

Redes y Sistemas Complejos

Monografía del curso: fundamentos, modelos y aplicaciones

Cristian Candia, Ph.D.

Computational Research in Social Science Lab

Instituto de Data Science, Facultad de Ingeniería, Universidad del Desarrollo

Northwestern Institute on Complex Systems (NICO), Northwestern University

Versión docente / material de curso

Semestre 1, 2026

Edición 0.9 (borrador)

26 de febrero de 2026

Contacto: crcandiav@gmail.com

© 2026 Cristian Candia, Ph.D. Todos los derechos reservados. No distribuir sin autorización.

Índice general

1. Redes como representación de sistemas complejos	6
1.1. Introducción	6
1.2. ¿Qué es (y qué no es) una red?	6
1.3. Redes como modelos de sistemas complejos	7
1.4. Escalas de análisis: micro, meso y macro	7
1.5. Tipos de redes y ejemplos	8
1.6. Importancia de la representación en redes	8
1.7. Conexión con el aprendizaje en redes	9
1.8. Redes como objetos epistemológicos	9
1.9. Conclusión	10
2. Estructura básica de las redes	11
2.1. Introducción	11
2.2. Distribución de grados	11
2.3. Matriz de adyacencia y grafos dirigidos	11
2.4. Redes ponderadas	12
2.5. Redes bipartitas	12
2.6. Coeficiente de clustering y estructura local	12
2.7. Caminos, distancias y pequeña-mundo	13
2.8. Centralidades y medidas de importancia	13
2.8.1. Centralidad de grado	13
2.8.2. Centralidad de intermediación	13
2.8.3. Centralidad eigenvectorial	13
2.8.4. Centralidad de cercanía	14
2.9. Laplaciano de grafos y operadores espectrales	14
2.9.1. Definición	14
2.9.2. Interpretación física (difusión, suavizado)	14
2.9.3. Relación con conectividad y cortes	14
2.9.4. Laplaciano normalizado	14
2.10. Conclusiones	14
3. Centralidad, poder y desigualdad en redes complejas	15
3.1. Introducción	15
3.2. Definiciones generales y clasificación de centralidades	15
3.2.1. Centralidad de grado	15
3.2.2. Centralidad de cercanía	16
3.2.3. Centralidad de intermediación	16
3.2.4. Centralidad vector propia y variantes	16
3.2.5. Medidas emergentes	17
3.3. Correlaciones y redundancia entre centralidades	17

3.4.	Centralidad, poder e inequidad	17
3.5.	Conclusiones y perspectivas	18
4.	Modelos generativos en la ciencia de redes	19
4.1.	Introducción	19
4.2.	Redes aleatorias de Erdős–Rényi	19
4.3.	Modelo de Watts–Strogatz y fenómeno de mundo pequeño	20
4.4.	Modelo de Barabási–Albert y attachment preferencial	20
4.5.	Discusión y perspectivas	21
5.	Organización mesoestructural: comunidades, core–periphery, block models	22
5.1.	Introducción	22
5.2.	Detección de comunidades	22
5.3.	Modularidad y su maximización	22
5.4.	Estructura núcleo-periferia	23
5.5.	Roles estructurales	24
5.6.	Discusión y perspectivas	24
6.	Inferencia estadística en redes complejas	26
6.1.	Dependencia estructural y límites del paradigma i.i.d.	27
6.1.1.	Violación del supuesto i.i.d.	27
6.1.2.	Dependencia inducida por enlaces	27
6.1.3.	Consecuencias para estimación y testing	27
6.2.	Modelos nulos y contrastes estructurales	27
6.2.1.	Erdős–Rényi como hipótesis base	27
6.2.2.	Modelo de configuración como control de grado	27
6.2.3.	SBM como hipótesis mesoestructural	27
6.2.4.	Significancia en redes no es igual a p-values clásicos	27
6.3.	Pruebas estadísticas en grafos	27
6.3.1.	Permutation tests sobre aristas	27
6.3.2.	Significancia de modularidad	27
6.3.3.	Testing de centralidad	27
6.3.4.	Bootstrap de métricas de red	27
6.4.	Incertidumbre y sesgos de observación	27
6.4.1.	Redes egocéntricas	27
6.4.2.	Missing links y missing nodes	27
6.4.3.	Sensibilidad estructural	27
6.4.4.	Robustez de inferencias	27
6.5.	Límites inferenciales en ciencia de redes	27
6.5.1.	Descriptivo vs inferencial	27
6.5.2.	Correlación estructural no implica mecanismo	27
6.5.3.	Advertencias metodológicas explícitas	27
7.	Procesos sobre redes	28
7.1.	Difusión y contagio en redes heterogéneas	28
7.1.1.	Intuición: por qué la topología cambia la dinámica	28
7.1.2.	Modelos SIS y SIR en redes: definición microscópica	28
7.1.3.	Campo medio heterogéneo (HMF): umbral en función de $\langle k^2 \rangle$	29
7.1.4.	Formulación espectral: umbral como función del radio espectral	29
7.1.5.	Más allá del caso “ideal”: correlaciones, modularidad y redes temporales	30
7.2.	Influencia de la centralidad en la difusión	30

7.2.1.	Intuición: “importante” depende del proceso	30
7.2.2.	Influencia como problema de intervención: inmunización e influencia máxima	30
7.2.3.	Centralidades “para contagio”: grado vs. coreness (k-core)	31
7.2.4.	Centralidad como modelo de flujo: cuándo sirve betweenness	31
7.3.	Redes con estructura de mundo pequeño	31
7.3.1.	Intuición: atajos reducen tiempos característicos	31
7.3.2.	Watts–Strogatz y variantes analíticas	31
7.3.3.	Redes temporales vs redes agregadas	32
7.3.4.	Caminos temporales	32
7.3.5.	Consecuencias prácticas (orden importa)	32
7.4.	Discusión y perspectivas	32
8.	Causalidad, interferencia y efectos de red	33
8.1.	Causalidad más allá de unidades independientes	33
8.1.1.	Violación de SUTVA	33
8.1.2.	Interferencia directa e indirecta	33
8.1.3.	Spillovers estructurales	33
8.2.	Descomposición de efectos causales en redes	33
8.2.1.	Efectos ego	33
8.2.2.	Efectos de pares	33
8.2.3.	Efectos de segundo orden	33
8.2.4.	Dependencia topológica	33
8.3.	Diseños causales en presencia de redes	33
8.3.1.	Randomización por clúster	33
8.3.2.	Diff-in-diff con redes	33
8.3.3.	Intervenciones parciales	33
8.3.4.	Limitaciones prácticas	33
8.4.	Límites de la inferencia causal basada en grafos	33
8.4.1.	Ambigüedad estructural	33
8.4.2.	Confusión latente	33
8.4.3.	Necesidad de diseño y teoría	33
9.	Redes multilayer y multiplex [Opcional]	34
9.1.	Grafos multilayer: definiciones y ejemplos	34
9.1.1.	Capas	34
9.1.2.	Tipos de enlaces	34
9.1.3.	Casos reales (movilidad, educación, RRSS)	34
9.2.	Supra-matrices y operadores multilayer	34
9.2.1.	Supra-adjacency matrix	34
9.2.2.	Tensores	34
9.2.3.	Generalización de Laplaciano	34
9.3.	Procesos dinámicos en múltiples capas	34
9.3.1.	Centralidad inter-capas	34
9.3.2.	Difusión acoplada	34
9.3.3.	Umbrales epidémicos multilayer	34
9.4.	Desafíos metodológicos y abiertos	34
9.4.1.	Identificabilidad	34
9.4.2.	Escalabilidad	34
9.4.3.	Interpretabilidad	34

10. Aprendizaje en redes	35
10.1. Representaciones y tareas en grafos	35
10.1.1. Intuición: ¿por qué aprender sobre grafos?	35
10.1.2. Fundamentos matemáticos de las representaciones	35
10.2. Embeddings basados en paseos aleatorios	36
10.2.1. Intuición: analogía entre grafos y lenguaje	36
10.2.2. Formulación matemática: DeepWalk y skip-gram	36
10.3. Redes neuronales de grafos y message passing	36
10.3.1. Intuición: operadores locales y acoplamiento a la estructura	36
10.3.2. Formulación general de message passing	37
10.3.3. Consideraciones avanzadas	37
10.4. Tareas supervisadas y no supervisadas	38
10.5. Conclusiones y perspectivas	38
11. Evaluación y validación de modelos en grafos [Completar]	39
11.1. Introducción: evaluación no trivial en redes	39
11.2. Dependencia estructural y leakage	39
11.2.1. Leakage estructural en grafos	39
11.3. Esquemas de partición de datos	39
11.3.1. Splits por nodos (clasificación de nodos)	40
11.3.2. Splits por aristas (predicción de enlaces)	40
11.3.3. Splits por subgrafos y tiempos (redes dinámicas)	40
11.4. Métricas de evaluación por tarea	40
11.4.1. Clasificación de nodos	40
11.4.2. Predicción de enlaces	40
11.5. Riesgos y métricas engañosas	41
11.6. Buenas prácticas para evaluación en grafos	41
11.7. Conclusiones	41
12. Estudios integrados en ciencia de redes [REHACER - SUJETO A FACTIBILIDAD]	42
12.1. Motivación: de la teoría a sistemas reales	42
12.2. Caso I: Redes XDR y propagación de anomalías	42
12.2.1. Construcción de la red	42
12.2.2. Estructura y heterogeneidad	42
12.2.3. Procesos dinámicos	42
12.2.4. Aprendizaje sobre grafos	43
12.3. Caso II: Redes de interacción social inferidas vía WiFi	43
12.3.1. Inferencia de vínculos	43
12.3.2. Mesoestructura y comunidades	43
12.3.3. Procesos sociales	43
12.3.4. Reflexión metodológica	43
12.4. Caso III: Red de carreras y trayectorias educativas	43
12.4.1. Red bipartita y proyecciones	43
12.4.2. Estructura de oportunidades	43
12.4.3. Conexión con causalidad	43
12.5. Caso IV: Redes de compras públicas	43
12.5.1. Estructura bipartita y concentración	43
12.5.2. Patrones anómalos	44
12.6. Síntesis transversal	44

13.Ética y privacidad en redes	45
13.1. Medir redes como acto de poder	45
13.1.1. Intuición: de lo local a lo global	45
13.1.2. De la re-identificación a la manipulación de algoritmos	45
13.2. Privacidad, homofilia y filtrado	45
13.2.1. Homofilia y fuga de información	45
13.2.2. Ataques de inferencia y modelos de filtrado	46
13.3. Buenas prácticas y principios éticos	46
13.3.1. Anonimización y minimización de daños	46
13.3.2. Consentimiento informado y transparencia	46
13.3.3. Equidad algorítmica y mitigación de sesgos	46
13.3.4. Directrices institucionales y deontología	47
13.4. Ética del aprendizaje automático en grafos	47
13.5. Conclusiones	47

Capítulo 1

Redes como representación de sistemas complejos

1.1. Introducción

La ciencia de redes ha emergido en las últimas décadas como un lenguaje común para describir fenómenos tan diversos como redes sociales, sistemas biológicos, infraestructuras tecnológicas y procesos cognitivos. Su éxito radica en que las redes proporcionan una representación general y poderosa de relaciones de distinto tipo: desde interacciones entre personas hasta vínculos químicos o flujos de información. Según Menczer et al. (2020), las redes aparecen en ámbitos que van desde el marketing hasta la medicina y su estudio requiere combinar herramientas de física estadística, informática, sociología y biología. Este módulo tiene por objetivo construir una base conceptual rigurosa para entender qué es una red, qué información captura y por qué es útil representarla de esa forma.

1.2. ¿Qué es (y qué no es) una red?

Una red o *grafo* es una estructura matemática definida como un par $G = (V, E)$, donde V es el conjunto de *nodos* o vértices y E es el conjunto de *enlaces* o aristas. Cada enlace $e \in E$ conecta un par de nodos y puede ser *dirigido* (si distingue una fuente y un destino) o *no dirigido*. El grafo puede representarse mediante su *matriz de adyacencia* $A \in \mathbb{R}^{n \times n}$, donde $n = |V|$ es el número de nodos y

$$A_{ij} = \begin{cases} 1, & \text{si existe una arista de } i \text{ a } j, \\ 0, & \text{en caso contrario,} \end{cases} \quad (1.1)$$

para un grafo no ponderado; si los enlaces llevan pesos w_{ij} , entonces $A_{ij} = w_{ij}$. La *red* se diferencia de un grafo puramente abstracto en que se utiliza para modelar un sistema real: los nodos representan entidades concretas (personas, moléculas, ciudades) y los enlaces representan interacciones, dependencias o flujos entre ellas. Esta distinción es importante, pues el significado del enlace determina qué interpretaciones son válidas.

