Redes, Interacciones, Emergencia

Lucas Lacasa, Universidad Politécnica de Madrid, España

Tenemos la física de lo muy grande, la física de lo muy pequeño ... y la física de lo mucho. Y en este último caso, parece fundamental el conocer cómo está conectado ese mucho. Pues finalmente, el todo es más que la suma de sus partes porque las partes no suman, simplemente. Se agregan en un enjambre de interconexiones. La arquitectura de la complejidad.

1. Introducción

La tradición en ciencia, ya sea en física o en biología, ha sido hasta la fecha eminentemente reduccionista: romper el problema en sus ladrillos básicos y encontrar la solución como una sencilla extrapolación lineal del comportamiento individual. Como se ha puesto de manifiesto en la parte tercera de este libro, este enfoque reduccionista, aún cuando los logros y el avance científico asociado son incuestionables, es ineficiente cuando el sistema bajo estudio está formado por un alto número de elementos que interaccionan de forma no lineal entre sí. El comportamiento complejo *emerge* de la agregación de estos elementos y para una correcta descripción de las propiedades macroscópicas del sistema se antoja necesario estudiar el sistema en su totalidad, describiendo tanto las propiedades locales e individuales de cada elemento como la arquitectura que emerge de las interconexiones y agregación del conjunto. Ejemplos existen por doquier: así como el cambio dramático que evidencia el agua al reducir la temperatura de la misma por debajo de $0°$ Celsius no puede explicarse únicamente a partir del estudio de las propiedades físico-químicas de la molécula H_2O, la extinción de especies en un ecosistema no puede entenderse a partir de las relaciones de predación o mutualismo entre dos especies separadas de su entorno. Es la *arquitectura* formada por las interacciones entre las moléculas de agua, o la formada por las interacciones entre cada especie que constituye un ecosistema, la que nos abre las puertas al entendimiento de los fenómenos colectivos (transiciones de fase, extinciones en cascada) que acaecen en el seno de esos sistemas. En los últimos años, un auténtico cambio de paradigma en la forma de entender los sistemas complejos está emergiendo, al constatar que una forma natural de describir dicha arquitectura es mediante un aparato matemático denominado red. Una red, formada por elementos (nodos) conectados entre sí por enlaces que cuantifican la interacción entre ellos. Aunque las herramientas para describir estos objetos -denominados grafos en la comunidad matemática- fechan de mediados del siglo pasado, no es hasta el principio del siglo XXI cuando esta aproximación se consolida como fundamental para el estudio y la descripción de sistemas complejos que muestran comportamiento emergente. La razón de este desfase es bien sencilla: el acceso a datos empíricos (experimentos) con los que construir dichas redes no fue posible hasta hace pocos años, cuando la capacidad de procesamiento y cómputo de los ordenadores experimentase un crecimiento explosivo. Vivimos pues en la era de los datos. Con la ayuda de los ordenadores, modelos teóricos de interacción en red se están viendo comprobados diariamente. Y como no, cada respuesta plantea muchos otros interrogantes. ¿Qué información de un sistema complejo puede extraerse de la arquitectura que emerge de su red de interacciones?, ¿Qué relevancia tiene esta arquitectura en el comportamiento dinámico del sistema?, ¿Qué tipo de arquitecturas podemos encontrarnos en las redes de interacción?, ¿Por qué? En este capítulo trataremos de dar respuesta a algunas de estas cuestiones, haciendo un buceo en el mundo de las redes complejas y su aplicación en diferentes problemas, desde la transmisión de epidemias o los algoritmos de búsqueda hasta la extinción de especies. Acabaremos el capítulo planteando cuál puede ser el devenir de esta rama científica, cuyo progreso actual es exponencial y cuyo potencial está prácticamente limitado únicamente por nuestra imaginación.

2. Redes complejas: definiciones y ejemplos

Los inicios

Los orígenes de la teoría de redes están relativamente desperdigados en el tiempo. El trabajo seminal, denominado el los siete puentes de Königsberg, data del siglo XVIII y fue planteado por el gran matemático Leonard Euler. En aquel entonces, la ciudad de Königsberg (el antiguo nombre que recibía la actual ciudad rusa de Kaliningrado), que durante el siglo XVIII formaba parte de Prusia Oriental, era atravesada por el río Pregolya, el cual se bifurcaba generando una pequeña isla en el centro de la ciudad y dividiendo a la misma en cuatro partes separadas por agua y únicamente conectadas por un total de siete puentes. El problema que Euler se planteaba era saber si era posible el encontrar un circuito que pasara por cada uno de los puentes de esta ciudad, de tal forma que el caminante regresara al mismo punto habiendo atravesado cada uno de los siete puentes una vez y una sola. La respuesta fue negativa: no existe una ruta con estas características. Para hallar esta solución, Euler necesitó formalizar el problema en términos de un conjunto abstracto de nodos conectados entre si por otro conjunto de enlaces, que caracterizan las regiones terrestres y las conexiones entre ellas (ver figura 1), y analizar las propiedades de este constructo. Su demostración no solo dio lugar al nacimiento de una nueva rama en matemática discreta, la teoría de grafos, sino que la generalización del resultado de Euler en poliedros convexos dio lugar de la mano de Cauchy a la topología.

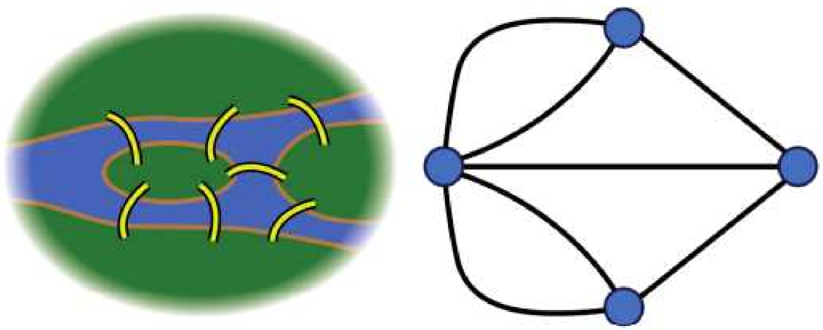

Figura 1: A la izquierda vemos el problema de los siete puentes de Königsberg, y a la derecha la abstracción matemática del mismo en términos de nodos (regiones de tierra) conectados por enlaces.

