c11d.html"

href="../../../../articles/r/e/c/Especial%7ERecentChanges_ce1f.html"

href="../../../../articles/c/o/n/Ayuda%7EContenidos_3c64.html"

href="http://wikimediafoundation.org"

href="../../../../bg/articles/%D0%BB/%D0%B0/%D0%BC/%D0%9B%D0%B0%D0%BC%D0%B1%D0%B0%D0%B4%D0%B0.html"

href="../../../../ru/articles/%D0%BB/%D0%B0/%D0%BC/%D0%9B%D0%B0%D0%BC%D0%B1%D0%B0%D0%B4%D0%B0.html"

[href="../../../../articles/p/u/g/Usuario%7EPuglianini_9bc8.html"](../../../../articles/p/u/g/Usuario%7EPuglianini_9bc8.html)
[href="../../../../articles/b/i/g/Usuario%7EBigsus-bot_95bf.html"](../../../../articles/b/i/g/Usuario%7EBigsus-bot_95bf.html)
[href="../../../../articles/a/r/r/Usuario%7EArrobando_0de1.html"](../../../../articles/a/r/r/Usuario%7EArrobando_0de1.html)
[href="../../../../articles/b/a/t/Usuario%7EBattroid_bbbd.html"](../../../../articles/b/a/t/Usuario%7EBattroid_bbbd.html)
[href="../../../../articles/c/e/m/Usuario%7ECEM-bot_a631.html"](../../../../articles/c/e/m/Usuario%7ECEM-bot_a631.html)
[href="../../../../articles/p/a/i/Usuario%7EPaintBot_96de.html"](../../../../articles/p/a/i/Usuario%7EPaintBot_96de.html)
[href="../../../../articles/m/a/t/Usuario%7EMatdrones_dd78.html"](../../../../articles/m/a/t/Usuario%7EMatdrones_dd78.html)
[href="../../../../articles/s/i/e/Usuario%7ESieBot_59da.html"](../../../../articles/s/i/e/Usuario%7ESieBot_59da.html)
[href="../../../../articles/b/o/t/Usuario%7EBotMultichill_0428.html"](../../../../articles/b/o/t/Usuario%7EBotMultichill_0428.html)
[href="../../../../articles/d/o/r/Usuario%7EDorieo_6266.html"](../../../../articles/d/o/r/Usuario%7EDorieo_6266.html)
[href="../../../../articles/b/o/t/Usuario%7EBOTpolicia_63cf.html"](../../../../articles/b/o/t/Usuario%7EBOTpolicia_63cf.html)
[href="../../../../articles/t/a/i/Usuario%7ETaichi_317d.html"](../../../../articles/t/a/i/Usuario%7ETaichi_317d.html)
[href="../../../../articles/a/m/a/Usuario%7EAmad%C3%ADs_5662.html"](../../../../articles/a/m/a/Usuario%7EAmad%C3%ADs_5662.html)
[href="../../../../articles/r/o/b/Usuario%7ERobotQuistnix_6739.html"](../../../../articles/r/o/b/Usuario%7ERobotQuistnix_6739.html)
[href="../../../../articles/p/a/t/Usuario%7EPatron_4e32.html"](../../../../articles/p/a/t/Usuario%7EPatron_4e32.html)
[href="../../../../articles/f/l/a/Usuario%7EFlaBot_24af.html"](../../../../articles/f/l/a/Usuario%7EFlaBot_24af.html)
[href="../../../../articles/v/i/t/Usuario%7EVitamine_9f54.html"](../../../../articles/v/i/t/Usuario%7EVitamine_9f54.html)
[href="../../../../articles/e/l/_/Usuario%7EEl_Pitifo_81b9.html"](../../../../articles/e/l/_/Usuario%7EEl_Pitifo_81b9.html)
[href="../../../../articles/a/n/t/Usuario%7EAntur_0d4e.html"](../../../../articles/a/n/t/Usuario%7EAntur_0d4e.html)
[href="http://wikimediafoundation.org"](http://wikimediafoundation.org)
[href="../../../../articles/a/c/e/Wikipedia%7EAcerca_de_959f.html"](../../../../articles/a/c/e/Wikipedia%7EAcerca_de_959f.html)
[href="../../../../articles/l/i/m/Wikipedia%7ELimitaci%C3%B3n_general_de_responsabilidad_a0a2.html"](../../../../articles/l/i/m/Wikipedia%7ELimitaci%C3%B3n_general_de_responsabilidad_a0a2.html)

Como se puede observar la página `lambada.html` se enlaza con 46 páginas de las cuales hay 2 páginas repetidas

[href="../../../../index.html"](../../../../index.html)
[href="http://wikimediafoundation.org"](http://wikimediafoundation.org)

Entonces el número de enlaces que se toma para estadística es 44.

Se ha visto un poco de donde están las redes complejas reales y como se ha evolucionado en el descubrimiento de como crear información y herramientas matemáticas que nos ayuden a estudiar mejor las redes.

Pero ¿para qué sirven las redes?.

El estudio de las redes sirve para conocer estos mecanismos microscópicos que dan lugar a los comportamientos de una red, y así poder desarrollar modelos más cercanos a la realidad.

Como por ejemplo: desarrollar mejores métodos para diseñar y suministrar vacunas y otros medicamentos; ser capaz de manejar mejor las epidemias en desarrollo.

Un ejemplo real que no se ha podido solucionar es el crecimiento exponencial de la epidemia del covid19. También puede servir para desarrollar mejores algoritmos para administrar el Internet.

Capítulo 2

Principales estudios de fenómenos y redes reales

Antes y en plena proposición de modelos de crecimiento de red complejas, ha habido varios estudios acerca de redes reales, de las cuales se han obtenido resultados como la distribución de conectividad en forma de ley de potencia de esas redes, de las cuales nos interesa el valor de su exponente γ , como su principal propiedad. También nos interesa conocer como segunda instancia, su diámetro y su coeficiente de agrupación.

De acuerdo con nuestra investigación el estudio más antiguo que se encuentra y que se relaciona con la propiedad distribución ley de potencias, fue la ley del Ing. Pareto en el año 1896.

2.1. El principio o ley de Pareto

La regla 80-20 es conocida como el principio de Pareto, basándose en el conocimiento empírico en el libro «Cours d'économie politique» de 1896, El Ing. Pareto comprobó que el 80 % de la población se reparte el 20 % de la riqueza y el otro 20 % se queda con el 80 % de la riqueza [40] [58].

Pareto estudió la propiedad de la tierra en Italia y lo que descubrió fue que el 20 % de los propietarios poseían el 80 % de las tierras, mientras que el restante 20 % de los terrenos pertenecían al 80 % de la población restante.

Estas cifras son arbitrarias, no son exactas; su aplicación reside en la descripción de un fenómeno y, como tal, es aproximada y adaptable a cada caso particular.

En abril 2021, la Revista Forbes publica los 50 hombres más ricos del mundo 2021, en total todos los billonarios en conjunto valen 13.1 billones de dólares.

La propia Organización de las Naciones Unidas (ONU) había documentado que, tan sólo en 2019, la pobreza mundial rebasaba 1,300 millones de personas que viven por debajo de 1.90 dólares al día.

Se estima que durante el año 2021 puede aumentar en 100 millones más las personas en pobreza extrema, es decir personas que viven con menos de dos dólares al día. “Todo el progreso realizado durante la década de 2010 se ha borrado”, se argumenta en el documento del BM.

Del fenómeno del reparto de la riqueza se puede tener una distribución diferente a la de poisson, ya que sólo 50 hombres concentran 13.1 billones de dólares y más de 1,300 millones de hombres sólo tienen acceso a 2 dólares diarios para vivir.

2.2. Estudio de géneros y Especies del Dr. Willis

El Dr. Willis había estudiado el número de especies entre diferentes géneros en la clasificación de plantas y animales [74]. Los datos que él había compilado demostraron que la mayoría de los géneros contenían solo una especie, que cada vez menos géneros contenían un mayor número de especies y que todavía había algunos géneros que contenían una gran cantidad de especies ver figura 2.1.

2.3. Estudio de Zipf

Zipf [6] hace un estudio sobre el lenguaje, analizando que frecuencia tiene cada palabra con respecto de un texto completo, su sorpresa fue encontrar que la frecuencia de aparición de cada palabra sigue una distribución que puede aproximarse a:

$$P_n \sim \frac{1}{n^a}$$

Donde P_n representa la frecuencia de la n -ésima palabra más frecuente, y el exponente a

número de Especies	número de géneros			
	Chrysomelidae	Cerambycinae	Snakes	Lizards
1	215	469	131	105
2	90	152	35	44
3	38	82	28	23
4	35	61	17	14
5	21	33	16	12
6	16	36	9	7
7	15	18	8	6
8	14	17	8	4
9	5	14	9	5
10	15	11	4	5
11-20	58	74	10	17
21-30	32	21	12	9
31-40	13	15	3	3
41-50	14	8	1	2
51-60	5	4	0	0
61-70	8	3	0	1
71-80	7	0	1	0
81-90	7	1	0	0
91-100	3	1	1	0
101-	16	4	0	2
total	627	1024	293	259

Tabla 2.1: Tabla de Géneros y Especies del Dr. Willis.

es un número real positivo, $a \geq 1$.

La ley de Zipf se cumple para la mayoría de las lenguas.

Si la ley de Zipf se aplica para las ciudades de la República Mexicana, la población de la ciudad de Guadalajara es menos de la mitad de la ciudad de México, pero la Ciudad de Monterrey está creciendo muy rápido, y hasta cerca de la cantidad de población de Guadalajara.

La realidad es que existen pocas ciudades muy grandes y muchas ciudades muy pequeñas, por lo tanto se cumple la ley de Zipf [58].

Existen pocos fenómenos de alta intensidad y muchos fenómenos de baja intensidad, por ejemplo existen pocos terremotos de alta intensidad y muchos terremotos de baja intensi-

dad, lo cual también cumple la ley de Zipf.

2.4. Estudio de Solla-Price

Solla-Price [64] hizo un estudio sobre cita de artículos, en el que comprobó que la frecuencia de citar artículos no sigue la forma de distribución de una red aleatoria, sino que hay artículos más citados que otros, él proponía una distribución diferente a la de poisson, Propuso una distribución de ventajas acumulativas, lo que hoy se conoce como una distribución ley de potencia.

Propuso un modelo de ventaja acumulativa: la probabilidad con que un artículo citado es directamente proporcional a la cantidad de citas que ya tiene, y que modelan estadísticamente la situación en la que el éxito genera éxito.

Es interesante notar que el número de papers con n referencias decae en forma rápida como $\frac{1}{n^2}$ hasta muchos cientos de referencias por paper.

2.5. Estudio sobre el Internet.

Los hermanos Faloutsos[34] hicieron un estudio del Internet, en el cual se tomaron 3 muestras del Internet, las cuales fueron nov de 1997 con 3015 nodos y 5156 enlaces, abril 1998 con 3530 nodos y 6432 enlaces, nov de 1998 con 4389 nodos y 8256 enlaces y con ellas descubrieron que el Internet muestra una ley de potencia en su distribución, también observaron que la conectividad promedio de cada nodo se incrementa y se modifica la pendiente de su distribución de grado.

2.6. Estudio del diámetro de la WWW

La world wide web (www), es una red formada por hiperlinks entre páginas web. Reka Albert et al. [3] estudian la www, obtienen como resultado una distribución de enlaces entrantes en forma de ley de potencia con exponente $\gamma \approx 2.1$, y de enlaces salientes un

exponente $\gamma \approx 2.45$, el diámetro promedio entre dos documentos como una función del tamaño del sistema, para verificar la validación de sus predicciones, determinan el diámetro d para documentos en el dominio nd.edu(Dominio de la Universidad de Notre Dame). La medición fue $\langle d_{nd.edu} \rangle \geq 11.2$.