El *grado* k_i de un nodo i en un grafo no dirigido se define como el número de enlaces incidentes en i y puede calcularse a partir de la matriz de adyacencia como

$$k_i = \sum_{j=1}^n A_{ij} = \sum_{j=1}^n A_{ji}, \quad (1.2)$$

para grafos no dirigidos. En grafos dirigidos se distingue entre *grado de entrada* $k_i^{\text{in}} = \sum_j A_{ji}$ y *grado de salida* $k_i^{\text{out}} = \sum_j A_{ij}$. El vector (k_1, \dots, k_n) captura la heterogeneidad estructural: en

muchas redes reales aparecen nodos con grados muy altos (*hubs*) y la distribución de grados presenta colas pesadas, lo que implica la presencia de nodos dominantes (Rodrigues, 2018; Barabási and Pósfai, 2016). Esta heterogeneidad tiene importantes consecuencias dinámicas, como veremos en secciones posteriores.

1.3. Redes como modelos de sistemas complejos

Los sistemas complejos son aquellos formados por muchos componentes interactuantes cuya dinámica colectiva no puede inferirse fácilmente a partir del comportamiento individual. Ejemplos incluyen ecosistemas, mercados financieros y redes neuronales. Una característica fundamental de estos sistemas es la *no linealidad*: pequeñas perturbaciones pueden provocar efectos macroscópicos. Modelar un sistema complejo como una red permite representar explícitamente sus patrones de interacción y estudiar cómo estos patrones influyen en procesos dinámicos como la difusión, la sincronización o la cooperación.

En una red heterogénea, la distribución de grados $P(k)$ (probabilidad de que un nodo tenga grado k) puede aproximarse como una ley de potencias $P(k) \propto k^{-\gamma}$ para $k \geq k_0$ en muchas redes reales. Esta propiedad, observada por primera vez en la estructura de la Web y conocida como *escala libre*, implica que la varianza del grado puede ser divergente para $\gamma \leq 3$ y conlleva que el *umbral epidémico* (la tasa crítica de transmisión por encima de la cual se produce una epidemia) tiende a cero en el límite de redes grandes (Rodrigues, 2018). Matemáticamente, si λ es la tasa de transmisión y μ la tasa de recuperación, el umbral en un grafo aleatorio con distribución de grado general se puede aproximar como $\lambda_c = \mu / \langle k \rangle$ en el límite de redes homogéneas, pero en redes de escala libre la divergencia de $\langle k^2 \rangle$ hace que $\lambda_c \rightarrow 0$.

Más allá de estas propiedades estadísticas, representar un sistema como red implica una decisión epistemológica: se asume que las relaciones binarias capturan la esencia del sistema. Las notas de Clauset (2022) subrayan que medir una red implica convertir información local en información global: registrar interacciones entre individuos otorga a los investigadores un poder de observación y análisis, pero también genera riesgos de privacidad, ya que atributos sensibles pueden inferirse indirectamente a través de los vecinos (Clauset, 2022). Por tanto, la modelización en términos de redes exige considerar tanto los beneficios analíticos como las implicaciones éticas.

1.4. Escalas de análisis: micro, meso y macro

Representar un sistema como red permite estudiar fenómenos a distintas escalas:

1. **Escala micro.** Se centra en propiedades locales de los nodos y los enlaces. Aquí se analizan métricas como el grado, la fuerza (suma de pesos en redes ponderadas) o la *centralidad de intermediación*, que mide la fracción de caminos mínimos que pasan por un nodo. Para un nodo i , la centralidad de intermediación se define como

$$C_B(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}, \quad (1.3)$$

donde σ_{st} es el número total de caminos geodésicos entre s y t y $\sigma_{st}(i)$ es el número de estos caminos que pasan por i . Esta medida sirve para identificar nodos que actúan como *puentes* entre comunidades distintas.

2. **Escala meso.** Analiza subestructuras emergentes, como *comunidades* (grupos de nodos densamente conectados), *motivos* (patrones recurrentes de pocas aristas) o *núcleo-periferia*. Las comunidades permiten inferir organización funcional y roles sociales; sin embargo,

no toda partición comunitaria es significativa y su interpretación depende del contexto (Fortunato, 2010). En grafos bipartitos (p.ej., autores y artículos) es frecuente proyectar a un grafo unipartito; dicha proyección puede inducir sesgos al sobrerrepresentar nodos muy activos.

3. **Escala macro.** Considera propiedades globales, como la distribución de grados, la fracción de nodos conectados en el *componente gigante*, el diámetro (la longitud máxima de un camino geodésico) o la *longitud de camino medio*. Otra medida global es el *coeficiente de clustering* promedio, definido como

$$C = \frac{1}{n} \sum_{i=1}^n \frac{2T_i}{k_i(k_i - 1)}, \quad (1.4)$$

donde T_i es el número de triángulos que incluyen al nodo i . El coeficiente C mide la tendencia de los amigos de mis amigos a ser amigos entre sí y es elevado en redes sociales reales, a diferencia de grafos aleatorios en los que $C \sim \langle k \rangle / n$.

El análisis a múltiples escalas proporciona una comprensión integral: identifica actores clave, patrones de organización y propiedades emergentes que no son evidentes al examinar los componentes por separado.

1.5. Tipos de redes y ejemplos

La diversidad de aplicaciones de la ciencia de redes implica que las estructuras pueden clasificarse en varios tipos según el dominio:

- **Redes sociales.** Representan interacciones entre personas o grupos, como relaciones de amistad, colaboración académica o comunicación en redes sociales en línea. Los nodos suelen ser individuos y las aristas, interacciones. Estas redes presentan alto clustering y distribuciones de grado heterogéneas.
- **Redes tecnológicas.** Incluyen redes de telecomunicaciones (Internet, telefonía), redes eléctricas, redes de transporte o infraestructuras logísticas. Su análisis permite estudiar eficiencia, robustez frente a fallos y optimización de rutas.
- **Redes biológicas.** Comprenden redes de interacción entre proteínas, redes metabólicas, sinapsis neuronales y redes de ecosistemas. Estas redes sirven para entender procesos biológicos y desarrollar aplicaciones en bioinformática.
- **Redes de información.** Conectan conceptos o documentos mediante relaciones semánticas o de citación: por ejemplo, la red de citas académicas es un grafo dirigido en el que cada artículo cita a otros trabajos. Los nodos con alto grado entrante suelen actuar como hubs de conocimiento.

Cada categoría tiene sus propias particularidades (tipo de enlaces, peso, direccionalidad), pero comparten la idea de que el comportamiento del sistema emerge de la estructura de las interacciones.

1.6. Importancia de la representación en redes

Construir una red a partir de datos no es un paso trivial ni neutro. Decidir qué constituye un nodo, qué constituye un enlace y qué atributos se asignan puede condicionar los resultados. Como

señala Hamilton (2020), las redes introducen un *sesgo inductivo relacional*: los modelos que aprenden sobre datos con estructura de grafo explotan la premisa de que entidades conectadas tienen propiedades relacionadas. Esta idea es fundamental en el aprendizaje automático sobre grafos, donde técnicas como *embeddings* de nodos y *redes neuronales de grafos* (GNN) aprenden representaciones vectoriales de los nodos aprovechando la topología (Hamilton, 2020; Sánchez-Lengeling et al., 2021). En particular, un modelo de mensajes en una GNN actualiza el vector de estado de un nodo i mediante una agregación de los mensajes de sus vecinos $j \in \mathcal{N}(i)$:

$$\mathbf{h}_i^{(\ell+1)} = \sigma \left(\mathbf{W}_0 \mathbf{h}_i^{(\ell)} + \sum_{j \in \mathcal{N}(i)} \mathbf{W}_1 \mathbf{h}_j^{(\ell)} \right), \quad (1.5)$$

donde $\mathbf{h}_i^{(\ell)}$ es la representación del nodo i en la capa ℓ , $\mathcal{N}(i)$ es el conjunto de vecinos de i , \mathbf{W}_0 y \mathbf{W}_1 son matrices de pesos compartidas y σ es una función de activación no lineal. Esta fórmula encarna el sesgo relacional: la representación de cada nodo se actualiza combinando su estado actual con la información de sus vecinos. Para que esta operación tenga sentido, es crucial que la red esté bien construida; un enlace mal definido puede inducir el modelo a aprender correlaciones espurias.

Además de su valor analítico, la representación en redes plantea retos éticos. El acto de registrar interacciones transforma información privada en pública: como recuerdan las notas de Clauzet (2022), los atributos personales pueden filtrarse a través de los vecinos por homofilia. Por ejemplo, si las personas con cierta ideología tienden a conectarse entre sí, el simple hecho de observar sus conexiones puede revelar sus preferencias políticas aunque no se hayan declarado explícitamente. De ahí la importancia de la anonimización y del consentimiento informado en estudios de redes humanas.

1.7. Conexión con el aprendizaje en redes

Aunque este módulo se centra en los fundamentos conceptuales de las redes, es importante esbozar la conexión con el aprendizaje automático. Las técnicas modernas de *graph representation learning* buscan mapear los nodos a vectores en espacios de dimensión reducida de manera que preserven la estructura. Los métodos basados en paseos aleatorios, como *DeepWalk* o *node2vec*, generan secuencias de nodos mediante procesos de Markov sobre el grafo y entrenan modelos de tipo word2vec para obtener representaciones; las GNN generalizan las convoluciones a grafos mediante agregaciones de vecinos como en la ecuación (1.5). Estas herramientas permiten realizar tareas de clasificación de nodos, predicción de enlaces y detección de comunidades (Hamilton, 2020; Sánchez-Lengeling et al., 2021).

Comprender la estructura y representación de las redes es, por tanto, el primer paso para aplicar con éxito estas técnicas. Un modelo de aprendizaje sólo puede ser tan bueno como la red que se le suministre; una red mal construida o mal interpretada puede conducir a inferencias erróneas o a sesgos injustificados.

1.8. Redes como objetos epistemológicos

Las redes no son entidades observables directamente, sino construcciones analíticas que resultan de decisiones de modelado. Definir qué constituye un nodo, una arista, su dirección o su peso implica adoptar supuestos teóricos y normativos que condicionan los resultados posteriores. En este sentido, una red no es un reflejo neutral de la realidad social, biológica o tecnológica, sino una representación parcial que privilegia ciertos tipos de interacción y excluye otros.

Desde una perspectiva epistemológica, el análisis de redes plantea al menos tres desafíos centrales. Primero, el problema de la *reificación*: una vez construida, la red tiende a tratarse como un objeto “real”, ocultando las decisiones que la originaron. Segundo, el problema de la *dependencia estructural*: las observaciones no son independientes, lo que cuestiona directamente muchos supuestos estadísticos clásicos. Tercero, el problema de la *interpretación causal*: la estructura observada puede ser tanto causa como consecuencia de los procesos que se desean explicar.

Esta monografía adopta una postura explícita: las redes son herramientas conceptuales potentes, pero su análisis requiere una reflexión constante sobre qué se mide, qué se infiere y qué se asume. Los capítulos posteriores incorporan esta perspectiva al discutir inferencia, dinámica, aprendizaje automático y ética, enfatizando que el rigor en ciencia de redes no es sólo técnico, sino también conceptual.

1.9. Conclusión

En este primer módulo hemos establecido una base rigurosa para el estudio de redes. Definimos formalmente qué es una red, discutimos la importancia de distinguirla del grafo abstracto y presentamos conceptos matemáticos esenciales como la matriz de adyacencia, el grado y el coeficiente de clustering. Vimos que las redes son modelos adecuados para sistemas complejos porque capturan la heterogeneidad y la interdependencia de sus elementos, y que su análisis debe considerar distintas escalas (micro, meso y macro) para revelar patrones de organización y actores clave. También exploramos la variedad de tipos de redes y subrayamos la importancia, tanto analítica como ética, de la representación en redes. Finalmente, apuntamos hacia el aprendizaje en redes, señalando que la construcción correcta de la red es un prerequisite para técnicas modernas como las GNN. Con estos fundamentos, estamos preparados para abordar los módulos siguientes, donde estudiaremos modelos generativos, procesos dinámicos y técnicas de aprendizaje sobre grafos.

Capítulo 2

Estructura básica de las redes

2.1. Introducción

Tras haber presentado la noción general de red como modelo de sistemas complejos en el primer módulo, el siguiente paso es comprender su estructura interna. La estructura de una red determina cómo se propaga la información, cómo responden los sistemas a perturbaciones y qué tan robustos son frente a fallos o ataques. En este módulo abordamos las herramientas básicas para describir y cuantificar la estructura: distribución de grados, matrices de adyacencia, redes ponderadas, grafos bipartitos y coeficientes de clustering. También introducimos las principales medidas de centralidad, que permiten identificar la importancia relativa de los nodos y los enlaces.