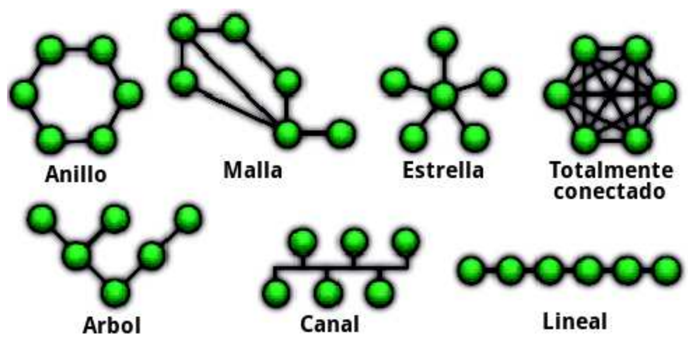

Figura 2: Grafos con diferentes topologías.

Tuvo que pasar casi un siglo, para que el matemático Cayley, que se hallaba estudiando ciertos problemas de cálculo diferencial, se topase en sus pesquisas con ciertas estructuras parecidas a las abstracciones de Euler: un tipo concreto de grafos denominados árboles (grafos acíclicos). Estos resultados tuvieron muchas aplicaciones en la química de la época. A partir de entonces, la teoría de grafos como disciplina matemática con derecho propio tuvo un auge importante, de la mano de científicos como Polya, Sylvester (quien introdujo la palabra *grafo*), Jordan o Kuratowski.

La teoría de grafos, hasta mediados del siglo pasado, estudiaba objetos que no constaban más que de un puñado de nodos (ver ejemplos en la figura 2). Con el desarrollo de las teorías de la probabilidad y la estadística, una nueva vía de estudio en la teoría de grafos tuvo lugar en los años sesenta del siglo pasado, de la mano del gran matemático Paul Erdös. Para los intereses de la subsiguiente teoría de redes este fue el punto de arranque, pues los métodos probabilísticos permitieron estudiar por primera vez las propiedades de grafos arbitrariamente grandes, cambiando ligeramente el enfoque y adquiriendo una

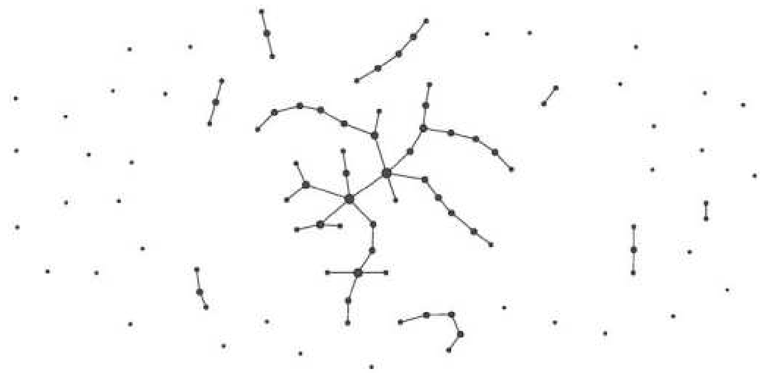

Figura 3: Proceso de formación de redes bajo el modelo Erdös-Renyi.

aproximación estadística. Como concepto opuesto al grafo regular (grafo donde cada nodo se conecta a k vecinos), emerge de la mano de Erdös y su colaborador Renyi el concepto de grafo aleatorio, o grafo de Erdös-Renyi, como aquel grafo generado por un proceso estocástico donde en cada paso dos nodos cualesquiera se conectan con cierta probabilidad p (ver figura 3). Estas aproximaciones teóricas se vieron muy beneficiadas a finales de los años noventa con la llegada de los ordenadores modernos, capaces de manejar una gran cantidad de datos. El estudio estadístico de grafos ya no residía únicamente en desarrollos formales. Los grafos reales, tales como la red de redes (internet) o las redes sociales o biológicas pudieron, por primera vez, ser examinadas y sus propiedades estadísticas calculadas. Y los resultados no pudieron ser más inesperados: se encontraron con que estas arquitecturas eran extremadamente más complejas de lo que la teoría de Erdös describía. Sin embargo, dentro de ese caos de conexiones, ciertos patrones de orden emergían. La nueva era de la teoría de grafos, el análisis de grafos enormes que no eran ni regulares ni aleatorios, se llamó la teoría de redes complejas. Pasen y vean.

Ejemplos de redes complejas

Existen multitud de tipos de redes: algunas son no dirigidas (donde los enlaces no tienen dirección preferente), otras que si lo son, otras donde existen realmente dos conjuntos bien diferenciados de nodos (redes bipartitas), otras donde cada enlace ha de ser pesado según la importancia del mismo, y así sucesivamente. En este capítulo no vamos a discriminar estas características y por sencillez nos centraremos en el caso más general de redes no dirigidas. A continuación, y a modo de casos ilustrativos, enumeramos una lista no exhaustiva de algunas de las redes complejas que podemos observar a nuestro alrededor.

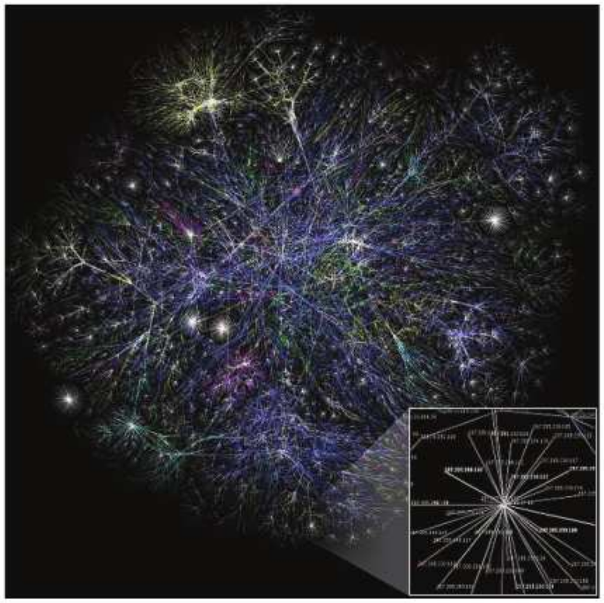

Figura 4: Red de internet.

Redes de información

- WWW: es la red más grande que ha sido analizada hasta el momento. Los nodos de esta red los conforman las páginas web, y dos nodos están conectados por un enlace si existe un hyperlink en una página que apunta a la otra. Esta red ha tenido un proceso de crecimiento descentralizado y autoorganizado, por lo que su descripción es de especial interés bajo el paraguas de los sistemas complejos.