2.7. Estudio de las publicaciones de Physical Review D, de 1981 a 1997

Redner.[61] examinó los datos numéricos de 20 años de publicaciones de Physical Review D, (24296 documentos), con los documentos y referencias de cada documento, obtuvo una distribución de citas en forma de ley de potencia con exponente $\gamma \approx 3$.

2.8. Distribución de algunas Redes Reales

Las redes reales se distinguen por tener una distribución de su conectividad de la forma:

$$P(k) = k^{-\gamma} \tag{2.1}$$

Donde k es la variable aleatoria y γ es el exponente o pendiente de la distribución de grado en forma ley de potencia.

Una distribución de enlaces entrantes se representa como $P(k) = k^{-\gamma_{in}}$ y una distribución de enlaces salientes se representa como $P(k) = k^{-\gamma_{out}}$.

El grado de conectividad de un nodo se denomina con la letra k para conectividad no dirigida, y para conectividad dirigida, con conectividad entrante k_{in} y con conectividad saliente es k_{out} .

Con el fin de visualizar mejor la distribución de grado, se procesó información de redes reales y se obtuvo la distribución de enlaces entrantes, y salientes, como se muestra en la tabla 2.2, en la que se muestran las tres propiedades más importantes de una red, se muestra el exponente de la distribución de enlaces entrantes y la de los enlaces salientes,

Modelo	Tamaño	γ_{out}	γ_{in}	Coeficiente de Agrupamiento	Diámetro	Referencia
Citas de patentes	3,774,768	3.2	4	0.0757	22	
Soc-Pokec	1,632,803	0.7	2	0.1094	11	
Email-EuAll	81,306	2.5	2.5	0.0671	14	
Web-Stanford	685,230	2.3	2.3	0.5976	674	
cit-hepPh	34,546	3	2.7	0.2848	12	
wiki-vote	7,115	1.3	0.5	0.1409	7	
web-Google	875,713	2.7	2.7	0.5143	21	

Tabla 2.2: Lista de algunas propiedades de Redes Reales, se tomaron los datos de la base de datos de redes de Universidad de Stanford que están disponibles en Internet.

de ahí se puede tomar nota acerca de cuál es el exponente más pequeño y el más grande. Se puede observar que el exponente de la distribución de enlaces salientes va de 0.7 hasta 5.1.

El exponente de la distribución de enlaces entrantes va del 0.5 hasta 4.6 [10].

Una red compleja tiene propiedades no triviales, que la hacen diferente de las redes aleatorias.

Las redes complejas son conjuntos de nodos conectados que interactúan de alguna forma, y se distinguen por tener una cola larga en su distribución de conectividad, un alto coeficiente de agrupamiento, una estructura comunitaria a diferentes escalas, un diámetro de red pequeña entre otras propiedades.

Entonces se puede decir que la red compleja es la columna vertebral de un sistema complejo.

Capítulo 3

Principales modelos ya propuestos

A raíz de los estudios que se habían hecho de redes como Internet, cita de artículos, la www, motivaron a algunos investigadores para desarrollar modelos de crecimiento de redes.

La mayoría de los modelos realizados para redes hasta 1999, parten de n nodos aislados. Posteriormente, a esos nodos se le añaden enlaces con una probabilidad. Es decir, son sistemas cerrados.

Sin embargo, la mayoría de las redes reales son sistemas abiertos que crecen con el nacimiento de nuevos nodos. Por ejemplo, al Internet que se le añaden nuevos sitios con el tiempo.

En la mayoría de los modelos, la adición de enlaces no depende del grado de los nodos entre los que se establece dicho enlace. Los enlaces se colocan aleatoriamente. En la realidad, esto no es así por ejemplo es más probable al escribir un artículo que cité artículos más conocidos y que tengan una gran cantidad de enlaces.

La conectividad que tiene un nodo entrante de la red crea una preferencia de que los nuevos nodos quieran enlazarse con él, fue el primer principio que se aplicó.

Para eso, se empieza del modelo más antiguo.

3.1. Modelo de Yule

El Dr. estadista de apellido Yule, cuando revisó el estudio del Dr. Willis de géneros y especies con plantas y animales, observó que un género se puede dividir en una o varias especies, o en un nuevo género, ver tabla 2.1.

Por lo tanto, logró encontrar un modelo que crea géneros y especies y obtuvo una distribución de géneros y especies [75] en forma de ley de potencia, pero como era un modelo muy futurista para esa época, se quedó ahí guardado y después fue verificado por el Dr. Simon [66], quien de ese modelo, genera otro modelo al cual se le llamó Yule-Simon [36]. Simon introdujo un modelo elegante de copia y crecimiento [3], en espíritu equivalente al modelo de Yule, para explicar la distribución observada de la ley de potencia de las frecuencias de las palabras en los textos [4-6]. En el modelo de crecimiento de Simon, se agregan nuevas palabras a un texto (más generalmente un flujo) con probabilidad constante p en cada paso de tiempo, mientras que con probabilidad complementaria $\bar{p} = 1 - p$, una palabra ya ocurrida se elige uniformemente dentro del texto ya formado (Arroyo). Este modelo produce una distribución de la ley de potencia para las frecuencias de palabras $P(k) \sim k^{-\beta}$ con $\beta = 1 + 1/\bar{p}$ [18].

3.2. Modelo de red aleatoria Mundo Pequeño

Con el objetivo de hacer una transición de una red regular (lattice) a una red aleatoria, Watts y Strogats [73] propusieron un modelo de red mundo pequeño, en el que dado n nodos aislados

1. Se colocan todos los nodos en un círculo;
2. Se conecta cada nodo a su primer $k = 2$ vecinos más cercanos a la derecha así como a su $k = 2$ vecinos más cercanos a la izquierda;
3. Con probabilidad p , se recablean algunos de los enlaces en la red obtenida antes.

La red construida en los pasos 1 y 2 es una red anillo (un grafo en forma de círculo), ver Figura 3.1.



Figura 3.1: En la figura de la Izquierda Se tienen 10 nodos conectados, cada nodo tiene a su primeros $k = 2$ vecinos más cercanos a la derecha así como a sus $k = 2$ vecinos más cercanos a la izquierda, y se enlazan con probabilidad p , se recablean algunos de los enlaces en la red obtenida antes y se muestran en la figura de la derecha.

3.3. Conectividad Preferencial

La conectividad preferencial es una selección aleatoria con cierta preferencia, y fue propuesta en un modelo por Price en 1976 [64] , mientras estudiaba la redes de citas entre artículos científicos. Tratando de explicar la razón de que esta red producía una distribución de grado ley de potencia y que más tarde fue afinada por Barabási [8].

Es decir, que al escribir un artículo su autor preferirá citar artículos bien conocidos, en lugar de escoger artículos desconocidos.

Su explicación es actualmente aceptada por la comunidad científica y se conoce con el nombre de conectividad preferencial (preferential attachment).

3.4. Modelo Barabási-Albert (BA)

El estudio realizado acerca de artículos científicos por Redner[61] , el de Internet por Faloutsos et al.[34] y el estudio del diámetro de la WWW por Barabási-Albert [8], le forman a Barabási y Albert la idea acerca de que estas redes presentan una distribución diferente a las redes aleatorias, lo cual le da elementos suficientes para poder crear el concepto Conectividad preferencial con crecimiento por agregación, y junto con su alumna, la Srita. Albert, proponen el modelo Barabási-Albert, cuyo algoritmo de dicho modelo es:

La fase de crecimiento se inicia con un número dado de nodos aislados n_0 . En cada paso del tiempo nace un nuevo nodo y se conecta a un número constante m de nodos ya existentes en la red.

Cuando se añade un nuevo nodo se elige a qué nodo existente se enlazara, se asume que la probabilidad de que un nuevo nodo se conecte a otro nodo depende del grado k del nodo i existente y se representa como k_i .

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (3.1)$$

Después de t pasos de tiempo, se tienen $n = t + n_0$ nodos y mt enlaces.

Siguiendo este procedimiento se pueden obtener redes que tienen una distribución $P(k) = k^{-\gamma}$.

Donde $P(k)$, es la probabilidad de que un nodo tenga k enlaces, $k^{-\gamma}$ es la ley de potencia, donde γ es el exponente o pendiente de la distribución, ver Figura 3.2.

Hoy en día este modelo es conocido como modelo BA (Barabási Albert), de acuerdo con el modelo, dado un conjunto n_0 de nodos y sin enlaces entre ellos, en cada paso de tiempo nace un nodo nuevo y se conecta aleatoriamente a m nodos existentes, donde el grado de conectividad que tenga el nodo existente genera preferencia.

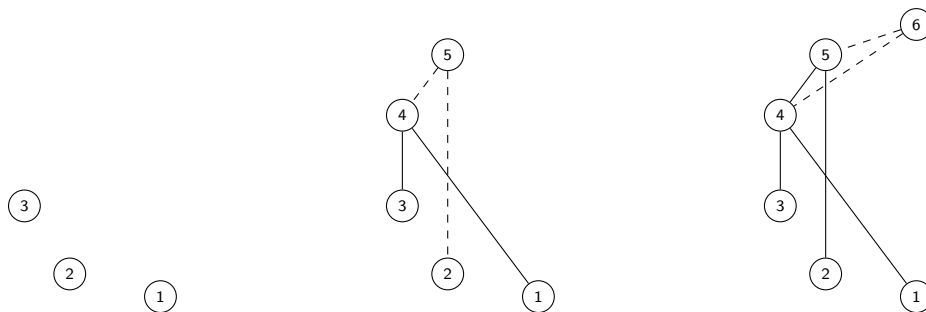


Figura 3.2: Modelo BA. para $n_0 = 3$ y $m = 2$. En $t = 0$, el sistema consiste en $n_0 = 3$ nodos aislados, En cada paso del tiempo, un nodo sera añadido y sera conectado a $m = 2$ nodos existentes, preferentemente a nodos con alta conectividad, en $t = 2$ hay $n_0 + t = 5$ nodos y $mt = 4$ enlaces. En $t = 3$ el sexto nodo es añadido(enlaces añadidos en discontinuo, conectadas a los nodos de mayor grado).

Este modelo generó una revolución en la investigación por crear modelos de crecimiento

de redes complejas, y en los años subsecuentes aparecen nuevos modelos, en los que se busca acercarse más a la realidad, ver Tabla 3.1.

Es así como aparece el modelo de Redner, el cual se explica más adelante.