2.2. Distribución de grados

Sea $G = (V, E)$ un grafo no dirigido con $n = |V|$ nodos. El *grado* de un nodo i se define como en la ecuación 1.2, es decir, $k_i = \sum_{j=1}^n A_{ij}$, donde A es la matriz de adyacencia. La *distribución de grados* $P(k)$ es la probabilidad de que un nodo elegido al azar tenga grado k . En términos discretos, si N_k es el número de nodos con grado k , entonces $P(k) = N_k/n$, con $\sum_k P(k) = 1$ (Barabási and Pósfai, 2016). Esta distribución constituye un resumen de la heterogeneidad del grafo: redes homogéneas tienen distribuciones concentradas alrededor de un valor medio, mientras que redes heterogéneas exhiben colas pesadas, con hubs de alto grado. La forma de $P(k)$ influye en la conectividad global, la robustez y la propagación de procesos en la red (Barabási and Pósfai, 2016).

La distribución de grados es relevante, por ejemplo, para comprender por qué algunas redes reales son robustas a fallos aleatorios pero vulnerables a ataques dirigidos. En redes de *escala libre* con $P(k) \propto k^{-\gamma}$ y $\gamma \in (2, 3)$, la presencia de hubs hace que la eliminación aleatoria de nodos raramente afecte a la conectividad, pero un ataque selectivo a los nodos de mayor grado puede fragmentar la red.

2.3. Matriz de adyacencia y grafos dirigidos

La descripción completa de un grafo puede codificarse en su *matriz de adyacencia* $A \in \mathbb{R}^{n \times n}$. Para un grafo no dirigido y no ponderado se define

$$A_{ij} = \begin{cases} 1, & \text{si existe una arista entre } i \text{ y } j, \\ 0, & \text{en caso contrario,} \end{cases} \quad (2.1)$$

de modo que A es simétrica. En grafos dirigidos se utiliza una convención similar, pero el hecho de que la arista vaya de i a j se refleja en la entrada A_{ij} ; por tanto A puede no ser

simétrica. La matriz de adyacencia permite expresar muchas propiedades locales y globales de la red de forma algebraica. Por ejemplo, el grado de salida de un nodo i en un grafo dirigido es $k_i^{\text{out}} = \sum_j A_{ij}$ y el grado de entrada es $k_i^{\text{in}} = \sum_j A_{ji}$ (Barabási and Pósfai, 2016).

El uso de A facilita el cómputo de caminos de longitud l : la entrada $(A^l)_{ij}$ cuenta el número de caminos de longitud l que conectan i con j . Además, muchas medidas de centralidad se definen a partir de funciones de A , como la centralidad eigenvectorial (véase la Sección 2.8).

2.4. Redes ponderadas

En numerosas aplicaciones los enlaces llevan asociada una intensidad o peso w_{ij} que cuantifica la fortaleza de la relación (p.ej., número de correos entre dos personas, capacidad de una línea de transmisión). Un *grafo ponderado* se representa mediante una matriz W cuyas entradas W_{ij} son los pesos. Para redes ponderadas, el grado generalizado de un nodo puede definirse como la *fuerza* $s_i = \sum_j W_{ij}$, que mide la suma de los pesos de las aristas incidentes. La decisión de ignorar o no los pesos tiene consecuencias analíticas: simplificar una red ponderada a un grafo no ponderado pierde información sobre la intensidad de las interacciones. Barabási and Pósfai (2016) advierten que muchas redes reales son inherentemente ponderadas y que el análisis puramente topológico puede ser insuficiente (Barabási and Pósfai, 2016).

2.5. Redes bipartitas

Un *grafo bipartito* es una red cuyos nodos pueden dividirse en dos conjuntos disjuntos U y V de modo que cada arista conecta un nodo de U con uno de V , y no existen aristas entre nodos del mismo conjunto. Ejemplos típicos son las redes de colaboración autor-artículo y las redes de recomendación, donde U representa usuarios y V objetos. Una característica importante es que los grafos bipartitos permiten construir *proyecciones* sobre cada conjunto: por ejemplo, dos autores están conectados en la proyección si han coescrito al menos un artículo. Sin embargo, estas proyecciones pueden introducir sesgos; nodos con alta actividad en el bipartito generan conexiones spuriamente altas en la proyección. Barabási and Pósfai (2016) presentan varios ejemplos de redes bipartitas, como el grafo de actores de Hollywood y la red de enfermedades humanas y genes (Barabási and Pósfai, 2016).

2.6. Coeficiente de clustering y estructura local

El *coeficiente de clustering* mide la propensión de los vecinos de un nodo a estar conectados entre sí, es decir, la densidad de triángulos alrededor de un nodo. Para un nodo i de grado $k_i \geq 2$, el coeficiente de clustering local se define como

$$C_i = \frac{2T_i}{k_i(k_i - 1)}, \quad (2.2)$$

donde T_i es el número de triángulos (tripletas completas) que incluyen a i (Barabási and Pósfai, 2016). El factor 2 contabiliza cada triángulo una sola vez. El coeficiente de clustering medio de la red es $\langle C \rangle = \frac{1}{n} \sum_{i=1}^n C_i$. En grafos aleatorios clásicos, $\langle C \rangle$ tiende a cero cuando $n \rightarrow \infty$, mientras que en redes sociales reales suele ser alto, reflejando la regla informal de “los amigos de mis amigos tienden a ser amigos”. La coexistencia de altos coeficientes de clustering con pequeñas distancias promedio se denomina *fenómeno de mundo pequeño*. El modelo de Watts–Strogatz muestra cómo una pequeña probabilidad de re-cableado β en una red regular reduce drásticamente la longitud de camino medio mientras mantiene el clustering alto (Watts and Strogatz, 1998a).

2.7. Caminos, distancias y pequeña-mundo

La *distancia* entre dos nodos i y j es la longitud del camino más corto (número de aristas) que los conecta; se denota $d(i, j)$. El *diámetro* de una red es la distancia máxima entre pares de nodos conectados y la *longitud de camino media* es $\langle d \rangle = \frac{1}{n(n-1)} \sum_{i \neq j} d(i, j)$. En muchas redes reales se observa que $\langle d \rangle$ crece lentamente con el tamaño del grafo (de forma logarítmica o sublogarítmica), lo que se denomina la propiedad de mundo pequeño. La coexistencia de un alto clustering con distancias cortas fue formalizada por Watts y Strogatz en su modelo, en el que se parte de un anillo regular y se re-cablean aristas con probabilidad β (Watts and Strogatz, 1998a). Para β pequeño, el clustering se mantiene cercano al del anillo regular pero la longitud de camino media disminuye rápidamente, explicando la paradoja de las “seis personas de distancia”.

2.8. Centralidades y medidas de importancia

En el análisis de redes es crucial identificar nodos o aristas importantes desde el punto de vista estructural o funcional. Las *medidas de centralidad* cuantifican diferentes nociones de importancia. Presentamos brevemente las más usadas y su interpretación.

2.8.1. Centralidad de grado

La forma más simple de centralidad es el grado k_i ya definido: nodos con alto grado tienen muchas conexiones y, en contextos sociales, pueden interpretarse como individuos populares. Su principal limitación es que no considera la calidad ni la posición de las conexiones. Rodrigues (2018) destacan que aunque el grado es informativo, no existe una única medida de importancia válida para todas las redes (Rodrigues, 2018).

2.8.2. Centralidad de intermediación

La *centralidad de intermediación* $C_B(i)$ se define como:

$$C_B(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}, \quad (2.3)$$

Un nodo con alta intermediación actúa como puente en muchos caminos geodésicos, por lo que su eliminación puede desconectar la red o retrasar la difusión. El cálculo exacto de $C_B(i)$ tiene un costo computacional alto para grafos grandes, ya que para cada nodo requiere analizar todos los caminos de la red, lo que motiva variantes basadas en muestreo o paseos aleatorios (Rodrigues, 2018).

2.8.3. Centralidad eigenvectorial

La *centralidad eigenvectorial* asigna un valor alto a los nodos que están conectados a otros nodos importantes. Se define como el vector propio principal de la matriz de adyacencia: sea \mathbf{v} un vector no nulo y λ un escalar tales que

$$A \mathbf{v} = \lambda \mathbf{v}, \quad (2.4)$$

con λ el mayor valor propio de A . La componente v_i de \mathbf{v} da la centralidad eigenvectorial del nodo i . Intuitivamente, un nodo es importante si se conecta con otros nodos importantes. Esta medida está relacionada con el número de caminos de todas las longitudes que comienzan en un nodo (Rodrigues, 2018).

2.8.4. Centralidad de cercanía

Otra medida clásica es la *centralidad de cercanía*, definida como el inverso de la suma de distancias desde i a todos los demás nodos:

$$C_C(i) = \frac{n-1}{\sum_{j \neq i} d(i, j)}. \quad (2.5)$$

Esta centralidad identifica nodos que están, en promedio, a poca distancia del resto de la red. Es útil para encontrar actores que pueden difundir información rápidamente o recolectar recursos de manera eficiente.

Ninguna de estas centralidades es universalmente la mejor. Su relevancia depende del contexto y del proceso de interés; a menudo conviene considerar varias en conjunto y analizar su correlación con observables empíricos (Rodrigues, 2018).

2.9. Laplaciano de grafos y operadores espectrales

2.9.1. Definición

2.9.2. Interpretación física (difusión, suavizado)

2.9.3. Relación con conectividad y cortes

2.9.4. Laplaciano normalizado

2.10. Conclusiones

En este módulo se han introducido las herramientas básicas para describir la estructura de una red. Hemos definido la distribución de grados y discutido su impacto en la heterogeneidad y la robustez de las redes. La matriz de adyacencia proporciona una representación algebraica compacta que permite computar grados y caminos, y su extensión a grafos ponderados resalta la importancia de los pesos en aplicaciones reales. Los grafos bipartitos constituyen un caso especial que exige cautela al proyectarlos. El coeficiente de clustering y la distancia media capturan propiedades locales y globales cuya combinación da lugar al fenómeno de mundo pequeño. Finalmente, las medidas de centralidad—grado, intermediación, eigenvectorial y cercanía—permiten evaluar la importancia estructural de nodos y aristas, aunque ninguna es definitiva para todas las redes. Estas nociones serán fundamentales para los módulos posteriores dedicados a modelos generativos y dinámicas, así como para el aprendizaje sobre grafos.

Capítulo 3

Centralidad, poder y desigualdad en redes complejas

3.1. Introducción

La teoría de redes proporciona un lenguaje unificado para modelar interacciones en sistemas complejos. Una de las ideas clave es que no todos los nodos contribuyen por igual a la dinámica global: existen nodos *centrales* que concentran conexiones o flujos y por ello ejercen influencia estructural. La centralidad está estrechamente relacionada con conceptos de poder e inequidad: en redes sociales o corporativas, los *hubs* acumulan acceso a recursos e información, mientras que en redes económicas la estructura por sí sola puede generar desigualdad salarial comparable a la de países enteros (DeCanio and Watkins, 2025). El objetivo de este módulo es presentar rigurosamente las medidas de centralidad más utilizadas, discutir su interpretación y limitaciones, y conectar estos conceptos con fenómenos de poder e inequidad.

3.2. Definiciones generales y clasificación de centralidades

Sea $G = (V, E)$ un grafo no dirigido y simple con $|V| = N$ nodos. Su estructura se describe mediante la matriz de adyacencia $A \in \{0, 1\}^{N \times N}$, donde $A_{ij} = 1$ si existe una arista entre i y j y $A_{ij} = 0$ en caso contrario. Una función de centralidad es un operador $c : V \rightarrow \mathbb{R}$ que asigna a cada nodo un valor real indicando su importancia relativa. De forma general, las centralidades se clasifican en *locales* (dependen sólo del vecindario inmediato) y *globales* (toman en cuenta caminos o información de toda la red) (Bendahman and Lotfi, 2024; Rodrigues, 2018). A continuación se presentan las medidas clásicas.

3.2.1. Centralidad de grado

La medida más sencilla es el *grado* k_i de un nodo i , definido como el número de vecinos que tiene. Utilizando la matriz de adyacencia, el grado se calcula como la suma de la fila correspondiente (Rodrigues, 2018)

$$k_i = \sum_{j=1}^N A_{ij}. \quad (3.1)$$

El grado cuantifica conectividad local y en redes *scale-free* presenta una distribución pesada: pocos nodos (*hubs*) tienen grados muy grandes mientras la mayoría posee pocos enlaces. Esta heterogeneidad facilita procesos como la propagación epidémica o la sincronización (Rodrigues, 2018), haciendo que los hubs se conviertan en nodos de poder estructural.