- Internet: dependiendo del nivel de estudio, se define como una red donde los nodos son (i) computadoras o routers enlazados entre sí físicamente, o (ii) sistemas autónomos (compuestos de cientos de routers y computadoras) enlazados entre sí. De especial interés para estudiar problemas de propagación de información (donde "información" puede ser: virus informáticos, publicidad, etc).

Redes sociales

- Redes de amistad y trabajo: donde los nodos son individuos y los enlaces constituyen relaciones de amistad (Facebook), interés profesional (LinkedIn), algún grado de interés a nivel informacional (Twitter), etcétera. Interesante a la hora de estudiar comunidades de nodos con cierta afinidad, y en general en problemas de difusión de información.

- Redes de coautoría y colaboración: donde los nodos son individuos y los enlaces entre dos nodos caracterizan individuos que han colaborado. En el caso de redes de actores, la colaboración sería haber trabajado en la misma película (datos a través del Internet Movie Database), mientras que en el caso de redes de coautoría científica el enlace caracteriza el haber participado en la publicación de un artículo.

- Redes de email: donde los nodos son individuos digitales y dos individuos están enlazados si uno de ellos posee el email del otro.

- Redes de contactos sexuales: donde los nodos son individuos y los enlaces denotan un contacto sexual. Interesante a la hora de elaborar estrategias de planificación y prevención de epidemias asociadas a enfermedades de transmisión sexual.

Redes naturales

- Redes de regulación genética: Cada nodo es una función booleana (cuya entrada es un conjunto de números binarios $\{1/0\}$ y salida es a su vez un número binario). Son modelos de expresión genética: la expresión de un gen, es decir la producción por transcripción y traslación de la proteína que el gen codifica, es controlada a su vez por la presencia de otras proteínas tanto activadoras $\{1\}$ como inhibidoras $\{0\}$. De esta forma, el genoma mismo no es sino un conjunto de elementos en constante cambio *encendido/apagado*, donde los nodos representan las proteínas y los enlaces (dirigidos) representan la dependencia en la producción de cierta proteína respecto a otras proteínas (sus nodos vecinos). Estas dependencias se integran a través de las funciones booleanas anteriormente descritas. Los primeros modelos son bastante anteriores a la explosión de las redes en la primera década del siglo, y fueron planteados por el biólogo teórico Stuart Kauffmann [1]

- Redes neuronales y el cerebro: existen diferentes aproximaciones a este problema, según la escala de estudio, desde redes formadas por nodos (neuronas) conectadas entre si, hasta las llamadas redes funcionales, donde los nodos son áreas neuronales y los enlaces indican una implicación funcional.

- Redes de plegamiento de proteínas: Cuando una proteína se pliega (paso previo a la expresión de la misma), toma secuencialmente diferentes conformaciones. Se representa cada nodo como una de estas configuraciones, de tal forma que dos nodos se conectan si existe un movimiento elemental que pasa de una a otra configuración. Se emplea este enfoque en problemas de dinámica de polímeros y en el estudio de proteínas.

- Redes tróficas: donde los nodos representan especies de un ecosistema y los enlaces, dirigidos, indican predación (alternativamente, el enlace puede ir en sentido con-

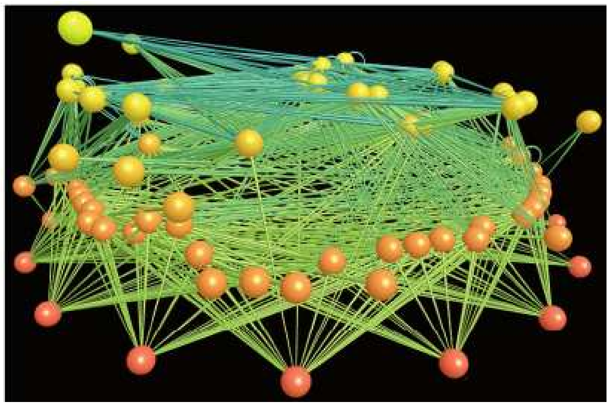

Figura 5: Ejemplo de una red trófica.

trario si lo entendemos como flujo de energía de la presa al predador). Un ejemplo puede visualizarse en la figura 5.

- Redes mutualistas: similar a la red trófica, caracteriza relaciones de mutualismo (donde ambas especies se benefician de la interacción) en lugar de predación.

Redes de infraestructura

- Red eléctrica: formada por nodos (centrales eléctricas, puntos de consumo) conectadas por enlaces que caracterizan a los cables de transmisión eléctrica. De fundamental importancia para la prevención de apagones en cascada.

- Redes de transporte aéreo: donde los nodos son aeropuertos y los enlaces son rutas aéreas. De interés para estudiar el problema de la propagación de retrasos, y en problemas de transmisión de epidemias mediante mecanismos superdifusivos.

Medidas clave

Como hemos comentado anteriormente, la teoría de redes complejas abarca el estudio de redes con un alto número de nodos, de tal forma que sus propiedades estructurales suelen estudiarse de un punto de vista estadístico, ya sea mediante técnicas matemáticas, mediante análisis por ordenador, o como una mezcla de ambas. En esta sección presentamos algunas de las medidas estadísticas clave en la descripción de las propiedades más relevantes de una red.

Camino medio $L(N)$. Es el camino mínimo entre dos nodos cualesquiera de una red, medido en saltos de nodo a nodo a través de enlaces, promediado a todos los pares de nodos

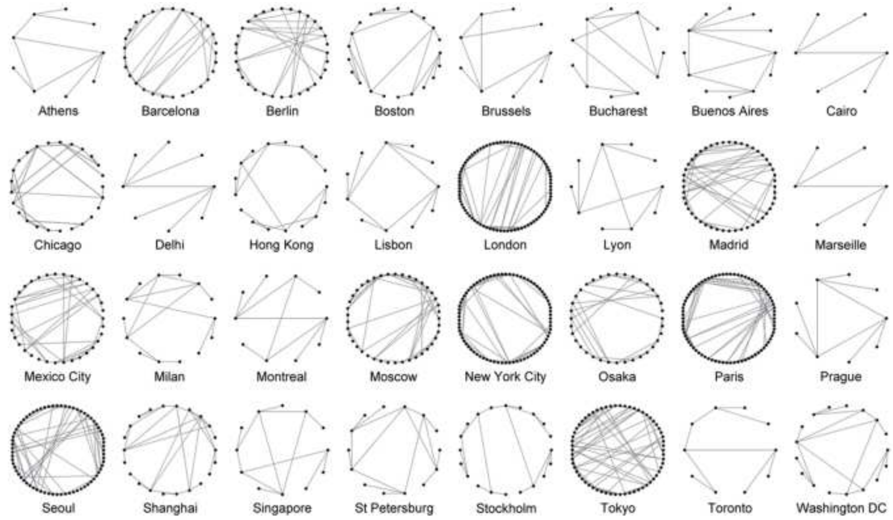

Figura 6: Topología de redes de metro de diferentes ciudades. Adaptado de Derrible [2].

de una red. Es una cantidad que caracteriza la distancia topológica a la que se encuentran, en promedio, cada uno de los nodos de una red de todos los demás. Este valor depende de la cantidad de nodos N que tenga una red, y es una medida central a la hora de entender el famoso dicho *el mundo es un pañuelo* (ver sección *redes Small-World*).