Propuesta	γ	Referencia	Año	Procesos
Rewiring, ? Topology of Evolving Networks: Local Events and Universality	$2 < \gamma < \infty$	[9]	2000	Recableado
Competition and multiscaling in evolving networks	$2 < \gamma < \infty$	[14]	2001	Aptitud
Growing Scale-Free Networks with Tunable Clustering	$\gamma \approx 3$	[39]	2002	Agrupacion ajustable
Emergence of tempered preferential attachment from optimization	$\gamma \approx 2$	[70]	2006	Optimizacion
Decay of Relevance in Exponentially Growing Networks	$2 < \gamma < \infty$	[41]	2018	Decae relevancia aptitud
Modelo de crecimiento que agrega nacimiento y muerte de enlaces	$\gamma_{out} \geq 1, \gamma_{in} \geq 1$	[17]	2003	Nac. y muerte enlaces
Structure of Growing Networks: Exact Solution of the Barabási-Albert's Model	$2 < \gamma < \infty$	[27]	2000	Atractivo Inic.
Evolution of networks with aging of sites	$2 < \gamma < \infty$	[20]	2000	Edad
Evolution of Networks: From Biological Nets to the Internet and WWW.	$2 \leq \gamma \leq 3$	[22]	2003	Evolución
Consideraron que pueden nacer mayor cantidad de enlaces que nodos nuevos (accelerating growth).	$3/2 < \gamma < 2$	[24]	2001	Aceleración
Growing network with local rules: Preferential attachment, clustering hierarchy and degree correlations	R.Walks. $\gamma \approx 1 + \frac{1}{q_e}$ nearest-neighbors γ desde 0 hasta 2 dup. divergence $\gamma \approx 1 + \frac{1}{1-\mu_1}$	[5]	2002	reglas locales Clustering

Tabla 3.1: Exponentes obtenidos con algunos modelos de crecimiento de Redes Complejas Variantes del Modelo Barabási.

3.5. Un modelo general de crecimiento geométrico para web libre de escalas pseudofractal

Zhang et al.[77] Proponen un modelo de crecimiento geométrico general para web sin pseudo fractal, que está controlado por dos parámetros ajustables.

Derivan exactamente las principales características de las redes: distribución de grado, correlaciones de grado, distribución del coeficiente de agrupación, así como el diámetro,

Propuesta	γ	Referencia	Año	Procesos
Epidemic spreading in complex networks with degree correlations	$2 < \gamma < 3$	[16]	2003	Propagacion epidemias
Preferential attachment mechanism of complex network growth: "rich-gets-richer" or "fit-gets-richer"?	$\gamma \approx 3$ cuando $\alpha \approx 1$	[37]	2018	Conectividad Pref.
A generalised model for asymptotically-scale-free geographical networks fitness		[19]	2019	aptitud
A general geometric growth model for pseudofractal scale-free web	$\gamma \approx 2 \text{ to } 1 + \frac{\ln 3}{\ln 2}$	[77]	2007	Crec geometrico
Complex growing networks with intrinsic vertex fitness	$\gamma(p) \approx 1 + \frac{2}{p}$ Cuando $p = 1$ $\gamma = 3$	[12]	2006	aptitud vertice intrinseco
Growing networks with mixed attachment mechanisms	$\gamma \approx 1 + \frac{2}{q}$	[76]	2006	conectividad mezclada
Growing scale-free networks with tunable distributions of triad motifs,	$\gamma \approx 2.3$	[69]	2015	Triadas de Motivos
Describen un modelo que evita que aparezcan enlaces múltiples. On the growth of directed complex networks with preferential attachment: effect upon the prohibition of multiple link	$1 < \gamma_{out} < \infty$	[31]	2015	Prohibe multiple enlace
On the Emergence of Islands in Complex Networks	$2 < \gamma < \infty$	[33]	2017	Islas
Growing networks with communities: A distributive link model		[42]	2020	Crecimiento Comunidad

Tabla 3.2: Continuación de exponentes obtenidos con algunos modelos de crecimiento de Redes Complejas Variantes del Modelo Barabási.

que están parcialmente determinados por los parámetros.

Los resultados analíticos muestran que las redes resultantes siguen distribuciones de grado de ley de potencia, con un exponente de grado más general ajustado de 2 a $1 + \frac{\ln 3}{\ln 2}$; el coeficiente de agrupamiento de cada nodo individual es inversamente proporcional a su grado y el coeficiente de agrupamiento promedio de todos los nodos se acerca a un valor grande distinto de cero en el orden infinito de la red; el diámetro crece logarítmicamente con la cantidad de nodos de red.

Detectan que las redes descritas por su modelo tienen efecto de pequeño mundo y topología libre de escala.

3.6. Modelo de Conectividad de redes aleatorias en crecimiento

Krapivsky et al. [46] presentan una solución para la distribución de conectividad dependiente del tiempo y la edad de una red aleatoria en crecimiento. La red se crea agregando sitios que enlazan a sitios anteriores con una probabilidad A_k que depende de la cantidad de enlaces k existentes en ese sitio.

Para los núcleos de conexión homogéneos, $A_k \sim k^\gamma$, surgen comportamientos diferentes para $\gamma < 1$, $\gamma > 1$ y $\gamma = 1$. Para $\gamma < 1$, el número de sitios con k enlaces, N_k , varía como un exponencial estirado.

Para $\gamma > 1$, un sitio único se conecta a casi todos los demás sitios. En el caso límite $A_k \sim k$, se encuentra la ley de potencia $N_k \sim k^{-\nu}$, donde el exponente ν puede ajustarse a cualquier valor en el rango $2 < \nu < \infty$.

Cuando $\gamma = 1$ se obtiene:

$$n_k = \frac{4}{k(k+1)(k+2)}. \quad (3.2)$$

$$n_k \approx k^{-3}$$

3.7. Modelo de Organización de Crecimiento de redes aleatorias

Krapivsky [45], propone un modelo de crecimiento de redes aleatorias con probabilidad r , el cual se planteó de la siguiente manera:

En cada paso del tiempo se agrega un nodo nuevo n , y se selecciona un nodo aleatorio existente x , con probabilidad $1 - r$ el nodo nuevo se enlaza al nodo existente x , y con probabilidad r el nodo nuevo se conecta un nodo ancestro del nodo existente x seleccionado. Se obtiene una distribución de enlaces entrante en forma de ley de potencia con exponente en el rango de $2 \leq \gamma < \infty$

El componente de enlaces salientes tiene un tamaño de orden típico en t y proporciona información básica sobre la genealogía de la red.

Krapivsky desarrolla un análisis matemático y obtiene la ecuación 3.3

$$\gamma = 1 + \frac{1}{r} \tag{3.3}$$

Para comprobar el modelo analítico, se desarrolló un algoritmo computacional y se obtuvieron las distribuciones de conectividad de acuerdo a como se muestran en la Figura 3.3.

Se puede observar que la distribución de conectividad con probabilidad $r = 0.25$ se obtiene un exponente $\gamma = 5$ y si se aplica la fórmula se obtiene ese mismo resultado $\gamma = 1 + \frac{1}{r} = 1 + \frac{1}{0.25} = 1 + 4 = 5$.

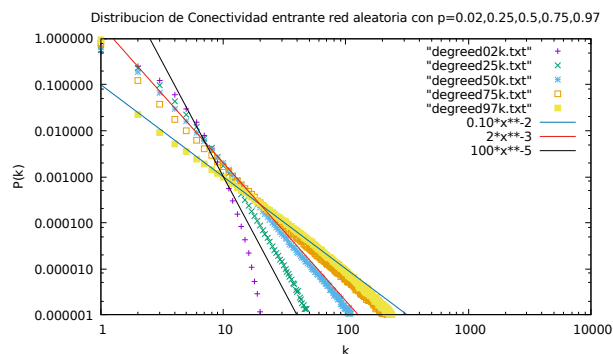


Figura 3.3: Distribución de enlaces salientes de red aleatoria, con $r = 0.02, 0.25, 0.5, 0.75, 0.97$ de Krapivsky. Cuando $r \rightarrow 0.02$ el exponente $\gamma \approx \infty$, $r \rightarrow 0.25$ el exponente $\gamma \approx 5$, y cuando $p \rightarrow 0.97$ exponente $\gamma \approx 2$.

Existen algunos modelos que utilizan el principio de funcionamiento del modelo Barabási-Albert [8] y/o el de copiado de krapivsky [47], de los cuales se da una pequeña explicación más del concepto de redes.

3.8. Modelo de aptitud Bianconi-Barabási

Bianconi y Barabási, revisaron el modelo BA, verificaron que el modelo generaba enlace preferencial con los nodos más antiguos, pero la realidad es que hay nodos jóvenes, que tienen un atractivo, así como hay algunas personas que son más carismáticas que otras o hay páginas web que son más visitadas que otras, entonces ellos decidieron que se puede agregar una variable de aptitud al modelo BA, Bianconi et al. [14], con la finalidad de hacer un modelo de crecimiento de redes menos rígido y un poco más acercado a la realidad.

Dada la ecuación del modelo BA.

$$\Pi(k) = \frac{k_i}{\sum_j k_j} \quad (3.4)$$

A esta ecuación, Bianconi le agrega la variable de aptitud η de la siguiente manera:

$$\Pi(k) = \frac{\eta k_i}{\sum_j \eta k_j} \quad (3.5)$$

Logrando con esto que la distribución de conectividad en forma de ley de potencia, muestre otros resultados, donde el exponente $2 \leq \gamma \leq \infty$

3.9. Topología de redes en evolución: Eventos locales y universalidad

Las redes crecen y evolucionan por eventos locales, tales como la adición de nuevos nodos y enlaces, o recableado de un nodo a otro [9].

Se genera un modelo extendido del modelo original BA, con una probabilidad p , se suman m ($m \leq m_0$) nuevos enlaces, se selecciona aleatoriamente un nodo existente como origen, y se selecciona aleatoriamente un nodo destino, y se genera un enlace del nodo origen al nodo destino.

Con una probabilidad q , se recablean m enlaces.

Con una probabilidad $1 - p - q$, se añade un nuevo nodo, el nuevo nodo tiene m enlaces con probabilidad $\Pi(k)$ son conectados a nodos i ya presentes en la red.

Con este modelo se puede obtener una distribución γ entre 2 y ∞

3.10. Modelo de estructura de crecimiento de red con atractivo inicial.

Dorogovtsev [26] propone un modelo con atractivo inicial, este modelo se basa en el modelo Barabási-Albert, al cual se le agrega un atractivo inicial.

Y funciona de la manera siguiente:

En cada paso de tiempo aparece un nuevo sitio (ver Figura.3.4). Simultáneamente, se introducen m nuevos enlaces dirigidos que provienen de sitios no especificados. Deja que la conectividad sea la cantidad de enlaces entrantes q_s a un sitio, es decir, a un sitio agregado en el momento s .

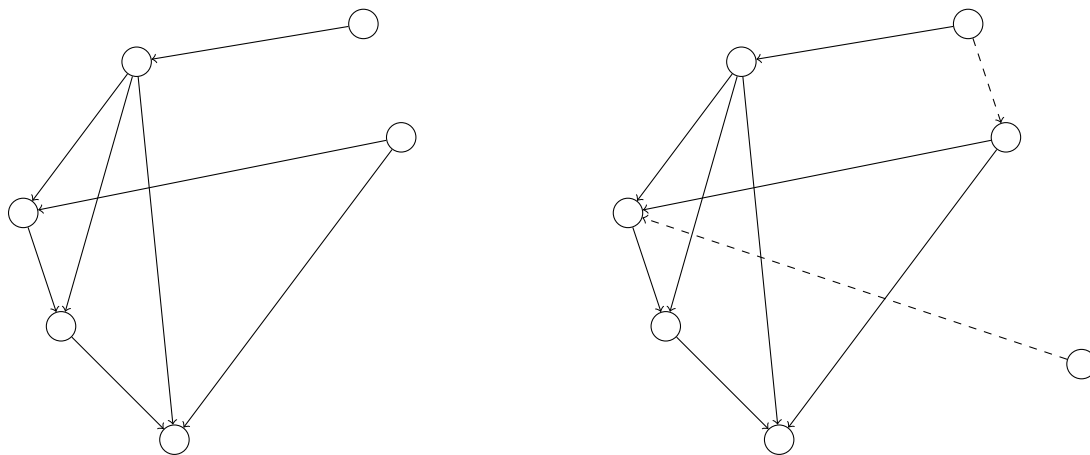


Figura 3.4: Ilustración del crecimiento de baja consideración. Cada instante un nuevo nodo y m (aquí, $m = 2$) nuevos enlaces dirigidos son añadidos. Estos enlaces son distribuidos entre los nodos de acuerdo a las reglas introducidas en el texto.

