



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

Un análisis comparativo de redes sociales

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIA

PRESENTA:

OLIVIA SASHIKO SHIRAI REYNA

TUTORA

DRA. BIBIANA OBREGÓN QUINTANA
2015



México, D.F.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno.

Shirai

Reyna

Olivia Sashiko

56 88 73 94

Universidad Nacional Autónoma de México

Facultad de Ciencias

Actuaría

309098267

2. Datos del Asesor.

Dra.

Bibiana

Obregón

Quintana

3. Datos del sinodal 1

Dra.

Claudia Orquídea

López

Soto

4. Datos del sinodal 2

M. en I.

Adrián

Girard

Islas

5. Datos del sinodal 2

Dra.

Mucuy-Kak del Carmen

Guevara

Aguirre

6. Datos del sinodal 2

M. en I.

Amalia

Maldonado

Rosas

7. Datos del trabajo escrito

Un análisis comparativo de redes sociales

62 p

2015

UN ANÁLISIS COMPARATIVO DE REDES SOCIALES

“Invertir en conocimientos produce siempre los mejores beneficios.”
Benjamin Franklin

“La disciplina es la parte más importante del éxito.”
Truman Capote

Agradecimientos

De manera muy especial quiero agradecer a la Dra. Bibiana Obregón Quintana, por su orientación, su tiempo, su persistencia, su paciencia, su disponibilidad, sus conocimientos, su esfuerzo para que todo estuviera bien y en tiempo, y sobre todo por todo su apoyo tanto para la realización de esta tesis como el servicio social, no hubiera podido realizar este trabajo sin su ayuda. Muchas gracias.

A todos los miembros del jurado por leer la tesis, por sus correcciones y sus acertados comentarios, en especial a la Dra. Claudia Orquídea López Soto, M. en I. Adrián Girard Islas, Dra. Mucuy-kak del Carmen Guevara Aguirre y M. en I. Amalia Maldonado Rosas.

Además, quiero dedicar esta tesis a mi familia, la cual siempre ha estado conmigo en todo momento, en especial a mi mamá Olivia, a mi papá José Akio y a mi hermana Yoshi, por darme todo su apoyo siempre. Este trabajo también va para mis abuelitos Raúl (q.e.p.d), Rosita, Ojichan (q.e.p.d) y Obachan, a todos mis tíos y primos, quienes forman una parte esencial de mi vida, siempre han creído en mí y me han apoyado siempre de manera incondicional. En especial agradezco a Mariana por leer esta tesis y ayudarme en la corrección de estilo.

Por último quiero agradecer a todos mis amigos, mis compañeros de vida, que me han soportado por tanto tiempo y sobre todo me han apoyado en todas mis locuras, en especial Andrea (Yeya), Héctor, Saúl, Diego, Gerardo, Karla, Lalo, Leo, Ian, Mariana, Raúl, Cintia, Brenda, y demás amigos (perdón que no los ponga a todos), gracias por estar ahí.

Índice general

Agradecimientos	I
Lista de figuras	V
Introducción	1
Capítulo 1	5
1.1 Contexto Histórico	5
1.2 Definiciones	8
1.3 Modelos de Redes Complejas	14
1.3.1 Modelos de Redes Aleatorias (Modelo ER)	14
1.3.2 Modelos de Redes de Mundo Pequeño (Modelo WS)	14
1.3.3 Modelos de Redes de Libre Escala (Modelo BA)	15
Capítulo 2	17
Estado del Arte de Redes Sociales	17
Capítulo 3	27
Datos, Metodología y Resultados	27
Conclusiones	45
Anexos	47
Referencias	49

Índice de figuras

1.	Representación gráfica del problema de los siete puentes de Königsberg. . .	5
2.	Descripción gráfica de una red con nodos y arcos.	8
3.	Matriz de Adyacencia	9
4.	Ejemplo de camino, ruta, ciclo, trayectoria, circuito	9
5.	Modelo de Redes Aleatorias	14
6.	Modelo de Mundo Pequeño	15
7.	Modelo de Libre Escala	16
8.	a) Red de Mundo Pequeño b) Red de Libre Escala c) Red Aleatoria	16
9.	Representación del estudio de J. Moreno a un grupo escolar en la década de 1930.	18
10.	Representación enlaces débiles de Granovetter.	20
11.	Conectividad homogénea y heterogénea	22
12.	Distribución de grado de las estaturas y de los amigos de los alumnos . .	23
13.	Tres Grados de Influencia	25
14.	Red de Karate	30
15.	Histograma distribución de grado de la red Karate	30
16.	Red de comunidad de delfines	31
17.	Histograma distribución de grado de la red de delfines	31
18.	Red de coautores de Teoría de Redes	32
19.	Histograma distribución de grado de la red de coautores	32
20.	Red de jugadores de futbol americano	33
21.	Histograma distribución de grado de la red de jugadores de futbol americano	33
22.	Red de correos electrónicos de Enron con un 30% de la red original	34
23.	Histograma distribución de grado de la red de correos electrónicos de Enron	34
24.	Red de miembros facultad UK	35
25.	Histograma distribución de grado de la red de miembros facultad UK	35
26.	Red de papalote	36
27.	Histograma distribución de grado de la red de papalote	36
28.	Red de pacientes geriátricos con un 80% de la red original.	37
29.	Histograma distribución de grado de la red de pacientes geriátricos	37
30.	Red de abogados	38
31.	Histograma distribución de grado de la red de abogados	38
32.	Red de miembros OJN	39
33.	Histograma distribución de grado de la red de OJN	39
34.	Plano log-log distribución de grado de la red de OJN	39

Introducción

Para la mayoría de la gente, hoy en día cuando escucha el término “*red social*”, piensa que se refiere a los servicios de redes sociales en línea como Facebook, Twitter, Instagram, entre otros, pero en realidad a estos se les conoce como *social media services*. Sin embargo, para formar una red social solo es necesaria una relación entre individuos, por ejemplo, amistad, cariño, relación laboral o de negocio, familiar, entre otras. El estudio de las redes sociales, se remonta mucho más allá de las redes informáticas que usamos a diario. Entre los investigadores que estudian las redes, los sociólogos tienen los estudios más antiguos, pero ellos hacían sus trabajos de forma empírica. Hay antecedentes claros de análisis de redes sociales que se encuentran en la literatura, que se remontan a finales del siglo XIX.

En las redes sociales, los vértices representan personas o grupos de personas, y los arcos o enlaces representan alguna forma de interacción social entre ellos, tales como la amistad. Los sociólogos han desarrollado su propio lenguaje para la discusión de las redes: se refieren a los vértices o nodos, como actores y a los arcos como lazos o vínculos, (algunas veces usaremos estos términos cuando se habla de redes sociales). Por lo tanto, una red social está integrada por individuos que por supuesto pueden ser de otras especies. En particular, aquellas que forman comunidades, por ejemplo: delfines, pájaros, peces, lobos, entre otros, además de las redes sociales formadas por personas. Algunas redes en particular pueden llegar a ser tan elaboradas y complejas que es imposible no preguntarse ¿cómo se forman?, ¿cómo funcionan?, ¿para qué sirven?, o hasta ¿por qué pertenecemos a ella?

En este trabajo, haremos un análisis comparativo de distintas redes sociales utilizando la teoría de redes, para lo cual consideramos distintas bases de datos. Dentro de las diferentes bases que estudiaremos, la base de datos que más nos interesa es la que esta conformada por los miembros de la Organización de Jóvenes Nikkei¹ (OJN) de la Asociación México Japonesa A.C. (AMJ), la cual fue fundada en septiembre de 1997, posterior a la IX Convención Panamericana Nikkei (COPANI IX), realizada en México en 1997, debido a que durante la convención se visualizó la gran falta de participación joven durante la última década (desde la desaparición del grupo Sansei). Durante este magno evento se contó con la valiosa colaboración de jóvenes de distintos países, que ayudaron a la elaboración de las mesas de trabajo realizadas durante la estancia de los jóvenes visitantes. El resultado del evento, fue la inmediata motivación e interés de los jóvenes nikkei mexicanos para crear un grupo juvenil, objetivo que se logró un par de meses después bajo el nombre de Organización Jóvenes Nikkei de la AMJ, bajo la dirección de Lydia Okuma Otori. La OJN es una organización representativa nikkei a nivel nacional dependiente de

¹El término nikkei se refiere a las personas que son descendientes de japoneses que radican fuera de Japón.

la Asociación México Japonesa que proyecta, en todo momento, los valores y tradiciones de la sociedad mexicana-japonesa al mundo. El lema es “MIRAI” (futuro en japonés): Máximo(a), Iniciativa, Respeto, Amistad, Integración. Las actividades que realizan son:

- Organizar el intercambio internacional Vibra Joven, el cual se realiza cada año a mediados del mes de julio desde el año 2006, se cuenta con la participación de jóvenes de toda la república mexicana y del extranjero.
- Reunir jóvenes y ayudar en su formación personal y social, relacionándolos con las colonias mexicana y japonesa formando un escalón generacional para el desarrollo, bienestar, unión y fortalecimiento de la sociedad.
- Apoyar la difusión cultural de las raíces mexicana y japonesa por medio de congresos, cursos, eventos y talleres.
- Realizar dinámicas de integración entre los miembros para generar lazos de amistad más fuertes, así como actividades que ayudan al desarrollo personal y profesional como conferencias, capacitaciones y lo más importante, experiencias que los ayuden a tomar decisiones para ser líderes de sus vidas.
- Animar y ayudar en la formación de jóvenes líderes que se sientan comprometidos con las colonias mexicano-japonesa y la sociedad que los rodea.
- Formar un equipo representativo nikkei nacional, el cual proyecta en toda ocasión los valores y las tradiciones del joven nikkei mexicano al mundo.
- Iniciar, estrechar y fortalecer los lazos amistosos entre los jóvenes de cada comunidad nacional e internacional nikkei, con la finalidad de ser parte de la maquinaria que impulse y aliente el espíritu de unión entre las distintas generaciones y las diferentes organizaciones involucradas en la trasmisión de la cultura japonesa.
- Crear y fomentar espacios libres dentro de los cuales los jóvenes tengan la oportunidad de expresar y realizar sus proyectos de vida. Los beneficios que se ofrecen a los jóvenes son: experiencia y amistad, así como poder desarrollarse en un ambiente sano adquiriendo experiencias en las cuales aprenden a ser mejores individuos.
- Participar en diversos intercambios internacionales como lo son: el Dale en Argentina, Movi-mente en Brasil, Fénix en Paraguay, Lidercambio en Perú, entre otros.

Actualmente, la organización cuenta con proyectos propios que poco a poco han tomado lugar dentro de las actividades tradicionales de cada año. El crecimiento de la organización no solo se ha limitado a la zona metropolitana del Distrito Federal, hoy en día el trabajo y dedicación de los miembros ha dado como resultado la creación y anexión de otras agrupaciones juveniles, las cuales siguen los mismo objetivos en varios estados de la república mexicana (Jalisco, Chihuahua, Morelos, Chiapas y Puebla). Dentro del ámbito internacional, se cuenta actualmente con una amplia red de comunicación panamericana, la cual nos ha permitido el desarrollo de planes de intercambio con Argentina, Bolivia, Brasil, Colombia, Paraguay, Perú, Uruguay, entre otros países, que cuentan con una gran infraestructura y desarrollo de la colonia japonesa.

Entonces, una vez que realicemos el estudio de las demás redes, podremos compararla con la red de la OJN, con la finalidad de tratar de identificar si existe alguna problemática específica debido a que, dicha organización ha presentado una baja en el número de miembros en los últimos años, por lo que nos interesa conocer las fortalezas, así como las debilidades de esta red. Además, nos interesa conocer si existe un nodo central dentro de la organización, así como el número de miembros, familias, enlaces, entre otras propiedades que nos pueden ayudar a comprender cómo es que esta red funciona y cómo es su estructura.

Por lo tanto, los capítulos a desarrollar son los siguientes, dentro del capítulo 1 encontraremos los antecedentes históricos de la teoría de redes, además se darán las definiciones de algunos conceptos importantes; en el capítulo 2, nos enfocaremos en específico en toda la parte de redes sociales, con un breve resumen de los orígenes y la investigación de éstas, estudiaremos distintas características, así como algunas propiedades; por último, en el capítulo 3 platicaremos sobre las bases de datos que utilizaremos, también, sobre el software para el análisis de las distintas redes sociales. Por último, se explicarán las conclusiones que se obtuvieron en la realización de este trabajo.

Capítulo 1

1.1 Contexto Histórico

La sociedad, el desarrollo de las ciudades, los avances tecnológicos, incluso, nosotros mismos, dependemos del funcionamiento de una gran diversidad de sistemas, los cuales pueden ser modelados mediante redes. Una red es un conjunto de entidades (nodos), unidos por una relación específica (arcos).

Los primeros estudios datan de 1736, con un problema conocido como “Los Puentes de Königsberg” en Prusia (actualmente, Kaliningrado, Rusia). El problema consistía en encontrar un camino que recorriera los siete puentes que atravesaban el río Pregel, sin repetir ninguno y regresando al punto de origen. Dicho problema fue resuelto por el matemático suizo Leonard Euler, quien estableció que el problema no tenía solución, sin embargo, se dio el inicio de la teoría de gráficas porque Euler representó el problema de una forma abstracta, donde la tierra la representó con puntos y los puentes con líneas, ver fig. 1. Podemos observar que la gráfica tiene 3 nodos de grado 3 y un nodo de grado 5.

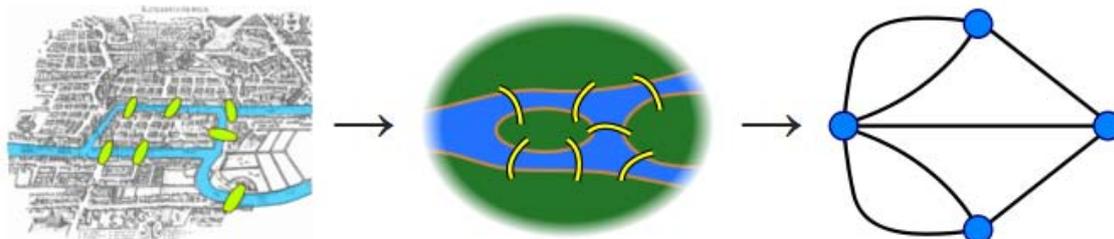


Figura 1: Representación gráfica del problema de los siete puentes de Königsberg.

Euler estableció que el problema de los Puentes de Königsberg no tenía solución, pero no fue hasta 1873 cuando Carl Hierholzer prueba que este problema tiene solución. Él determinó que este tipo de caminos existen sí y solo sí, cada vértice de la gráfica tiene un número par de aristas [12].

En general, el progreso en el estudio de las gráficas fue muy lento, en parte debido a la dificultad de trabajar con redes grandes. En particular, se hicieron varios trabajos empíricos en áreas sociales, los cuales aunque no fueron considerados formales, dieron guías para el estudio y análisis de las relaciones en una red.

En 1950, dos grandes matemáticos Paul Erdős y Alfred Rényi, propusieron un modelo de redes aleatorias [14], lo que da inicio a la “*Teoría de Redes*”. La aportación principal que hicieron Erdős y Rényi fue que dada una probabilidad de interconexión entre nodos, al irse incrementando, llega un momento en que la mayoría de los nodos están unidos en un componente gigante. Posteriormente, en 1967 Stanley Milgram, quien realizó un experimento en la Universidad de Harvard [21], en el cual envió 296 cartas a diferentes personas ubicadas en Wichita, Kansas y en Omaha, Nebraska. En dichas cartas se pedía a los receptores que reenviaran la carta a una persona en específico, que se ubicaba en Boston, Massachusetts. En las cartas se les indicaban las instrucciones, las cuales consistían en que se tenía que enviar la carta directamente a la persona en Boston en caso de que lo conocieran y, en caso contrario, se tenía que reenviar la carta a alguna persona que consideraran que tuviera una mayor probabilidad de conocer a dicha persona. El objetivo de este experimento consistía en encontrar cuántas personas en promedio se requerían para que la carta llegara a su destinatario. Al final del experimento, solo llegaron 64 cartas, aun así pudo establecerse que en promedio se necesitaron 6 personas, de ahí surge el concepto de *6 grados de separación*, que se refiere a que existe una distancia pequeña entre cualquier par de nodos en una red.

En 1973, otro estudio importante dentro de la redes, fue realizado por Mark Granovetter, quien publica un artículo en el cual desarrolló la idea de que los enlaces significativos en una red, no son las conexiones más fuertes, en cambio lo son las conexiones más débiles y más tenues. En particular, Granovetter se interesó en el mercado laboral, es decir, las relaciones que incrementa la posibilidad de encontrar un trabajo. El artículo en el cual establece sus resultados es: *The Strength of Weak Ties* (La Fuerza de los Enlaces Débiles) [15].

En 1998, los matemáticos Duncan Watts y Steven Strogatz llevaron a cabo un estudio en el cual se centraron en propiedades de conectividad, con lo que mostraron que las redes se podían clasificar con base en dos propiedades: la distancia y el coeficiente de aglomeración (en inglés Clustering). Cabe mencionar que, la distancia media es pequeña en la mayoría de las redes, esta propiedad recibe el nombre de *Mundo Pequeño*. Además, introducen un nuevo modelo, el *Modelo de Mundo Pequeño*. Estas investigaciones fueron publicadas en el artículo *Collective Dynamics of Small-World Networks* (Dinámica colectiva de las redes de mundo pequeño) [36]. Un año más tarde, en 1999, Albert-Lászlo Barabási y Reka Albert estudiaron redes que presentan nodos con mayor conectividad que los demás nodos, por lo tanto tienen un mayor grado de conexión que el grado promedio de los nodos de la red. A estos nodos con una conectividad muy alta, se les llaman hubs. Estas investigaciones fueron publicadas en el artículo *Emergence of scaling in random networks* (La emergencia de la escala en las redes aleatorias) [3].