Coeficiente de clustering C. Caracteriza la probabilidad que tienen los nodos conectados a un tercero de conectarse entre si mismo. Coloquialmente, esta cantidad mide la tendencia de que *mis amigos sean amigos entre si*. Si el clustering es alto, la cantidad de motivos triangulares en la red es alto, como suele suceder en redes de tipo social.

Distribución de grado $P(k)$. Caracteriza la probabilidad de que un nodo al azar escogido de la red tenga k enlaces. La forma de calcular esta cantidad es mediante un histograma: empezamos contando cuántos enlaces tiene cada nodo, y finalmente contamos la cantidad de nodos con un enlace, con dos, etcétera. Las redes homogéneas tienden a tener la misma cantidad de enlaces por cada nodo (en la terminología de grafos, aquellas redes donde todos los nodos se conectan con exactamente k vecinos se denominan grafos k-regulares, y sus propiedades son bien conocidas). Las redes inhomogéneas tienden a tener una distribución de nodos asimétrica. Son de especial interés, como veremos más adelante, aquellas redes cuya $P(k)$ tengan una caída lenta, concretamente en ley de potencias $P(k) = Ak^{-\gamma}$.

Matriz de adyacencia A_{ij}. Es una matriz que aglomera la relación de enlaces entre nodos. Si la red tiene N nodos, la matriz de adyacencia será una matriz $N \times N$ donde cada posición (i, j) de la matriz será, respectivamente, 1 si los nodos i y j comparten un enlace,

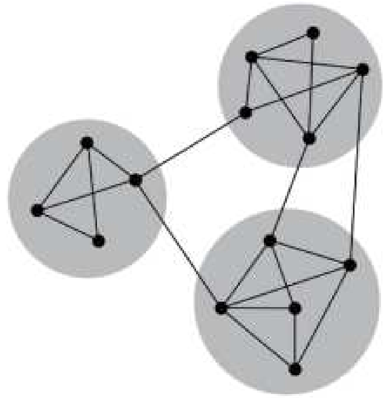

Figura 7: Red donde una estructura en comunidades evidente.

y 0 en caso contrario. Esta matriz engloba toda la información de la red, de tal forma que se puede definir una relación biunívoca entre una red y su matriz de adyacencia. Por ejemplo, estudiando las propiedades espectrales de la matriz de adyacencia A_{ij} (es decir, haciendo un estudio de los autovalores y autovectores de esa matriz), podemos, sorprendentemente, responder a preguntas tales como ¿cuán enmarañada está una red? Además, estudiando las potencias de A_{ij} podemos calcular la cantidad de maneras de circular entre el nodo i y el nodo j (o, dicho de otra forma, conocer el número de caminos que unen sendos nodos, siendo esta una medida de la redundancia en la comunicación de estos).

Existen muchas otras medidas que caracterizan una red, que por su complejidad, sólo describiremos someramente. Los *motivos* del grafo (donde un motivo es un subgrafo de la red, entendiendo a esta última como la agregación de diversos motivos) son los ladrillos básicos de una red, y su análisis es muy relevante en el estudio de las propiedades funcionales de redes biológicas. Por otro lado, la *asortatividad* caracteriza la tendencia con la que nodos con alto número de enlaces tienden a estar conectados entre sí (en el caso opuesto, se habla de *disasortatividad*), y da una medida de las correlaciones de grado entre nodos. Otras propiedades incluyen medidas de meso-escala como son las comunidades (conjuntos de nodos que comparten ciertas características -ver figuras 7 y 8-), la modularidad, la centralidad, y un largo etcétera. Encarecemos al lector interesado en recurrir a textos específicos como [4].

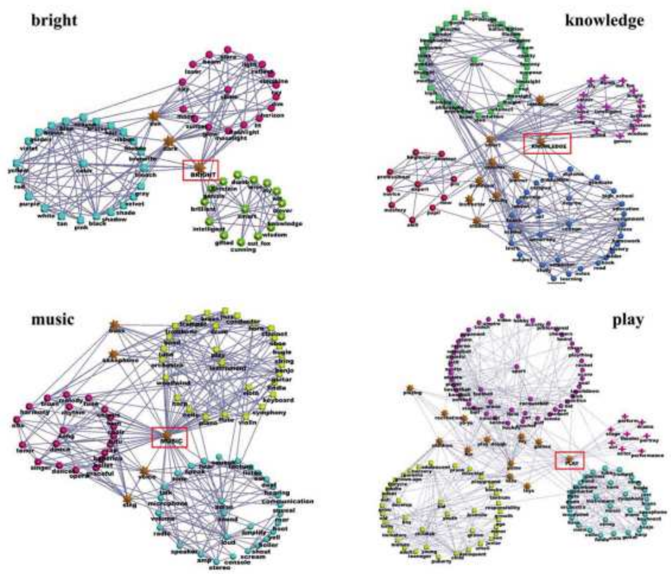

Figura 8: Redes complejas de diferente índole donde se resaltan las estructuras de comunidades, mediante un algoritmo de optimización (OSLOM). Adaptado de Lancichinetti y coautores [3]

3. Estructura y dinámica: propiedades

Redes Small-World: el mundo es un pañuelo

El experimento de Milgram

En los años 60, Stanley Milgram, un sociólogo que trabajaba en la universidad de Yale, planteándose qué estructura tendría la red social americana, diseñó el siguiente experimento. Milgram envió 160 paquetes a personas desconocidas escogidas al azar que habitaban en el estado de Nebraska. Junto con cada paquete, las instrucciones: "*el propósito del experimento es que este paquete llegue a manos de XXX*" -un corredor de bolsa desconocido que vivía en Boston, Massachussets-. "*No puede enviarlo directamente a su dirección, en cambio, tienen que enviárselo a alguna persona que usted conozca, que piense pueda conocer a su vez al objetivo, o al menos, tener una probabilidad mayor de conocer al objetivo que usted. Le enviará este paquete junto con estas mismas instrucciones*".