Los nuevos enlaces se distribuyen entre los sitios de acuerdo con la siguiente regla. La probabilidad de que un nuevo enlace apunte a un sitio dado s es proporcional a la siguiente característica del sitio:

$$A_s = A + q_s, \tag{3.6}$$

A partir de esto, fue llamado su atractivo. Todos los sitios nacen con cierto atractivo inicial A_0 , pero luego aumenta debido al término q_s . El parámetro A introducido es el atractivo inicial que rige la probabilidad de que los sitios “jóvenes” obtengan nuevos enlaces.

3.11. Modelo de edad de Dorogovtsev

En las redes reales de referencia, como es la cita de artículos, el envejecimiento de los documentos generalmente ocurre, sin embargo, rara vez se cita a documentos antiguos.

Uno puede preguntarse, ¿cómo cambia la estructura de la red si se introduce el envejecimiento de los sitios, es decir, si la probabilidad de conexión del nuevo sitio con alguno antiguo es proporcional no sólo a la conectividad del sitio antiguo sino también a la potencia de su edad, τ^α , por ejemplo? [por supuesto, las funciones que disminuyen más rápidamente, por ejemplo, son la exponencial, $e(-const X \tau)$, están incluidas en el caso particular de $\alpha \rightarrow \infty$ [21]. Esta pregunta es bastante razonable: de hecho, la Fig. 3.5, demuestra que la estructura de la red depende obviamente de α .

Se muestra, tanto numérica como analíticamente, que este cambio es dramático: la escala desaparece cuando $\alpha > 1$ (por lo tanto, está ausente también para cualquier función de disminución más rápida que una potencia), y los exponentes de las leyes de escala dependen fuertemente de α en el rango $\alpha < 1$.

3.12. Efecto del crecimiento acelerado en la estructura de las redes de comunicaciones

Dorogovtsev [24] llamó a las redes con crecimiento acelerado a aquellas redes que tienen crecimiento real, la cantidad media de conexiones por nodo aumenta con el tiempo. Entre los ejemplos de redes grandes que presentan este tipo de crecimiento se encuentran Internet, World Wide Web, redes de colaboración y muchas otras.

Muestra que el crecimiento acelerado influye en la distribución de los enlaces y, como consecuencia, puede determinar la estructura de una red.

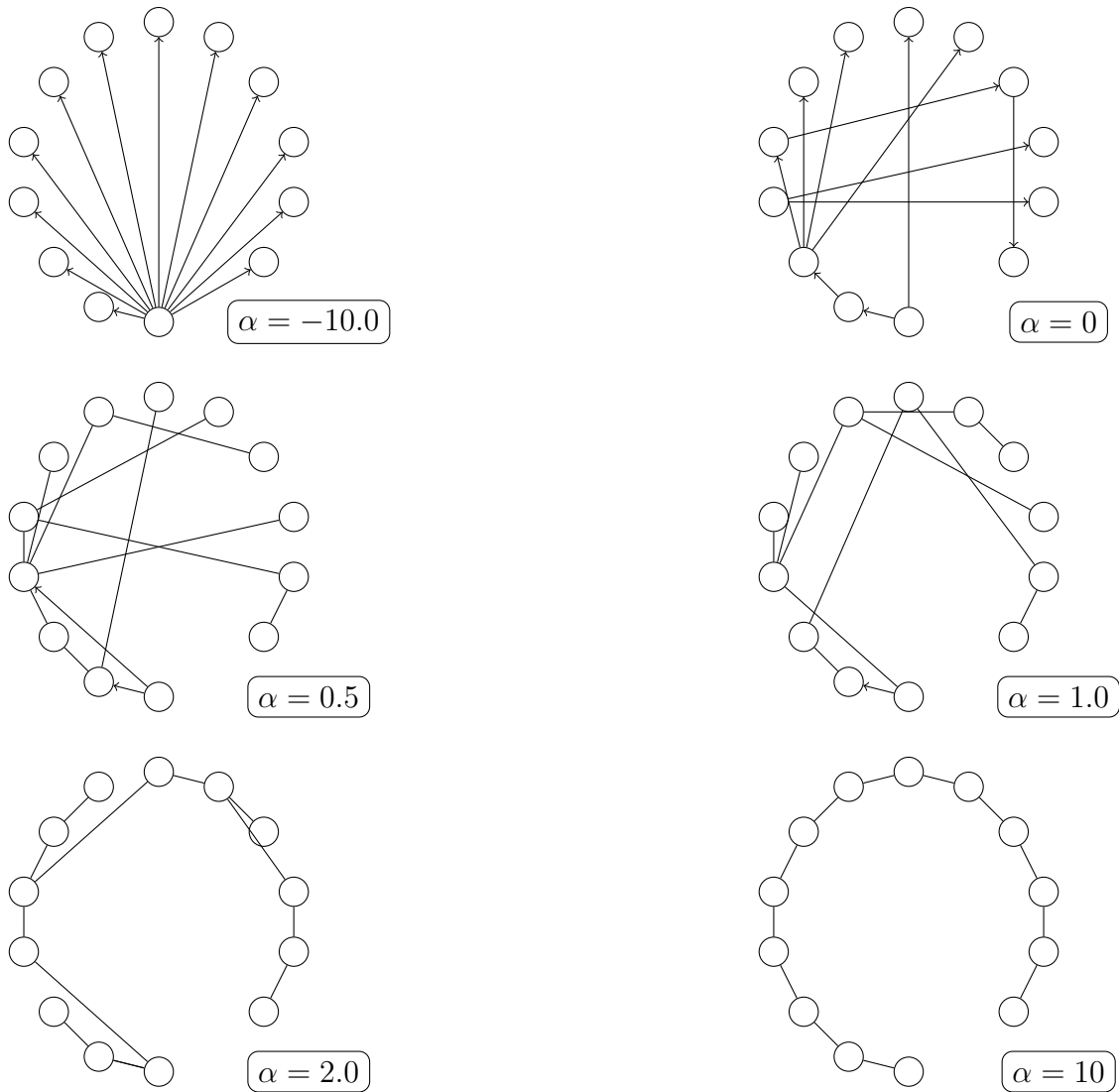


Figura 3.5: Cambios de la estructura de la red con el aumento del exponente α son mostradas. La probabilidad de que un sitio nuevo se una a uno viejo, es proporcional a $\tau^{-\alpha}$, donde τ es la edad del sitio viejo.

Para las redes en crecimiento con enlaces preferenciales y una mayor densidad de enlaces, son posibles dos escenarios. En uno de ellos, el valor del exponente γ de la distribución de conectividad $P(q, t) \propto q^{-\gamma}$ está entre $3/2$ y 2 . En el otro, el exponente es, $\gamma > 2$, y la distribución no es necesariamente estacionaria.

3.13. Crecimiento de redes libres de escala en el mundo pequeño con coeficiente de variante ajustable

Qiang et al. [60] propone un modelo simple que genera redes de mundo pequeño sin escala con coeficiente de variante ajustable.

Estas redes se construyen mediante un proceso de adición en dos etapas para cada nuevo nodo. El modelo puede reproducir distribuciones de grado sin escala y efectos de mundo pequeño. Los resultados de la simulación son consistentes aproximadamente con las predicciones teóricas.

Curiosamente, se obtiene el coeficiente de agrupamiento no trivial y la clasificación de grado ajustable, ajustando el parámetro: el exponente preferencial b . El modelo puede unificar la caracterización de las redes clasificadas como (ordenadas) y desordenadas.

De acuerdo al valor de β se obtiene una distribución de conectividad en forma ley de potencia con exponente σ entre 2.95 y 3.11.

3.14. Modelo de crecimiento de red compleja con distribuciones de motivos triada ajustables

Para descubrir sus principios de diseño estructural, se define “motivos de red”, patrones de interconexiones que ocurren en redes complejas en números que son significativamente más altos que aquellos en redes aleatorias [56].

Shuguang et al. [69] Proponen un modelo para desarrollar redes libres de escala con distribuciones de motivos ajustables a través de una operación combinada de unión preferencial y pasos de siembra de motivos de triada.

Los experimentos numéricos muestran que las redes construidas tienen distribuciones de los motivos triada local, al tiempo que conservan las características globales de las distribuciones de grado ley de potencia, las longitudes de camino corto promedio de los nodos y estructuras altamente agrupadas.

Se obtiene una distribución en forma ley de potencia con exponente $\gamma \approx 2.3$

3.15. Modelo de grafos dirigidos libres de escala.

El modelo fue propuesto por Bollobas et al.[17] e incorpora tres operaciones distintas:

- (i) con probabilidad α , se crea un nuevo nodo v y se conecta a un nodo destino existente w , donde w se selecciona de acuerdo con su grado.
- (ii) con probabilidad β , se origina un nuevo enlace entre dos nodos existentes, el enlace se origina en v y termina en w , donde v se selecciona según su grado de enlaces salientes, mientras que w se selecciona de acuerdo con su grado de enlaces entrantes.
- (iii) con la probabilidad γ , se origina un nuevo nodo w y se convierte en el destino de un nuevo enlace que comienza en un nodo existente v , donde v se selecciona de acuerdo con su grado saliente. La suma de las probabilidades α , β y γ es uno.

Los exponentes χ de distribución de enlaces entrantes y enlaces salientes resultantes que usan este modelo varían entre 2 y ∞ .

3.16. Modelo de Crecimiento de redes aleatorias por copiado de Krapivsky

Krapivsky [47] propuso un modelo de crecimiento de red compleja por copiado de enlaces, el cual funciona de la siguiente manera:

En cada paso del tiempo nace un nodo nuevo n , se conecta con el nodo seleccionado x y con todos los ancestros del nodo seleccionado x , a este concepto se le llama copiado, debido a que el nodo nuevo se enlaza con todos los ancestros del nodo seleccionado (copia todos sus enlaces salientes).

Como resultado se obtiene una distribución de conectividad entrante en forma de ley de potencia con exponente $\gamma \approx 2$, y de conectividad saliente en forma exponencial.

Hay algunos artículos publicados que utilizan el copiado con variantes, ver tabla 3.2.

Propuesta	γ	Refer	Año	
Organization of Growing Random Networks	$\gamma_{in} \approx 2$	[45]	2001	Copia 1 enlace
Network growth by copying	$\gamma_{in} \approx 2$	[47]	2005	Copia m enlaces
Growing networks with superjoiners	$\gamma_{out} \approx \alpha, \alpha > 2$ $\gamma_{in} \approx 1 + \frac{1}{r}$ $\alpha < 2$ $\gamma_{in} \approx 0$	[2]	2014	superjoiners
Activity ageing in growing networks		[49]	2007	
The Accelerating Growth of Online tagging Systems		[51]	2011	
Densification and Structural Transitions in Networks that Grow by Node Copying		[13]	2016	
Network growth models and genetic regulatory networks		[35]	2018	
A growth model for directed complex networks with power-law shape in the out-degree distribution	$\gamma_{out} \approx 1$	[30]	2015	Copia m enlaces
On a growth model for complex networks capable of producing power-law out-degree distributions with wide range exponents	$1 \leq \gamma \leq \infty$	[32]	2015	copia m enlaces

Tabla 3.3: Exponentes obtenidos con algunos modelos de crecimiento de redes complejas variantes del Modelo de krapivsky.