En 2003, Duncan Watts junto con Peter Dodds y Roby Muhamad, decidieron hacer una réplica del experimento de Milgram, pero ahora a una escala global, en la cual emplearon el correo electrónico en vez de cartas. En este experimento utilizaron a poco más de 98,000 personas, en su mayoría estadounidenses, en el cual se les pidió a estas personas que enviaran un correo a una de las 18 personas objetivo situadas en 13 países distintos, con la condición de que tenían que enviar dicho correo a alguno de sus conocidos que ellos consideraran tenían más probabilidades de conocer a dicha persona objetivo. Los resultados arrojaron que la información corría por enlaces débiles, por lo que no se

requirieron grandes conexiones como hubs para pasar la información, se estimó que los pasos necesarios están en una media de entre 5 y 7 pasos (además los estudios realizados por Milgram, 6 grados de separación) [10].

1.2 Definiciones ²

Sea una red³ $G = (N, A)$, donde N es el conjunto de nodos o vértices, y A es el conjunto de arcos o aristas o enlaces. En el contexto de una red social, los nodos reciben el nombre de actores y los arcos de vínculos, mientras que en teoría de gráficas a los nodos se les conoce como vértices y a los arcos como aristas. Se han desarrollado una gran cantidad de métricas y herramientas para clasificar y analizar las redes, una medida básica es el grado de un nodo k , el cual consiste en el número de enlaces incidentes en dicho nodo. En ocasiones una red puede tener una dirección privilegiada donde los arcos se representan con flechas (ver figura 2), en este caso, el grado se divide en grado interior (el arco entra al nodo), y grado exterior (el arco sale del nodo). La distribución de grado $P(k)$ es fundamental en el conocimiento de la topología de una red, se determina graficando el grado de los nodos vs el número de nodos. Las redes con topología homogénea presentan una distribución tipo Poisson para N grande, debido a que todos los nodos poseen más o menos el mismo grado. En cambio, las redes heterogéneas presentan una distribución libre de escala, la cual en un plano log-log representa una línea recta con pendiente negativa, debido a que existen nodos con un alto grado de conectividad llamados hubs, es decir, unos cuantos nodos atraen la mayoría de las conexiones y la gran parte de los nodos restantes tienen que compartir los enlaces que sobran.

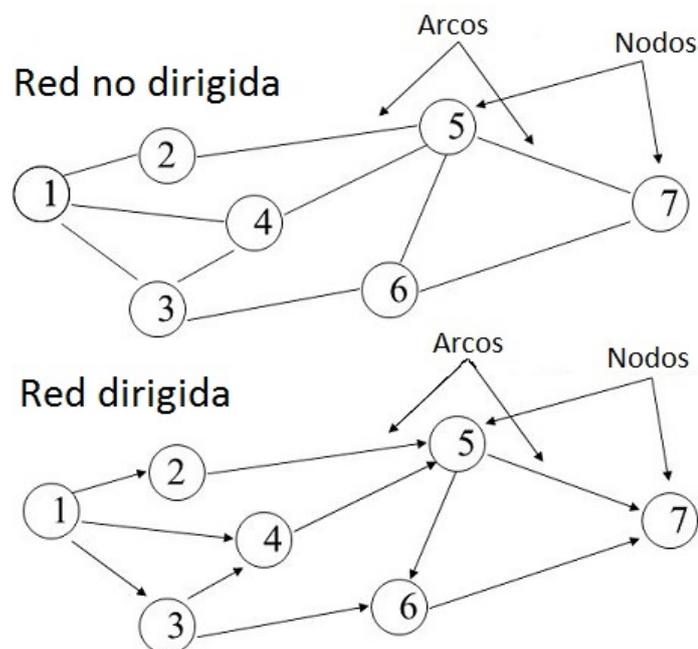


Figura 2: Descripción gráfica de una red con nodos y arcos.

Una de las maneras que tenemos para representar a las redes es por medio de matrices, la matriz de adyacencia nos da una representación de la relación nodo-nodo que se tiene en una red, de forma que tenemos una matriz de $N \times N$, en donde la entrada ij toma el valor de 1 si existe un arco del nodo i al nodo j , en otro caso es 0. En la figura 3 se muestra una red con su respectiva matriz de adyacencia. Este tipo de matrices representa quién

²Para mayor detalle sobre las definiciones y conceptos revisar las referencias [5], [7], [24] y [34].

³Cuando hablamos de red o gráfica nos estamos refiriendo al mismo concepto, dentro de este trabajo utilizaremos ambos conceptos de igual forma.

está cerca de quién, de hecho dentro del análisis de redes sociales a estas matrices se les conoce como sociomatriz.

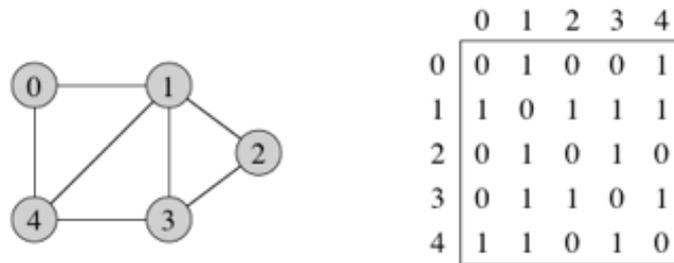


Figura 3: Matriz de Adyacencia

La pareja ordenada (i, j) , representa el arco que va del nodo inicial i al nodo final j , además, es común representar el arco con una flecha, $i \rightarrow j$. Un camino entre el nodo i y el nodo j , es una secuencia alternada de nodos y arcos, la cual comienza en el nodo i y termina en el nodo j . De hecho, hemos escuchado las palabras camino, trayectoria, ruta y podríamos pensar que se refieren a lo mismo o que son sinónimos, pero en el caso de las redes estas palabras nos llevan a conceptos distintos, es decir, cuando hablamos de una ruta nos referimos a un camino en el cual los nodos son distintos, en cambio cuando nos referimos a una trayectoria⁴ estamos hablando de un camino en el cual los arcos son distintos. Además, una ruta o trayectoria cerrada, en las cuales no es posible identificar el nodo inicial del nodo final, se conocen como ciclo y circuito, respectivamente. Se dice que una red es conexa si existe un camino entre cada par de nodos que conforman la red, es decir, todos los nodos están conectados, no hay nodos aislados.

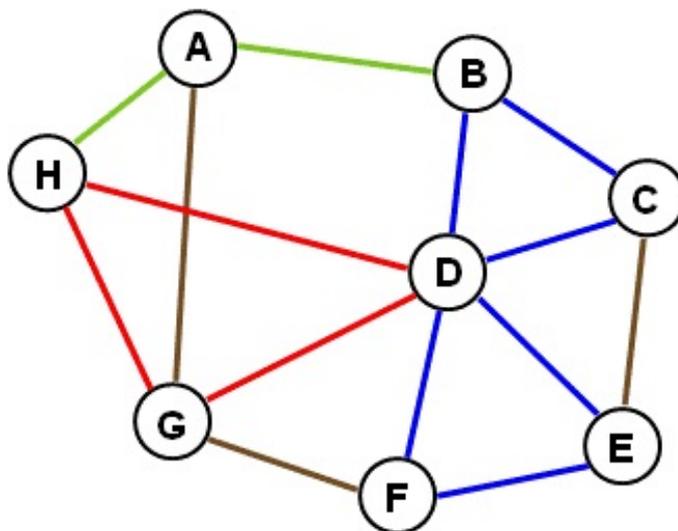


Figura 4: Ejemplo de camino, ruta, ciclo, trayectoria, circuito

⁴En teoría de gráficas se conoce como trayectoria a la ruta y como paseo a la trayectoria.

En la figura 4, mostramos gráficamente las definiciones anteriores, por ejemplo, si seguimos los nodos $A \rightarrow B \rightarrow D \rightarrow G \rightarrow H \rightarrow D \rightarrow F \rightarrow E \rightarrow C$ ⁵ tenemos un camino, pero en el caso que siguiéramos la secuencia de nodos $A \rightarrow G \rightarrow H \rightarrow D \rightarrow F \rightarrow E \rightarrow C \rightarrow B$ tenemos una ruta, si tuviéramos esta misma ruta, solo que al final se conectará con A, es decir, $A \rightarrow G \rightarrow H \rightarrow D \rightarrow F \rightarrow E \rightarrow C \rightarrow B \rightarrow A$ estaríamos hablando de un ciclo. Ahora bien, en el caso de que tuviéramos lo siguiente $A \rightarrow H \rightarrow G \rightarrow D \rightarrow F \rightarrow E \rightarrow D \rightarrow C \rightarrow B \rightarrow D$ estamos hablando de una trayectoria, pero si tuviéramos la misma trayectoria pero añadiendo de $D \rightarrow G \rightarrow A$, es decir, quedaría como $A \rightarrow H \rightarrow G \rightarrow D \rightarrow F \rightarrow E \rightarrow D \rightarrow C \rightarrow B \rightarrow D \rightarrow G \rightarrow A$ estamos hablando de un circuito.

Si queremos conocer un poco sobre las métricas de una red, podemos empezar por calcular la longitud de ruta, esta nos dice el número de arcos que contiene una ruta, sin embargo, no solo podemos saber cuántos arcos están involucrados en nuestra ruta, sino que además podemos conocer tanto la distancia como la distancia media (p). La distancia entre dos nodos i y j , se define como el mínimo número de arcos que conectan a los nodos i y j , en tanto que la distancia media nos da la distancia de forma más general, ya que nos dice el promedio de las distancias entre todo par de nodos. Por último, podemos hablar del diámetro de una red, el cual es la ruta más larga de entre todas las rutas más cortas entre todo par de nodos.

Por otro lado, cuando hablamos de la vecindad ⁶ de un nodo, nos referimos a los nodos que se encuentran a la misma distancia de un determinado nodo, es decir, existe un número de arcos fijo de separación hacia un determinado nodo. La eficiencia de una red mide el intercambio de información dentro de la red, se calcula mediante el inverso de la ruta más corta ($\epsilon_{ij} = \frac{1}{d_{ij}}$) y se refiere, a cómo se mueve la información dentro de una red. Se puede calcular la eficiencia global en la cual se considera toda la red o se puede calcular la eficiencia local de un nodo en específico, en el cual se considera la primera vecindad del nodo en cuestión. En el caso de que no exista la ruta entre el nodo i y el nodo j entonces, la distancia entre el nodo i y j es infinito, por lo tanto, la eficiencia es 0.

Además, existen diferentes tipos de redes como son redes regulares, redes vacías y, redes completas. Las redes regulares, son aquellas en las que los nodos que conforman la red tienen el mismo número de arcos o grado, mientras que las redes vacías, son aquellas redes las cuales no tienen arcos, por lo que es una red no conexa, y por último, las redes completas, son aquellas redes en las cuales existe un arco entre cada par de nodos. Por otra parte, los árboles tienen características especiales ya que son acíclicos (no presentan ciclos), son conexos, es una red no dirigida, cuenta con N nodos y $N - 1$ arcos, existe una ruta única entre cada par de nodos y, además, si eliminamos un arco se desconecta, pero si añadimos un arco se forman ciclos, y en ambos casos ya no estaríamos hablando de un árbol. La densidad de una red, es una medida expresada en porcentaje del cociente entre el número de relaciones existentes con los posibles enlaces, es decir, nos muestra la alta o baja conectividad de la red.

⁵En gráficas se utiliza la siguiente notación (A,B), (B,D), (D,G), (G,H), (H,D), (D,F), (F,E), (E,C). Aunque la notación que estamos utilizando es más utilizada en teoría de redes.

⁶Cuando hablamos de vecindad en gráficas, nos referimos a los vértices adyacentes que tiene un vértice en específico.

Una red está constituida por un conjunto de componentes, estos componentes constan de nodos y arcos, incluso podemos hablar de un solo nodo y decir que es un componente, utilizamos el término componente gigante cuando la mayoría de los nodos se encuentran conectados. Es importante definir los componentes fuertemente conexos y débilmente conexos. El componente fuertemente conexo se define en una red dirigida como, el subconjunto de nodos en que para cualquier par de nodos (i, j) existe una ruta directa desde el nodo i hasta el j y viceversa, además, puede consistir de un solo nodo. Cuando hablamos de componente débilmente conexo, nos referimos a que dos o más nodos están conectados por una o más rutas sin importar la dirección.

Por otro lado, los cliques ⁷ son importantes dentro de las redes, si aplicamos la definición formal, ésta nos dice que es un subconjunto máximo de nodos dentro de una red no dirigida, en la que cada uno de los nodos está conectado a todos los demás nodos, es decir, es una subred o subgráfica completa dentro de la red. El k -core de una red, es la subgráfica más grande en la cual todos sus nodos tienen al menos grado k . La colección completa de los k -cores, es decir, $k=1, 2, 3, \dots$, nos da una descripción más detallada de la red, que cuando únicamente se consideran los componentes, en el caso de un árbol (red acíclica y conexa), solo existe el 1-core.

Un concepto que utilizamos con frecuencia es el coeficiente de aglomeración (*Clustering*), el cual es una forma de calcular la tendencia a formar triángulos dentro de la red, también conocido como transitividad de un nodo. Para determinarlo, se calcula el cociente de el número de arcos entre los vecinos del nodo i , y el número máximo de los posibles enlaces entre los vecinos del nodo i , esto se puede ver como la cantidad de triángulos conectados al nodo dividido entre los triángulos incidentes a él, es decir, nos indica la tendencia que tiene la red de formar triángulos. Podemos calcular el coeficiente de aglomeración local (la vecindad del nodo) y general (toda la red). Cuando el clustering es igual a 0 nos dice que hay nodos aislados, es decir, que los nodos no están conectados por ningún arco. Si el clustering es igual a 1, esto nos indica que todos los nodos están conectados por un arco directo. Los nodos que presentan un grado bajo tienden a pertenecer a comunidades locales bien interconectadas, mientras que los hubs se conectan con muchos nodos que no están directamente interconectados. Un ejemplo rápido del coeficiente de aglomeración dentro de las redes sociales es, si dos personas A y B se conocen y B conoce a C , entonces existe una probabilidad alta de que A y C se conozcan, por lo que se dice también que se tiene una relación transitiva y se forma un triángulo.

Clustering local:

$$C_i = \frac{2|C_3(i)|}{k_i(k_i - 1)}$$

Clustering global:

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i$$

Donde:

$C_3(i)$ = número de triángulos que conectan al nodo i

⁷En gráficas los cliques se conocen como clanes.

k_i = grado del nodo i
 n = número de nodos

La centralidad intermedia (*Betweenness centrality*), es una medida de distancia por medio de la cual se calcula la centralidad de cada uno de los nodos, es decir, nos dice qué tan importante es un nodo dentro de la red. La idea es que un nodo es central siempre y cuando esté situado en la ruta más corta para llegar a todo par de nodos, si existen varias rutas más cortas, se asume que todos tienen la misma probabilidad de ser usadas. Los nodos con alta intermediación juegan el rol de conectar a los diferentes grupos. La siguiente fórmula nos muestra la manera en la que se define:

$$C_B(i) = \sum_{j,k \neq i} \frac{g_{jk}(i)}{g_{jk}}$$

Donde:

$g_{jk}(i)$ = longitud de las rutas más cortas que conectan al nodo j y al nodo k , que pasan a través del nodo i

g_{jk} = ruta más corta entre los nodos j y k

Una medida de centralidad más sofisticada es el grado de cercanía (*Closeness centrality*), debido a que mide el inverso de la distancia media desde un nodo hacia los demás nodos de la red. El grado de cercanía puede ser considerado como una forma de medir cuánto tiempo tomará para que la información se propague de un determinado nodo a los otros nodos de la red, razón por la cual es muy utilizado en el análisis de redes sociales. Se calcula de la siguiente manera:

$$C_C(n_i) = \sum_j \frac{1}{p_{ij}}$$

Donde:

p_{ij} = distancia media entre el nodo i y el nodo j

n_i = nodo i

Mark Newman propuso una clasificación importante conocida como asortatividad (*Assortativity*), la cual nos indica la preferencia de los nodos de una red por unirse a otros nodos que son similares en alguna característica. Es decir, los nodos con un grado alto tienden a conectarse a otros nodos con un grado alto, por lo que la comunicación fluye más rápido. En caso contrario, una red es disortativa (*Disassortative*) cuando nodos de bajo grado se conectan con hubs, por lo que hay una tendencia a la generación de cuellos de botella, esto significa que la comunicación no fluye tan rápido por lo que se puede llegar a saturar la red. Estudios realizados han encontrado que las redes tecnológicas y biológicas, presentan una estructura negativa, es decir, son disortativas, en cambio en las redes sociales presentan un número positivo, por lo tanto son asortativas. El coeficiente de asortatividad depende de la correlación, se calcula con la siguiente fórmula:

Coeficiente de Asortatividad:

$$r = \frac{\sum_{jk} jk(e_{jk} - q_j q_k)}{\sigma_q^2}$$

Donde:

e_{jk} = es el exceso de la probabilidad conjunta del exceso de grado de j y del exceso de grado de k (el exceso de grado de un nodo también conocido como el grado restante, es igual al grado del nodo menos 1)

$q_k = \frac{(k+1)p_{k+1}}{\sum_j j p_j}$ es la distribución normalizada del exceso del grado D_r de un nodo elegido aleatoriamente

σ_q = desviación estándar de q_k

La fórmula anterior nos da valores de r para la cual $-1 < r < 1$, esto significa que cuando r nos da un valor de 1 o muy cercano, se dice que la red es asortativa, cuando r es igual a 0 esto nos indica que la red es neutra y cuando r nos da un valor de -1 o muy cercano, nos dice que la red es disortativa.