Milgram monitoreo el flujo de paquetes a través de Estados Unidos. Estos paquetes fueron saltando entre personas conocidas, difundiéndose en la red social igual que la tinta se difunde en el agua. Finalmente, computó, de entre aquellas caminatas que alcanzaron el objetivo, la cantidad de intermediarios, es decir, la longitud (en pasos de red) entre el individuo inicial y el objetivo. La conclusión de este trabajo fue inesperado. Pese a que la población estadounidense rozaba los 250 millones de personas, lo que *a priori* sugería que

Figura 9: Esquema gráfico del modelo de interpolación entre red regular y red aleatoria de Watts-Strogatz. Adaptado de Humphries y Gurney [5].

el camino medio de la red (el número de intermediarios promedio) fuese muy grande, el camino medio encontrado por Milgram fue únicamente de seis pasos (los llamados seis grados de separación). Seguramente, sorprendido, concluyese en primera instancia ¡el mundo es un pañuelo! Aunque, ¿cómo es posible que en una red tan inmensa, cualquier par de nodos estén, en promedio, tan próximos? La respuesta llegó en 1999 de la mano de un modelo ingenioso y de extremada sencillez, propuesto por el físico estadounidense Steven Strogatz y su estudiante de doctorado, Duncan Watts.

El modelo de Watts-Strogatz

El modelo de mundo pequeño (Small-World) de Watts y Strogatz es aquella idea tan genial como sencilla, que por su simplicidad, uno se da de cabezas contra la pared pensando: ¿por qué no se me ocurrió a mi primero? La situación era la siguiente. En la teoría de grafos, existían dos modelos canónicos y antagónicos. El primero, un grafo llamado *lattice*, un modelo de red totalmente regular. En esta red, el camino medio escala linealmente con el tamaño del sistema, por lo que no funciona para explicar los seis grados de separación de Milgram, ni dar cuenta de esa medida en las redes biológicas o tecnológicas estudiadas. Sin embargo el clustering en la red tipo lattice es especialmente alto, parecido al que se encontraban en las redes sociales. El modelo de lattice, por tanto, sólo funcionaba a medias para describir la realidad. El segundo modelo canónico, ampliamente usado en epidemiología, era la red aleatoria de tipo Erdös-Renyi, de la que hablamos al principio del capítulo. Esta red posee un camino medio especialmente bajo, que escala logarítmicamente con el número de nodos, en acuerdo con lo encontrado en las redes reales. Sin embargo el clustering es también extremadamente bajo, y tiende a cero al aumentar el tamaño del sistema. Tampoco funcionaba. Y si tenemos dos modelos que funcionan a medias, ¿por qué no montarnos un proceso que interpole sendas redes? El modelo de Watts-Strogatz comienza con una red regular (con alto clustering y bajo camino medio). Con cierta probabilidad p, se escogen dos nodos al azar de esta red y se conectan (este proceso se denomina *rewiring*). Este proceso se repite, de tal forma que el número total de enlaces de la red se conserve (es decir, si un nodo sufre rewiring pierde su enlace anterior en favor del nuevo). Es fácil observar que para valores de p bajos este efecto es irrelevante, mientras que si p es muy grande (cercano a uno), la red original queda aleatorizada (y se convierte en una Erdös-Renyi). El efecto del rewiring es tender puentes o atajos entre pares de nodos, lo que hace disminuir el camino medio de la red (ver figura 9). Ahora bien, el clustering también se reduce mediante rewiring, aunque esa reducción es relativamente más lenta que la del camino medio. Por tanto, todo parecía sugerir que existiría una situación intermedia, en donde para algún valor de p, la red resultante estaría dotada de un camino medio bajo, y un clustering relativamente alto. ¡Igual que las redes reales! Watts y Strogatz publicaron este mecanismo en la revista *Nature* [6], acompañado por una lista de redes reales cuyas medidas de camino medio y clustering coincidían con las arrojadas por su modelo. A la red teórica resultante, como un engendro a medio camino entre el orden y el desorden, la denominaron red small-world, un tipo de red compleja.

Efectos dinámicos asociados a la propiedad de mundo pequeño

Una consecuencia muy importante del llamado efecto small-world, es que la navegación por una red con estas características es muy rápida (recuerde el mecanismo de ”atajo” planteado en el modelo de Watts-Strogatz). De hecho, propiedades dinámicas como puede ser la difusión de información en este tipo de redes es extremadamente eficaz. Ahora bien, téngase en cuenta que por información puede ser cualquier cosa: desde un virus que se difunde entre los individuos de una red social, un apagón en un nodo de una red eléctrica que genera otro apagón en los nodos vecinos, una especie que al desaparecer perturba al ecosistema a través de la red de interacciones tróficas, incluso el retraso de un avión en un aeropuerto, que genera un retraso en la salida de otro vuelo, lo que transmite otros retrasos en los aeropuertos vecinos (red de transporte aéreo). Por tanto, para una misma dinámica local, la dinámica global que emerge en una estructura de mundo pequeño es radicalmente diferente a la que emerge de una red regular.

Redes libres de escala

En la sección anterior explicamos el llamado efecto small-world como resultado del balance entre dos propiedades importantes de una red: el camino medio $L(N)$ y el clustering C. Otra propiedad fundamental para entender el efecto que tiene la estructura de una red en las propiedades del sistema complejo asociado es la llamada distribución de grado (ver definición). Recordamos que en una red homogénea todos los nodos de la red tienen, en promedio, una cantidad de enlaces parecida. Ejemplos de redes de este tipo son las redes regulares (donde todos los nodos tienen exactamente la misma cantidad de enlaces), o las redes aleatorias Erdös-Renyi (donde la mayoría de nodos tienen una conectividad igual a la conectividad promedio, mientras que la cantidad de nodos altamente conectados es exponencialmente pequeña). Por el contrario, en una red inhomogenea existirá toda una

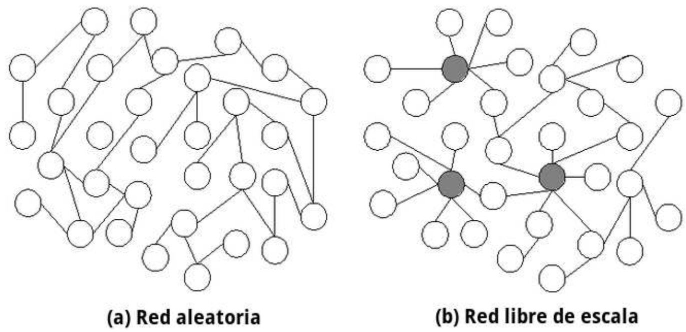

Figura 10: (a) Red aleatoria, cuya distribución de grado con escala característica: la mayoría de los nodos tienen grado similar. (b) Red libre de escala, donde existen ciertos nodos con alto grado (en gris).)