3.17. Modelo de crecimiento de red con Superjoiners

El modelo de Ameerah Jabr-Hamdan et al. [2] considera la distribución del grado de salida. Este modelo se basa en el modelo de Organization del crecimiento aleatorio de redes de Krapivsky y Redner [45]. En el modelo Ameerah, cualquier nuevo nodo se conectará a través de un enlace dirigido a un nodo existente, que se llama *ancestro* del nuevo nodo porque existía antes.

Este modelo funciona de la siguiente manera:

Se crea un nuevo nodo n_{new} en cada incremento de tiempo. Este nodo se conecta a m nodos destino que se seleccionan siguiendo el método de Krapivsky-Redner: Un nodo objetivo x se selecciona al azar de la red y con probabilidad $1 - r$, n_{new} se conecta a x . De lo contrario, con probabilidad r , n_{new} se conecta a un antecesor seleccionado al azar de x . El

modelo original de Krapivsky-Redner sólo considera que $m = 1$, mientras que el modelo de Ameerah cambia el valor de m en cada paso de tiempo usando una distribución de ley de potencia. Como m se obtiene de una distribución de ley de potencia, este modelo realmente genera red compleja dirigida, con una distribución de ley de potencia sólo para su grado. m se obtiene de manera artificiosa, lo cual no es natural que se obtenga una distribución de enlaces salientes en modo ley de potencia.

3.18. Modelo de Esquivel

Esquivel et al. [30] propusieron un modelo de crecimiento simple para CN (Redes Complejas) dirigido que es capaz de generar distribuciones de grado que decaen como una ley de potencia con exponente entre 0, y 1.

El modelo se basa en el modelo de krapivsky [45], con una pequeña modificación en la forma de que el nodo n_{new} se enlaza a los nodos destino, ya que en este modelo se conecta a los ancestros del nodo destino, y en el Krapivsky se conecta al destino y los ancestros del destino.

De esta manera, con probabilidad p se copia los enlaces de un nodo seleccionado y con probabilidad $1 - p$ se enlaza a un número aleatorio $[1 : N]$ de nodos existentes, de este modelo se obtiene la distribución de enlaces salientes, que de acuerdo con la probabilidad p , cuando la probabilidad $p=0.1$, el exponente $\gamma = 0.1$, cuando el $p = 0.9$, el exponente $\gamma = 0.9$, cuando $p = 0.99$, el exponente $\gamma \approx 0.99$ [30]

En el modelo que propusieron, el crecimiento de la red se realiza mediante la adición de nodos de uno en uno. Al principio, solo los n_0 nodos existen en la red y su grado saliente es 0. Entonces se considera que el grado de salida de cualquier nodo nuevo que se agregue a la red se determina de la siguiente manera:

Con probabilidad p donde $0 < p < 1$, n_{new} copia el grado de salida de un nodo seleccionado al azar de la red. Es importante tener en cuenta que a medida que aumenta la cantidad de Q_s de nodos con grado de salida, la probabilidad de que el nodo n_{new} tiene grado de salida s también aumenta a $\frac{Q-s}{N}$, donde N es la cantidad total de nodos en la red.

Con probabilidad complementaria $1 - p$, n_{new} selecciona aleatoriamente un grado de salida

uniformemente distribuido de 0 a N .

Donde el nodo n_{new} tiene un grado 0, 1, 2, ... N . Es considerable tener en cuenta que esta regla produce un grado de salida poco realista del nuevo nodo casi todas las veces que se aplica.

Es decir, los nuevos nodos pueden tener un grado superior al orden N .

Aplicando las consideraciones previas y usando el método continuo, se puede escribir la siguiente ecuación diferencial:

$$\frac{dQ_s(N)}{dN} = \frac{Q_s(N)}{N} + (1-p)\frac{1}{N+1}, \quad (3.7)$$

$$\frac{dQ_s(N)}{dN} + \frac{-p}{N}Q_s(N) = \frac{(1-p)}{N+1}, \quad (3.8)$$

$$\frac{Q_s}{N} \sim N^{p-1}s^{-p} \text{ para } 1 \ll s \ll N, \quad (3.9)$$

Con exponente $\gamma_{out} \sim p$.

Sin embargo, no se puede obtener la distribución de enlaces entrantes en modo ley de potencia.

Capítulo 4

Problema y Objetivo

A finales de los años 90's, fue cuando se hicieron los estudios de las redes reales como Internet [34], la www [3], cita de artículos [61], donde se observó que estas mostraban propiedades diferentes a las redes aleatorias, tales como su distribución de conectividad en forma de ley de potencia, diámetro pequeño, agrupación, etc. Lo cual demostró que estas redes eran diferentes a las que se les consideró como redes aleatorias.

Barabási y Albert motivados por este comportamiento, lograron proponer un modelo de crecimiento de red muy sencillo, del cual se obtiene una distribución ley de potencia con exponente $\gamma \sim 3$, no obstante en redes reales, el exponente de la distribución ley de potencia varía de 1 a 5 [10] después de cierto tiempo de crecimiento. Sin embargo, el modelo fue bien recibido por los demás investigadores de redes, y fue utilizado como base para poder encontrar otras propiedades de las redes reales, por ejemplo, los nodos que se hacen viejos ya no son citados, algunos nodos nuevos nacen con cierto atractivo que atraen enlaces entrantes.

Por ejemplo, el atractivo inicial explica que en la distribución de enlaces entrantes pueda aparecer ley de potencia, sin embargo, no se explica como lograrlo en los enlaces salientes. Tanto las redes reales que soportan nacimiento y muerte de enlaces como sucede en la www. y las redes reales que no tienen nacimiento ni muerte de enlaces como es la cita de artículos, cita de patentes, etc., Pueden mostrar una distribución de enlaces entrantes y salientes en forma ley de potencia.

Cuando un artículo es publicado, a medida que pasa el tiempo puede recibir citas (Enlaces

entrantes), y por conectividad preferencial originar que haya pocos artículos con muchas citas y muchos poco citados, no obstante en el caso de las citas que se hacen en el artículo (Enlaces salientes) ya publicado no se pueden aumentar o disminuir al pasar el tiempo, y a pesar de ello, hay pocos artículos que hacen muchas citas y muchos que hacen pocas citas, en otras palabras, sus enlaces salientes decaen en forma de ley de potencia y este fenómeno no se puede explicar por conectividad preferencial.

Se buscaron modelos que pudieran este problema, se encontró que el que mas acercaba al objetivo, era el modelo de copiado de Krapivsky [47], ya que el copiado podría servir para obtener ese resultado, sin embargo, el resultado que se obtuvo fue una distribución exponencial de los enlaces salientes, también se revisó el modelo propuesto por Ameerah [2], el cual se basa en krapivsky [45] y si se obtiene la distribución ley de potencia en el outdegree, pero artificialmente se obtiene el número de enlaces salientes del nodo nuevo, lo cual no es natural.

De manera que, eso es lo que yo traté de resolver en esta tesis.

Por eso mi objetivo general es:

Proponer un modelo de crecimiento de red compleja donde se pueda obtener una distribución ley de potencia de enlaces entrantes y salientes y que la red no acepte nacimiento y ni muerte de enlaces en nodos existentes.

Capítulo 5

Modelo de crecimiento red propuesto.

5.1. Introducción

Muchos sistemas y sus interacciones pueden describirse utilizando Redes Complejas Dirigidas (DCN) que comparten propiedades similares [15],[10], [54]. Para modelar un sistema como una red compleja, un conjunto de componentes en el sistema se definen como nodos y las relaciones entre ellos como enlaces. Por ejemplo, las redes de citas de artículos científicos representan artículos como nodos y citas como enlaces que los unen.

Las citas en un artículo se dirigen externamente a los artículos que citan. Cuando los enlaces dirigidos son necesarios para representar una red, se llama DCN. Sin embargo, cuando todos los enlaces son bidireccionales o no direccionales, la red se considera una red compleja no dirigida (NDCN). En una red compleja dirigida, el número de enlaces que salen de un nodo se denomina grado saliente, mientras que el número de enlaces que ingresan a un nodo se denomina grado entrante.

La motivación detrás de este trabajo es que, para el caso de redes complejas que no permiten agregar y eliminar enlaces, no hay modelos capaces de producir distribuciones de enlaces salientes y entrantes que decaigan como ley de potencia.

Se presenta un nuevo modelo de crecimiento acelerado de una red compleja dirigida que, sin agregar nuevos enlaces o volver a cablear entre los nodos existentes en la red, es capaz

de generar redes en las que las distribuciones de nodos de enlaces entrantes y salientes decaen como una ley de potencia.

El crecimiento acelerado es un comportamiento disponible en algunas redes complejas, donde la proporción a la que se crean nuevos enlaces es mayor que la proporción a la que se agregan nuevos nodos [24]

5.2. Modelo de crecimiento

El modelo de crecimiento de red compleja dirigida propuesto en este documento [59], se basa en el modelo Krapivsky-Redner [47]. Inicialmente, la red tiene n_0 nodos aislados y en cada paso de tiempo se crea un nuevo nodo n y ocurre una de las siguientes dos operaciones:

1. Con probabilidad $1 - p$, se selecciona un número aleatorio m , donde m es el número de enlaces salientes para n . El número m tiene un rango entre 1 y N , donde N es la cantidad de nodos en la red antes de que se creara n . El nodo de origen n selecciona aleatoriamente m nodos en la red y se conecta a cada uno de estos m nodos a través de un enlace dirigido.
2. Con probabilidad p , n selecciona aleatoriamente un nodo existente x y luego se conecta n a todos los antecesores de x , donde los enlaces dirigidos se originan en n y terminan en cada antecesor de x .

La Figura 5.1 muestra un ejemplo para el modelo propuesto. En este ejemplo, la red tiene un conjunto de nodos $N \in [0, 1, \dots, 4]$ y $p = 0.8$. Luego, en el primer paso, el nodo 5 se genera y selecciona aleatoriamente un número real entre 0 y 1 que determina si necesita para ejecutar la operación que corresponde a la probabilidad p o la que corresponde a la probabilidad $1 - p$. Por ejemplo, si el número elegido es 0.1, entonces la operación corresponde a la probabilidad $1 - p(1 - 0.8)$ como se muestra en la Figura (5.1.A). Luego, se selecciona un número aleatorio entre 1 y N , que determina el grado de salida del nuevo nodo, m . Observe que el rango de m para este ejemplo es de 1 a 5.

Suponga un valor de $m = 2$ para este ejemplo, luego se crean dos enlaces salientes desde el nodo 5 a dos nodos diferentes seleccionados al azar de la red: nodos 0 y 3.

La figura (5.1.B) muestra un ejemplo de la operación que corresponde a la probabilidad p : se crea un nuevo nodo, 6, y selecciona aleatoriamente un número entre 0 y 1.

Suponga que este número es 0.35, que es mayor que $1 - p$ y, por lo tanto, esta debería ser una operación que corresponde a p . Luego, se elige un nodo aleatorio de la red existente, por ejemplo 5, y el nuevo nodo copia todos los enlaces salientes de este nodo. Esto también se expresa cuando el nodo 6 se conecta a los antepasados del nodo 5.