3.2.2. Centralidad de cercanía

La centralidad de cercanía (c_v) evalúa hasta qué punto un nodo está cerca del resto en términos de distancias geodésicas. Para un nodo v , la distancia d_{uv} es la longitud del camino más corto entre u y v . La cercanía se define como el inverso de la distancia promedio a todos los demás nodos (Rodrigues, 2018):

$$\frac{1}{c_v} = \frac{1}{N-1} \sum_{u \neq v} d_{uv}, \quad (3.2)$$

donde un mayor valor de c_v indica que el nodo está, en promedio, más próximo al resto. La cercanía es ampliamente utilizada para evaluar accesibilidad en redes de transporte, biológicas o sociales, y se ha correlacionado con indicadores socioeconómicos y regulatorios (Evans and Chen, 2022). Sin embargo, recientes trabajos han mostrado que la cercanía está fuertemente correlacionada con el grado: la *farness* (recíproca de la cercanía) es lineal en el logaritmo del grado (Evans and Chen, 2022). Esto implica que, salvo que se elimine explícitamente la dependencia con el grado, la cercanía añade poca información nueva y su cómputo (costoso) puede ser redundante (Evans and Chen, 2022).

3.2.3. Centralidad de intermediación

La centralidad de intermediación (*betweenness*) caracteriza nodos que actúan como puentes entre pares de nodos. Para cada par de nodos s, t , sea σ_{st} el número de caminos más cortos entre s y t y $\sigma_{st}(i)$ los que pasan por i . La intermediación de i se define como (Rodrigues, 2018; Wasserman and Faust, 1994)

$$B_i = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}. \quad (3.3)$$

Esta medida refleja la capacidad del nodo para controlar el flujo de información y ha sido usada para identificar brokers o cuellos de botella. Sus desventajas son el alto coste computacional, $\mathcal{O}(N^3)$ en su forma exacta (Rodrigues, 2018), y el hecho de que sólo considera caminos más cortos, lo cual puede no reflejar rutas reales en redes donde la información viaja por caminos alternativos (Rodrigues, 2018).

3.2.4. Centralidad vector propia y variantes

Mientras que el grado sólo cuenta vecinos, la centralidad *vector propia* incorpora la importancia de los vecinos: un nodo es central si se conecta a otros nodos centrales. Esta definición conduce al problema de autovector para la matriz de adyacencia (Rodrigues, 2018):

$$\lambda x = xA, \quad (3.4)$$

donde x es el vector de centralidades. La solución correspondiente al mayor autovalor λ (teorema de Perron–Frobenius) produce valores positivos de centralidad. El método de potencia muestra que x coincide con el número de caminos de longitud grande que llegan a cada nodo (Rodrigues, 2018). Esta centralidad tiende a concentrarse en hubs y puede sufrir fenómenos de *localización* donde la mayor parte de los pesos se acumulan en unos pocos nodos (Rodrigues, 2018).

Una variante es la centralidad de PageRank, utilizada por el motor de búsqueda Google. Definida sobre grafos dirigidos, asigna a cada nodo una puntuación c satisfaciendo (Rodrigues, 2018)

$$c = \alpha A^T D^{-1} c + \beta \mathbf{1}, \quad (3.5)$$

donde A^T es la matriz de adyacencias transpuesta, D la matriz diagonal con los grados de salida, $0 < \alpha < 1$ un factor de amortiguamiento y β controla el reinicio aleatorio. Este método combina caminos de longitud arbitraria con un salto aleatorio que evita que la importancia se concentre sólo en hubs.

El *rank* de Katz es otra generalización que suma caminos de todas las longitudes con un parámetro de decaimiento μ , resolviendo $(I - \mu A^T)k = \mathbf{1}$. Estudios recientes muestran que la distribución de la centralidad de Katz en grafos aleatorios se decompone en picos asociados a nodos de distinto grado (Bartolucci et al., 2024), lo que evidencia su fuerte dependencia con el grado.

3.2.5. Medidas emergentes

La revisión de centralidades ha impulsado medidas que combinan información local y global. Por ejemplo, *centrality degree paths* (CDP) integra la conectividad inmediata con la longitud de los caminos que conectan el nodo con el resto. Este método evalúa no sólo los vecinos directos sino también las trayectorias indirectas, capturando mejor la influencia estructural en redes sociales (Bendahman and Lotfi, 2024). Los autores de la propuesta muestran que las medidas clásicas pueden ser inestables en redes grandes y que los hubs no siempre coinciden con los nodos más influyentes; CDP ofrece un equilibrio entre coste computacional y fidelidad estructural (Bendahman and Lotfi, 2024).

3.3. Correlaciones y redundancia entre centralidades

Aunque cada medida pretende capturar un aspecto diferente, diversos estudios han mostrado que las centralidades están fuertemente correlacionadas. Evans y Chen demostraron que la inversa de la centralidad de cercanía (*farness*) se relaciona linealmente con el logaritmo del grado en una amplia gama de redes aleatorias y reales (Evans and Chen, 2022). El mismo trabajo sugiere que medir la cercanía no aporta información adicional a menos que se elimine su dependencia con el grado (Evans and Chen, 2022). Análisis más amplios muestran que muchas parejas de centralidades (grado, intermediación, PageRank, eigenvector) presentan coeficientes de correlación de Pearson elevados (Evans and Chen, 2022). Por ello, la elección de la medida debe basarse en la interpretación sustantiva y en su relación con el fenómeno de interés.

Además, el estudio de Bartolucci et al. analizó la distribución de la centralidad de Katz en redes aleatorias mediante el método de cavidad. Encontraron que la distribución presenta picos asociados a clases de nodos con distinto grado y que, al aumentar la conectividad media, estos picos se fusionan (Bartolucci et al., 2024). Este resultado refuerza la idea de que muchas centralidades son reflejo de la heterogeneidad de grados y que la estructura de la red condiciona la desigualdad en los valores de centralidad.

3.4. Centralidad, poder e inequidad

La relación entre centralidad y poder se manifiesta en múltiples contextos: en redes sociales, los individuos con mayor grado o intermediación controlan el flujo de información y actúan como líderes; en redes económicas, las empresas situadas en posiciones centrales pueden movilizar recursos y acceder a oportunidades privilegiadas. Esta concentración de centralidad se traduce en desigualdad: las distribuciones de centralidad suelen seguir leyes de potencia, reflejando que una minoría de nodos acapara la mayor parte de la influencia.

DeCanio y Watkins proponen un modelo organizacional donde la estructura de la red por sí sola genera desigualdad salarial (DeCanio and Watkins (2025)). Consideran que la compensación de cada empleado puede basarse en su coste de procesamiento (vertex cost) o en su centralidad

(betweenness o PageRank). Su estudio muestra que el índice de Gini asociado a la compensación basada únicamente en la posición en la red oscila entre 0.3 y 0.5, mientras que la compensación basada en el coste del nodo supera 0.6 (DeCanio and Watkins, 2025). Estos valores son comparables a los índices de desigualdad de países reales, demostrando que la desigualdad puede surgir endógenamente de la arquitectura de la red. Además, la definición de intermediación y PageRank utilizados en su modelo se formaliza mediante las ecuaciones (3) y (4) anteriores (DeCanio and Watkins, 2025).

En términos sociológicos, las medidas de centralidad se relacionan con capital social, prestigio y poder simbólico. Wasserman y Faust discuten cómo el grado se asocia con popularidad, la intermediación con control de recursos y la cercanía con acceso rápido a información. Sin embargo, las conexiones no son lineales y existen críticas al uso acrítico de estas métricas para justificar jerarquías. Por ejemplo, la dependencia de la centralidad de eigenvector respecto a hubs puede invisibilizar a nodos periféricos con roles clave en cohesión.

3.5. Conclusiones y perspectivas

El estudio de la centralidad en redes complejas permite comprender cómo la posición estructural de los nodos influye en procesos dinámicos y en la distribución desigual de poder y recursos. Hemos revisado definiciones formales y discutido su interpretación. Las medidas clásicas como el grado, la cercanía, la intermediación, la centralidad de eigenvector y PageRank capturan distintos aspectos de conectividad, pero muestran correlaciones significativas y limitaciones. La heterogeneidad inherente a las redes —especialmente las de tipo *scale-free*— conduce a distribuciones de centralidad altamente sesgadas y facilita la aparición de desigualdad y poder concentrado. Estudios recientes demuestran que la cercanía y el grado están no linealmente relacionados, que la distribución de centralidades puede descomponerse en clases de grado y que la estructura organizacional puede producir desigualdad salarial comparable a la observada en economías nacionales. Nuevas métricas, como la centralidad degree paths, intentan capturar influencias combinando caminos directos e indirectos y buscan superar la inestabilidad de las medidas clásicas.

Una comprensión crítica de las centralidades requiere, por tanto, no sólo conocimiento técnico de sus definiciones y algoritmos, sino también sensibilidad a sus implicancias sociales y económicas. En módulos posteriores se profundizará en la inferencia estadística, la dinámica sobre redes y el aprendizaje automático sobre grafos, herramientas que permitirán extender el análisis de centralidad hacia la explicación causal y la predicción.

Capítulo 4

Modelos generativos en la ciencia de redes

4.1. Introducción

El análisis estructural de redes reales revela patrones que no son propios de los grafos aleatorios clásicos. Para comprender el origen de estas regularidades y explicar por qué ciertas propiedades emergen con frecuencia en diferentes dominios (sociales, biológicos o tecnológicos), la ciencia de redes recurre a *modelos generativos*. Estos modelos formalizan mecanismos de construcción de grafos y permiten reproducir distribuciones de grado, coeficientes de clustering o longitudes de camino medio observados en datos empíricos.

En este módulo presentamos tres paradigmas de generación: las *redes aleatorias* de Erdős–Rényi, el *modelo de re-cableado* de Watts–Strogatz y el *crecimiento con attachment preferencial* (modelo de Barabási–Albert). Cada uno captura diferentes características: la aleatoriedad pura explica la aparición de distancias cortas, el re-cableado introduce clustering sin sacrificar la pequeña longitud media, y el attachment preferencial genera heterogeneidad extrema en forma de distribuciones de grado en ley de potencia Barabási and Pósfai (2016).

4.2. Redes aleatorias de Erdős–Rényi

Sea $G(n, p)$ un grafo construido sobre un conjunto de n nodos en el que cada una de las $\frac{n(n-1)}{2}$ posibles aristas se incluye de forma independiente con probabilidad p . Alternativamente, en el modelo $G(n, M)$ se escogen uniformemente al azar M aristas entre todas las posibles combinaciones Barabási and Pósfai (2016).

En el régimen $G(n, p)$, la probabilidad de observar un grafo concreto con m aristas es

$$p^m(1-p)^{\binom{n}{2}-m}.$$

La *distribución de grados* de un nodo cualquiera es binomial,

$$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}, \quad (4.1)$$

la cual converge a una distribución de Poisson con media $\langle k \rangle = p(n-1)$ cuando n es grande y p pequeño.

Intuición estructural

En un grafo aleatorio, cada nodo “ve” esencialmente el mismo entorno estadístico. No existen nodos privilegiados ni jerarquías estructurales: la homogeneidad es una consecuencia directa de la independencia entre aristas.

Un argumento de ramificación explica la aparición del fenómeno de *mundo pequeño*. Si cada nodo tiene en promedio z vecinos, el número esperado de nodos alcanzables a distancia d crece como z^d . Para cubrir toda la red se requiere $z^D \sim n$, lo que implica

$$D \sim \frac{\log n}{\log z},$$

es decir, distancias que crecen logarítmicamente con el tamaño del sistema Newman (2000a).

Sin embargo, los grafos de Erdős–Rényi carecen de estructura local significativa. El coeficiente de clustering satisface

$$C \approx p = \frac{\langle k \rangle}{n-1},$$

lo que implica que $C \rightarrow 0$ cuando $n \rightarrow \infty$. Este resultado está en fuerte contraste con redes sociales reales, donde el clustering permanece alto incluso en redes grandes Barabási and Pósfai (2016).

4.3. Modelo de Watts–Strogatz y fenómeno de mundo pequeño

El modelo de *Watts–Strogatz* fue propuesto para reconciliar dos propiedades empíricas aparentemente incompatibles: *altos coeficientes de clustering y pequeñas distancias promedio* Watts and Strogatz (1998a).

La construcción parte de un anillo regular con n nodos, donde cada nodo se conecta a sus z vecinos más cercanos. Sobre esta red ordenada se realiza un proceso de *re-cableado*: cada arista se reconecta, con probabilidad p , a un nodo elegido al azar.

Intuición matemática del re-cableado

El efecto crucial del re-cableado es la introducción de *atajos*. Un número muy pequeño de enlaces largos es suficiente para reducir drásticamente la distancia media, ya que estos enlaces conectan regiones del grafo previamente lejanas.

Watts y Strogatz mostraron que para valores pequeños de p ,

- el clustering permanece cercano al de la red regular,
- la longitud de camino media cae rápidamente hacia valores del orden $\log n$.