Por otro lado, la robustez es un indicador el cual nos ayuda a saber qué tan vulnerable es nuestra red a perturbaciones. Existen dos tipos de perturbaciones, la primera se refiere a fallas internas que se presentan dentro de la red, esto es, debido a una posible falla o a un problema interno, a esto se le conoce como error interno, mientras que el segundo tipo de perturbaciones son los ataques, en este caso la falla o el problema está fuera de la red, se pueden tener ataques dirigidos a un componente específico. Si el componente gigante continúa, sin ningún daño grave se dice que la red resiste estos errores o ataques por lo que la red es robusta.

Como vimos en el contexto histórico, uno de los principales resultados del experimento realizado por Milgram fueron los 6 grados de separación, lo que llevó a que posteriormente se estudiara la propiedad de mundo pequeño. La propiedad de mundo pequeño consiste en el hecho de que la distancia media entre cualesquiera dos nodos es muy pequeña, es decir, si un nodo dentro de una red tiene pocos nodos conectados a él y pocos alejados de él, entonces la mayoría de los nodos están a una distancia media que es relativamente corta. Lo anterior se refiere a que a partir de un determinado nodo, casi todos los nodos están a muy pocos pasos de él, por lo que aunque añadamos muchos nodos a la red con sus respectivos arcos, la distancia media no aumentaría mucho, ya que existe una ruta más corta promedio entre los nodos, es decir, que aunque no todos los nodos son vecinos entre sí, se puede llegar a la mayoría de los nodos en un número de pasos pequeño.

1.3 Modelos de Redes Complejas

1.3.1 Modelos de Redes Aleatorias (Modelo ER)

El primer modelo de redes complejas que se conoció fue el modelo de redes aleatorias, el cual fue propuesto por los matemáticos húngaros Erdős y Rényi. En este modelo se inicia con un número de nodos N , los cuales se enlazan con la misma probabilidad, por lo que en este tipo de redes todos los nodos tienen más o menos el mismo número de enlaces. Las redes que poseen un número de nodos N pequeño siguen una distribución de grado binomial, pero en el caso en el que el número de nodos N es grande siguen una distribución de grado Poisson, es decir, presentan una conectividad media porque la mayoría de los nodos tienen el mismo grado. En estas redes la distancia media tiende a $p \sim \log N$, mientras que el clustering generalmente es bajo, ya que tiende a k/N , donde k es el grado promedio de los nodos. En redes aleatorias, los vecinos de un determinado nodo se eligen al azar, por lo que no hay correlación entre el grado de nodos vecinos. Por último, estas redes son más robustas a ataques dirigidos, pero a su vez son vulnerables ante errores internos. A continuación en la figura 5 se da un ejemplo gráfico del modelo de redes aleatorias.

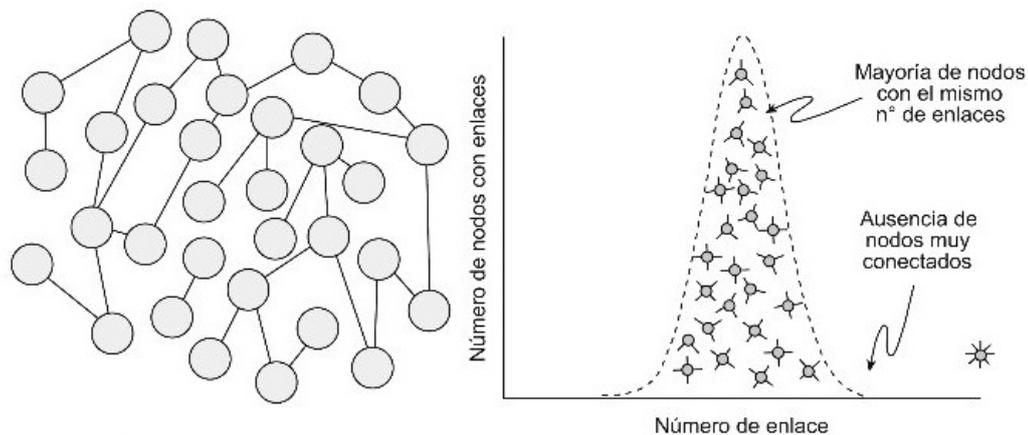


Figura 5: Modelo de Redes Aleatorias

1.3.2 Modelos de Redes de Mundo Pequeño (Modelo WS)

El modelo de Mundo Pequeño fue introducido por Watts y Strogatz en 1998, en este tipo de redes la distancia entre dos nodos elegidos aleatoriamente crece proporcionalmente al logaritmo del número de nodos N que contiene la red. El modelo consiste en la interpolación de una red regular y una red aleatoria (ver figura 6), el ejemplo más sencillo consiste en un anillo con N nodos que se conectan con sus primeros k vecinos, los cuales tienen una misma probabilidad de conexión, en este procedimiento se evita tener multiarcos y rizos o lazos (multiarcos se refiere a dos o más arcos que son incidentes a dos nodos, rizos o lazos son arcos que conectan un nodo consigo mismo). Este modelo ha sido observado en redes, que presentan a nivel global una baja distancia media (su distancia media tiende a $p \sim \log N$), y a nivel local presentan un clustering alto. Además, estas

redes siguen una distribución Poisson, es decir, presentan una conectividad media porque la mayoría de los nodos tienen el mismo grado. Las redes de mundo pequeño son muy eficientes con respecto al intercambio de información local y globalmente. Por otro lado, son más robustas a ataques dirigidos, pero a su vez son vulnerables ante errores internos. Este tipo de redes se han observado en redes sociales.

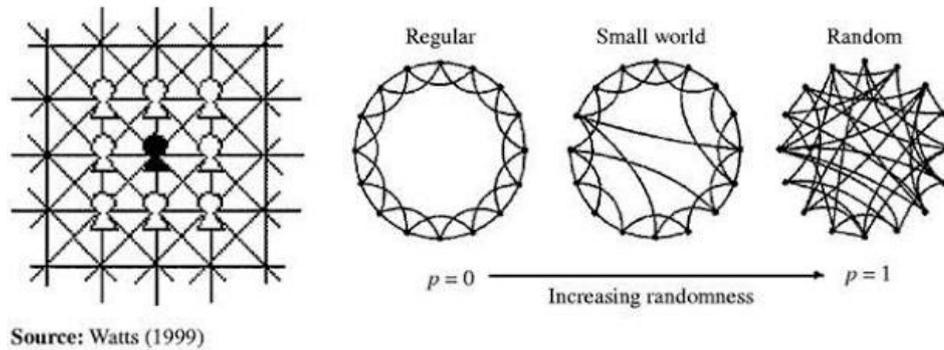


Figura 6: Modelo de Mundo Pequeño

1.3.2 Modelos de Redes de Libre Escala (Modelo BA)

En 1999, Barabási y Albert plantearon un modelo en el cual buscaban la representación de redes reales, a este modelo se le conoce por el nombre de modelo de redes de libre escala. Este es un modelo del crecimiento de una red, en el cual desarrollaron el concepto de preferencia de enlace, la cual empieza con un conjunto pequeño de nodos (incluso sólo dos o tres nodos) conectados de forma aleatoria. Posteriormente se añaden nuevos nodos con un ritmo constante, en el cual cada uno de estos nuevos nodos lleva un número determinado de nuevos enlaces. La regla para añadir los nuevos nodos es muy simple y es la siguiente, los nodos entrantes prefieren conectarse con los nodos antiguos los cuales ya tienen muchos enlaces (ver figura 7). En principio los nuevos nodos se pueden conectar a cualquier nodo, pero cuanto mayor sea el grado de un nodo antiguo, mayor es la probabilidad de que atraiga a un nuevo nodo, generando una tendencia a la formación de hubs. Cabe señalar que los nodos menos conectados recibirán nuevos enlaces, pero en la mayoría de los casos los hubs serán mucho más atractivos. Por lo tanto, surge una jerarquía de nodos, la cual va desde los menos conectados hasta los hubs, dando como resultado una red heterogénea, con una distribución de grado la cual corresponde a la ley de potencia, $P(k) \sim k^{-\gamma}$ donde γ es el exponente del grado, esto indica que un número pequeño de nodos tiene alto grado de conectividad, en un plano log-log se observa que la gráfica tiene una pendiente negativa. Además, debido a la presencia de los hubs, la distancia media de estas redes es ultra pequeña y tiende a $p \sim \log \log N$. Por último estas redes son más vulnerables a ataques dirigidos a hubs, pero a su vez son más robustas a errores internos.

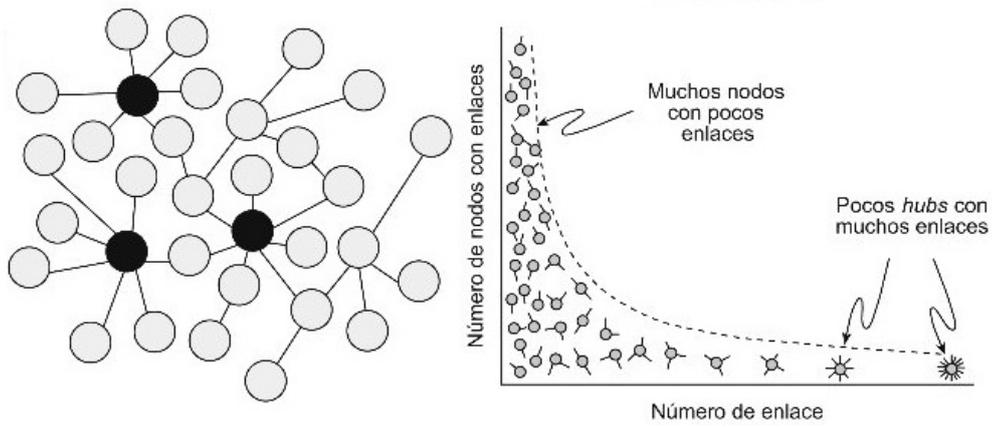


Figura 7: Modelo de Libre Escala

Finalmente, en la figura 8 mostramos una forma sencilla de visualizar los 3 diferentes modelos.

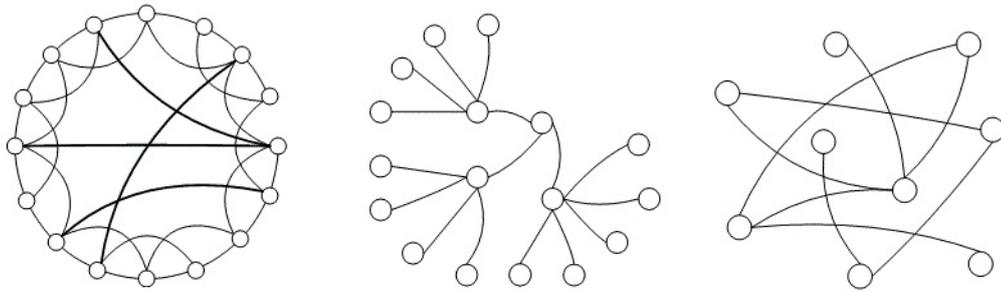


Figura 8: a) Red de Mundo Pequeño b) Red de Libre Escala c) Red Aleatoria

Capítulo 2

Estado del Arte de Redes Sociales

En las redes, la forma de una red también llamada la estructura o la topología, es más importante que las métricas geográficas, es decir, es más importante cómo están conectados los elementos que su posición y distancia. Por lo que en las redes sociales, la relevancia de la topología significa que la estructura social es lo que más cuenta.

Una gran parte de la investigación en redes sociales, se dedica a estudiar cómo los distintos tipos de vínculos se influyen entre sí. Dos personas pueden tener distintos tipos de relaciones, ya que pueden compartir gustos, ideas, actitudes, religión, orientación sexual, además pueden ser amigos, pareja, compañeros de trabajo, familiares, jugar en un mismo equipo, entre otros. Por lo que en general tendremos relaciones de transacción de recursos materiales y no materiales (intercambios comerciales), relaciones de comunicación (transmisión de mensaje entre los actores), relaciones instrumentales (los actores se relacionan buscando algún tipo de beneficio), relaciones afectivas (se trata del afecto, admiración, amistad que existen entre los individuos) y, relaciones de poder y autoridad (redes de organizaciones con distintas jerarquías), relaciones de parentesco (relaciones que expresan lazos familiares). Pero simultáneamente se podría tener dos o más relaciones, mientras que en otros casos hay una parte en la que se ignora una parte de la relación de y, otra parte de la relación es conocida, por ejemplo, en el caso de artistas o personas famosas que tienen un grupo de fans, los fans conocen al artista, pero el artista no conoce a cada uno de sus fans. La sociología ha clasificado una gran variedad de relaciones entre individuos, la tendencia a tener varios tipos de relaciones en las redes sociales se denomina multiplexidad.

Los sistemas sociales parecen estar estrechamente conectados en un grupo suficientemente grande de desconocidos, por lo que no es tan difícil encontrar pares de personas con cadenas de relaciones bastante cortas que los unen, ejemplos sencillos de este hecho son, conocemos a alguien nuevo en nuestro grupo, pero al platicar con esta persona descubrimos que va en la misma escuela que nosotros o es primo de un amigo o toma clases con alguien cercano a nosotros y este descubrimiento siempre se toma con sorpresa y se llega a comentar frases como ¡qué pequeño es el mundo! o ¡el mundo es un pañuelo!

Jacob Moreno fue un psiquiatra que se interesó por el estudio de las redes sociales, lo que lo llevó a que, en 1932 hiciera un trabajo en el cual considera como *átomo social* no tanto al individuo, sino al individuo y las relaciones interpersonales de aceptación y rechazo que mantiene a su alrededor formando una red. Es decir, J. Moreno se enfoca en relaciones de amistad, sin embargo, sus estudios son dirigidos hacia el análisis social.

La Figura 9, por ejemplo, muestra un diagrama dibujado a mano del libro de Moreno, *Who shall survive? (¿Quiénes sobrevivirán?)* [22], que representa a las amistades dentro de un grupo escolar. Los triángulos representan a los niños y los círculos a las niñas, la cifra revela, entre otras cosas, que hay muchas amistades entre dos niños o dos niñas pero pocas entre un niño y una niña.

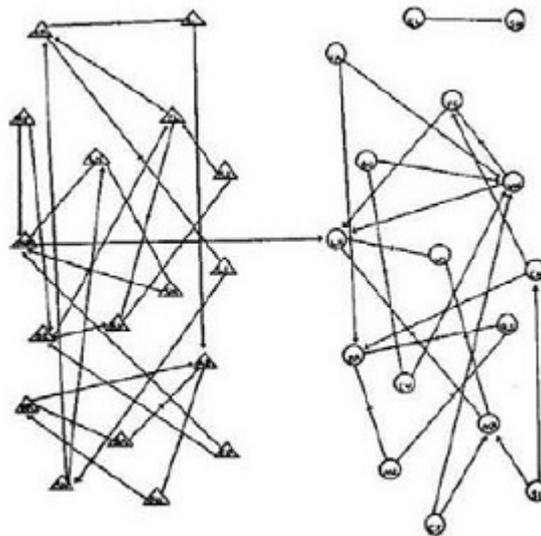


Figura 9: Representación del estudio de J. Moreno a un grupo escolar en la década de 1930.

Sin embargo, cabe señalar otra de sus aportaciones, ya que es el inventor del sociograma, que se define como la investigación sobre la evolución y organización de grupos y, la posición de los individuos dentro de dichos grupos. El sociograma es un método sistemático para la representación gráfica de un grupo mediante un diagrama bidimensional, en el que los actores son representados mediante puntos y las relaciones existentes entre ellos mediante líneas dibujadas entre pares de puntos con conexiones directas. Mediante el sociograma se busca:

- Conocer el nivel de aceptación que una persona tiene en su grupo.
- Evaluar el grado de cohesión entre personas de un grupo.
- Localizar a los individuos más rechazados y más valorados (líderes potenciales).
- Localizar a los sujetos aislados, que no despiertan ni admiración ni rechazo.
- Comprobar las consecuencias de la incorporación de nuevas personas al grupo.
- Verificar el grado de aceptación e incorporación de personas a un nuevo lugar de trabajo.

De igual forma, Moreno introdujo el término de *estrella sociométrica*, que es aquel individuo que goza de una mayor popularidad.