jerarquía de nodos. Se denominan redes libres de escala las redes con una distribución de grado que sigue una ley de potencias $P(k) = Ak^{-\gamma}$ (donde γ es el exponente característico y A la constante de normalización). En una red con esta cualidad, existirán muchos nodos poco conectados (con "poca importancia") pero existirán una cantidad de nodos no despreciable con una conectividad inmensa. De hecho, no existirá una escala característica: la proporción de nodos con ciertas conexiones se repetirá a todos los niveles (como recordará el lector de las propiedades de las leyes de potencias, que aparecen en fenómenos críticos).

Estos nodos muy conectados, denominados *hubs*, van a resultar clave, y como ejemplo pongamos la siguiente situación: imagínese el lector que descubre una información muy importante que considera ha de hacerse pública, y únicamente dispone del tiempo suficiente para mandar un email con esa información a una única persona de su lista de conocidos. Ha de ser inteligente y buscar, para maximizar su probabilidad de éxito, aquel conocido que a su vez tenga la lista de amigos más grande, ya que de esa forma mandará el email a la persona que es capaz de llegar, con un solo clic, al mayor número de gentes. Ese individuo constituye uno de los hubs de su red social, querido lector.

Sorprendentemente (o quizás no tanto), la gran mayoría de redes reales son libres de escala -recordamos que las redes generadas de forma aleatoria, como las tipo Erdös-Renyi, tiene una distribución de grado con una escala característica del orden de la conectividad media-. Así como el modelo de Watts-Strogatz explicó que el fenómeno small-world podía verse como una red a medio camino entre el orden y el desorden, el afamado mecanismo de *preferential attachment*, postulado por Barabasi y Albert en un artículo en la revista *Science* [7], explica la existencia de redes cuyo grado no tenga una escala característica mediante un sencillo proceso multiplicativo. En este proceso, cuando un nuevo nodo es creado en la red (ya sea un individuo que llega al barrio o una página web de nueva creación), es estadísticamente más probable que este nuevo nodo se relacione con nodos ”importantes” del sistema (aquel personaje popular y extrovertido que te ayuda a integrarte en el barrio, aquella web importante a la que vas a mandar un hyperlink desde tu portal). Este mecanismo, redescubierto a lo largo de las décadas en diferentes contextos (modelo de Simon-Yule en lingüística, efecto Mateo por el cual ”los ricos se vuelven más ricos” en sociología), lo redescubren Barabasi y Albert llamándolo enlace preferencial.

Nótese que si combinamos la propiedad de mundo pequeño (por la que la navegación en la red es rápida) con la propiedad de red libre de escala (por la que la cantidad de hubs -nodos muy conectados- es no despreciable), los mecanismos de propagación asociados al primer mecanismo se ven amplificados. Parece por tanto muy deseable, si deseásemos maximizar intercomunicación en un sistema de este tipo (por ejemplo, para reducir la energía empleada en transportar información de una parte a otra de la red), diseñar una red subyacente que posea sendas cualidades. En los últimos años se ha venido resaltando que de hecho tanto la naturaleza como la sociedad genera tales tipos de arquitecturas... ¡de forma autoorganizada! No existe ningún plan de diseño global de la www, ni de la red metabólica, ni de la arquitectura de la red social: únicamente mecanismos de beneficio mutuo a nivel local. Sin embargo, el proceso evolutivo tiende a optimizar a nivel global dicha arquitectura: pues a través del tiempo las páginas web aparecen, desaparecen y algunas -las mejor adaptadas al entorno- cristalizan, igual que las amistades, igual que las especies.

Asortatividad: un problema de máxima entropía

Un hecho bien conocido desde los análisis iniciales de las propiedades básicas de redes complejas reales constataba que existía una diferencia muy clara en redes tipo social y el resto de arquitecturas: mientras que las primeras tenían un alto grado de asortatividad, donde los nodos con conectividad alta tendían a estar conectados entre sí, las segundas tenían una asortatividad significativamente baja (o equivalentemente, una disasortatividad alta). En los últimos años ha habido mucha especulación en relación al motivo de tal diferencia: mecanismos evolutivos? No existe un plan de diseño predefinido para todas estas redes, luego era necesario un mecanismo lo más general posible. En un trabajo publicado en 2010 en la prestigiosa *Physical Review Letters* [8], Sam Johnson (actualmente en el Imperial College británico) y colaboradores dieron con una solución elegante y matemáticamente rigurosa: el tipo de correlaciones más probable (es decir, el que tiene un número de posibles configuraciones mayor a una distribución de grado fija, o dicho de otra forma, el estado de máxima entropía), coincide con una situación disasortativa en ausencia de mayores restricciones. Este es desde luego un resultado extremadamente general, pues no presupone mecanismos adicionales, para el origen general de la disasortatividad de redes biológicas, tecnológicas o de información. En el mismo trabajo, explican que el caso de las redes sociales introduce una restricción adicional, el llamado mecanismo homofílico por el que grandes personalidades tienden a interactuar, dando lugar a redes sociales asortativas. Este constituye un ejemplo notable de la forma de pensar de los físicos: si existe un comportamiento universal, por muy complejo que este sea, es muy probable que sea debido a un conjunto mínimo de razones.

En resumen, hemos visto que las propiedades de las redes complejas son en alto grado universales y emergen por razones evolutivas o por principios de máxima entropía. Además, estas propiedades parece que le otorgan cierto grado de optimización, al ser redes con cualidades bastante buenas para el transporte de información. En la próxima sección ahondaremos en estas características, para describir y comprender la emergencia de fenómenos colectivos en red, como son la propagación de una epidemia o la extinción masiva de especies. También veremos que estas redes, que parecen óptimamente diseñadas, son muy robustas ante perturbaciones naturales (aleatorias) pero que son extremadamente frágiles frente a ataques dirigidos.