Esta separación de escalas explica por qué sistemas sociales reales pueden ser simultáneamente altamente clusterizados y globalmente accesibles Watts and Strogatz (1998a); Newman (2000a).

4.4. Modelo de Barabási–Albert y attachment preferencial

Muchas redes reales exhiben distribuciones de grado con colas pesadas. El modelo de *Barabási–Albert* explica este fenómeno mediante dos mecanismos fundamentales: *crecimiento y attachment preferencial* Barabási and Albert (1999).

Partiendo de un pequeño grafo inicial, en cada paso se añade un nuevo nodo con m enlaces que se conectan a nodos existentes con probabilidad proporcional a su grado:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \quad (4.2)$$

Derivación de la ley de potencia

El análisis continuo del crecimiento del grado de un nodo muestra que

$$\frac{dk_i}{dt} \propto \frac{k_i}{t},$$

lo que conduce a una solución

$$k_i(t) \sim \left(\frac{t}{t_i}\right)^{1/2},$$

donde t_i es el tiempo de incorporación del nodo.

De esta dinámica se obtiene una distribución estacionaria

$$P(k) = \frac{2m^2}{k^3}, \quad (4.3)$$

correspondiente a un exponente universal $\gamma = 3$ Barabási and Albert (1999); Barabási and Pósfai (2016).

Interpretación

El attachment preferencial formaliza el principio de “los ricos se hacen más ricos”. Los nodos antiguos acumulan ventaja estructural simplemente por haber estado más tiempo en el sistema, sin necesidad de atributos intrínsecos.

Extensiones posteriores introducen fitness, aging y attachment no lineal, mostrando que la heterogeneidad observada en redes reales puede surgir por múltiples mecanismos Barabási and Pósfai (2016).

4.5. Discusión y perspectivas

Los modelos generativos proporcionan un puente entre mecanismos locales simples y patrones globales complejos. Cada modelo captura un aspecto distinto de las redes reales, pero ninguno es completo por sí solo.

Las redes reales combinan crecimiento, clustering, heterogeneidad, modularidad y dinámica temporal. El desafío contemporáneo consiste en integrar estos mecanismos en modelos más realistas, manteniendo al mismo tiempo interpretabilidad analítica y poder explicativo.

Estos modelos constituyen la base conceptual para los módulos posteriores dedicados a dinámicas, inferencia estadística y aprendizaje automático sobre grafos.

Capítulo 5

Organización mesoestructural: comunidades, core–periphery, block models

5.1. Introducción

El análisis de redes no sólo se interesa en propiedades de nodos individuales (micro) o en características globales del grafo (macro). En un nivel intermedio, denominado *mesoestructural*, aparecen patrones de organización que reflejan la interacción de conjuntos de nodos. Dos ejemplos paradigmáticos de estas estructuras son las *comunidades*, es decir, grupos de nodos con muchas conexiones internas y pocas conexiones externas, y las *estructuras núcleo–periferia*, donde existe un conjunto central densamente conectado y un conjunto periférico con pocas conexiones entre sí. Otro concepto relacionado es el de *roles estructurales*, que agrupa nodos según patrones de conectividad similares en lugar de densidades de enlace. Comprender y cuantificar estas mesoestructuras es fundamental para explicar fenómenos emergentes en redes sociales, biológicas, tecnológicas y económicas.

5.2. Detección de comunidades

En una red no dirigida $G = (V, E)$ con $|V| = N$ nodos y $|E| = m$ aristas, una comunidad o *módulo* es un subconjunto $C \subseteq V$ de nodos que posee una alta densidad de aristas internas en comparación con las aristas que lo conectan con el resto de la red. La detección de comunidades es crucial para descomponer el grafo en unidades funcionales que suelen corresponder a grupos de actores con intereses comunes, módulos funcionales en redes biológicas o tópicos en redes de co-ocurrencia. Fortunato (2010) muestran que la identificación de comunidades es esencial para interpretar la organización de sistemas reales y proporcionan una revisión exhaustiva de métodos basados en la densidad, el flujo y la dinámica de procesos sobre redes.

Existen numerosos algoritmos para la detección de comunidades: desde enfoques divisivos basados en la eliminación de aristas con alta intermediación (algoritmo de Girvan–Newman) hasta métodos jerárquicos y heurísticos que maximizan funciones de calidad. A continuación se presenta la función de calidad más empleada, la *modularidad*.

5.3. Modularidad y su maximización

La *modularidad* Q es una métrica que cuantifica la diferencia entre la fracción de aristas que caen dentro de las comunidades y la fracción esperada bajo un modelo nulo que conserva

el grado de cada nodo. Consideremos una partición de V en comunidades $\{C_1, \dots, C_k\}$ e introduzcamos la matriz de adyacencia $A \in \{0, 1\}^{N \times N}$ y los grados $k_i = \sum_j A_{ij}$. El modelo nulo más habitual es el modelo de *configuración*, en el que el número de aristas entre i y j se modela como $\frac{k_i k_j}{2m}$. La modularidad se define como (Fortunato, 2010)

$$Q = \frac{1}{2m} \sum_{i,j} (A_{ij} - \frac{k_i k_j}{2m}) \delta(C_i, C_j), \quad (5.1)$$

donde $\delta(C_i, C_j)$ es igual a 1 si los nodos i y j pertenecen a la misma comunidad y 0 en caso contrario. Este valor se puede reescribir en términos de propiedades de cada comunidad c como

$$Q = \sum_c \left[\frac{l_c}{m} - \left(\frac{d_c}{2m} \right)^2 \right], \quad (5.2)$$

donde $l_c = \sum_{i,j \in C_c} A_{ij}/2$ es el número de aristas internas de la comunidad c y $d_c = \sum_{i \in C_c} k_i$ es la suma de los grados de sus nodos. La modularidad valora positivamente las comunidades que contienen más aristas de las esperadas al azar y penaliza aquellas que sólo reflejan la distribución de grados global.

Interpretación y limitaciones. La modularidad fue popularizada por Newman (2006) como criterio para detectar comunidades maximizando Q . Según el planteamiento de Viles and O'Malley (2023), la modularidad “maximiza la diferencia entre la fracción de aristas dentro de las comunidades y la fracción esperada si las aristas se asignaran aleatoriamente conservando la distribución de grados”. Su simplicidad ha motivado algoritmos eficientes como el método de Louvain y sus variantes. Sin embargo, la modularidad presenta la *resolución límite*: puede no detectar comunidades muy pequeñas en redes grandes, ya que la contribución de una comunidad al valor total Q disminuye al crecer el tamaño de la red. Además, existen particiones degeneradas con valores similares de modularidad, lo que dificulta la interpretación única de los resultados.

5.4. Estructura núcleo-periferia

Otra mesoestructura relevante es la *estructura núcleo-periferia* (*core-periphery*). A diferencia de las comunidades, que buscan conjuntos internamente densos y mutuamente poco conectados, la estructura núcleo-periferia divide los nodos en dos clases: el *núcleo* y la *periferia*. Según Yanchenko and Sengupta (2023), en una red con estructura núcleo-periferia “la asignación de los nodos consta de dos grupos, un núcleo y una periferia. Los nodos del núcleo están densamente conectados entre sí y además se conectan frecuentemente con los nodos periféricos, mientras que los nodos periféricos tienen pocas conexiones entre ellos”. Otra forma de caracterizar el núcleo es que sus nodos están a corta distancia de todos los demás nodos.

Esta organización surge, por ejemplo, en redes de comercio mundial: los países con economías grandes (núcleo) comercian entre sí y con economías pequeñas, mientras que las economías pequeñas (periferia) comercian poco entre ellas. Un patrón similar se observa en redes de aeropuertos, donde los grandes aeropuertos tienen conexiones con otros hubs y con aeropuertos regionales, mientras que los aeropuertos regionales tienen pocos vuelos entre sí. La presencia de una estructura núcleo-periferia influye en la difusión de información y en la resiliencia del sistema: los nodos centrales suelen ser más influyentes o críticos para la conectividad global (Yanchenko and Sengupta, 2023).

Modelos y detección

Existen diversos enfoques para detectar estructuras núcleo-periferia. Algunos métodos combinan estimación estadística con modelos de bloques, en los que se asume que la matriz de

adyacencia puede aproximarse mediante bloques densos (núcleo) y bloques dispersos (periferia). Otros métodos emplean técnicas espectrales basadas en autovectores de la matriz de adyacencia o del Laplaciano. La detección de núcleo–periferia comparte similitudes conceptuales con el problema de *rich-club*, donde se estudia la sobreconectividad entre los nodos de grado elevado.

5.5. Roles estructurales

Las comunidades y las estructuras núcleo–periferia describen cómo los nodos se agrupan según la densidad de sus conexiones. Por el contrario, el concepto de *rol* agrupa nodos según su patrón de conexiones, independientemente del número total de aristas que comparten entre sí. Siguiendo a Rossi and Ahmed (2015), el descubrimiento de roles fue definido inicialmente como “cualquier proceso que divide los nodos de un grafo en clases de nodos *estructuralmente equivalentes*”. Dos nodos son estructuralmente equivalentes si mantienen las mismas conexiones con el resto de la red. En la práctica, esta definición estricta se relaja buscando patrones de conectividad similares.

Equivalencias y tipos de roles. Los roles pueden formalizarse mediante tres nociones de equivalencia (Rossi and Ahmed, 2015):

1. *Equivalencia estructural*: dos nodos i y j son estructuralmente equivalentes si $A_{i*} = A_{j*}$, es decir, tienen exactamente los mismos vecinos. Esta equivalencia es muy estricta y rara en redes grandes.
2. *Equivalencia regular*: dos nodos son regularmente equivalentes si sus vecinos pertenecen a las mismas clases de equivalencia. Permite agrupar, por ejemplo, a actores que ejercen roles de “intermediario” aunque conecten con diferentes individuos.
3. *Equivalencia estocástica*: dos nodos son estocásticamente equivalentes si las probabilidades de conectarse a otros nodos dependen únicamente de sus clases, lo cual es la base del modelo de bloques estocásticos.

Los roles capturan patrones como nodos estrella (centros de estrella), nodos periféricos, nodos puente o nodos casi clique, entre otros. A diferencia de las comunidades, los roles no necesitan conformar subconjuntos densos, sino que reflejan funciones o posiciones relativas en la red. Métodos recientes proponen descubrir roles a partir de representaciones de atributos o de incrustaciones (embeddings) aprendidas sobre la red (Rossi and Ahmed, 2015).

5.6. Discusión y perspectivas

La organización mesoestructural de una red tiene implicancias profundas en su dinámica y en la interpretación de sus datos. Las comunidades permiten segmentar la red en grupos cohesionados, la estructura núcleo–periferia revela jerarquías de conectividad e influencia, y los roles estructurales identifican patrones funcionales que no dependen de la densidad de enlaces. La modularidad ha sido un criterio popular para detectar comunidades, pero presenta limitaciones como la resolución y la existencia de máximos locales; se han desarrollado alternativas y extensiones para mitigar estos problemas. Las estructuras núcleo–periferia son complementarias a las comunidades y requieren métodos específicos basados en modelos de bloques o técnicas espectrales. Por su parte, la detección de roles abre un campo distinto enfocado en la equivalencia y la similitud estructural, con aplicaciones en identificación de anomalías y análisis funcional.

En redes reales es común encontrar la coexistencia de varias mesoestructuras: comunidades que exhiben subestructuras núcleo–periferia o nodos que pertenecen a un mismo rol dentro

de diferentes comunidades. El estudio de estas interacciones sigue siendo un área activa de investigación. Además, incorporar información exógena y considerar restricciones demográficas, funcionales o de equidad en la identificación de comunidades —como propone Viles and O'Malley (2023)— es fundamental para análisis aplicados más robustos y responsables.