A principios de la década de los años 60, la socióloga Harriet Zuckerman entrevistó a científicos los cuales habían sido ganadores del Premio Nobel, ella quería descubrir qué había de especial en su forma de trabajar y también, cuál era el secreto de su éxito dentro de la investigación, para lo cual encontró una respuesta recurrente dentro de los premiados, un Nobel de Física respondió que el mundo es raro a la hora de reconocer los méritos, ya que suelen reconocer a la gente que ya es famosa. Un Nobel de Química añadió que cuando la gente veía su nombre en un artículo, solían recordar su nombre, pero no el de los demás colaboradores. Mientras que un Nobel de Fisiología y Medicina dijo que, normalmente cuando lees un artículo científico te fijas en los nombres que ya conoces y no en la larga lista de los otros autores, además, la persona más conocida es la que se lleva el mayor reconocimiento. Por lo tanto, estas observaciones llevaron a que en 1968 el sociólogo Robert Merton, esposo de Harriet Zuckerman, propusiera una ley brillante, bajo el efecto de un mecanismo social, a la cual llamó efecto Mateo [20]- “Los ricos se vuelven más ricos y los pobres se vuelven más pobres” - (el nombre procede del Evangelio de San Mateo 25:29, “Porque al que tiene, le será dado, y tendrá más; y al que no tiene, aun lo que tiene le será quitado”), ya que este mecanismo desempeña una función en el reparto de premios, financiamiento, prestigio, visibilidad, etc. Los científicos que cuentan con un buen número de activos adquieren más con facilidad, al contrario, los que no tienen tanto prestigio pasan grandes apuros para poder obtener y conservar los apoyos.

En 1973, el sociólogo Mark Granovetter entrevistó a una serie de profesionistas de Boston, los cuales habían conseguido recientemente un trabajo a través de contactos personales, les preguntó que con qué frecuencia tenían contacto con la persona la cual les había informado de la oferta de trabajo antes de obtener dicho empleo, la mayoría de las respuestas fueron ocasionalmente, mientras que otra parte respondió que raramente. De hecho, las ofertas de trabajo llegaban por parte de viejos amigos de la universidad o preparatoria, antiguos compañeros de trabajo, entre otros, más que por los amigos íntimos o cercanos. Las amistades mutuas eran los canales por los cuales se redescubrían estas conexiones, a este fenómeno Granovetter lo denominó como *la fuerza de los enlaces débiles* [15] (ver figura 10). Granovetter explicaba este fenómeno dibujando el círculo de conocidos que tenía un individuo hipotético llamado Ego, a estas redes se les conoce como redes ego o egocéntricas. Entonces, Ego se relaciona todos los días con familia y amigos cercanos, en particular estas personas también pueden estar en contacto entre sí, por lo que la información viaja muy rápido dentro de ese grupo, sin embargo, las conexiones débiles lo conectan con personas alejadas, ya que estas personas no pertenecen al entorno social de Ego, por lo tanto, se abren grupos de nuevos amigos los cuales cuentan con información que de otra forma sería inaccesible. Si se desaprovechan las oportunidades que se tienen con los enlaces débiles, se pueden causar muchos problemas en organizaciones o empresas o instituciones, ya que la información queda atrapada dentro del grupo sin llegar a los que la necesitan. Las ideas de Granovetter hicieron que surgiera la teoría del capital social, la cual supone que los contactos de una persona le permiten acceder a recursos que en última instancia proporcionan mejores trabajos y promociones más rápidas. Por lo que la posición de una persona en su red social, es crucial para determinar oportunidades, limitaciones y resultados futuros.

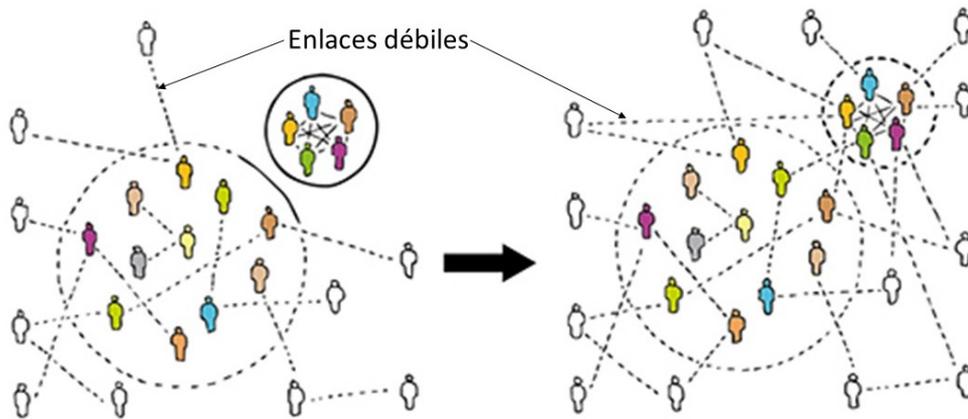


Figura 10: Representación enlaces débiles de Granovetter.

En 1976, el físico e historiador de la ciencia Derek de Solla Price, analizó un gran número de artículos científicos que estaban conectados por citas mutuas, comprobó que los artículos que contaban con un buen número de citas en un momento determinado tendían a adquirir más citas después, mientras que aquellos que solo tenían unas pocas al principio no incrementaban tanto el número de citas. Para su estudio, demostró matemáticamente que esta sencilla regla podía explicar la aparición de artículos muy citados (hubs de la red de citas), más aún, mostró que este mecanismo podía explicar por qué la distribución del número de citas por artículo, mostraba la cola larga característica de una ley de potencia.

En 1993, los sociólogos John F. Padgett y Christopher K. Ansell [28], analizaron la información sobre bodas, relaciones económicas y relaciones de patronazgo, que conectaban a la familia Medici con otras familias importantes de la ciudad. La familia Medici era una familia importante dentro de Italia, Cosme de Medici fue el hombre que en el siglo XV consiguió para su familia el poder en Florencia, lo describían como una esfinge inescrutable. Estos sociólogos encontraron que, la familia de Cosme de Medici era el centro de una red de vínculos con los linajes más importantes, sino fuera por esta conexión, muchas de las familias estarían débilmente relacionadas o incluso enfrentadas. La red de los Medici es un ejemplo de las redes ego descritas anteriormente (red compuesta de un conjunto de nodos con enlaces directos a uno central, además de los enlaces que los vinculan entre ellos), cuando dentro de la red falta uno de estos enlaces débiles que nos conectan a un grupo de amigos con otro grupo de amigos, la red presenta agujeros estructurales (por ejemplo, dos vecinos del ego no son vecinos entre sí). Padgett y Ansell, se dieron cuenta de que la red de los Medici estaba llena de agujeros estructurales y su familia los ocupaba como estrategia «divide y vencerás», se consideraban parte interesada en muchos conflictos y las familias tenían que pedir su intermediación en las relaciones con otras familias.

Por otra parte, Paul Lazarsfeld y Robert Merton [17] proponen los conceptos de homofilia y heterofilia, para entender cómo se transmiten los mensajes en las redes sociales. El concepto se basa en el principio básico de la comunicación, que sostiene que el intercambio de ideas ocurre con mayor frecuencia entre iguales, por lo que homofilia es el grado de similitud entre dos individuos que se comunican. La similitud ocurre porque la comunicación es más efectiva cuando emisor y receptor son parecidos y comparten diferentes atributos como edad, educación, creencias, clase social, entre otros. Mientras más

parecidos son dos individuos, más efectiva será la comunicación entre ellos, al contrario, comunicarse con personas diferentes implica un mayor esfuerzo para lograr una buena comunicación. Por su parte, la heterofilia es el grado de diferencia en ciertos atributos de dos individuos que interactúan. Tiene un potencial especial ya que conecta a dos mundos distintos. La homogamia es un caso concreto de la homofilia, la homogamia es la tendencia de las personas a casarse con personas muy parecidas a ellas, es un fenómeno muy fuerte incluso en las sociedades en las que el matrimonio no está concertado y teóricamente, cualquier persona podría casarse con cualquier persona. Un estudio realizado a mediados de los años noventa por el sociólogo Edward O. Laumann, mostró que tres cuartas partes de las parejas casadas de Estados Unidos comparten muchos rasgos similares, por ejemplo, pertenecen a la misma clase social, mismo grupo étnico y nivel educativo, comparten ideas políticas, actitudes, creencias religiosas e incluso niveles parecidos de atractivo físico. Cuando el vínculo es otro tipo de relación, la proporción es mucho menor, aunque sigue siendo alta.

Los científicos han diseñado numerosas estrategias para representar o dibujar las redes sociales, estas van desde cuestionarios hasta muestreos de bola de nieve (snowball sampling), un sistema en el cual el sujeto entrevistado sugiere a alguien de su círculo como sujeto de la siguiente entrevista. Estas estrategias han permitido recoger datos como el mapa de las relaciones sexuales entre personas de distintos grupos, el conocimiento de estas redes permite comprender mejor cómo se difunden las enfermedades de transmisión sexual.

Un estudio realizado durante los años noventa, mostró que por cada estadounidense de raza blanca con una enfermedad de transmisión sexual, hay hasta veinte afroamericanos con la misma enfermedad en algunas zonas de Estados Unidos, esta cifra se consideró el resultado de desigualdades raciales, sin embargo, los mecanismos de contagio que generaron esta gran diferencia solo se conocían parcialmente. En 1999, los sociólogos Edward O. Laumann y Yoosik Youm encontraron evidencias interesantes, los afroamericanos con menos actividad sexual (los que solo tuvieron una pareja el año anterior), tenían cinco veces más posibilidades de tener relaciones con afroamericanos más activos sexualmente (aquellos que habían tenido cuatro o más parejas el año anterior), que los blancos en la misma situación. Es decir, en la red de interacción sexual de los blancos, la vecindad de gente menos activa estaba parcialmente separada del núcleo de individuos activos, en el caso de la red de afroamericanos, estos grupos estaban más conectados. En la primera red, las enfermedades de transmisión sexual proliferan fundamentalmente en el núcleo, mientras que en la red de afroamericanos también se desbordan hacia la periferia. En este caso, el grado de los nodos de la red no nos es muy útil, ya que individuos con el mismo número de conexiones sexuales tienen diferente exposición a la enfermedad y esto también depende de si son blancos o afroamericanos, por lo tanto, no basta saber cuántos amigos tienes, sino que también es necesario conocer cuántos amigos tienen tus amigos. En este caso, la distribución de grado nos proporciona mucha información de cómo es la estructura global de la red, pero no es suficiente, es necesario contar con medidas que nos describan más a fondo la red. En la red sexual de los blancos, los nodos de grado bajo tienden a conectarse con nodos de grado bajo, y los nodos de grado alto con nodos de grado alto, por lo tanto la red es asortativa, ya que los nodos tienden a conectarse con otros nodos similares en el número de conexiones, mientras que en la red de afroamericanos, los nodos tanto de alto como de bajo grado están conectados entre sí, por lo que la red es disortativa. Laumann

y Youm también demostraron que los afroamericanos suelen tener parejas de sus mismas comunidades, lo que ocasiona que cuando una enfermedad entra a la comunidad ésta se queda atrapada en ella, esto hace que los afroamericanos tengan una probabilidad de tener una enfermedad de transmisión sexual 1.3 veces mayor que la de los estadounidenses blancos.

Las redes, tanto de la naturaleza como de la sociedad, pueden estar divididas en partes diferenciadas, en particular, en el caso de sistemas sociales, algunos grupos humanos pueden estar completamente separados, sin embargo estos grupos desconectados o componentes aislados son una minoría muy pequeña, ya que en todas las redes, casi todos los elementos forman parte de una gran estructura conectada llamada componente gigante conectado. En general, este componente incluye como mínimo entre el 90 % y el 95 % del sistema en casi todas las redes. Por ejemplo en las redes de interacción sexual, las relaciones pasadas y actuales, pueden conectarnos con personas con las que no imaginamos que podríamos tener algún tipo de relación. Otro ejemplo es la red de colaboraciones entre científicos, ya que tenemos una gran estructura de cooperación, en la cual son pocos los científicos que trabajan en solitario. Dentro de la red de los consejos de administración de una empresa, la interconexión garantiza que existe un tipo de relación, en la que se engloba a la mayoría de las empresas.

En general, en las redes existen diferentes tipos de conectividad, en las redes con conectividad heterogénea presentan hubs que son muy reconocibles. Cuando una red es pequeña es fácil saber si una red tiene conectividad heterogénea u homogénea (como se muestra en la figura 11), ya que en el primer caso es fácil ubicar unos cuantos hubs, pero en el segundo caso todos los nodos tienen más o menos el mismo número de enlaces. En cambio, cuando la red es muy grande no es tan sencillo saber qué tipo de conectividad presenta.

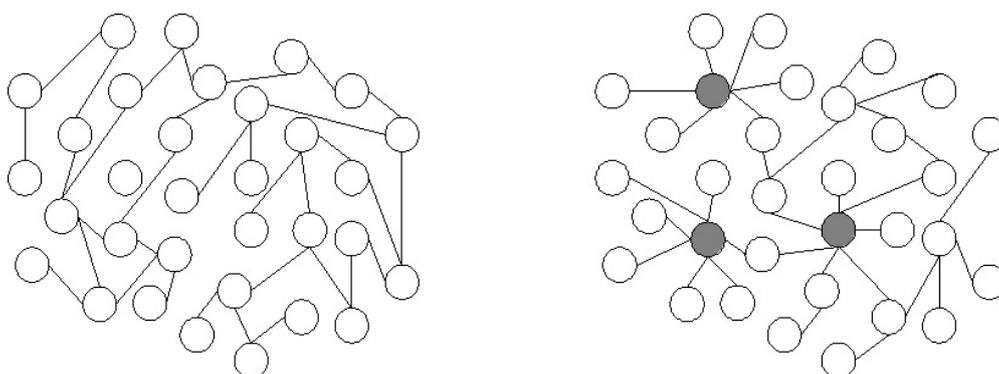


Figura 11: Conectividad homogénea y heterogénea

Dentro de las redes homogéneas, la distribución de grado es una curva de campana, mientras que en las redes heterogéneas siguen una ley de potencia, como vimos anteriormente, la ley de potencia supone que hay muchos más hubs en redes heterogéneas que en redes homogéneas. Un ejemplo de esto sería el siguiente; si tomamos las alturas de los alumnos de una clase, los formamos en una fila por estaturas, primero los que miden entre 1.50 metros y 1.55 metros, después formamos una fila paralela con los que miden entre 1.55 metros y 1.60 metros, en la siguiente fila paralela tenemos a los alumnos que

miden entre 1.60 metros y 1.65 metros, y así sucesivamente vamos aumentando el rango en cada fila de 5 cm en 5 cm, al final la distribución de los alumnos en las filas nos dará como resultado una curva de campana, ya que el número de alumnos aumenta al aumentar la altura, pero son pocos los alumnos que son o muy bajos o muy altos, la mayoría de los alumnos se encuentra en el centro, en la media. Si seguimos trabajando con el mismo grupo de alumnos pero ahora queremos el número de amigos que tiene cada alumno, entonces ahora las filas serán ordenadas por el número de amigos, la primera fila corresponde a los alumnos que tienen entre 0 y 20 amigos, la siguiente fila a los que tienen entre 20 y 40 amigos, la siguiente fila será entre 40 y 60 amigos y así sucesivamente, el resultado de este procedimiento nos dará la distribución de la conectividad de los nodos de la red social, en otras palabras la distribución de grado, la forma que tendrá esta distribución de grado es una curva denominada ley de potencia. Por lo tanto, observar las distribuciones de grado es la mejor forma para comprobar si una red es heterogénea, es decir, la distribución tendrá una cola larga, ya que la red tiene hubs y heterogeneidad. A continuación en la figura 12 se muestra gráficamente las distribuciones de grado de las alturas y de los amigos de los alumnos.

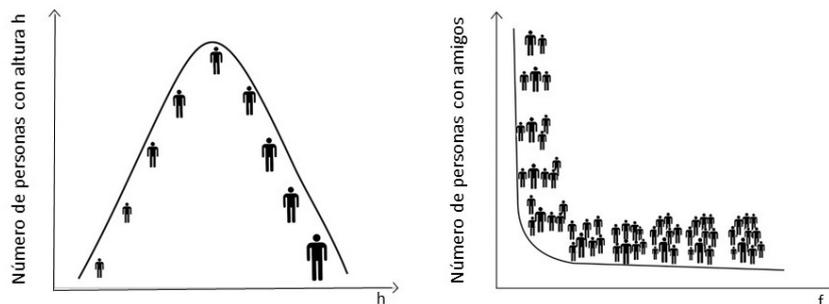


Figura 12: Distribución de grado de las estaturas y de los amigos de los alumnos

Cuando nuestros amigos y nuestra familia se conectan mejor, aumenta nuestro nivel de conexión con el conjunto de la red social. Nosotros decimos que eso nos centra porque contar con amigos bien conectados, nos aleja literalmente de los márgenes hacia el centro de la red social. La centralidad se puede medir contando no solo el número de amigos y de otros contactos que tenemos, sino contando también el número de amigos de nuestros amigos, el número de amigos de los amigos de nuestros amigos, etc. Por lo que al final no importa cuántos amigos tengas o quiénes sean, lo que de verdad importa es quiénes son los amigos de tus amigos, es decir, si tus amigos son amigos entre sí, esto es el denominado clustering. Un ejemplo rápido, si pensamos en una persona con dos amigos, entonces forman un tríptico conectado, si los dos amigos son amigos entre sí forman un tríptico triangular, de tal manera que, visto desde la forma más simple, el clustering es la cantidad de triángulos en una red en comparación con el número global de trípticos conectados, por lo tanto, mide la densidad de los triángulos en la red, su aglomeración o agrupamiento global.

En las redes sociales, los nodos con alta intermediación son los agentes y conectores que mantienen juntos a los demás. Esto significa que, un nodo tiene la capacidad de controlar el flujo de información entre la mayoría de los demás nodos. Los individuos con

alta intermediación son los pivotes en el flujo de conocimiento de la red. Los nodos con mayor intermediación, también resultan en el mayor incremento en la distancia entre los demás cuando estos son retirados. El concepto de centralidad intermedia, fue introducido por el sociólogo Linton C. Freeman, se consideran todos los pares de nodos de una red y se cuentan las rutas más cortas que los conectan, es básicamente la proporción de rutas que atraviesan a un determinado nodo, cuánto más grande es esta proporción, más central es el nodo. Los nodos centrales suelen funcionar como puentes o como obstáculos, ya que son paradas casi obligatorias dentro de una red.