4. Aplicaciones

Fenómenos de propagación: epidemias, apagones y extinciones

Imagine el lector que un buen día abre su periódico preferido, y entre los titulares, lee:

-"*El 10 de Agosto de 1996, el fallo de dos plantas eléctricas en Oregon (Estados Unidos de América) generó un apagón en cascada que dejó sin electricidad a once estados enteros y dos provincias canadienses, dejando a unos siete millones de personas a oscuras durante unas 16 horas. Más información en página 24.*"

Y sigue leyendo:

-"*El virus informático* I love you*, introducido en Internet en el año 2000, ha infectado millones de ordenadores en el mundo entero, generando pérdidas por un valor incalculable. A fecha de hoy sigue siendo un virus no erradicado contra el que uno no puede más que protegerse, nunca matarlo. Más información en la sección de tecnología.*" .

Quizás, la información extendida sobre sendas noticias, versa sobre los problemas asociados a un fallo humano en el caso de la red eléctrica, mientras que en el caso del virus informático el periódico se centra en detallar los antivirus más novedosos. Y a usted lector, que ya se ha familiarizado con los conceptos básicos de las redes, muy seguramente le rechinaría una idea en la cabeza:

-"*Son el efecto combinado de las propiedades de mundo pequeño y de grado libre de escala en sendas redes complejas*", reflexiona.

Y prosigue la reflexión:

-"*En el momento en que un nodo se ve afectado (ya sea una central eléctrica con un incendio o un ordenador donde un programador malicioso introduce un virus), esta infección se propagará, con cierta tasa de infección nodo-nodo. En una red homogénea, esta propagación es lenta y el efecto* *puede ser controlado a tiempo. Pero no en una red compleja. Debido al efecto de mundo pequeño, por la que cualquier nodo de la red es rápidamente alcanzable (incluidos los hubs), este virus llegará a un hub con mucha celeridad. Y ese es el final. Pues una vez que llega al hub, este nodo infectará a una porción enorme de los nodos de la red, a los que está conectado. Si la red fuera aleatoria, la existencia de estos hubs no estaría garantizada, por lo que uno aún tendría tiempo de proteger localmente al resto de nodos. Pero no en una red compleja. La propiedad de red libre de escala garantiza la presencia de nodos con muchas conexiones, lo que dificulta hasta casi echar por tierra cualquier intento de mitigar la propagación del apagón, o de la epidemia cibernética. Con las redes complejas, mucho mejor prevenir que curar."*

Nótese que las noticias anteriores son reales. La razón por la cual un virus informático no puede "erradicarse" fue explicada en un trabajo publicado en *Physical Review Letters* [9] por Romualdo Pastor Satorras y Alessandro Vespignanni. Estudiaron mediante un modelo matemático clásico de difusión de epidemias en qué situaciones cierta infección pasa a ser endémica (no desaparece, convirtiéndose en una pandemia). Si la dinámica corre sobre redes regulares o aleatorias, existe un margen de acción bastante amplio -y ahí es donde tienen efecto las estrategias clásicas de curación- antes de que la infección se vuelva pandémica. Sin embargo, estos investigadores demostraron matemáticamente que en una red libre de escala la transición entre infección controlada y la pandemia es mucho más rápida, por la existencia de hubs.

Este comportamiento no se limita a apagones o virus informáticos. El conocer cómo se propaga la información en este tipo de redes es clave a la hora de planificar estrategias de prevención para la propagación de enfermedades (donde la red social es una red compleja), o a la hora de entender qué efecto tiene la extinción de una especie dentro de un ecosistema (donde la red trófica asociada también es compleja). Ha sido solo muy recientemente, en el seno de los científicos que trabajan en la teoría de redes complejas, donde estos efectos han sido constatados. La estructura afecta muy dramáticamente a la dinámica, y para entender el comportamiento dinámico, cuando está embebido en una red, es necesario conocer la arquitectura subyacente.

Robustez frente a ataques: el caso de una red trófica

La mayoría de las redes complejas que podemos observar en la naturaleza (biología, sociedad, tecnología) son redes autoorganizadas: no existe ningún plan de diseño previo a las mismas. Por otro lado, como ha quedado de manifiesto en las secciones anteriores, los rasgos característicos de las redes complejas (redes de mundo pequeño y libres de escala) parecen universales, en tanto en cuanto redes reales de diferentes ámbitos presentan en su estructura patrones similares. Estos dos hechos complementarios sugieren que el crecimiento de dichas redes y su evolución a lo largo del tiempo atiende a criterios Darwinistas. Siguiendo esta linea de pensamiento, es de interés plantearse cuán robustas son estas redes. Para que una red sea robusta, su funcionamiento ha de ser relativamente independiente de perturbaciones aleatorias externas. Por ejemplo, en un ecosistema es posible que una especie se extinga por un hecho puntual, generando la desaparición aleatoria de un nodo en la red trófica. Bien, pues es posible demostrar que la robustez de esta red trófica es muy alta si la misma tiene la propiedad de ser libre de escala. La naturaleza por tanto es ingeniosa, y a través de la evolución organiza una arquitectura de interconexiones que no sufra dramáticamente si uno de los nodos de la red es eliminado al azar. Denominamos a esta situación robustez frente a ataques aleatorios, y es una propiedad que comparten la gran mayoría de redes complejas que observamos a nuestro alrededor.

Cabe preguntarse qué escenario sucede si la perturbación no es aleatoria (como suele suceder en problemas biológicos) sino dirigida. Dicho de otro modo, si yo deseara causar el máximo impacto en una red con un único ataque, ¿ese impacto puede llegar a ser importante? ¿y en caso afirmativo, cómo he de efectuarlo? En contraste con las perturbaciones aleatorias, se ha demostrado que las redes libres de escala son especialmente vulnerables ante ataques dirigidos: existen nodos clave cuya desaparición puede generar una cascada de desapariciones sucesivas, o puede sencillamente romper la red compleja en dos redes separadas. Estos nodos clave son típicamente hubs del sistema o en general nodos con una alta centralidad. Nótese que este fenómeno altamente no lineal es de fundamental importancia para entender porqué en algunas situaciones, la extinción de una única especie puede derrumbar el ecosistema entero. Y para conocer qué especies pueden jugar este rol, se antoja necesario aplicar el enfoque de red al problema ecológico que estamos comentando. Este tipo de aproximaciones de red en problemas de conservación biológica son muy recientes, y este campo está experimentando una verdadera revolución en los últimos años, de la mano del biólogo teórico Jordi Bascompte y su equipo.