Capítulo 6

Inferencia estadística en redes complejas

6.1. Dependencia estructural y límites del paradigma i.i.d.

6.1.1. Violación del supuesto i.i.d.

6.1.2. Dependencia inducida por enlaces

6.1.3. Consecuencias para estimación y testing

6.2. Modelos nulos y contrastes estructurales

6.2.1. Erdős–Rényi como hipótesis base

6.2.2. Modelo de configuración como control de grado

6.2.3. SBM como hipótesis mesoestructural

6.2.4. Significancia en redes no es igual a p-values clásicos

6.3. Pruebas estadísticas en grafos

6.3.1. Permutation tests sobre aristas

6.3.2. Significancia de modularidad

6.3.3. Testing de centralidad

6.3.4. Bootstrap de métricas de red

6.4. Incertidumbre y sesgos de observación

6.4.1. Redes egocéntricas

6.4.2. Missing links y missing nodes

6.4.3. Sensibilidad estructural

6.4.4. Robustez de inferencias

6.5. Límites inferenciales en ciencia de redes

6.5.1. Descriptivo vs inferencial

6.5.2. Correlación estructural no implica mecanismo

6.5.3. Advertencias metodológicas explícitas

Capítulo 7

Procesos sobre redes

7.1. Difusión y contagio en redes heterogéneas

7.1.1. Intuición: por qué la topología cambia la dinámica

En modelos de contagio (biológico, informacional o conductual), la red determina *quién* interactúa con *quién* y, por tanto, cuántas oportunidades de transmisión existen. En una población bien mezclada (mean-field clásico), todos los individuos “ven” al mismo entorno promedio; en una red, en cambio, los individuos con alto grado (*hubs*) tienen muchos más contactos y actúan como amplificadores. Esta asimetría topológica genera dos consecuencias clave:

- **Umbrales y persistencia:** la transición entre extinción y endemidad puede depender de momentos altos de la distribución de grados, de modo que redes muy heterogéneas pueden reducir drásticamente (o anular, asintóticamente) el umbral epidémico.
- **Heterogeneidad de riesgo e impacto:** la probabilidad de infectarse y de infectar a otros es altamente desigual; una fracción pequeña de nodos puede explicar gran parte de la propagación.

Para formalizar estas ideas, introducimos modelos SIS y SIR sobre grafos, y discutimos aproximaciones analíticas (campo medio heterogéneo y formulaciones espectrales) que conectan dinámica con estructura (Pastor-Satorras et al., 2015; Barrat et al., 2008; Newman, 2010).

7.1.2. Modelos SIS y SIR en redes: definición microscópica

Sea $G = (V, E)$ un grafo no dirigido, simple, con $|V| = N$ nodos y matriz de adyacencia A . Denotemos por $X_i(t) \in \{0, 1\}$ el estado del nodo i en tiempo t (0 susceptible, 1 infectado).

SIS (Susceptible–Infectado–Susceptible). En tiempo continuo, un nodo susceptible i se infecta con tasa

$$\beta \sum_j A_{ij} X_j(t), \quad (7.1)$$

y un nodo infectado se recupera con tasa μ (volviendo a susceptible). Este proceso es Markoviano y posee un estado absorbente libre de infección ($X_i \equiv 0$) y, para ciertos parámetros, un régimen endémico estacionario (Pastor-Satorras et al., 2015; Van Mieghem et al., 2009).

SIR (Susceptible–Infectado–Recuperado). Aquí, la recuperación es absorbente: $I \rightarrow R$ con tasa μ y R no vuelve a infectarse. SIR describe brotes finitos y su tamaño final, y se conecta elegantemente con percolación y métodos de message passing (Newman, 2002; Karrer and Newman, 2010; Pastor-Satorras et al., 2015).

7.1.3. Campo medio heterogéneo (HMF): umbral en función de $\langle k^2 \rangle$

En redes grandes, una aproximación estándar para SIS es el *campo medio heterogéneo* (HMF), que agrupa nodos por grado k y asume equivalencia estadística dentro de cada clase de grado (Pastor-Satorras et al., 2015; Barrat et al., 2008).

Sea $\rho_k(t)$ la fracción de nodos de grado k infectados en t . En redes *no correlacionadas* (sin assortatividad por grado), la probabilidad de que un vecino de un nodo cualquiera esté infectado se aproxima por

$$\Theta(t) = \sum_k \frac{kP(k)}{\langle k \rangle} \rho_k(t), \quad (7.2)$$

donde $\frac{kP(k)}{\langle k \rangle}$ es la distribución de grado “vista desde una arista”. El balance de tasas SIS en HMF es entonces

$$\frac{d\rho_k}{dt} = -\mu \rho_k + \beta k (1 - \rho_k) \Theta. \quad (7.3)$$

Linealización y umbral. Cerca del estado libre de infección, $\rho_k \ll 1$ y $(1 - \rho_k) \approx 1$, de modo que

$$\frac{d\rho_k}{dt} \approx -\mu \rho_k + \beta k \Theta. \quad (7.4)$$

Buscando soluciones estacionarias no triviales, se obtiene que existe un umbral en el cociente β/μ dado por

$$\left(\frac{\beta}{\mu}\right)_c^{\text{HMF}} = \frac{\langle k \rangle}{\langle k^2 \rangle}. \quad (7.5)$$

Equivalentemente, un “número reproductivo efectivo” a nivel de red puede definirse como

$$R_* = \frac{\beta}{\mu} \frac{\langle k^2 \rangle}{\langle k \rangle}, \quad (7.6)$$

y la condición $R_* > 1$ marca endemividad en HMF (Pastor-Satorras et al., 2015).

Caso scale-free y desaparición del umbral (HMF). Si $P(k) \propto k^{-\gamma}$ con $2 < \gamma \leq 3$, entonces $\langle k^2 \rangle$ diverge con N (bajo supuestos estándar de corte superior), lo cual lleva a $(\beta/\mu)_c^{\text{HMF}} \rightarrow 0$ cuando $N \rightarrow \infty$ (Pastor-Satorras et al., 2015; Barrat et al., 2008). Intuitivamente: la presencia de hubs crea “reservorios” que sostienen la infección incluso para transmisiones muy débiles.

7.1.4. Formulación espectral: umbral como función del radio espectral

HMF enfatiza momentos de $P(k)$, pero ignora detalles de conectividad. Una visión complementaria proviene de aproximaciones tipo *quenched mean-field* / NIMFA, que trabajan directamente con la matriz de adyacencia (Pastor-Satorras et al., 2015; Van Mieghem et al., 2009).

Sea $p_i(t) = \mathbb{P}[X_i(t) = 1]$. Una ecuación de balance (con cierre de independencia) es

$$\frac{dp_i}{dt} = -\mu p_i + \beta (1 - p_i) \sum_j A_{ij} p_j. \quad (7.7)$$

Linealizando alrededor de $p_i \equiv 0$ se obtiene

$$\frac{d\mathbf{p}}{dt} \approx (\beta \mathbf{A} - \mu \mathbf{I}) \mathbf{p}. \quad (7.8)$$

La estabilidad del estado libre de infección se determina por el mayor autovalor Λ_1 de A :

$$\beta_c^{\text{spec}} = \frac{\mu}{\Lambda_1}. \quad (7.9)$$

Este resultado hace explícito que la propagación está controlada por el *modo dominante* del grafo (autovector principal), típicamente localizado en hubs y su vecindario en redes muy heterogéneas (Pastor-Satorras et al., 2015; Van Mieghem et al., 2009).

Comentario conceptual. HMF y el umbral espectral coinciden en ciertos regímenes, pero pueden diferir cuando existen correlaciones, localización espectral o estructura modular fuerte. En práctica, (7.9) suele ser una buena guía operacional, pues computar Λ_1 es factible incluso para redes grandes usando métodos iterativos (Pastor-Satorras et al., 2015; Van Mieghem et al., 2009).

7.1.5. Más allá del caso “ideal”: correlaciones, modularidad y redes temporales

En redes reales, tres extensiones son particularmente relevantes:

Correlaciones de grado (assortatividad). Si nodos de alto grado se conectan preferentemente con alto grado (assortative mixing), se modifica el umbral y la prevalencia; en general, el análisis requiere matrices de mezcla por grado o enfoques espectrales más finos (Newman, 2010; Pastor-Satorras et al., 2015).

Modularidad y cuellos de botella. En redes comunitarias, la propagación puede ser rápida dentro de comunidades pero lenta entre ellas; esto induce *dos escalas* de crecimiento y sugiere que ciertos nodos “puente” controlan la transmisión inter-comunidad.

Redes temporales y de contacto. Cuando las aristas aparecen/desaparecen en el tiempo, el orden temporal de contactos importa (camino temporal). El umbral y la velocidad de difusión pueden cambiar drásticamente respecto a la red agregada.

Estas extensiones motivan modelos y métodos más avanzados (percolación dependiente del tiempo, aproximaciones de pares, message passing dinámico), pero los principios base siguen siendo: *heterogeneidad + estructura* reconfiguran umbrales, velocidades y tamaños finales (Pastor-Satorras et al., 2015; Barrat et al., 2008).

7.2. Influencia de la centralidad en la difusión

7.2.1. Intuición: “importante” depende del proceso

Un nodo puede ser “importante” por múltiples razones: tener muchos contactos (grado), conectar comunidades (intermediación), estar cerca de todos (cercanía), o estar conectado a otros importantes (eigenvector/PageRank). Pero el mejor predictor de influencia depende del mecanismo de difusión (contagio simple vs complejo, presencia de modularidad, redundancia de caminos, etc.). Por eso, no existe una centralidad universal (Rodrigues, 2019; Borgatti, 2005).

7.2.2. Influencia como problema de intervención: inmunización e influencia máxima

En epidemiología de redes, una pregunta operativa es: *¿qué nodos conviene inmunizar (o bloquear) para minimizar la propagación?* En difusión de información: *¿qué nodos conviene activar como “semillas” para maximizar alcance?*

Inmunización dirigida. En redes heterogéneas, inmunizar al azar es ineficiente: la enfermedad “sobrevive” en hubs. Estrategias dirigidas (vacunar hubs) reducen fuertemente la prevalencia y pueden restaurar un umbral efectivo (Pastor-Satorras and Vespignani, 2002; Cohen et al., 2003).

Maximización de influencia. En modelos de difusión monotónica (p. ej., independent cascade o linear threshold), seleccionar un conjunto de semillas S para maximizar el alcance esperado es un problema combinatorio. Un resultado clásico es que para funciones de alcance submodulares, un algoritmo voraz logra una aproximación $(1 - 1/e)$ (Kempe et al., 2015).

7.2.3. Centralidades “para contagio”: grado vs. coreness (k-core)

Un hallazgo importante es que el grado no siempre identifica a los mejores propagadores. Kitsak et al. muestran que, en diversos grafos, nodos ubicados en el *núcleo* (alto *k-shell index* o *coreness*) pueden ser más influyentes que hubs periféricos: la razón es que el núcleo está densamente interconectado y sostiene cascadas de forma robusta (Kitsak et al., 2010).

Formalmente, la descomposición *k-core* elimina iterativamente nodos con grado menor que k ; el mayor k para el cual un nodo permanece define su *coreness*. En difusión, el núcleo suele coincidir con regiones donde existen múltiples caminos redundantes, reduciendo cuellos de botella y aumentando el tamaño del brote.

7.2.4. Centralidad como modelo de flujo: cuándo sirve betweenness

Si el proceso efectivamente viaja por caminos geodésicos (o cuasi-geodésicos), la intermediación puede capturar “control” de flujo. Sin embargo, si la difusión se parece a un paseo aleatorio o contagio con múltiples rutas, betweenness puede sobrevalorar puentes que no son realmente críticos. Borgatti formaliza esta idea distinguiendo procesos por *transferencia* (walk-like) versus *serial* (path-like), lo cual guía qué centralidad es conceptualmente coherente (Borgatti, 2005).

7.3. Redes con estructura de mundo pequeño

7.3.1. Intuición: atajos reducen tiempos característicos

La estructura small-world enfatiza un trade-off topológico: redes regulares tienen alto clustering pero caminos largos; redes aleatorias tienen caminos cortos pero clustering bajo. El re-cableado de Watts–Strogatz introduce pocos “atajos” que colapsan las distancias manteniendo clustering alto (Watts and Strogatz, 1998b; Newman, 2000b).

Desde la perspectiva de procesos, reducir distancias reduce *tiempos de mezcla* y acelera difusión: si el diámetro efectivo pasa de orden N a orden $\log N$, entonces el número de “saltos” necesarios para alcanzar una fracción macroscópica de la red se reduce dramáticamente.

7.3.2. Watts–Strogatz y variantes analíticas

El modelo WS parte de un anillo regular donde cada nodo conecta con z vecinos, y re-cablea aristas con probabilidad p (Watts and Strogatz, 1998b). Newman discute variantes (p. ej. Newman–Watts) donde se *añaden* enlaces aleatorios, lo que facilita cálculos analíticos de longitudes promedio y muestra la escala característica de atajos (Newman, 2000b). El mensaje clave para difusión es: incluso $p \ll 1$ puede mover la red al régimen de distancias tipo aleatorio.

7.3.3. Redes temporales vs redes agregadas

7.3.4. Caminos temporales

7.3.5. Consecuencias prácticas (orden importa)

7.4. Discusión y perspectivas

Los procesos sobre redes exhiben una dependencia estructural fuerte: la topología no sólo “modula” la dinámica, sino que puede cambiar cualitativamente sus transiciones (umbral epidémico), su velocidad (small-world) y quiénes son los actores dominantes (hubs vs núcleo). Para SIS, HMF vincula el umbral a $\langle k^2 \rangle / \langle k \rangle$, mientras que el enfoque espectral lo conecta al radio espectral Λ_1 , ofreciendo herramientas complementarias para análisis y control (Pastor-Satorras et al., 2015; Van Mieghem et al., 2009). Para intervención, la literatura muestra que la selección de nodos críticos debe alinearse con el mecanismo de difusión: estrategias de inmunización dirigida y medidas como coreness pueden superar heurísticas basadas sólo en grado (Kitsak et al., 2010; Pastor-Satorras and Vespignani, 2002; Cohen et al., 2003).