En 2008, los sociólogos Nicholas Christakis y James Fowler, realizaron un estudio [9] en el cual preguntaron a más de 12,000 personas de Framingham, Massachusetts, sobre su sensación subjetiva de la felicidad y, además, hicieron un mapa en el cual se mostraba de quiénes eran amigos, cónyuge, hermano y familia. Al dibujar esta red, se dieron cuenta que la gente que se conecta comparte los mismos sentimientos, es decir, la gente feliz se conecta, mientras que la gente infeliz suele hacer lo mismo. Más de veinte años antes, en 1984, se hizo un estudio que arrojó los siguientes resultados, ganar 5,000 dólares más al año, aumenta las posibilidades de ser feliz tan solo un 2%, mientras que tener un amigo feliz, la aumenta en un 15%. Christakis y Fowler encontraron resultados aún más interesantes, la felicidad de un individuo, también está influida por la felicidad de las personas que no son sus vecinos de al lado, por lo que el efecto felicidad a dos pasos de distancia (amigos de amigos) es de aproximadamente 10%, a tres pasos de distancia (amigos de amigos de amigos) es de aproximadamente 6%, el efecto disminuye en el cuarto paso. Estos dos científicos con ayuda de otros científicos, encontraron recomendaciones e influencia para la obesidad, hábitos de fumadores y las clásicas recomendaciones de boca en boca, por lo tanto, la influencia y la información llegan a la persona desde tres grados de distancia.

Esta regla de tres grados de influencia, que se da en procesos sociales dice que la influencia de nuestros comportamientos, opiniones, etc., no se extiende tan lejos, sino que es similar a las ondas que se crean al lanzar una piedra en un lago, ya que la influencia se va disipando conforme se aleja del nodo emisor. Esta regla como se dijo anteriormente, demuestra que en las redes sociales solo se alcanzan tres grados de distancia, a nuestros amigos (grado 1), amigos de amigos (grado 2), amigos de los amigos de nuestro amigos (grado 3) y a partir del cuarto grado se comienza a disipar el efecto, es decir, no influimos ni somos influenciados por las personas que están a cuatro grados de distancia. Esto se explica por tres razones:

1. Por la explicación de la decadencia intrínseca de los efectos, disipación
2. Por la explicación de la propia dinámica de las redes en constante reconfiguración, donde se agregan y se eliminan nodos, explicación de la inestabilidad de la red
3. Por la explicación a efectos evolutivos, una explicación biológica, por la cual la evolución de los grupos sociales estuvo restringida tradicionalmente a tres grados de separación

A continuación en la figura 13 se muestra un pequeño ejemplo de cómo funcionan los tres grados de influencia.

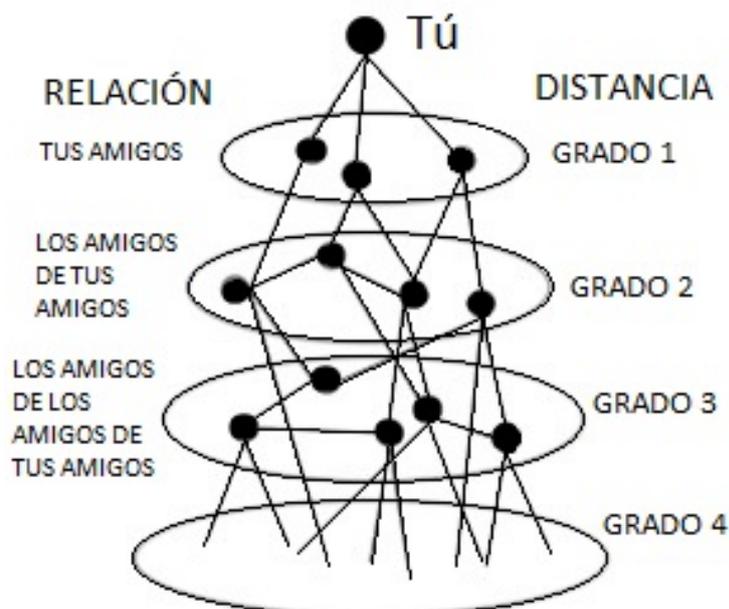


Figura 13: Tres Grados de Influencia

Finalmente, es interesante mencionar que a finales del año 2013, la Universidad Nacional Chiao Tung de Taiwán, publicó un estudio realizado por Eman Yasser Daraghmi y Shyan-Ming Yuan [37], en el cual se propone la reducción de seis grados de separación a cuatro grados de separación. Para la realización de este estudio, crearon una base de datos de Facebook con 950 millones de personas, tras eliminar duplicados, cuentas falsas y cuentas de celebridades, usaron métodos estadísticos para calcular la distancia entre las personas dentro de la red, lo que descubrieron fue que existía una distancia promedio de 3,9 grados. Por lo que los investigadores proponen que, el número promedio de conocidos que separan a dos personas cualesquiera sin importar dónde están, no es seis, sino 3,9, cifra que se redondea a 4 grados de separación. Más aún, si eres una persona que está en un trabajo donde convives con muchas personas, los grados de separación son todavía menos, llegando a un promedio de 3,2 personas.

Capítulo 3

Datos, Metodología y Resultados

Los datos utilizados en este trabajo fueron obtenidos de diferentes fuentes ([38], [39], [40], [41] y [42]). La principal fuente es la página de internet de uno de los investigadores más reconocidos dentro de Teoría de Redes, su nombre es Mark Newman. En esta página [39], se pueden encontrar diferentes bases de datos las cuales están abiertas a todo el público para fines de investigación.

Se eligieron las siguientes bases de datos que representan redes sociales reales, aunque tenemos una red ficticia, pero nos interesa ya que fue creada especialmente para la investigación. Nuestras bases de datos son muy variadas, por lo que nos es de gran interés estudiarlas, analizarlas y compararlas, aunque la red que más nos interesa es la red conformada por los miembros que forman parte de la Organización de Jóvenes Nikkei (OJN). Las características generales de cada una de las bases de datos, se describen a continuación.

1. La primera base que analizaremos es una base muy conocida utilizada por muchos científicos para diferentes publicaciones y una de las primeras bases de datos. En 1970, el señor Hi, profesor de karate de un club universitario de Estados Unidos, pidió al presidente del club de karate John A., que subiera el precio de las clases para tener un mejor salario, a lo cual recibió un no por respuesta. En 1972, dos profesores de karate de dicho club, tuvieron enfrentamientos por dicha cuestión, a tal grado que decidieron separarse y formar dos diferentes clubes, este hecho captó la atención del antropólogo Wayne W. Zachary, quien en 1977 publicó un estudio en el que daba su punto de vista sobre este suceso. Durante todo ese tiempo Zachary recolectó información sobre las clases de karate, las reuniones, las fiestas y las comidas de los miembros del club e identificó como buenos amigos, a aquellos miembros que también se reunían fuera del club. Zachary pudo dibujar una red en la cual se representaban las amistades del club, lo que resultó que dentro de la red se observaba claramente la estructura de los dos grupos que se habían formado por la disputa entre los profesores y por las amistades de ambos. Esta red consta de 34 miembros y 78 enlaces.
2. La segunda base de datos que utilizamos, corresponde a los datos obtenidos de las asociaciones de una comunidad de 62 delfines nariz de botella que se encuentran en Doubtful Sound, Nueva Zelanda. Esta base de datos fue realizada por David Lusseau en 2003. Esta red también fue estudiada por Mark Newman y encontraron

diferentes tipos de comunidad y subcomunidades, así como, una relación homofílica con el género y la edad en la cual se formaban clusters. Esta red es muy parecida a una red social de personas, pero difiere en ciertos aspectos, por ejemplo, en el nivel de asortatividad.

3. Otra de las bases de datos que utilizaremos, es una base realizada por Mark Newman en mayo del 2006, en la cual los datos corresponden a la red de coautores que trabajan en teoría de redes y en la experimentación. Esta versión de la base de datos cuenta con todos los componentes de la red, esto nos da un total de 1589 científicos. Para conocer la importancia que tiene cada científico dentro de la red, se toma en cuenta el número de artículos escritos y también el número de citas que tienen en otros artículos.
4. Por otro lado, otra de las bases de datos que utilizaremos es una red de jugadores de fútbol americano colegial (División IA) en Estados Unidos, durante la temporada regular de Otoño 2000, la cual contiene 115 jugadores. Esta red fue compilada por M. Girvan y M. Newman. Se tienen atributos que indican a cuál de las 12 conferencias (regiones) pertenecen.
5. Otra de las bases de datos que utilizaremos es la red de los correos electrónicos de Enron la cual fue hecha pública por el Departamento de Justicia de Estados Unidos en el 2001, esta red consta de 184 miembros y es importante para los investigadores ya que es una red real, con una buena organización, y además cubre un periodo de 3.5 años. Contiene gran información de las interacciones, comunicación, relaciones a nivel tanto individual como grupal. Todo esto nos ayuda a comprender cómo todos estos factores impactan a la red, por lo que tenemos que analizar y procesar la información de una forma eficiente y eficaz.
6. Una de las bases utilizadas en este análisis, es la red de amistad entre los miembros de una facultad de la UK University durante el año 2008, la cual consta de 81 individuos y 817 enlaces dirigidos. Las afiliaciones escolares de cada individuo están guardadas en forma de un vector. Esta base de datos sirve como base para probar algoritmos para detectar comunidades.
7. Otra de las bases de datos que estaremos utilizando es la red de papalote (“Kite Network”), la cual es una red especial, ya que es una red social ficticia realizada en 1990 por el científico David Krackhardt, la cual consta de 10 actores, pero tiene propiedades interesantes ya que el actor central no corresponde con las medidas de centralidad tradicionales como lo son el grado, el grado de cercanía y la centralidad intermedia. Además, recibe el nombre de red de papalote, por la forma que tiene esta red.

8. La siguiente base de datos corresponde al año 2010, entre el 1 de Diciembre y el 10 de Diciembre, los datos corresponden al área geriátrica de un hospital en Lyon, Francia. Fueron 75 personas que participaron en este estudio, las cuales utilizaron un sensor RFID en pequeñas tarjetas de identificación durante este periodo de tiempo, lo que hacían estas tarjetas era almacenar la información si dos de estos pacientes se encontraban cara a cara en un rango de entre 1 y 1.5 metros, durante un intervalo de tiempo de 20 segundos.
9. Por otro lado, utilizaremos la base de datos que corresponde al estudio de un bufete de abogados en el noroeste de Estados Unidos (Nueva Inglaterra), la firma conocida como *SG&R* participó en el estudio de 1988 a 1991. Esta red consta de 36 socios y asociados de la firma. Originalmente, esta base de datos se utilizó para identificar procesos sociales dentro de la red como el intercambio de conocimientos, amistad, solidaridad, control de calidad, entre otros. La base de datos que usaremos es sólo un subconjunto de la base original.
10. Por último, la red más importante que nos interesa analizar es una red que corresponde a un entorno conocido. Esta red corresponde a los miembros de la Organización Jóvenes Nikkei (OJN). En particular, nos interesa conocer las distintas propiedades que tiene esta red y poderla comparar con las redes que mencionamos anteriormente. Esta base de datos está actualizada al día 17 de Septiembre de 2015 y contiene 205 miembros.

Una vez que tenemos todas las bases de datos, usamos el software R [16], para realizar un programa con la paquetería “*igraph*”, la cual nos ayuda para el análisis de redes. El software R contiene una serie de paqueterías las cuales nos ayudan al análisis de redes como “*igraph*”, “*sna*”, “*network*”, “*statnet*”, “*sand*”, “*igraphdata*”, “*rsnns*”, “*neuralnet*”, entre otras, pero es importante tener en cuenta que si se corren todas las paqueterías a la vez podemos tener problemas, ya que depende de cómo está organizada la base de datos, las matrices de adyacencia, y de las distintas funciones. Es decir, dependiendo de la paquetería, algunas funciones pueden cambiar, por ejemplo si utilizamos *igraph* y *sna* al mismo tiempo, hay funciones que tienen el mismo nombre, pero la forma de declarar los argumentos y la forma de ingresar los datos es distinta.

Lo anterior nos servirá para, poder conocer la estructura o topología de las redes y poder compararlas con la red de la OJN, a fin de conocer las características de esta red, así como sus fortalezas y debilidades. Para conocer y analizar las redes descritas anteriormente, es decir, la topología y características específicas y poder compararlas, calculamos las siguientes métricas: los grados, el diámetro, la densidad, la distancia media, el número de cliques, el valor del clustering global y local, la media del grado de cercanía, la media de la centralidad intermedia, el valor de asortatividad y, por último, una de las medidas más importantes, la distribución de grado.

A continuación se mostrarán las redes de cada una de las bases de datos, además del histograma de distribución de grado. Posteriormente, todos los resultados serán presentados dentro de una tabla, la cual nos ayudará a visualizar numéricamente las propiedades

y características de cada una de las redes. Con todos estos resultados podremos realizar un buen análisis comparativo de nuestras redes.