Navegación en red y algoritmos de búsqueda

Una aplicación muy interesante de la teoría de redes complejas se basa en en el estudio de las propiedades de navegación. Como ejemplo curioso veamos el caso de las llamadas redes de recomendación musical. Sea una red donde cada nodo representa una canción, y dos nodos estén enlazados si las canciones asociadas aparecen de forma conjunta en la playlist de un individuo. Tómese en segundo lugar la información contenida en las playlists de muchos individuos. La red resultante crecerá en nodos y enlaces. También, muchas canciones se repetirán, hecho que podemos tener en cuenta dando un peso al enlace entre dos nodos (la cantidad de veces que esas dos canciones co-aparecen en una lista). Una vez construida esta red, lo único que tenemos que hacer para *recomendar* música es partir de un nodo inicial, una canción que el usuario propone, e ir navegando la red. Existen diferentes algoritmos para esta recomendación, desde caminatas aleatorias hasta métodos espectrales (el famoso PageRank de Google siendo un algoritmo que pertenece a esta última familia).

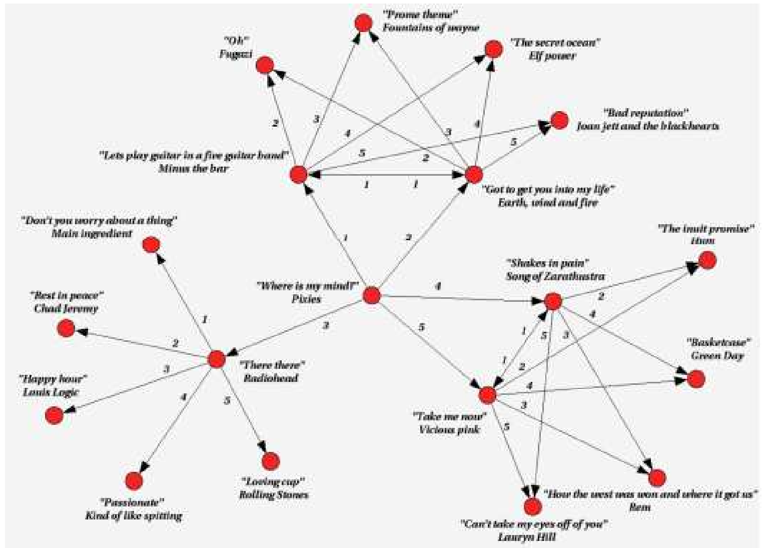

Figura 11: Red de recomendación musical. Adaptado de Buldu y colaboradores [10].

5. Perspectivas de futuro

En este capítulo hemos descrito, de forma somera, cómo el enfoque de red puede arrojar luz en el estudio de sistemas complejos con muchos elementos que interactúan. La teoría de redes complejas como disciplina científica fecha de finales del siglo XX, y en la actualidad es una rama en constante crecimiento que muy probablemente resultará clave en el desarrollo científico del siglo XXI, dentro de las denominadas ciencias de la complejidad. Sin embargo en la actualidad existen muchas más preguntas que respuestas y la teoría de redes complejas dista de ser un corpus científico cerrado y maduro. A continuación planteamos algunos de los hitos concretos a los que se enfrenta esta disciplina:

- Las medidas estadísticas que podemos realizar sobre una red compleja son muy diversas. ¿De qué modo se relacionan? ¿Cuáles son las medidas clave que caracterizan plenamente la estructura relevante, y cuáles son medidas secundarias?

- La descripción de procesos dinámicos embebidos en red es una parte fundamental de la teoría de redes que no hemos hecho más que empezar a descubrir. Se antoja necesario profundizar en las propiedades dinámicas para entender de qué forma emergen los comportamientos colectivos en sistemas complejos formados por muchos elementos.

- Mientras que la estructura de una red afecta a la dinámica que corre encima, en muchos casos estos problemas no pueden estudiarse por separado: los llamados escenarios de co-evolución, donde la estructura de una red se adquiere de forma dinámica, son problemas interesantes cuyo estudio no ha hecho más que comenzar.

- La modelización y el estudio de las condiciones bajo las cuales existen cambios dramáticos en una red pueden arrojar luz en fenómenos tales como las extinciones masivas, los crashes bursátiles o la propagación de fallos en redes de infraestructura. El conocimiento adquirido podrá dar lugar a un conjunto de estrategias de prevención orientadas a la conservación de la biodiversidad, la estabilidad financiera o la estabilidad de nuestras infraestructuras críticas (transporte, energía).

- Por último, es interesante plantear la teoría de redes como un método alternativo para estudiar problemas dinámicos y series temporales. La dinámica generada por un sistema complejo, como por ejemplo la turbulencia o la actividad cerebral, está hoy en día poco entendida. Se antojan nuevos métodos que describan, de manera profunda, la estructura que opera debajo de las series de tiempo generadas por estas dinámicas. Existen aproximaciones novedosas para estudiar dichas estructuras, que transforman una serie de tiempo a una red y trabajan sobre la red asociada para describir patrones de orden de la dinámica subyacente. Un ejemplo notable en esta dirección es el llamado algoritmo de visibilidad [11].

6. Referencias

[1] S. Kauffman, *The origins of order: Self-organization and selection in evolution*. Oxford University Press, USA, 1993.

[2] S. Derrible, "Network centrality of metro systems," *PloS one*, vol. 7, no. 7, p. e40575, 2012.

[3] A. Lancichinetti, F. Radicchi, J. Ramasco, and S. Fortunato, "Finding statistically significant communities in networks," *PloS one*, vol. 6, no. 4, p. e18961, 2011.

[4] M. Newman, "The structure and function of complex networks," *SIAM review*, vol. 45, no. 2, pp. 167–256, 2003.

[5] M. Humphries and K. Gurney, "Network 'small-world-ness': a quantitative method for determining canonical network equivalence," *PLoS One*, vol. 3, no. 4, p. e0002051, 2008.

[6] D. Watts and S. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, pp. 440–442, 1998.

[7] A. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.

[8] S. Johnson, J. Torres, J. Marro, and M. Munoz, "Entropic origin of disassortativity in complex networks," *Physical review letters*, vol. 104, no. 10, p. 108702, 2010.

[9] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical review letters*, vol. 86, no. 14, pp. 3200–3203, 2001.

[10] J. Buldú, P. Cano, M. Koppenberger, J. Almendral, and S. Boccaletti, "The complex network of musical tastes," *New Journal of Physics*, vol. 9, no. 6, p. 172, 2007.

[11] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. Nuño, "From time series to complex networks: The visibility graph," *Proceedings of the National Academy of Sciences*, vol. 105, no. 13, pp. 4972–4975, 2008.