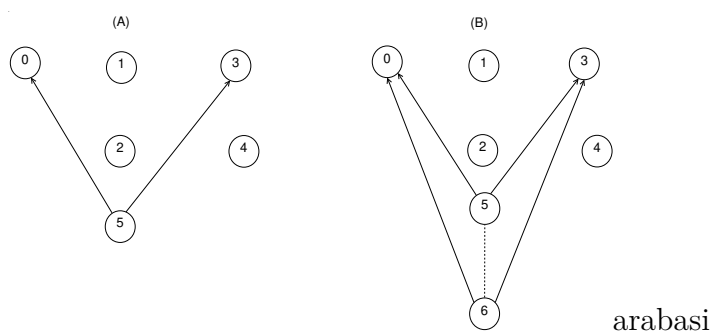


Figura 5.1: Un ejemplo del modelo propuesto. (A) El nuevo nodo 5 realiza la operación 1, donde elige aleatoriamente dos nodos y se conecta a ellos. (B) El nuevo nodo 6 realiza la operación 2, donde selecciona aleatoriamente el nodo 5 y se conecta a los antepasados de este nodo..

5.3. Experimentos y resultados

Los siguientes experimentos se diseñaron para encontrar el impacto que los parámetros del modelo propuesto tienen en la distribución de enlaces entrantes y salientes de las redes generadas y para determinar el rango del exponente en estas distribuciones.

El modelo propuesto fue probado usando simulaciones numéricas. Las redes generadas crecieron de $N = 1$ a $N = 10^4$. El rango para los enlaces salientes m se encuentra entre 1 y N . El valor de probabilidad p varió de 0 a 1. Los registros de estas simulaciones se emplearon para generar los gráficos que se muestran en la Figura 5.2. Para mayor claridad, esta figura sólo muestra las distribuciones para $p = 0.10, 0.80, 0.90, 0.97$ y 0.99 .

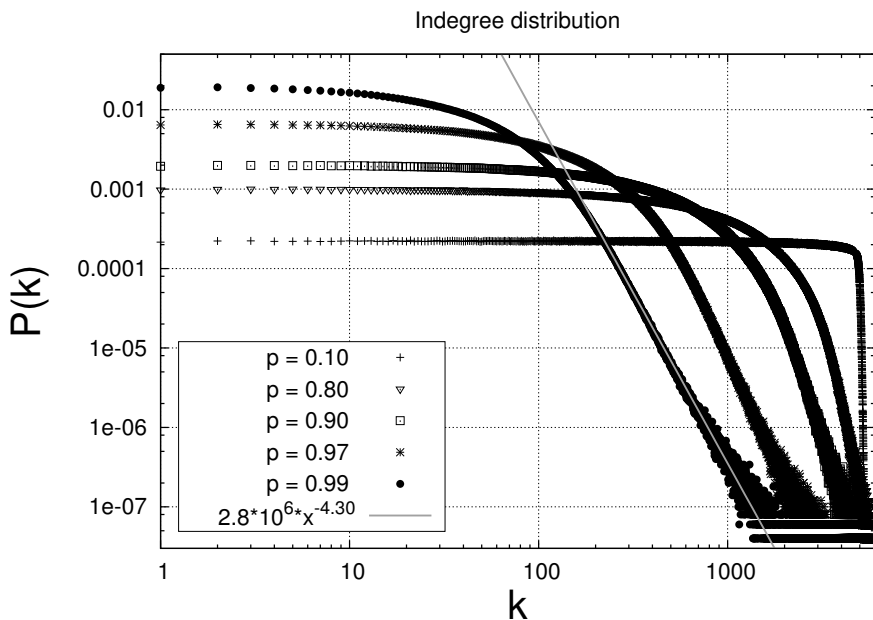


Figura 5.2: La distribución de conectividad de enlaces entrantes de las redes generadas usando el modelo propuesto. Esta figura muestra una familia de curvas en la que el γ_{in} varía de ∞ a aproximadamente 4.30.

La figura 5.2 muestra que, cuando $p = 0.10$, la cola de la distribución de enlaces entrantes decae como una función exponencial. En este caso, la probabilidad de que un nuevo nodo se conecte con m nodos existentes seleccionados al azar es 0.90. Para esta condición, el crecimiento de la red se rige por procesos aleatorios no sesgados. En otras palabras, cada nodo en la red tiene la misma probabilidad de obtener nuevos enlaces entrantes.

Cuando $p = 0.99$, es posible ver que la cola de la distribución tiene aproximadamente tres décadas en el eje y , que decaen como ley de potencia con el exponente $\gamma \approx 4.30$.

Para el modelo propuesto, los experimentos muestran que el promedio de enlaces entrantes aumenta a medida que la red crece.

Por ejemplo, cuando $p = 0.99$ y la red alcanza 10^3 nodos, el promedio de enlaces entrantes es 4.79. Por el contrario, cuando el crecimiento alcanza hasta 10^4 nodos y usa el mismo p , la red tiene un promedio de enlaces entrantes de 48.64.

Este incremento en el grado medio de enlaces entrantes indica que, a medida que la red crece, la velocidad a la que nacen los enlaces aumenta con respecto a la velocidad a la que nacen los nodos. En otras palabras, el modelo exhibe un crecimiento acelerado [24].

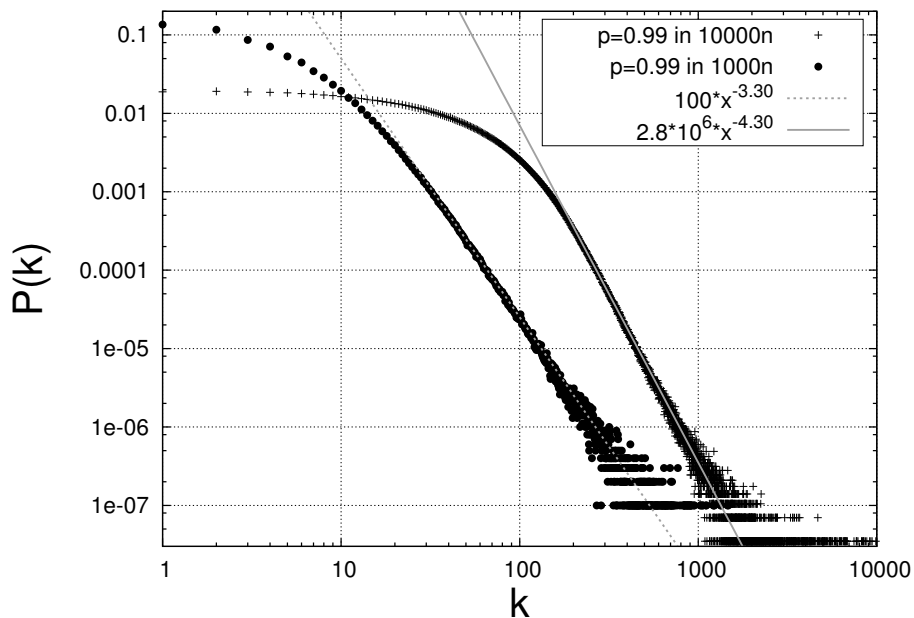


Figura 5.3: El impacto del crecimiento en el valor de γ_{in} en una red generada por el modelo propuesto. Cuando la red ha crecido a 10^3 nodos, la distribución de enlaces entrantes tiene una pendiente $\gamma_{in} \sim 3.3$, pero cuando la red alcanza 10^4 nodos, el valor de γ_{in} ha aumentado a aproximadamente 4.3

La figura 5.3 muestra que, para una red que ha crecido a 10^3 nodos, el exponente $\gamma_{in} \approx 3.30$, y cuando la red alcanza 10^4 nodos, $\gamma_{in} \approx 4.3$.

En otras palabras, γ_{in} aumenta a medida que la red crece. Dorogovtsev y sus colegas analizaron este comportamiento en [24], donde descubrieron analíticamente que γ_{in} oscila entre 2 y ∞ en este tipo de sistemas acelerados.

Esto coincide con los resultados del modelo propuesto: cuando $m = 1$ y permanece constante durante el crecimiento de la red, la aceleración es igual a cero y el modelo propuesto es idéntico al publicado por Krapivsky et al. [45]. Para este caso, γ_{in} también varía de 2 a ∞ .

La figura 5.4 muestra una familia de curvas de las distribuciones de enlaces salientes producidas por el modelo propuesto. Esta figura muestra que, cuando p se aproxima a 0, la distribución de enlaces salientes con exponente γ_{out} también se aproxima a 0. En otras palabras, cuando p se acerca más a 0, la distribución de enlaces salientes se aproxima a una distribución uniforme. Cuando p tiende a 1, el γ_{out} también tiende a 1. Por lo tanto,

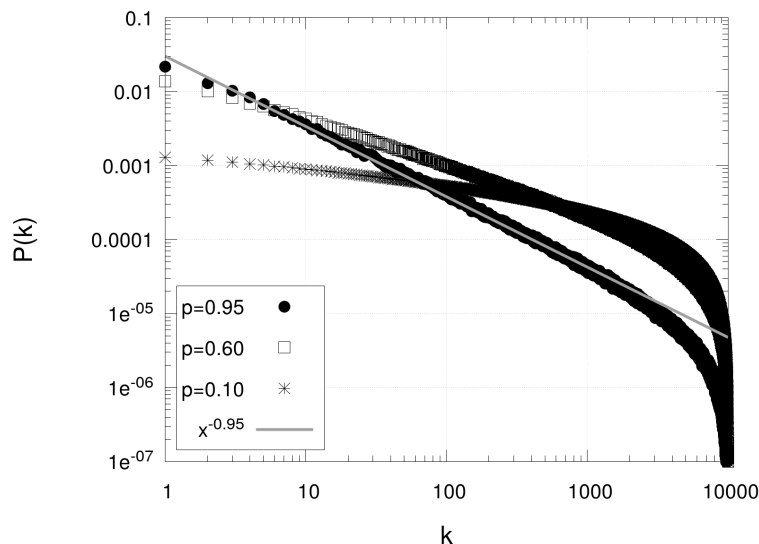


Figura 5.4: La distribución de conectividad de enlaces salientes de las redes generadas utilizando el modelo propuesto: cuando $p \rightarrow 0$, $\gamma_{out} \rightarrow 0$ y cuando $p \rightarrow 1$, $\gamma_{out} \rightarrow 1$.

este experimento numérico muestra que γ_{out} tiene un rango entre 0 y 1.

Este resultado coincide perfectamente con el modelo analítico publicado por Esquivel et al. [31], donde aplicaron el modelo de Krapivsky-Redner a una generación aleatoria de enlaces salientes para generar su propio modelo. El γ_{out} obtenido con el modelo propuesto en este documento es similar al obtenido en [31] porque es uno de los componentes del modelo propuesto.

Con el modelo propuesto, la aceleración no influye en el valor de γ_{out} y depende sólo de la probabilidad p , es decir $\gamma_{out} = p$ [31]. Por lo tanto, el rango de valores para γ_{out} está entre 0 y 1.

5.4. Patentes de Estados Unidos de América y su γ_{out}

Las referencias entre las patentes aprobadas en los Estados Unidos entre 1975 y 1999 [72] son un ejemplo de una red compleja que cambia su valor de exponentes γ a medida que crece, de manera similar al modelo propuesto. En esta red, cada nodo representa una patente y los enlaces dirigidos, las referencias entre patentes.

Este sistema se seleccionó porque hay un registro que permite reproducir el crecimiento de

la red, lo que permite analizar su estado y propiedades en diferentes intervalos de tiempo. Esto es imposible o considerablemente más complicado en otros sistemas, como WWW. O la red de citas de artículos, ya que no hay registros precisos sobre como evolucionan estos sistemas con el tiempo.