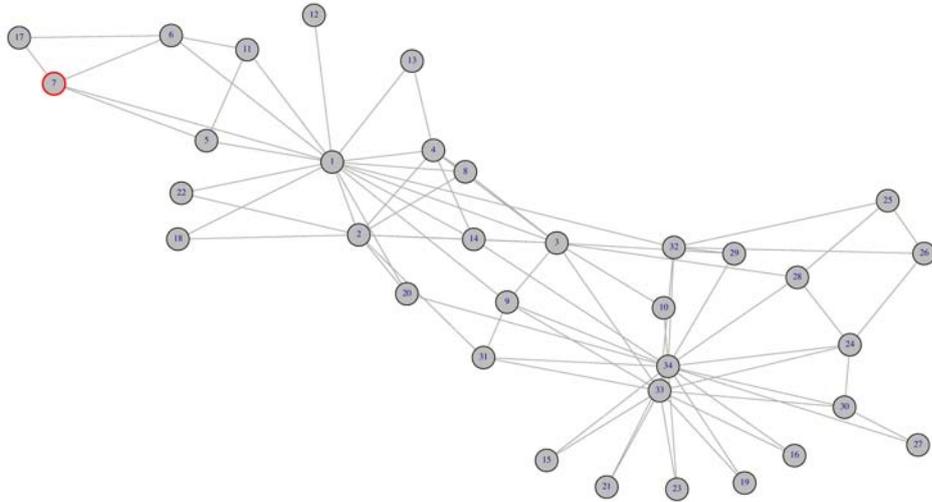


Figura 14: Red de Karate

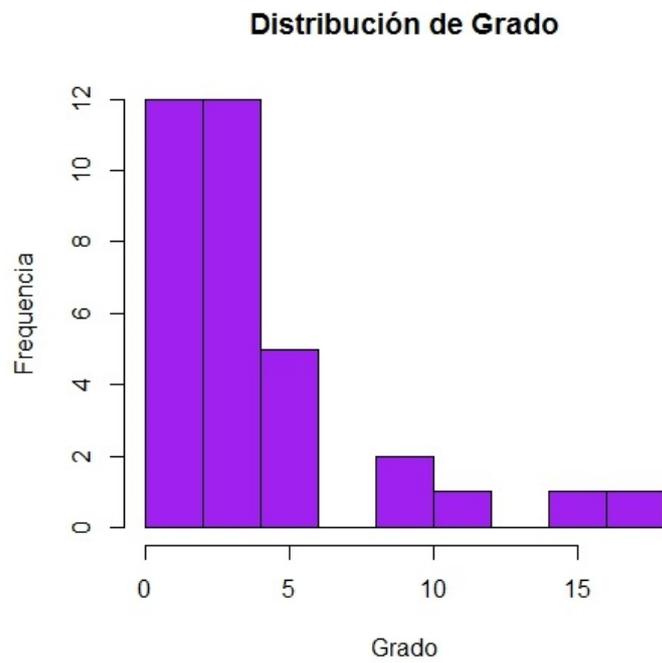


Figura 15: Histograma distribución de grado de la red Karate

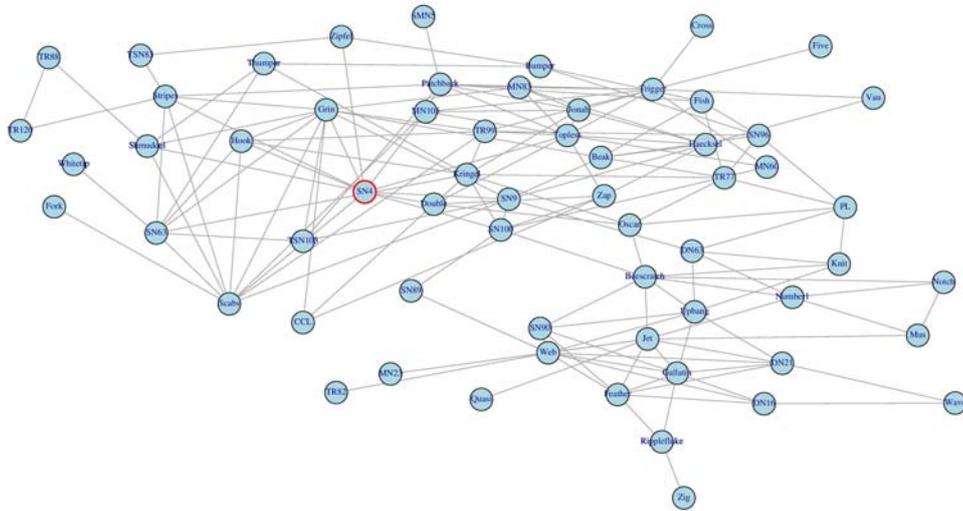


Figura 16: Red de comunidad de delfines

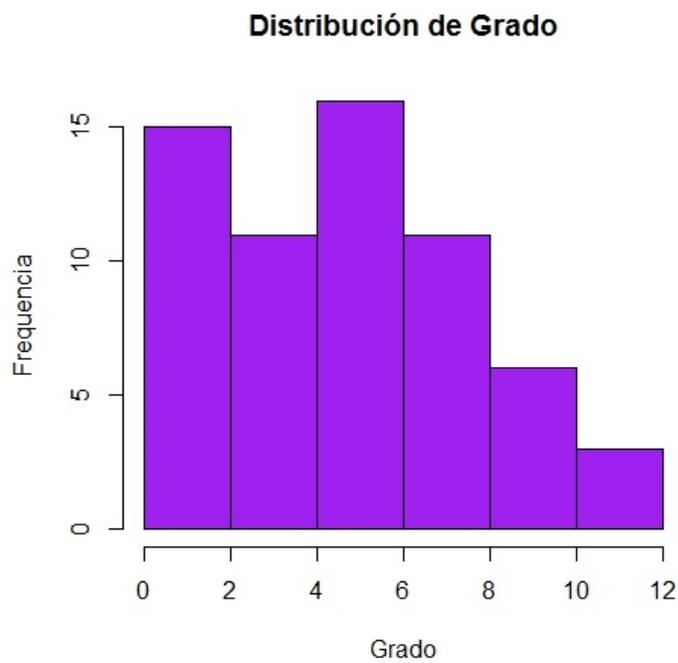


Figura 17: Histograma distribución de grado de la red de delfines

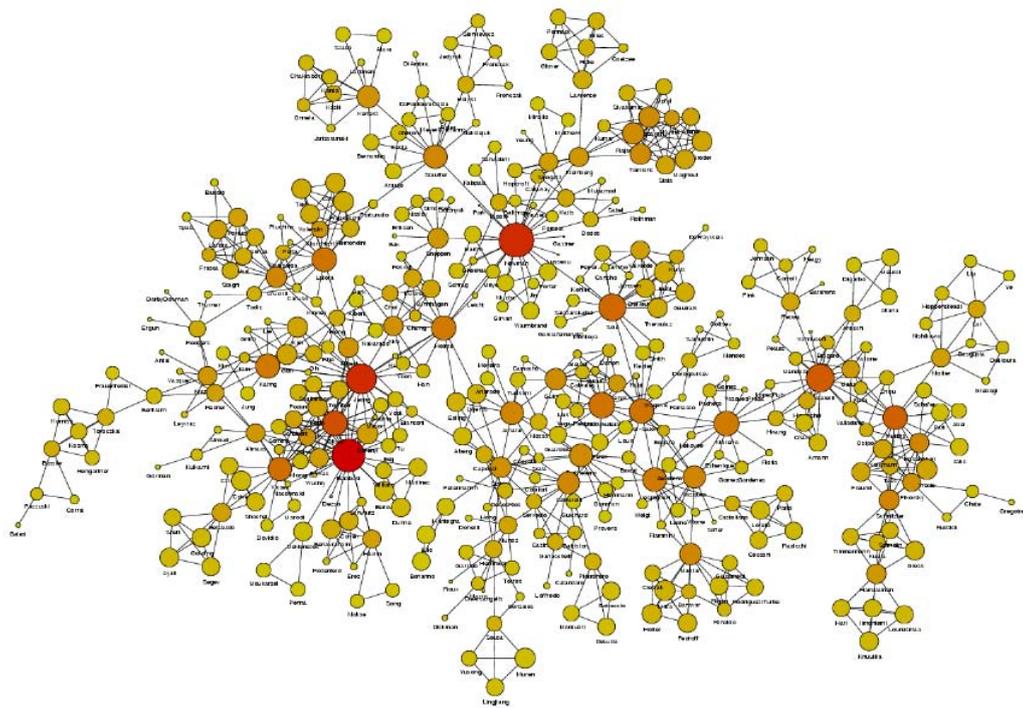


Figura 18: Red de coautores de Teoría de Redes

Distribución de Grado

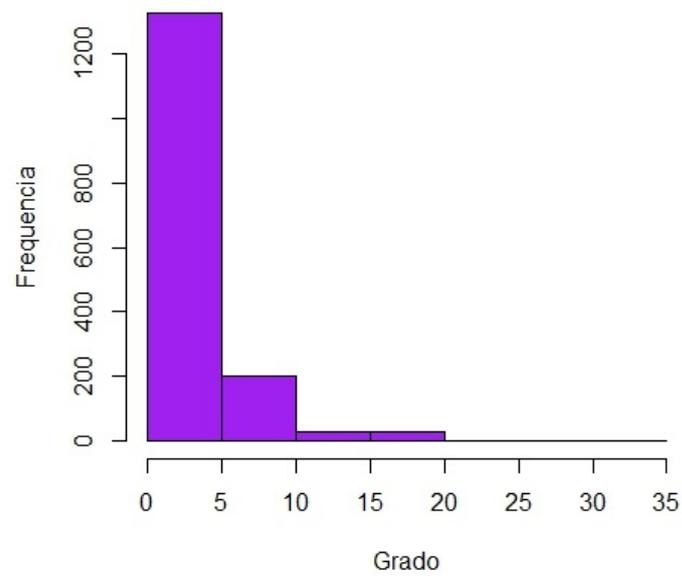


Figura 19: Histograma distribución de grado de la red de coautores

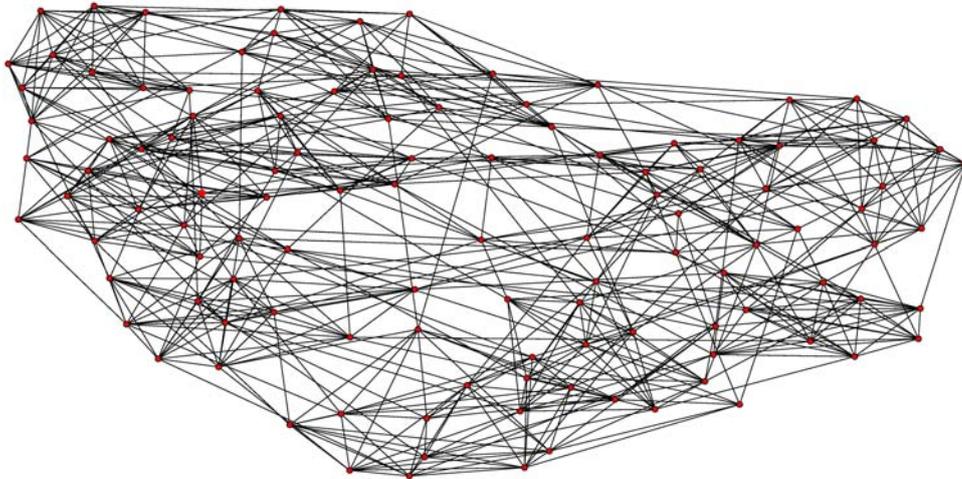


Figura 20: Red de jugadores de futbol americano

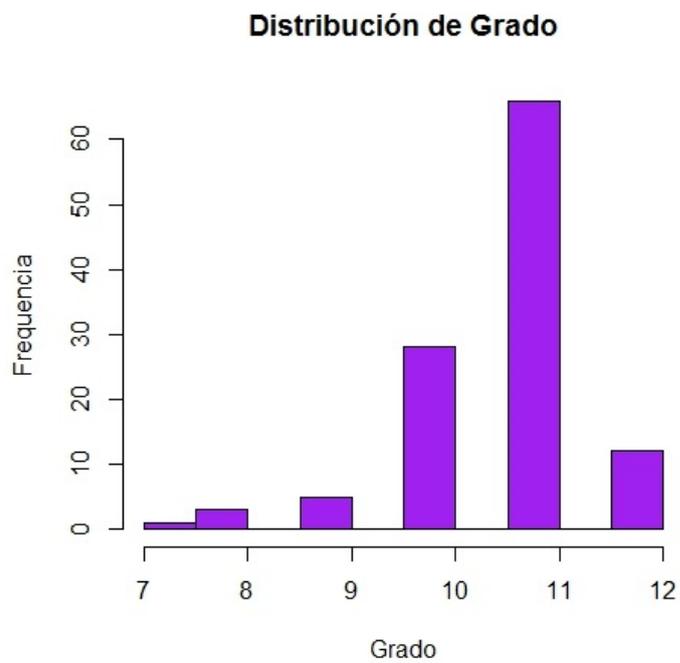


Figura 21: Histograma distribución de grado de la red de jugadores de futbol americano

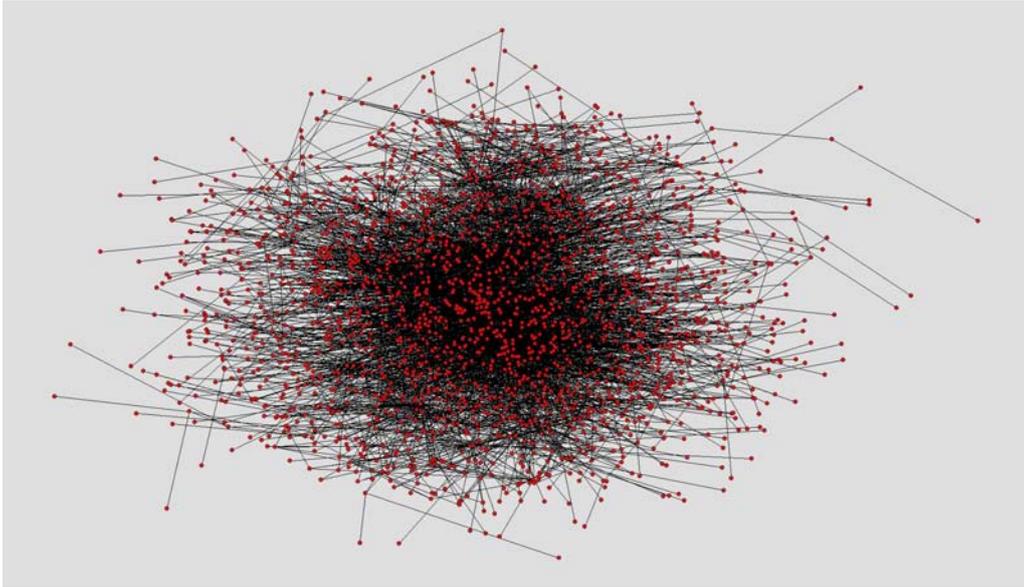


Figura 22: Red de correos electrónicos de Enron con un 30% de la red original

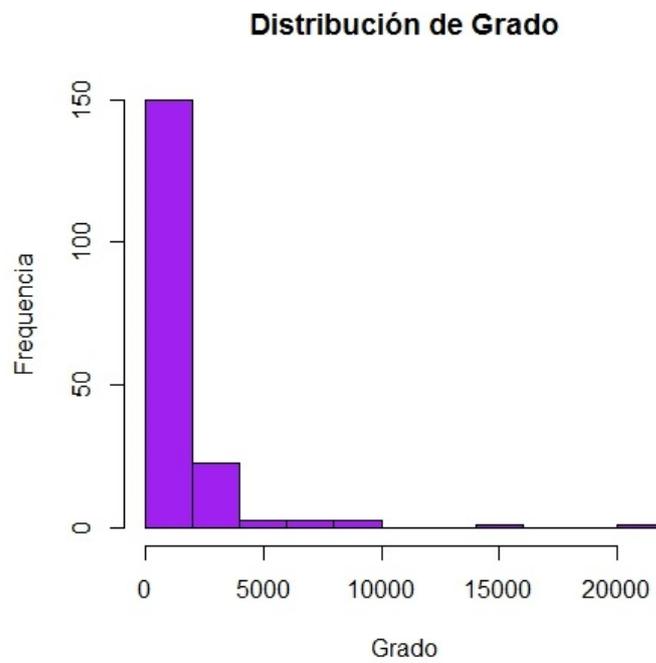


Figura 23: Histograma distribución de grado de la red de correos electrónicos de Enron