A nivel más avanzado, tres direcciones son centrales en investigación actual: (i) redes temporales y multilayer; (ii) contagios complejos (que requieren refuerzo social); y (iii) inferencia causal e identificación de efectos de red bajo interferencia. Sin embargo, el núcleo conceptual permanece: *estructura \Rightarrow oportunidades de transmisión \Rightarrow dinámica emergente*.

Capítulo 8

Causalidad, interferencia y efectos de red

8.1. Causalidad más allá de unidades independientes

8.1.1. Violación de SUTVA

8.1.2. Interferencia directa e indirecta

8.1.3. Spillovers estructurales

8.2. Descomposición de efectos causales en redes

8.2.1. Efectos ego

8.2.2. Efectos de pares

8.2.3. Efectos de segundo orden

8.2.4. Dependencia topológica

8.3. Diseños causales en presencia de redes

8.3.1. Randomización por clúster

8.3.2. Diff-in-diff con redes

8.3.3. Intervenciones parciales

8.3.4. Limitaciones prácticas

8.4. Límites de la inferencia causal basada en grafos

8.4.1. Ambigüedad estructural

8.4.2. Confusión latente

8.4.3. Necesidad de diseño y teoría

Capítulo 9

Redes multilayer y multiplex [Opcional]

9.1. Grafos multilayer: definiciones y ejemplos

9.1.1. Capas

9.1.2. Tipos de enlaces

9.1.3. Casos reales (movilidad, educación, RRSS)

9.2. Supra-matrices y operadores multilayer

9.2.1. Supra-adjacency matrix

9.2.2. Tensores

9.2.3. Generalización de Laplaciano

9.3. Procesos dinámicos en múltiples capas

9.3.1. Centralidad inter-capas

9.3.2. Difusión acoplada

9.3.3. Umbrales epidémicos multilayer

9.4. Desafíos metodológicos y abiertos

9.4.1. Identificabilidad

9.4.2. Escalabilidad

9.4.3. Interpretabilidad

Capítulo 10

Aprendizaje en redes

10.1. Representaciones y tareas en grafos

10.1.1. Intuición: ¿por qué aprender sobre grafos?

En visión por computador o procesamiento del lenguaje natural, los datos suelen estar organizados en estructuras regulares (imágenes, secuencias) donde cada elemento es independiente e idénticamente distribuido. En las redes, los datos vienen con *dependencias explícitas*: la presencia de un vínculo entre dos nodos implica que sus etiquetas y atributos pueden estar correlacionados (*homofilia*), o que desempeñan roles similares en la estructura (*equivalencia estructural*) (Hamilton, 2020). Estas dependencias rompen las suposiciones del aprendizaje supervisado convencional, de modo que los algoritmos deben incorporar *sesgos inductivos relacionales* para aprovechar la información topológica.

Los principales problemas a abordar incluyen:

- **Clasificación de nodos:** predecir la etiqueta de un nodo (por ejemplo, el tipo de usuario en una red social) dado un pequeño subconjunto de nodos etiquetados.
- **Predicción de enlaces:** estimar la probabilidad de que exista (o se forme) una arista entre dos nodos, útil en sistemas de recomendación y detección de anomalías.
- **Detección de comunidades:** identificar grupos de nodos densamente interconectados o con roles similares; estas tareas pueden resolverse mediante técnicas de clustering sobre representaciones aprendidas.
- **Clasificación de grafos:** cuando los objetos de interés son grafos completos (p. ej. moléculas), se busca asignar una etiqueta al grafo entero, lo que requiere mecanismos de *pooling* o *lectura* global.

10.1.2. Fundamentos matemáticos de las representaciones

Sea $G = (V, E)$ un grafo, donde V es el conjunto de nodos, E el conjunto de aristas y $X \in \mathbb{R}^{|V| \times d_0}$ la matriz de atributos iniciales. El objetivo del aprendizaje en grafos es obtener una función de embedding $f : V \rightarrow \mathbb{R}^d$ (o $f : V \times E \rightarrow \mathbb{R}^d$) que produzca vectores $h_i = f(v_i)$ que capturen relaciones de proximidad, roles y semántica. Estas representaciones se aprenden *end-to-end* optimizando objetivos supervisados o auto-supervisados. A diferencia de métodos basados en atributos estadísticos (grado, PageRank), los embeddings pueden codificar patrones no lineales y dependencia de atributos (Hamilton, 2020).

10.2. Embeddings basados en paseos aleatorios

10.2.1. Intuición: analogía entre grafos y lenguaje

Los métodos de incrustación inspirados en modelos de lenguaje, como *DeepWalk* (Perozzi et al., 2014), tratan a los nodos como “palabras” y a los paseos aleatorios como “oraciones”. La idea es que dos nodos que aparecen frecuentemente en contextos similares deberían tener vectores incrustados similares. Esta idea captura la homofilia local: un nodo y sus vecinos cercanos (por ejemplo, a distancia 2 o 3) tendrán contextos similares, por lo que sus embeddings serán similares.

10.2.2. Formulación matemática: DeepWalk y skip-gram

DeepWalk genera una colección de paseos $w_{1:t}^{(r)} = (v_{i_1}, v_{i_2}, \dots, v_{i_t})$ mediante un proceso de Markov sobre el grafo:

$$\Pr[v_{i_{k+1}} = u \mid v_{i_k} = v] = \begin{cases} \frac{1}{\deg(v)}, & \text{si } (v, u) \in E, \\ 0, & \text{en otro caso.} \end{cases} \quad (10.1)$$

Para cada paseo, se define un contexto de tamaño C alrededor de cada paso v_{i_k} y se maximiza la probabilidad de observar los nodos del contexto dado el nodo actual mediante un modelo *skip-gram*:

$$\mathcal{L} = \sum_{w_{1:t}^{(r)}} \sum_{k=1}^t \sum_{l=-C}^C \log \Pr(v_{i_{k+l}} \mid v_{i_k}), \quad (10.2)$$

donde $\Pr(v_j \mid v_i) \propto \exp(\mathbf{z}_i^\top \mathbf{z}_j)$. En la práctica se utiliza *negative sampling* o aproximaciones jerárquicas para eficiencia. El aprendizaje ajusta los embeddings \mathbf{z}_i para maximizar la co-ocurrencia de nodos cercanos en los paseos (Perozzi et al., 2014).

Generalizaciones: node2vec y variantes. *node2vec* (Grover and Leskovec, 2016) introduce dos parámetros (p, q) que controlan la tendencia de los paseos a regresar al nodo anterior (sesgo de profundidad) o explorar nuevos vecindarios (sesgo de amplitud). Ajustando p, q , se pueden aprender embeddings que reflejen homofilia (trayectorias que permanecen cerca del nodo) o equivalencia estructural (trayectorias que saltan a nodos con rol similar). Otras extensiones, como LINE (Tang et al., 2015), HOPE, o métodos heterogéneos, abordan grafos dirigidos, con atributos o múltiples tipos de nodos.

10.3. Redes neuronales de grafos y message passing

10.3.1. Intuición: operadores locales y acoplamiento a la estructura

A partir de 2017, los *Graph Neural Networks* (GNN) se consolidaron como un paradigma general para el aprendizaje en grafos (Kipf and Welling, 2017; Hamilton et al., 2017). Su idea central es definir una operación *local*, análoga a una convolución, que combina información del nodo y su vecindario inmediato. Al aplicar repetidas veces la operación, la representación de cada nodo incorpora información de nodos a distancias crecientes. El proceso es *permutación-invariante*: el orden de los vecinos no importa y la red aprende a “resumir” el vecindario.

10.3.2. Formulación general de message passing

Considere un nodo v_i con embedding $h_i^{(l)}$ en la capa l . El esquema de message passing define (Hamilton et al., 2017; Veličković et al., 2018):

$$m_i^{(l+1)} = \text{AGG}\left(\{\phi(h_i^{(l)}, h_j^{(l)}, e_{ij}) : j \in \mathcal{N}(i)\}\right), \quad (10.3)$$

$$h_i^{(l+1)} = \psi(h_i^{(l)}, m_i^{(l+1)}), \quad (10.4)$$

donde $\mathcal{N}(i)$ es el vecindario, e_{ij} representa atributos de la arista, ϕ es una función de mensaje (por ejemplo, una red neuronal), AGG es un operador agregador (suma, media, máximo, atenciones, etc.) y ψ es la función de actualización (normalmente una MLP con activación no lineal). La clave es que AGG sea *conmutativo* y *asociativo*, para garantizar invariancia al orden.

Ejemplo: Graph Convolutional Networks (GCN). Kipf y Welling (Kipf and Welling, 2017) propusieron una versión simplificada del message passing en el dominio espectral, que se puede escribir en forma matricial:

$$H^{(l+1)} = \sigma(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(l)} W^{(l)}), \quad (10.5)$$

donde $\tilde{A} = A + I$ agrega lazos propios, \tilde{D} es la matriz diagonal de grados de \tilde{A} , $H^{(l)}$ es la matriz de embeddings en la capa l y $W^{(l)}$ es una matriz de pesos. Esta capa lineal combina los atributos de un nodo con la media (normalizada) de sus vecinos.

Ejemplo: GraphSAGE. Hamilton et al. (Hamilton et al., 2017) generalizaron el concepto con agregadores paramétricos (media, LSTM, max pooling). Para cada nodo i se toma un subconjunto de vecinos $\mathcal{N}_S(i)$ y se define:

$$h_i^{(l+1)} = \sigma(W^{(l)} h_i^{(l)} \parallel \text{AGG}(\{h_j^{(l)} : j \in \mathcal{N}_S(i)\})), \quad (10.6)$$

donde \parallel denota concatenación. GraphSAGE opera de forma *inductiva*: puede generalizar a nodos no vistos durante el entrenamiento, pues el agregador no depende del tamaño de G .

Ejemplo: Graph Attention Networks (GAT). Veličković et al. (Veličković et al., 2018) introdujeron atención a nivel de aristas. Para cada vecino j se calcula un coeficiente de atención

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^\top [W h_i^{(l)} \parallel W h_j^{(l)}]))}{\sum_{k \in \mathcal{N}(i)} \exp(\text{LeakyReLU}(a^\top [W h_i^{(l)} \parallel W h_k^{(l)}]))}, \quad (10.7)$$

y luego

$$h_i^{(l+1)} = \sigma\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} W h_j^{(l)}\right). \quad (10.8)$$

La atención permite que la red aprenda a ponderar la importancia de cada vecino, lo que puede ser ventajoso cuando el grafo contiene ruidos o aristas irrelevantes.

10.3.3. Consideraciones avanzadas

La literatura reciente aborda retos como:

- **Heterofilia:** cuando los nodos tienden a conectarse con otros de diferente etiqueta, lo que rompe la suposición de homofilia. Redes como GAT o métodos basados en difusión ajustada pueden mejorar en estos casos (Bo et al., 2021).

- **Gráficos dinámicos o temporales:** se requiere incorporar el tiempo en la estructura y en los embeddings, mediante memorias o procesos recurrentes (Rossi et al., 2021).
- **Grafos heterogéneos y multiparadigma:** donde hay múltiples tipos de nodos y aristas; se emplean meta-aristas y redes atencionales específicas (Shi et al., 2020).

10.4. Tareas supervisadas y no supervisadas

Clasificación de nodos. En problemas de clasificación semisupervisada, se observa un conjunto de nodos etiquetados $V_L \subset V$ y se busca inferir etiquetas para $V \setminus V_L$. Dado un embedding h_i , un clasificador softmax produce una distribución $p(y | h_i)$ y la pérdida típica combina entropía cruzada sobre V_L con regularizadores de suavizado que penalizan diferencias de representaciones entre nodos conectados (Kipf and Welling, 2017).

Predicción de enlaces. En la tarea de link prediction se entrena un modelo que, a partir de embeddings h_i y h_j , devuelve una puntuación $\sigma(h_i, h_j)$ (por ejemplo, producto escalar o MLP). La pérdida se define sobre pares de aristas observadas y no observadas; esto se utiliza en redes de recomendación y detección de anomalías (Grover and Leskovec, 2016; Kipf and Welling, 2016).

Detección de comunidades. Aunque la detección de comunidades se aborda tradicionalmente mediante criterios de modularidad o blockmodels, los embeddings permiten aproximar una partición latente: se aplica clustering (por ejemplo, k -means) sobre h_i o se optimiza un autoencoder de grafos para maximizar la modularidad (Mucha et al., 2010). Los modelos variacionales (VGAE) implementan funciones de decodificación de aristas y generan agrupaciones implícitas (Kipf and Welling, 2016).