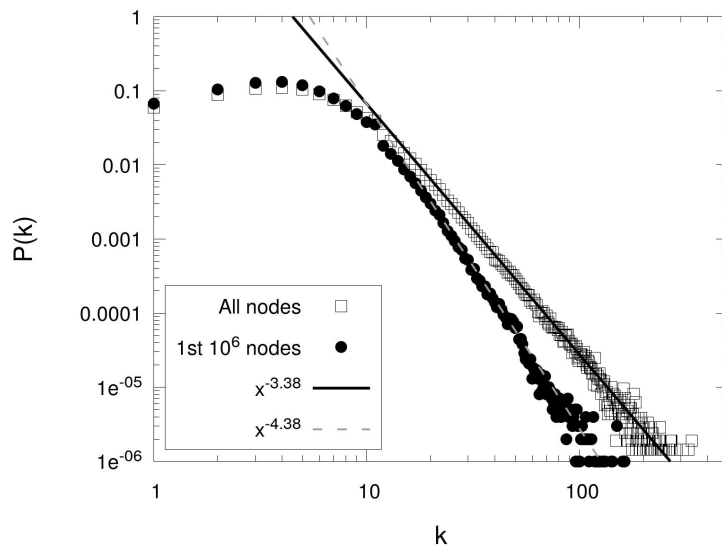


Figura 5.5: La distribución de conectividad de enlaces salientes para la red de citas de patentes de EE. UU. 1975-1999. Cuando esta red tenía 10^6 patentes, γ_{out} era aproximadamente 4.38; cuando la red alcanzó 2,089,345 (todas las patentes en el conjunto de datos) γ_{out} cambió a ≈ 3.38 .

La red de patentes analizada aquí tiene 2,089,345 patentes que tienen referencias de otras que pueden haber sido aprobadas antes o después de 1975. Esta red tiene un total de 16,518,948 enlaces. Para el análisis actual, el crecimiento de la red se ha dividido en dos etapas: ST1 se utiliza para representar cuando el sistema ha crecido a 10^6 nodos y ST2 cuando la red ha alcanzado su número máximo de nodos.

La Figura 5.5 muestra la distribución de enlaces salientes de este sistema cuando la red se encuentra en ambas etapas. Es posible observar que $\gamma_{out} \approx 4.38$ en ST1 y cambia aproximadamente a 3.38 en ST2.

Este es un claro ejemplo de un sistema real que tiene un comportamiento de crecimiento similar al observado en el modelo propuesto. En otras palabras, γ_{out} o γ_{in} pueden variar con el tiempo a medida que la red crece. Desafortunadamente, este análisis no puede

hacerse para γ_{in} porque el rango de datos se detiene en 1999, lo que oculta todas las citas recibidas de patentes creadas después de esta fecha.

5.4.1. ¿Las redes complejas son libres de escala?

En una red real, pocos nodos están altamente conectados, es decir, poseen un gran número de enlaces a otros nodos, aunque el grado de conexión de casi todos los nodos es bastante bajo.

En general, la probabilidad $p(k)$ de que un nodo de la red estuviera conectado con k nodos era proporcional a $k^{-\gamma}$, es decir, sigue una ley de potencias.

El exponente γ no es universal, sino que depende del tipo específico de red.

Para la mayor parte de los sistemas se encontró que el parámetro declarado se encontraba en el rango $2 < \gamma \leq 3$.

Libre de escala significa que la pendiente de la distribución ley de potencia es una línea infinita con un exponente γ que no varía con el tiempo, un ejemplo es la distribución con exponente $\gamma \sim 3$ que se produce del modelo de Barabási, y que no varía con el tiempo, lo cual muestra que es una red libre de escala.

Los resultados que muestra la red de citas de patentes no son libres de escala porque como se pudo observar la pendiente γ_{out} de la distribución de enlaces está variando con el tiempo, la pendiente γ_{out} de la etapa 1 es $\gamma_{out} \sim 4.38$ y la de la etapa 2, es $\gamma_{out} \sim 3.38$.

Los resultados que muestra el modelo propuesto en esta tesis presenta una situación similar. cuando se lleva un crecimiento de 1000 nodos se obtiene una distribución con pendiente $\gamma_{in} \sim 3.3$, y cuando el crecimiento son 10000 nodos, se obtiene una distribución con pendiente $\gamma_{in} \sim 4.3$.

Hasta ahora se tiene la idea de que las redes complejas son libres de escala, sin embargo, la variación del γ_{out} de la red de patentes podría indicar la existencia de otras redes complejas reales que podrían no ser libres de escala y que además es muy difícil saber cuántas y cuáles podrían ser, ya que no se puede conocer su γ_{out} y γ_{in} en diferentes estados de evolución. La presencia de redes complejas reales que no fueran libres de escala invalidarían los métodos analíticos desarrollados para su estudio, tales como el método continuo de Barabási[8] o el método de la ecuación maestra de Dorogovtsev [27], ya que estos métodos asumen que el

sistema de ecuaciones que describen al sistema en un tiempo T_1 , son válidas en un tiempo T_2 , sin embargo, si no son escalables, esto no se cumpliría.

arabasi

Capítulo 6

Discusión

Krapivsky en su Modelo de crecimiento de redes aleatorias [45], indica que en cada paso del tiempo nace un nodo nuevo y selecciona un nodo aleatorio x , con probabilidad r , el nodo nuevo se enlaza al ancestro del nodo seleccionado, y con probabilidad $1 - r$, se enlaza al nodo seleccionado, se obtuvo una distribución de enlaces en modo de ley de potencia.

Después en el modelo de crecimiento por copiado [47], indica que en cada paso de tiempo, nace un nodo nuevo, donde el nodo nuevo copia los enlaces del nodo seleccionado existente, y se enlaza al nodo seleccionado existente, obtiene una distribución de enlaces entrantes en forma ley de potencia, y una distribución de enlaces salientes en forma exponencial.

El modelo propuesto [59], con probabilidad p , se copia los ancestros del nodo seleccionado x , y con probabilidad $1 - p$, el nodo nuevo se enlaza a un número aleatorio m de nodos existentes de $1 : N$, donde cada nodo destino se selecciona aleatoriamente.

El modelo presentado tiene **un comportamiento de crecimiento acelerado que se esperaba, ya que cada nuevo nodo agregado a la red crea m enlaces, donde m puede ser mayor que uno. Sin embargo, el hecho de que el exponente γ cambia a medida que la red crece, no era un comportamiento esperado. Por lo tanto, es importante estudiar los mecanismos que producen tal comportamiento**, que para el caso de la red de citas de patentes, produce que el exponente γ_{out} toma valores más pequeños a medida que la red crece, mientras que este mismo exponente no cambia para el modelo propuesto.

También puede ser interesante estudiar qué podría suceder a medida que crece el sistema

de red de la cita de patentes: ¿Continuará disminuyendo el exponente γ_{out} ? ¿En qué punto dejará de cambiar? Tal vez el γ_{out} se detendrá cuando la red quede libre de escala (γ_{out} entre 2.0 y 3.0), donde puede convertirse en una red de mundo ultra pequeño.

Este estudio puede tener implicaciones al intentar modelar otros sistemas. Por ejemplo, si se puede usar una red para modelar la propagación de un virus biológico, es posible anticipar como crecerá la red, su diámetro esperado y otras propiedades estructurales, lo que puede ayudar a predecir o contener la propagación de tales virus. La respuesta a las preguntas anteriores puede estar disponible una vez que haya suficiente información sobre la evolución de este tipo de redes y una comprensión más profunda sobre los diferentes procesos que permiten producir y modelar este tipo de redes.

Capítulo 7

Trabajo Futuro

La cita de patentes es una red compleja con 3,774,768 nodos, 16,518,948 enlaces, y muchas islas, donde existe una isla que se considera el componente gigante que concentra la mayoría de citas, pero el conjunto de islas forma una distribución de islas en forma de ley de potencia, en las que se observa que una isla es un tema [71].

Se queda detenido este proyecto donde se buscaba que usando el modelo de Yule (Géneros y Especies), se pudiera proponer un modelo de crecimiento de red compleja, donde la cita de patentes nos sirviera como un ejemplo para comprobar que el modelo que se está experimentando se acercara a una red real, hasta el momento se han hecho experimentos y se han obtenido resultados satisfactorios, falta por hacer experimentos en los cuales se pudieran hacer comparaciones con la red cita de patentes.

Capítulo 8

Conclusiones.

En el trabajo relacionado se presenta un conjunto de estudios de redes reales, donde se le da una importancia especial a la red de Internet, ya que es más identificable por las personas, y a un conjunto de modelos de crecimiento de red donde se le da especial importancia al modelo de BA(Barabási Albert). Que es el primer modelo de crecimiento hecho por el hombre, muy sencillo que se basa en la conexión preferencial, donde cada nodo genera preferencia con base a la cantidad de conexiones entrantes que ya tiene antes de enlazar el nodo nuevo. Esto muestra los estudios de redes reales, como de creación de modelos de crecimiento de redes complejas que sirven de introducción para entender el modelo que se explican en esta tesis.

Esta investigación ha introducido un nuevo modelo de crecimiento DCN basado en modelos anteriores de Krapivsky-Redner [47] y de Esquivel et al. [30] El nuevo modelo ha resultado en un mecanismo de crecimiento que es capaz de generar DCN con una distribución de nodos de enlaces entrantes y salientes que decae como una ley de potencia y que también incluye un fenómeno de crecimiento acelerado, donde la velocidad a la que se crean los enlaces es mayor que la velocidad a la que se crean los nodos. Esto hace que el número medio de enlaces por nodo aumente a medida que la red crece y también exhibe un aumento en el exponente γ_{in} , pero no para γ_{out} .

Gracias a este estudio se pudieron obtener resultados interesantes, que puedan servir de motivación para las futuras investigaciones que se efectúen.

Espero que mi aportación sirva para continuar este camino de investigación y que pueda

servir de apoyo en los resultados de futuras investigaciones.

Sentí que se alargó el tiempo en encontrar el camino correcto en esta investigación debido a que mi línea de aprendizaje había sido enfocada en otras áreas, tal vez un poco retrasadas para esta época, donde me olvidé de revisar que lo escrito debe ser entendible por el lector. Mejorar el idioma Inglés ha sido un gran desafío, porque absorbió mucho tiempo, y hoy en día se ha tenido que acelerar varias tareas, como es el mejorar el Inglés, tratar de escribir mejor, y aprender la ciencia de las redes complejas.

Bibliografía

- [1] Acosta-Elias Jesus, Arjona-Villicaña P. David, Moldes-Navarro L., "The impact of time delay in the connectivity distribution of complex networks generated using the Barabási-Albert model", *Revista Mexicana de Física*, Vol. 60, pp. 145-148, 2014.
- [2] Ameerah Jabr-Hamdan, 1, Jie Sun, 2, and Daniel ben-Avraham, Growing networks with superjoiners
- [3] Réka Albert, Hawoong Jeong, Albert-László Barabási, Diameter of the World-Wide Web, *NATURE | VOL 401 | 9 SEPTEMBER 1999 | www.nature.com*
- [4] Reka Albert and Albert-Laszlo-Barabási, *Rev of Mod Phy*, Vol 74, Jan. (2002).
- [5] Alexei Vazquez, Growing network with local rules: Preferential attachment, clustering & nearest-neighbors hierarchy and degree correlations dup. divergence.
- [6] Lada A. Adamic, Bernardo A. Huberman, Zipf's law and the Internet, *Glottometrics* 3, 2002, 143-150
- [7] Andrewartha, H.G. *Introducción al estudio de poblaciones animales*. Editorial Alambre, 1987.
- [8] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, *Science* 286, 509-512 (1999).
- [9] R. Albert, A.-L. Barabási Topology of evolving networks: local events and universality *Physical Review Letters* 85, 5234-5237 (2000)
- [10] Albert-Laszlo-Barabási, *Network Science*, Cambridge University Press, (2016).