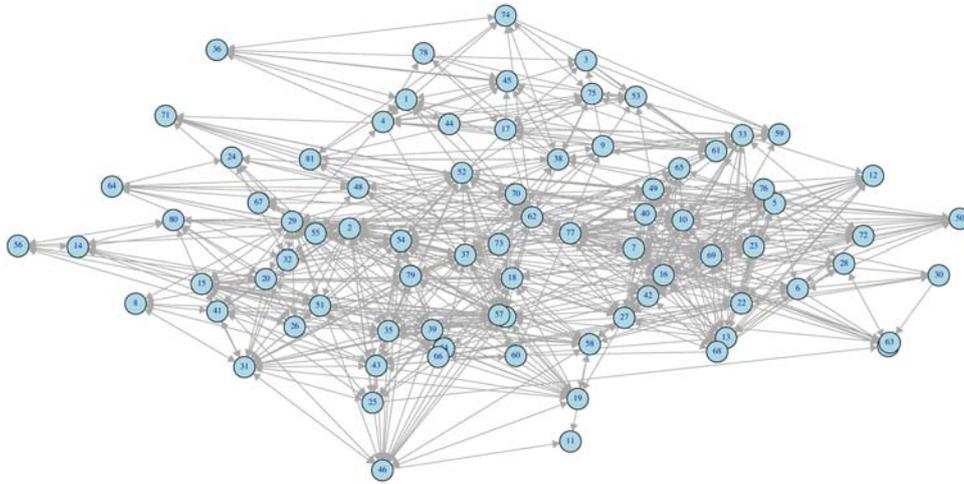


Figura 24: Red de miembros facultad UK

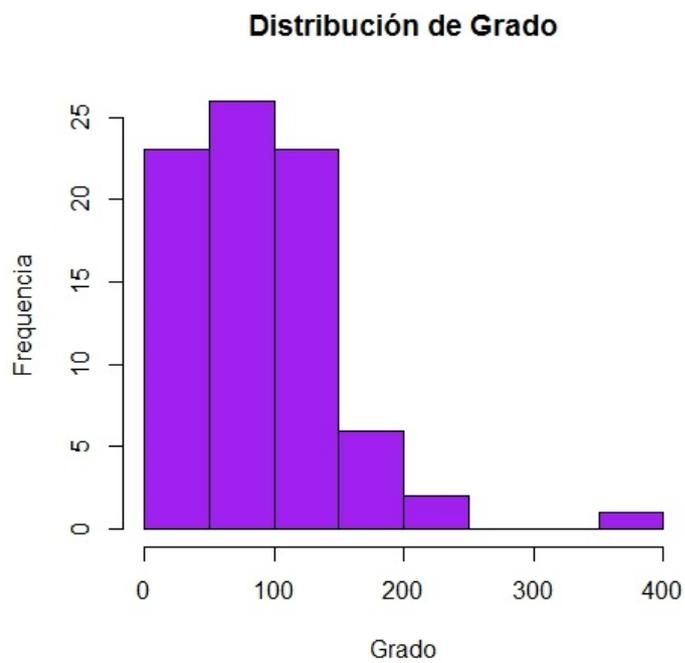


Figura 25: Histograma distribución de grado de la red de miembros facultad UK

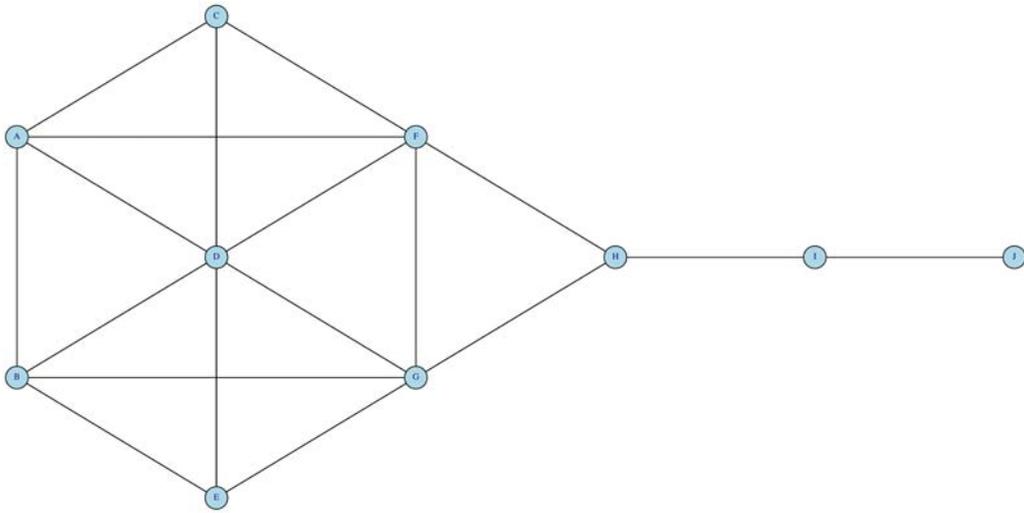


Figura 26: Red de papalote

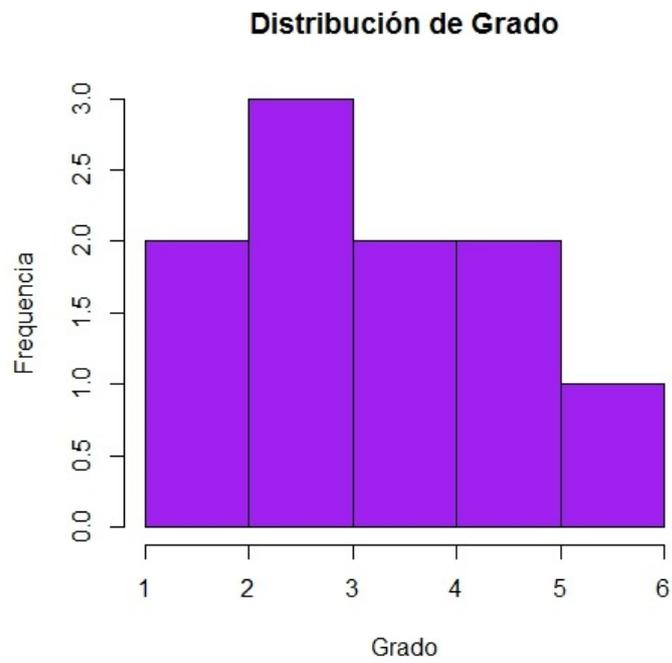


Figura 27: Histograma distribución de grado de la red de papalote

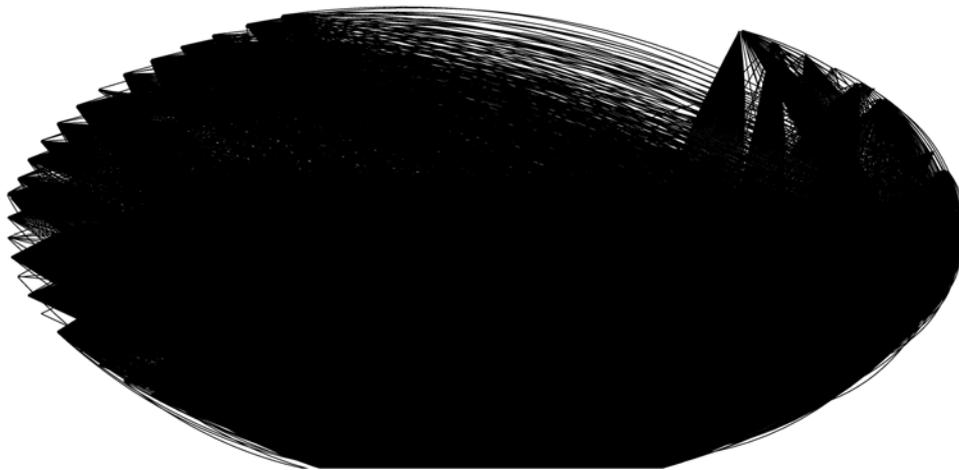


Figura 28: Red de pacientes geriátricos con un 80% de la red original.

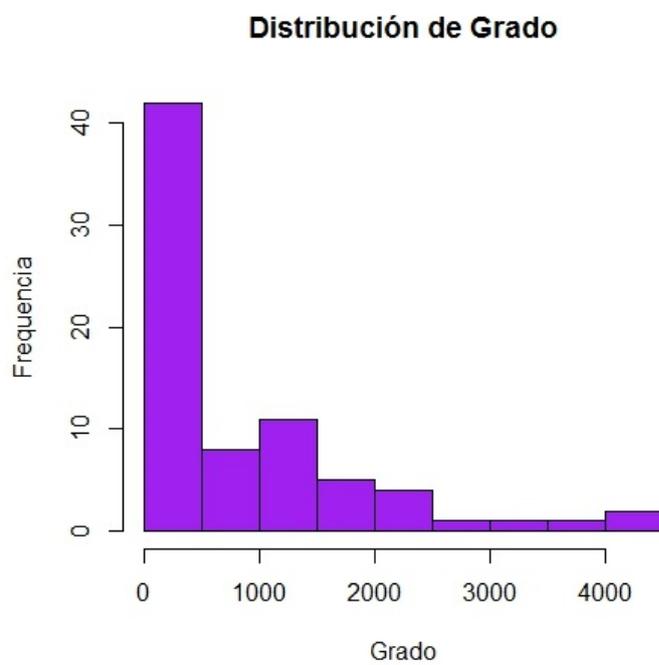


Figura 29: Histograma distribución de grado de la red de pacientes geriátricos

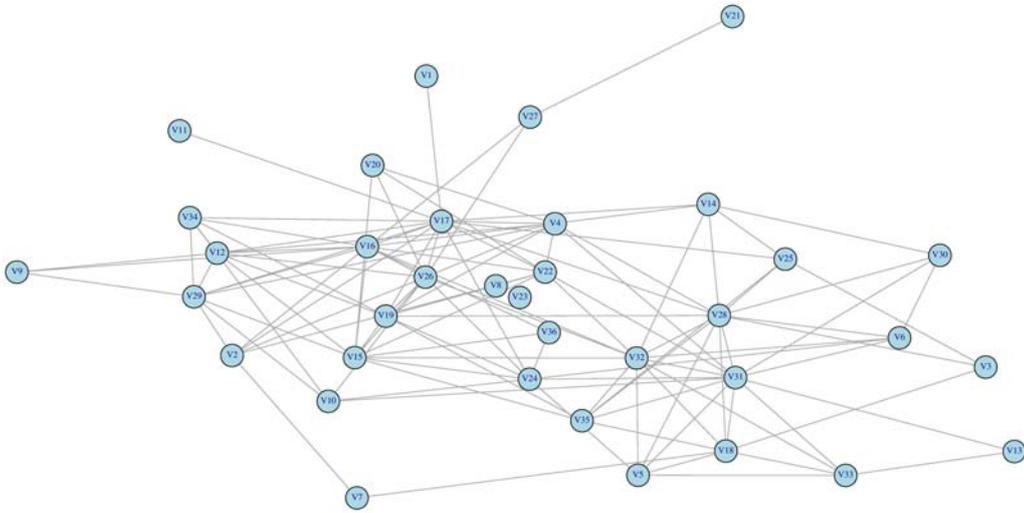


Figura 30: Red de abogados

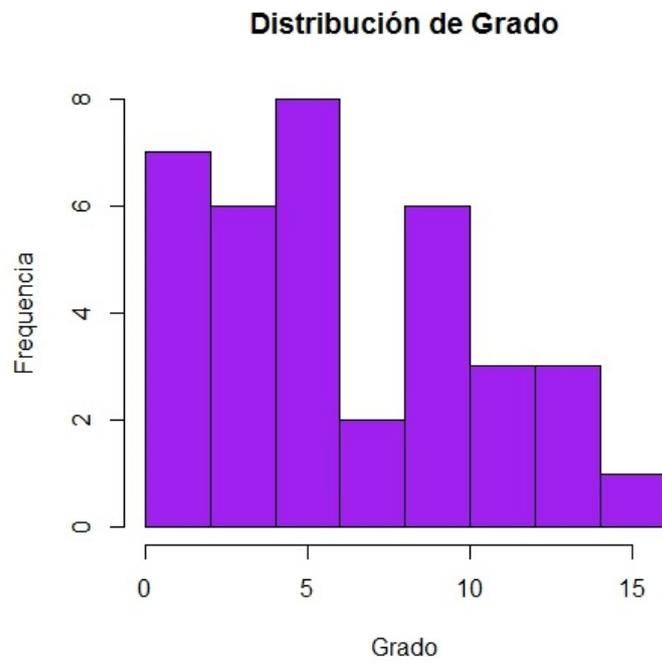


Figura 31: Histograma distribución de grado de la red de abogados

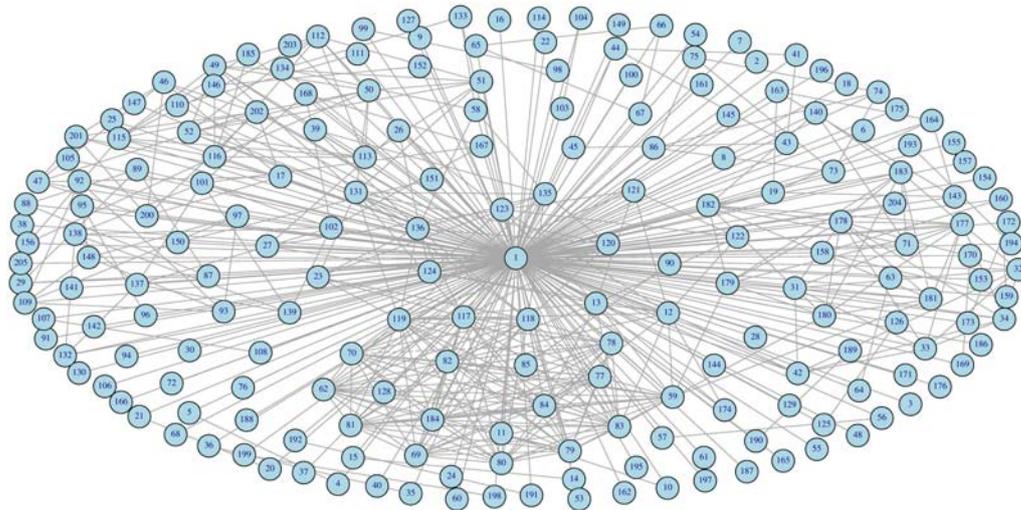


Figura 32: Red de miembros OJN

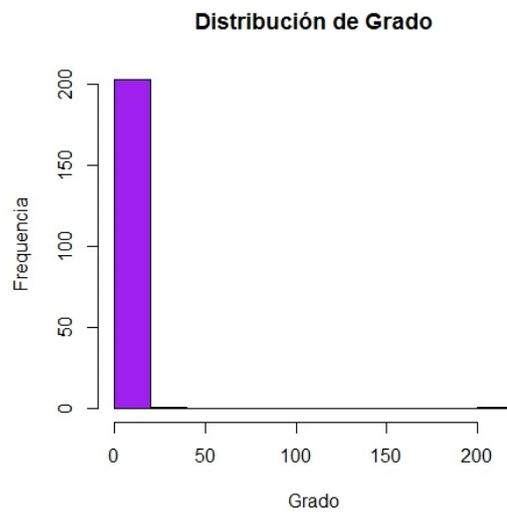


Figura 33: Histograma distribución de grado de la red de OJN

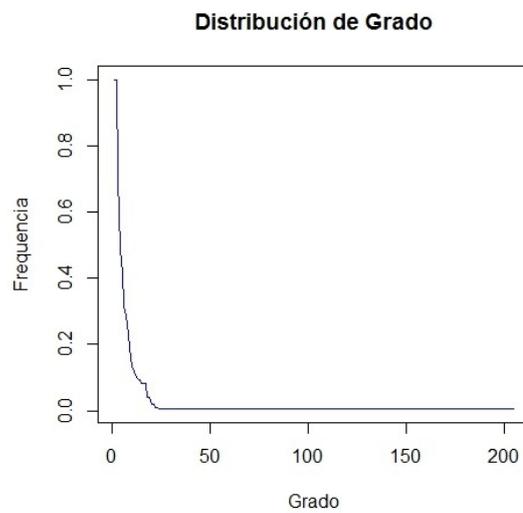


Figura 34: Plano log-log distribución de grado de la red de OJN

Resumen de Resultados

RESULTADOS	Karate	Delfines	Coautores	Football	Emails Enron	Facultad UK	Papalote	Hospital Geriátrico	Abogados	OJN
No. nodos	34	62	1589	115	184	81	10	75	36	205
No. arcos	78	159	2742	613	125409	817	18	32424	115	560
Tipo de red	No Dirigida	No Dirigida	No Dirigida	No Dirigida	Dirigida	Dirigida	No Dirigida	No Dirigida	No Dirigida	No Dirigida
Grados (máx.)	17	12	34	12	10392	24	6	4286	15	204
Grado int. (máx.)	-	-	-	-	11970	41	-	-	-	-
Grado ext. (máx.)	-	-	-	-	21560	62	-	-	-	-
Diámetro	5	8	17	4	6	22	4	3	5	2
Densidad	0.1390374	0.0840825	0.002173317	0.0935164	3.72443	0.1260802	0.4	11.68432	0.1825397	0.02678144
Distancia media	2.4082	3.356954	5.82324	2.508162	2.390464	2.511094	1.977778	1.597838	2.144385	1.973219
Cliqués	5	5	20	9	13	9	4	17	5	16
Clustering Global	0.2556818	0.3087757	0.6934414	0.4072398	0.3725138	0.4734084	0.5789474	0.5880037	0.3887689	0.1418663
Clustering local (prom.)	0.5879306	0.3029323	0.8782056	0.403216	0.5055302	0.5737133	0.5777778	0.6402797	0.4867335	0.9204121
Cercanía (prom.)	0.01292363	0.005036721	0.0000004271227	0.003502799	0.000976732	0.002655465	0.05870505	0.008617982	0.00670515	0.002490412
Betweenness (prom.)	23.23529	71.8871	231.1057	85.96522	238.0489	177.287	4.4	22.12	17.83333	99.26829
Asortatividad	-0.4756131	-0.04359403	0.4616225	0.1624422	0.6762668	0.03805201	0.1525424	0.08623212	-0.1681201	-0.278245

De la tabla anterior, empezamos comparando el número de nodos y arcos, y observamos que la red con más nodos es la red formada por los coautores de teoría de redes ya que cuenta con 1,589 nodos, mientras que la red con menos nodos, es la red ficticia formada por Krackhardt (red de papalote) que solo cuenta con 10 nodos, en tanto que las demás redes tienen a lo más 205 nodos. Ahora bien, si nos fijamos en el número de arcos, observamos que la red con mayor número es la red de correos electrónicos de Enron, la cual tiene 125,409 arcos, mientras que la de menor número, es nuevamente la red ficticia formada por Krackhardt (red de papalote), con solo 10 arcos. En las demás redes, el número de arcos varía mucho, tenemos redes con muchos arcos como la red formada por los pacientes geriátricos de un hospital (32,424) y la red de coautores (2,742), en tanto que las otras redes tienen un menor número de arcos. Por otro lado, si observamos el grado máximo, este varía entre las redes dirigidas y las no dirigidas, ya que en las primeras el grado interior y exterior es diferente, y en las últimas, el grado interior es igual al grado exterior, en total contamos con 2 redes dirigidas y 8 redes no dirigidas. Dentro de las redes dirigidas, la red con mayor grado es la red de correos electrónicos de Enron, con grado máximo general de 10,392, grado máximo interior de 11,970 y grado máximo exterior de 21,560. En cambio, en las redes no dirigidas encontramos que la red con mayor grado es la red formada por los pacientes geriátricos, que cuenta con grado de 4,286, de ahí la segunda red con grado alto, es la red formada por los miembros de la OJN, con un grado de 204, las demás redes tienen un grado de a lo más 34.

Además, calculamos el diámetro de cada una de las redes, esto como ya vimos anteriormente, es la ruta más larga de entre todas las rutas más cortas. En estas redes, la red que tiene el diámetro más pequeño es la red de OJN, con un valor de 2, y la red con el diámetro más grande es de 22, que es la red de miembros de la facultad de UK, mientras que las demás redes, excluyendo a la red de coautores que tiene un diámetro de 17, todas las demás redes tienen un diámetro de entre 3 y 8. Por otro lado, la densidad es importante, nos dice qué tan conectada está la red, y como estamos analizando redes sociales, esperaríamos un valor de densidad alto, sin embargo, los sistemas reales modelados con redes, en general, son poco densos, esto debido al costo de los enlaces. La red con mayor densidad, es la red formada por los pacientes geriátricos con 11.68432, esto nos dice que se tiene una conectividad muy alta dentro de la red, en cambio la red con la densidad más pequeña, es la red formada por los coautores con una densidad de 0.002173317, lo cual nos indica que la conectividad dentro de la red es muy baja, mientras que las demás redes también presentan una densidad baja, esta se encuentra entre 0.4 y 0.02678144; la única red que se sale de estos parámetros es la red de correos electrónicos de Enron, con una densidad de 3.72443, bastante grande en comparación con las anteriores. Otra de las métricas que utilizamos en esta comparación es la distancia media, que es el promedio de las distancias entre todo par de nodos, así que esperamos que las redes tengan una distancia media baja, tal como se vio anteriormente con la propiedad de mundo pequeño. Al observar nuestras redes, nos dimos cuenta que la red con mayor distancia media es la red de coautores, con una distancia media de 5.82324, al contrario la distancia media más baja es de 1.597838 correspondiente a la red de pacientes geriátricos. En general, todas nuestras redes tienen una distancia media pequeña, que es lo que queremos, como es el caso de la red de miembros de OJN, con una distancia media de 1.973219, la cual es una distancia pequeña y nos indica que nos podemos mover o comunicar fácilmente dentro de la red.