-
- [11] A.L. Barabási , H. Jeong,,Z. Néda, E. Ravasz, A. Schubert, T. Vicsek; Evolution of the Social Network of the scientific collaborations, University of Notre Dame, 2002
- [12] C. Bedogne and G. J. Rodgers, Complex growing networks with intrinsic vertex fitness
- [13] U. Bhat, P. L. Krapivsky, R. Lambiotte, S. Redner, Densification and Structural Transitions in Networks that Grow by Node Copying, *Physica Review*, 2016
- [14] G. Bianconi and A.-L. Barabási, Competition and multiscaling in evolving networks, *Europhys. Lett.*, 54 (4), pp. 436–442 (2001)
- [15] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang. *Physics Reports* 424, (2006) 175- 308.
- [16] Marián Boguña , R. Pastor-Satorras, and Alessandro Vespignani, Epidemic spreading in complex networks with degree correlations
- [17] Bollobás, B., Christian, B., Chayes, J. & Riordan, O. Directed scale-free graphs. SODA'03. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 132-139 (2003).
- [18] C. Cattuto 1,2 , V. Loreto 2 and V. D. P. Servedio 3,1, .^A Yule-Simon process with memory", *Europhysics* 2006
- [19] Nicola Cinardi, A generalised model for asymptotically-scale-free geographical networks, arXiv:1911.02494v1, 2019
- [20] S. N. Dorogovtsev and J. F. F. Mendes¹, Evolution of Networks with aging of sites, *Physical Review E*, Vol 62, number 2, August 2000
- [21] Dorogovtsev, S. N. & Mendes, J. F. F. Evolution of Networks. *Advances in Physics* 51, 1079– 1187 (2002).
- [22] Dorogovtsev S., Mendes J., and Dorogovtsev S. Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford University Press, New York, NY, 2003.

-
- [23] S.N. Dorogovtsev and J.F.F. Mendes, Exactly solvable small-world network, *Europhys. Lett.* 50, 1-7 (2000).
- [24] S.N. Dorogovtsev and J.F.F. Mendes, Effect of the accelerated growth of communications networks on their structure, *Phys. Rev. E* 63, 025101 1-4 (2001).
- [25] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, WWW and Internet models from 1955 till our days and the popularity is attractive"principle. *ArXiv:cond-mat/0009090*.
- [26] S. N. Dorogovtsev,^{1,2,*} J. F. F. Mendes,^{1, †} and A. N. Samukhin, Structure of Growing Networks with Preferential Linking, *VOLUME 85, NUMBER 21, PHYSICAL REVIEW LETTERS*(Nov 2000).
- [27] S.N. Dorogovtsev, J.F.F. Mendes, and A.N. Samukhin, Structure of Growing Networks: Exact Solution of the Barabási–Albert’s Model, *arXiv.org* 2000
- [28] Dongbo Bu¹, Yi Zhao¹, Lun Cai¹, Hong Xue², Xiaopeng Zhu², Hongchao Lu¹, Jingfen Zhang¹, Shiwei Sun¹, Lunjiang Ling², Nan Zhang², Guojie Li¹ and Runsheng Chen “Topological structure analysis of the protein–protein interaction network in budding yeast”.
- [29] Erdős, P. and Rényi, A. “On the Evolution of Random Graphs.” *Publ. Math. Inst. Hungar. Acad. Sci.* 5, 17-61, 1960.
- [30] Esquivel-Gomez, E. Stevens-Navarro , U. Pineda-Rico & J. Acosta-Eliás. ”A growth model for directed complex networks with power-law shape in the out-degree distribution”
- [31] J. Esquivel-Gómez, R. E. Balderas-Navarro, Edgardo Ugalde, J. Acosta-Elias, On the growth of directed complex networks with preferential attachment: effect upon the prohibition of multiple links, *Int. J. Mod. Phys. C*, Vol. 26, No. 6 (2014).
- [32] J. Esquivel-Gómez, P. D. Arjona-Villicaña, E. Stevens-Navarro, U. Pineda-Rico, R. E. Balderas-Navarro & J. Acosta-Elias, On a growth model for complex networks capable

- of producing power-law out-degree distributions with wide range exponents, *Scientific Reports* volume 5, Article number: 9067 (2015)
- [33] J. Esquivel-Gómez, R. E. Balderas-Navarro, P. D. Arjona-Villicaña, P. Castillo-Castillo, O. Rico-Trejo, and J. Acosta-Elias, On the Emergence of Islands in Complex Networks, *Hindawi Complexity* Volume 2017
- [34] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On Power-Law Relationships of the Internet Topology, *ACM SIGCOMM*, Cambridge, MA, September 1999.
- [35] D. V. Foster, S. A. Kauffman, and J. E. S. Socolar, Network growth models and genetic regulatory networks, *ARxiv*, 2018
- [36] Garcia Garcia, Juan Manuel (2011). “A fixed-point algorithm to estimate the Yule-Simon distribution parameter”. *Applied Mathematics and Computation*. 217 (21): 8560–8566.
- [37] Michael Golosovsky, Preferential attachment mechanism of complex network growth: “rich-gets-richer” or “fit-gets-richer”?, *Phys. Rev. E* 97, 062310 (2018)
- [38] Guare, J., *Six Degrees of Separation: A Play*, Vintage, New York (1990).
- [39] Petter Holme and Beom Jun Kim, *Growing Scale-Free Networks with Tunable Clustering*, 2002
- [40] Bunkley, Nick (March 3, 2008). “Joseph Juran, 103, Pioneer in Quality Control, Dies”. *The New York Times*.
- [41] Jun Sun, Steffen Staab, Fariba Karimi, Decay of Relevance in Exponentially Growing Networks, *WebSci '18: Proceedings of the 10th ACM Conference on Web Science*, 2018
- [42] Ke-ke Shang, Bin Yang, Jack Murdoch Moore, Qian Ji, and Michael Small, Growing networks with communities: A distributive link model, *Chaos* 30, 041101 (2020)
- [43] Paul, Clayton R. (2001). *Fundamentals of Electric Circuit Analysis*. John Wiley & Sons

-
- [44] D. E. Knuth, *The Stanford GraphBase: A Platform for Combinatorial Computing*, Addison-Wesley, Reading, MA (1993).
- [45] P. L. Krapivsky and S. Redner, Organization of growing random networks, *Phys. Rev. E* 63, 066123 – Published 24 May 2001
- [46] P. L. Krapivsky,^{1,2} S. Redner,¹ and F. Leyvraz, Connectivity of Growing Random Networks, *VOLUME 85, NUMBER 21 PHYSICAL REVIEW LETTERS* (Nov 2000).
- [47] P. L. Krapivsky, and S. Redner, *PHYSICAL REVIEW E* 71, 036118 (2005)
- [48] P.L.Krapivsky,S.Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics*(Cambridge University Press,Cambridge,2010).
- [49] R. Lambiotte, Activity ageing in growing networks ,*Arxiv*,2007
- [50] Fabrizio Leisen, Luca Rossini, and Cristiano Villa. A note on the posterior inference for the Yule-Simon distribution. *Journal of Statistical Computation and Simulation*, 87(6):1179–1188, 2017.
- [51] Lingfei Wu, The Accelerating Growth of Online tagging Systems, *EUROPEAN PHYSICAL JOURNAL B* 83 (2) , pp.283-287, 2011
- [52] Roberts Lucas, Robert Denisa,An Expectation Maximization Framework for Yule-Simon,arxiv,2018 Preferential Attachment Models
- [53] Markoff, John. “Armies of Expensive Lawyers, Replaced by Cheaper Software”. *New York Times* March 5, 2011. p A1.
- [54] Michael Small, Lvlín Hou, Linjun Zhang, *National Science Review* 1, (2014), 357-367.
- [55] Milgram, S., 1967, *Psychol. Today* 1, 60.
- [56] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U., Network motifs: Simple building blocks of complex networks, *Science* 298, 824–827 (2002).
- [57] Mark Newman, *Networks: An Introduction*, Oxford University Press, 2010

-
- [58] M. E. J. Newman 2007, Power laws, Pareto distributions and Zipf's law, Contemporary Physics
- [59] P-castillo,D-Arjona,J. Acosta,“An accelerated growth model to generate complex networks with connectivity distribution slope that varies with time”, Revista Mexicana de Fisica, March 2019
- [60] Qiang Guoa, Tao Zhoub, Jian-Guo Liuc, Wen-Jie Baid, Bing-Hong Wangb , Ming Zhaob,Growing scale-free small-world networks with tunable assortative coefficient,Physica A 371 (2006) 814–822
- [61] S. Redner,How popular is your paper? An empirical study of the citation distribution,Eur. Phys. J. B 4, 131–134 (1998)
- [62] M. Richardson and R. Agrawal and P. Domingos. Trust Management for the Semantic Web. ISWC, 2003.
- [63] Shiwei Sun, Lunjiang Ling, Nan Zhang, Guojie Li and Runsheng Chen: Topological structure analysis of the protein-protein interaction network in budding yeast. Nucleic Acids Research, 2003, Vol. 31, No. 9 2443-2450
- [64] Derek J. de Solla Price,“Networks of Cientific Papers”,Science, 149(3683) : 510-515, July 30, 1965
- [65] Shiwei Sun, Lunjiang Ling, Nan Zhang, Guojie Li and Runsheng Chen: Topological structure analysis of the protein-protein interaction network in budding yeast. Nucleic Acids Research, 2003, Vol. 31, No. 9 2443-2450
- [66] Simon, H. A. (1955). “On a class of skew distribution functions”. Biometrika. 42 (3–4): 425–440.
- [67] Daniel B. Stouffer,Scaling from individuals to networks in food webs,Functional Ecology 2010, 24, 44–51
- [68] Scott, J., Social Network Analysis: A Handbook, Sage Publications, London, 2nd ed. (2000).

-
- [69] Shuguang Li , Jianping Yuan, Yong Shi, Juan Cristóbal Zagal, Growing scale-free networks with tunable distributions of triad motifs, *Physica A* 428 (2015) 103?110
- [70] Raissa M. D'Souza, Christian Borgs, Jennifer T. Chayes, Noam Berger, and Robert D. Kleinberg, Emergence of tempered preferential attachment from optimization, *PNAS*, 2006
- [71] Vladimir Batagelj , Natasa Kejzar, Simona Korenjak-Cerne3 , and Matjaz Zaversnik, *Data Science Clasification, Analyzing the Structure of U.S. Patents Network*, Springer-Verlag Berlin · Heidelberg 2006
- [72] V. Batagelj, arXiv:cs/0309023v1 [cs.DL], (2003).
- [73] D. J. Watts and S. H. Strogatz, *Nature* 393, 440-442 (1998).
- [74] N. Bacaer, *A Short History of Mathematical Population Dynamics, Yule and evolution* (1924), DOI 10.1007/978-0-85729-115-8 15, © Springer-Verlag London Limited 2011
- [75] Yule, *Una teoría matemática de la evolución, sobre la base de las conclusiones del Dr. JC Willis, FRS*, 1925
- [76] Zhi-Gang Shao, Xian-Wu Zou, Zhi-Jie Tan and Zhun-Zhi Jin, Growing networks with mixed attachment mechanisms
- [77] Zhongzhi Zhang , Lili Rong, Shuigeng Zhou, A general geometric growth model for pseudofractal scale-free web
- [78] Zipf, G.K. (1949). *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.