Por otro lado, además de las métricas, nos interesa conocer la topología de las redes, para lo cual calculamos el coeficiente de aglomeración o clustering, empezamos con el clustering global, que nos indica cuál es la tendencia de la red a formar triángulos, encontramos que la red de coautores tiene el clustering global más alto de 0.6934414. Lo anterior nos indica que hay una tendencia alta a formar triángulos, en contra parte, el clustering global más bajo es de la red de los miembros de la OJN, con un valor de 0.1418663, por lo que tiene una baja tendencia a formar triángulos. En tanto que, si nos fijamos en el clustering local promedio, es muy similar al clustering global, aunque en la red de los miembros de OJN sucede algo peculiar, ya que el clustering global es muy bajo pero el clustering local promedio es muy alto, de hecho es el valor más alto de todas las redes, es de 0.9204121, indicándonos que la mayoría de los actores se conocen. El clustering local promedio más bajo es de la red de delfines, que tiene un valor de 0.3029323, y podemos decir que los delfines no tienen una tendencia a formar grupos pequeños, es decir, permanecen en el grupo completo. Por otro lado, la centralidad intermedia o betweenness, nos ayuda a identificar qué tan importante es un nodo dentro de una red, calculando cuántas rutas cortas pasan por el nodo en cuestión, así que calculamos el promedio de la centralidad intermedia de cada caso y obtuvimos que, la red con una centralidad intermedia mayor es la red de correos electrónicos de Enron con 238.0489, por su parte, la red con una menor centralidad intermedia es la red de papalote con solo 4.4. El grado de cercanía o closeness, se centra en calcular las rutas más cortas de cada nodo hacia todos los demás nodos de la red, obtenemos que el grado de cercanía es mayor en la red de papalote, el cual corresponde a 0.05870505, mientras que la red que tiene el menor grado de cercanía tiene 0.0000004271227 y esto corresponde a la red de coautores.

Con respecto a la correlación de nodos, tenemos el coeficiente de asortatividad, que como ya dijimos, nos da valores entre -1 y 1, por lo tanto podemos saber si una red es asortativa o es disortativa, si nos fijamos en nuestras redes, las redes disortativas son: la red de karate, la red de miembros de la OJN y la red de abogados, mientras que las que tienden a cero son neutras, como lo son: la red de delfines, la red de miembros de la facultad de UK y la red de pacientes geriátricos, mientras que el resto de las redes son asortativas, es decir, la red de coautores, la red de jugadores de futbol americano, la red de papalote y la red de correos electrónicos de Enron.

Por otro lado, daremos un pequeño análisis de las características específicas de cada una de las redes antes descritas. Comenzaremos con la red de karate, es una red no dirigida, el grado máximo es de 17, tiene diámetro de 5, densidad de 0.1390374, distancia media de 2.4082. En particular, tiene clustering local alto 0.5879306, que nos indica que la comunicación local se favorece y lo confirmamos con un grado de cercanía alto 0.01292363. La siguiente red es la comunidad de delfines de Nueva Zelanda, la principal cualidad que presenta es que tanto el clustering local (0.3087757) como el global (0.3029323) son valores bajos, además la correlación de nodos es neutra, esto significa que los delfines forman un grupo con toda la comunidad. Continuamos con la red de coautores de teoría de redes, en este caso es muy alta la tendencia a formar pequeños grupos, es decir, el clustering local (0.8782056) y global (0.6934414) son altos y la correlación es asortativa 0.4616225. La red de football tiene buena comunicación entre todos sus miembros, esto lo observamos en el clustering global 0.4072398 y en la centralidad intermedia 85.96522. Seguimos con la red de correos electrónicos de Enron, tiene el grado promedio más alto (10,392), la centralidad intermedia más alta (238.0489) y la asortatividad más alta (0.6762668),

esto se traduce en tendencias homofílicas y que presenta muy buena comunicación. Los miembros de la facultad de UK tiene una tendencia alta a formar grupos pequeños, ya que el clustering local es de 0.5737133, además presenta una correlación de nodos neutra de 0.03805201, lo interesante es que la tendencia a formar grupos locales no es ni heterogénea ni homogénea. El caso de la red ficticia de papalote es singular porque como ya se había mencionado fue creada para fines de investigación y aún cuando es una red muy pequeña (10 nodos), tiene el valor más alto de grado de cercanía (0.05870505). La siguiente red corresponde a los pacientes geriátricos del hospital de Lyon, Francia, la característica sobresaliente es la densidad (11.68432) y por supuesto, la tendencia a formar grupos es alta (clustering global 0.5880037 y clustering local 0.6402797) y la navegabilidad de la red es muy rápida (1.597838). La última red es el bufete de abogados sin características dominantes, únicamente tiene buena comunicación entre sus miembros (distancia media 2.144385 y grado de cercanía 0.00670515).

A continuación mostraremos una tabla con el top 5 de cada una de las métricas importantes, además que destacamos el lugar que ocupa la red de miembros de la OJN.

Resultados Top 5

Grado (prom)	Densidad	Dist. Media	Clust. global	Clust. local	Closeness	Betweenness	Asortatividad
1. Enron Email	1. Hospital G	1. Hospital G	1. Coautores	1. OJN	1. Papalote	1. Enron Email	1. Enron Email
2. Hospital G.	2. Enron Email	2. OJN	2. Hospital G.	2. Coautores	2. Karate	2. Coautores	2. Coautores
3. OJN	3. Papalote	3. Papalote	3. Papalote	3. Hospital G	3. Hospital G	3. Facultad UK	3. Football
4. Coautores	4. Karate	4. Abogados	4. Facultad UK	4. Karate	4. Abogados	4. OJN	4. Papalote
5. Facultad UK	5. Facultad UK	5. Enron Email	5. Football	5. Papalote	5. Delfines	5. Football	5. Hospital G
	9. OJN		10. OJN		8. OJN		9. OJN

Finalmente, realizamos un análisis más específico a la red de miembros de la OJN, en esta red el nodo principal es el nodo 1, ya que este nodo es la organización como tal, debido a que, si omitimos este nodo, la red sería una red no conexa y tendríamos problemas para trabajar con esta red, esto se debe a que los vínculos representan relación familiar y por supuesto, no se trata de una sola familia. Lo anterior, podemos observarlo en la formación de subredes, éstas representan a las familias, y observamos que hay familias mucho más numerosas que otras, por lo tanto, encontramos que la familia con más miembros dentro de la organización es la familia Kimura, las siguientes familias con un número importante de miembros son la familia Teramoto y Fukumoto, aunque dentro de la familia Kimura encontramos que Juan Carlos y Paulina Kimura son los miembros con mayor número de relaciones (cuentan con un mayor número de arcos). Sin embargo, dentro de esta red encontramos que Amy Murakami es el miembro que tiene el mayor número de vínculos de la red, es decir, es un nodo central, también podemos usar el término introducido por J. Moreno, Amy Murakami es una estrella sociométrica. Además, hay miembros que sólo están conectados con el nodo principal y que, además de ese arco no cuentan con otro, ya que no tienen más miembros de su familia dentro de la organización. Cabe mencionar que, aunque el clustering global es muy bajo, nos interesa más conocer el clustering local, el cual es muy alto, y es de esperarse, debido a que, es una comunidad más o menos pequeña en donde se convive frecuentemente y es fácil conocer y tener relaciones de amistad con la mayoría de los miembros, e incluso, cuando no se tiene una relación de amistad, casi toda la comunidad se conoce.

Conclusiones

Por último, hablaremos de forma general de las conclusiones a las que llegamos. Al hablar de redes sociales, no todos los individuos se comportan de la misma manera ni se relacionan igual, por eso tenemos redes tan diferentes, con distintas propiedades, características y estructuras. Por ejemplo, dentro de la topología o la estructura de las redes sociales, se dice que tienden a ser asortativas, pero con este estudio observamos que no es así, ya que las redes sociales también pueden ser neutras o disortativas, es por eso que no podemos generalizar que las redes sociales se comportan de manera asortativa. De manera específica, es bastante común encontrar relaciones homofílicas, sin embargo, los grupos sociales también pueden ser heterofílicos. Una característica generalizada, no solo de redes sociales, es la propiedad de mundo pequeño, en particular las redes que estudiamos, todas cuentan con esta propiedad, es decir, todas nuestras redes tienen una distancia media baja, claro unas más que otras. Por otra parte, observamos que todas nuestras redes tienen un clustering promedio local más o menos alto. Esta es una característica que se presenta en las redes sociales, debido a la tendencia a formar grupos pequeños.

Con respecto a la distribución de grado, hay una tendencia de las redes sociales hacia modelos de mundo pequeño, sin embargo, también pueden ser representados con modelos de libre escala. La red de los miembros de la OJN, es representada como un modelo de redes de libre escala, es decir, tiende a presentar hubs, además de una distancia media muy baja y un clustering local muy alto. También, podemos especificar que el funcionamiento de la red es muy bueno porque tiene buena comunicación y trabajo en equipo. Por otra parte, es muy claro que la red de miembros de la OJN es completamente homofílica, es decir, todos los miembros de la organización comparten características similares, representan a un rango de edades determinado, además que todos tienen más o menos el mismo nivel de educación y clase social, por lo tanto, la comunicación fluye de una manera más sencilla. Por último, podemos decir que la red de miembros de la OJN es una buena red social, ya que cuenta con características como que es una red de libre escala, presenta homofilia, tiene una distancia media pequeña y un clustering promedio local alto. Sin embargo, como habíamos mencionado, la OJN ha estado teniendo problemas en la disminución de miembros, con base en nuestro análisis llama la atención que el clustering local sea tan alto, por lo que la extremadamente alta tendencia a formar grupos pequeños o clusters, podría indicarnos que falta apertura de las subredes hacia el grupo en general, es decir, mayor comunicación entre todos los miembros de la red.

Cabe mencionar, antes de realizar este análisis ya conocía un poco esta red, de hecho pertenezco a ésta e incluso esperaba algunos de los resultados anteriores, como el hecho de que la familia más importante de la red es la familia Kimura, pero yo pensaba que los nodos centrales eran Juan Carlos y Paulina Kimura, ya que me encuentro más relacionada a esta familia, podía ver que eran las personas que contaban con un grado muy alto dentro

de la organización, sin embargo, después del análisis nos dimos cuenta que esto no era así, ya que el nodo central de la red es Amy Murakami Tsutsumi. Lo anterior tiene mucho sentido, en la organización Amy tiene un gran número de familiares, tanto de la parte materna como paterna, por lo que tiene una gran cantidad de vínculos en la comunidad.

Finalmente, este tipo de análisis no sólo nos sirve para redes sociales, sino que se pueden realizar diferentes tipos de análisis tanto a redes de áreas específicas como tecnología, biología, transporte, lingüística, entre otras, como análisis comparativos entre las distintas áreas. En general, La finalidad de los análisis y la investigación puede ser comprender mejor las estructuras y la topología, las propiedades locales y globales, así como características específicas de sistemas reales modelados como redes.

Anexos

En esta sección de la tesis se muestra el programa realizado en R, que se utilizó para obtener los resultados anteriores.

```
library(igraph) # Paquetería Igraph
# Base datos Karate 1977
karate = read.graph(C: \Users\Sashiko\Downloads\karate.gml, format = c(gml))
karate # En este paso se importa la base de datos a R y se visualiza la base
tkplot(karate, vertex.color = "gray", main = "Karate") # Se hace la red
graph.density(karate) # Obtenemos la densidad
transitivity(karate) # Clustering global
transitivity(karate, type = "average") # Clustering promedio
transitivity(karate, type = "local") # Clustering local
closeness(karate) # Grado de cercanía
mean(closeness(karate)) # Grado de cercanía promedio
betweenness(karate) # Centralidad intermedia
mean(betweenness(karate)) # Centralidad intermedia promedio
coreness = graph.coreness(karate) # k-core
coreness # Se muestra el resultado del k-core
make_k_core_plot = function (g)
{ lay1 = layout.fruchterman.reingold(g)
plot(g, vertex.color = graph.coreness(g), layout = lay1, edge.arrow.size = .5) }
make_k_core_plot(karate) # Gráfica del k-core
is.connected(karate) # Nos indica si la red es conexas
max(degree(karate, mode = "in")) # Grado máximo interior
max(degree(karate, mode = "out")) # Grado máximo exterior
max(degree(karate, mode = "all")) # Grado máximo total
diameter(karate) # Diámetro
assortativity.degree (karate, directed = FALSE) # Asortatividad
mean_distance(karate, directed = FALSE, unconnected = FALSE) # Distancia media
clique.number(karate) # Número de cliques
cliques(karate) # Nos muestra los cliques
dd.karate = degree.distribution(karate) # Se obtiene la distribución de grado
d = 1:max(d.karate)-1
ind = (dd.karate != 0)
plot(d [ind], dd.karate[ind], log = xy, col = blue ,type = l, xlab = c( Log-Grado ), ylab =
c( Log-Intensidad ), main = Log-Log Degree Distribution ) # Gráfica de la distribución
de grado
```


Bibliografía

- [1] Ball B., Newman M.E.J., Friendship networks and social status. *Network Science* **1,16-30**. (2013).
- [2] Barabási A.-L., *Linked: The New Science of Networks*. Cambridge, Mass.: Perseus Publishing. (2002).
- [3] Barabási A.-L., Albert R., Emergence of scaling in random networks. *Science* *286 (5439)*: **509-512**.(1999).
- [4] Barrat A. Barthlemy M., Vespignani A., *Dynamical Processes on Complex Networks*. Cambridge University Press, New York, NY, USA. (2008).
- [5] Bollobás B., *Modern Graph Theory*.Springer. (1998).
- [6] Buchanan M., *Nexus: Small Worlds and the Groundbreaking Science of Networks*. New York, N.Y.: W.W. Norton. (2002).
- [7] Caldarelli G., Catanzaro M., *Networks: A Very Short Introduction*. Oxford University Press, Oxford. (2012).
- [8] Caldarelli G., Vespignani A., *Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Science*. World Scientific Publishing Co., Inc., USA. (2007).
- [9] Christakis N.A., Fowler J.H., *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*. Little, Brown. (2009).
- [10] Dodds P. S., Muhamad R., Watts D.J., An Experimental Study of Search in Global Social Networks. *Science* *301*, **827-829**. (2003).
- [11] Doerr C., Tang S., Blenn N., Van Mieghem P., Are Friends Overrated?. *NETWORKING 2011 - 10th International. 6 Networking Conference*, Valencia, España. (2011).
- [12] Dorogovtsev S.N., *Lectures on Complex Networks*. Oxford University Press, Inc. New York, NY, USA. (2010).
- [13] Easley D., Kleinberg J., *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, Cambridge. (2010).

- [14] Erdős P., Rényi A., On Random Graphs. I. *Publicationes Mathematicae* 6: **290-297**. (1959).
- [15] Granovetter M.S., The Strength of Weak Ties. *The American Journal of Sociology* 6, **1360-1380**. (1973).
- [16] Kolaczyk E.D., Csárdi G., *Statistical Analysis of Network Data with R*. Springer, New York, USA. (2014).
- [17] Lazarsfeld P.F., Merton R.K., Friendship as Social Process: A Substantive and Methodological Analysis. *Freedom and Control in Modern Society*, New York, Octagon. (1964).
- [18] Lazega E., *The Collegial Phenomenon: The Social Mechanisms of Cooperation Among Peers in a Corporate Law Partnership*. Oxford University Press, Oxford (2001).
- [19] Martin T., Zhang X., Newman M.E.J., Localization and centrality in networks. *Physical Review Letters* 90,052808. American Physical Society. (2014).
- [20] Merton R.K., The Matthew Effect in Science. *Science* 159 (3810), **56-63** (1968).
- [21] Milgram, S., The Small World Problem. *Psychology Today* 2: **60-67**.(1967).
- [22] Moreno J.L., *Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy and Sociodrama*. Beacon House. (1953).
- [23] Nepusz T., Petroczi A., Negyessy L., Bazso F., Fuzzy communities and the concept of bridgeness in complex networks. *Physical Review E* 77:016107. (2008).
- [24] Newman M.E.J., *Networks: An Introduction*. Oxford University Press, Oxford (2010).
- [25] Newman M.E.J., The Structure and Function of Complex Networks. *SIAM Review* 45,**167-256**. (2003).
- [26] Newman M.E.J., Barabási A.-L., Watts D.J., *The Structure and Dynamics of Networks*. Princeton University Press, Princeton, NJ, USA. (2006).
- [27] Newman M.E.J., Peixoto T.P., Generalized Communities in Networks. *Physical Review Letters* 115, **08870**. **American Physical Society. (2015)**.
- [28] Padgett J.F., Ansell C.K., Robust Action and the Rise of the Medici, 1400-1434. *American journal of sociology*, **1259-1319**. (1993).
- [29] Richards W, Wormald N., The evolution and structure of social networks. *Network Science* 2, **326-340**. (2014).

- [30] Scott J.P., Carrington P.J., *The SAGE Handbook of Social Network Analysis*. Sage Publications Ltd. (2011).
- [31] Scott J.P., Carrington P.J., Wasserman S., *Models and methods in social network analysis*. Cambridge University Press, Cambridge. (2005).
- [32] Van Der Hofstad R., *Random graphs and complex networks*. In process. (2014).
- [33] Vanhems P. , Barrat A., Cattuto C., Pinton J.-F., N. Khanafer, Regis C., Kim B.-A., Comte B., Voirin N., Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PloS One* 8(9), e73970 306 (2013).
- [34] Wasserman S., Faust K., *Social network analysis: Methods and applications*. Cambridge University Press, Cambridge. (1994).
- [35] Watts D.J., *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton, N.J.: Princeton University Press. (1991).
- [36] Watts D.J., Strogatz S.H., Collective dynamics of 'small-world' networks. *Nature* 393, **440-442**. (1998)
- [37] Yuan S.M., Daraghmi E.Y., We are so close, less than 4 degrees separating you and me!. *Computers in Human Behavior*, 30, **273285**. (2014).
- [38] <http://www.network-science.org/> (27 Agosto 2015)
- [39] <http://www.personal.umich.edu/~mejn/netdata/> (27 Agosto 2015)
- [40] <http://www.cis.jhu.edu/~parky/Enron/> (2 Septiembre 2015)
- [41] <http://www.orgnet.com/sna.html> (27 Agosto 2015)
- [42] <http://creativecommons.org/licenses/by-sa/2.0/uk/> (2 Septiembre 2015)