Clasificación de grafos. Cuando las instancias son grafos completos (químicos, redes funcionales), se requiere una operación de *pooling* para resumir nodos en un vector fijo. Métodos como Set2Set, SortPool, DiffPool y Graph U-Nets aprenden a combinar representaciones nodales en un descriptor global (Ying et al., 2018; Yoo et al., 2020). La pérdida se calcula sobre etiquetas a nivel de grafo.

10.5. Conclusiones y perspectivas

El aprendizaje en redes constituye un puente entre la topología de grafos y modelos estadísticos modernos. Métodos basados en paseos (DeepWalk, node2vec) proporcionan representaciones de bajo coste computacional que capturan homofilia y roles. Las redes neuronales de grafos incorporan mecanismos de *message passing* que pueden adaptarse a diversas tareas y estructuras, permitiendo aprendizaje supervisado, no supervisado y auto-supervisado. Retos actuales incluyen la explicación de las predicciones (interpretabilidad), la escalabilidad a redes de miles de millones de aristas, el manejo de heterofilia, redes temporales y heterogéneas, y la integración con marcos de causalidad. Un curso de doctorado debe capacitar al estudiante para comprender y construir arquitecturas propias, elegir agregadores y criterios de pérdida adecuados, y evaluar críticamente las hipótesis sobre la estructura de los datos.

Capítulo 11

Evaluación y validación de modelos en grafos [Completar]

11.1. Introducción: evaluación no trivial en redes

La evaluación de modelos de aprendizaje automático aplicados a datos con estructura de red no se puede asumir que siga las mismas reglas que en escenarios i.i.d. clásicos. En tareas como predicción de enlaces o clasificación de nodos, la relación entre nodos genera dependencias significativas que, si no se controlan, conducen a estimaciones de desempeño infladas o engañosas. Además, las topologías heterogéneas de los grafos (grados no uniformes, comunidades, hubs) complican la generalización fuera de los conjuntos de entrenamiento.

Particularmente en predicción de enlaces, se ha documentado que prácticas comunes de evaluación —como uso de muestras negativas fáciles o particiones de datos inconsistentes— pueden sesgar los resultados y obstaculizar comparaciones justas entre métodos Li et al. (2023).

El objetivo de este capítulo es formular principios metodológicos sólidos para la evaluación de modelos sobre grafos, destacando riesgos, estrategias de partición de datos y métricas adecuadas para cada tarea.

11.2. Dependencia estructural y leakage

11.2.1. Leakage estructural en grafos

En aprendizaje convencional, se asume que las observaciones de entrenamiento y prueba son independientes. Sin embargo, en redes las etiquetas y las representaciones de nodos interrelacionados pueden filtrarse entre particiones si la partición no controla explícitamente la estructura de conexiones. Este fenómeno se conoce como *leakage estructural*, donde información derivada de la estructura (por ejemplo, conexiones con nodos de prueba) inadvertidamente influye en el entrenamiento del modelo, inflando métricas de desempeño aparente Zhu et al. (2023).

Por ejemplo, incluir en la representación de entrenamiento aristas que serán objetivo de predicción sin excluir apropiadamente conexiones incidentes puede generar correlaciones espurias que los modelos explotan, reduciendo la validez de la evaluación Zhu et al. (2023).

11.3. Esquemas de partición de datos

Diseñar esquemas de partición adecuados es crucial porque determina qué tipo de generalización se está evaluando.

11.3.1. Splits por nodos (clasificación de nodos)

La partición en nodos separa el conjunto de nodos en entrenamiento y prueba, pero mantiene toda la estructura de conexiones original. Este enfoque evalúa la capacidad del modelo para generalizar etiquetas a nodos no observados, manteniendo la topología restante como contexto. Sin embargo, no siempre controla la filtración de información si los vecinos de nodos de prueba aparecen en entrenamiento.

11.3.2. Splits por aristas (predicción de enlaces)

En la predicción de enlaces, el conjunto de aristas E se divide en entrenamiento y prueba. Este esquema debe elegirse con cuidado, ya que:

- Particiones aleatorias simples pueden generar negatividades fáciles que inflan métricas.
- Particiones realistas requieren técnicas de muestreo que produzcan negativos difíciles o heurísticos más representativos de escenarios reales Li et al. (2023).

Un desafío adicional es evitar la inclusión de aristas que conectan componentes de prueba durante el entrenamiento, lo cual puede introducir correlaciones espurias Zhu et al. (2023).

11.3.3. Splits por subgrafos y tiempos (redes dinámicas)

Dividir conjuntos por subgrafos completos o por instantes de tiempo es especialmente útil en redes dinámicas o temporales, donde el orden de aparición de aristas refleja procesos reales (por ejemplo, evolución de relaciones sociales). En estos casos, el entrenamiento debe respetar la flecha del tiempo para evitar información futura en la construcción de representaciones de entrenamiento.

11.4. Métricas de evaluación por tarea

Dependiendo de la tarea, las métricas apropiadas varían y pueden dar lugar a interpretaciones diferentes del desempeño.

11.4.1. Clasificación de nodos

Para la clasificación semisupervisada de nodos, las métricas comúnmente reportadas incluyen *accuracy*, *F1-macro* y *F1-micro*. En escenarios con clases desbalanceadas, las métricas basadas en F1 son preferibles porque penalizan desbalances que pueden pasar desapercibidos con medidas agregadas simples.

11.4.2. Predicción de enlaces

La métrica más utilizada en predicción de enlaces es el *AUC* (Area Under ROC Curve). Sin embargo, AUC puede ser engañosa si los negativos de prueba son triviales. Estudios recientes han señalado que:

- La elección de negativos difíciles o heurísticos más representativos (e.g., muestreo heurístico relacionado) cambia significativamente los rankings de desempeño de modelos avanzados frente a métodos simples Li et al. (2023).
- Métricas alternativas como *precision@k* o curvas de precisión-recuerdo pueden complementar la evaluación Poursafaei et al. (2022).

11.5. Riesgos y métricas engañosas

Algunos errores metodológicos comunes incluyen:

- Comparar modelos complejos con baselines estructurales simples sin aplicar esquemas de partición y métricas consistentes, lo que conduce a conclusiones sesgadas.
- Reportar únicamente AUC cuando los negativos son fáciles de distinguir, produciendo estimaciones excesivamente optimistas Li et al. (2023).

Estos problemas enfatizan la necesidad de diseñar evaluaciones que reflejen las dificultades reales de los escenarios de aplicación.

11.6. Buenas prácticas para evaluación en grafos

Para asegurar evaluaciones robustas y comparables:

- Definir claramente la tarea y el tipo de generalización objetivo.
- Diseñar esquemas de partición de datos que respeten la estructura y eviten filtraciones no deseadas de información.
- Utilizar múltiples métricas relevantes para la tarea (e.g., AUC, precision@k).
- Comparar contra baselines estructurales bien establecidos bajo las mismas condiciones de evaluación.
- Documentar explícitamente las particiones y procedimientos de muestreo.

Estas prácticas son esenciales para producir evaluaciones científicamente válidas y útiles en contextos profesionales de Data Science.

11.7. Conclusiones

Evaluar modelos en grafos exige un enfoque metodológico que considere la dependencia estructural entre observaciones y los efectos de esquemas de partición de datos. En predicción de enlaces, métricas estándar pueden inducir conclusiones erróneas si los esquemas de muestreo no son realistas. La literatura reciente provee evidencia clara de estas limitaciones y propone caminos para una evaluación más rigurosa y comparable entre métodos.

Capítulo 12

Estudios integrados en ciencia de redes [REHACER - SUJETO A FACTIBILIDAD]

12.1. Motivación: de la teoría a sistemas reales

Los capítulos anteriores introdujeron los conceptos fundamentales de la ciencia de redes: estructura, modelos generativos, mesoestructura, procesos dinámicos y aprendizaje automático. Sin embargo, estos elementos adquieren sentido pleno sólo cuando se aplican a sistemas empíricos con restricciones reales: datos incompletos, ruido, decisiones de modelamiento y consecuencias sociales.

En este capítulo se presentan cuatro estudios integrados, cada uno basado en un conjunto de datos reales utilizados por el autor. El objetivo no es agotar el análisis de cada sistema, sino mostrar cómo un mismo marco metodológico se adapta a dominios distintos y cómo las preguntas científicas guían la elección de herramientas.

12.2. Caso I: Redes XDR y propagación de anomalías

12.2.1. Construcción de la red

Los datos XDR (Extended Detection and Response) se modelan como una red temporal dirigida y ponderada. Los nodos representan entidades (usuarios, dispositivos o direcciones IP) y las aristas corresponden a eventos observados en el tiempo, con pesos asociados a frecuencia o severidad.

12.2.2. Estructura y heterogeneidad

El análisis estructural revela distribuciones de grado y fuerza altamente heterogéneas, con nodos que concentran una fracción desproporcionada de eventos. Esta estructura anticipa una dinámica dominada por pocos nodos críticos.

12.2.3. Procesos dinámicos

La propagación de anomalías se interpreta como un proceso de difusión sobre la red. La ausencia de un umbral claro, consistente con redes heterogéneas, implica que eventos locales pueden escalar rápidamente a incidentes globales.

12.2.4. Aprendizaje sobre grafos

Se formulan tareas de detección de anomalías mediante Graph Neural Networks, combinando atributos locales con información estructural. El caso ilustra cómo el aprendizaje en grafos extiende, pero no reemplaza, el análisis estructural.

12.3. Caso II: Redes de interacción social inferidas vía WiFi

12.3.1. Inferencia de vínculos

Las interacciones sociales entre estudiantes se infieren indirectamente a partir de conexiones WiFi temporales. La red resultante es una aproximación probabilística de la interacción social, no una observación directa.

12.3.2. Mesoestructura y comunidades

La detección de comunidades revela agrupamientos coherentes con cursos, horarios y espacios, validando parcialmente la inferencia, pero también mostrando ambigüedades estructurales.

12.3.3. Procesos sociales

La difusión de información o comportamientos sobre esta red depende críticamente de la temporalidad y del contexto, destacando los límites de modelos estáticos.

12.3.4. Reflexión metodológica

Este caso enfatiza la necesidad de distinguir entre red observada y red latente, y sirve como ejemplo central de inferencia bajo incertidumbre.

12.4. Caso III: Red de carreras y trayectorias educativas

12.4.1. Red bipartita y proyecciones

La relación estudiante–carrera se modela como una red bipartita. Las proyecciones inducen sesgos estructurales que deben ser tratados explícitamente.

12.4.2. Estructura de oportunidades

La red revela trayectorias educativas, barreras estructurales y efectos de especialización que no son visibles en análisis individuales.

12.4.3. Conexión con causalidad

Este sistema sirve como base para discutir interferencia y efectos de red en decisiones educativas, anticipando los problemas tratados en el capítulo de causalidad.

12.5. Caso IV: Redes de compras públicas

12.5.1. Estructura bipartita y concentración

La red empresa–institución exhibe una estructura de core–periphery marcada, con empresas centrales que concentran contratos y montos.

12.5.2. Patrones anómalos

El análisis estructural y el aprendizaje sobre grafos permiten identificar patrones atípicos, aunque no necesariamente ilegales, resaltando la diferencia entre detección y acusación.

12.6. Síntesis transversal

Los cuatro casos ilustran principios comunes:

- la estructura condiciona la dinámica;
- la centralidad relevante depende del proceso;
- el aprendizaje automático amplifica, pero no sustituye, la comprensión estructural;
- los datos de redes requieren decisiones epistemológicas explícitas.

Este capítulo muestra que la ciencia de redes es un marco integrador para analizar sistemas complejos reales con rigor y responsabilidad.

Capítulo 13

Ética y privacidad en redes

13.1. Medir redes como acto de poder

13.1.1. Intuición: de lo local a lo global

Recolectar datos relacionales no es un acto neutral: medir una red social, biológica o tecnológica transforma interacciones locales en información global. Esta transformación otorga al investigador o al proveedor del servicio un poder considerable: tiene acceso a patrones agregados (comunidades, hubs, flujos) que los propios participantes desconocen. Como observa Clauset, medir redes implica controlar la forma en que se modelan y se interpretan esas relaciones, lo que condiciona las narrativas y decisiones basadas en ellas (Clauset, 2022). A diferencia de datos tabulares, la interdependencia entre nodos significa que capturar un enlace no sólo revela información sobre sus extremos, sino también sobre su vecindario y la posición de ambos en el conjunto de interacciones. Esta característica amplifica riesgos de privacidad y confidencialidad.

13.1.2. De la re-identificación a la manipulación de algoritmos

Aunque se anonymicen los nombres y atributos de los nodos, la estructura de la red puede ser suficiente para re-identificar individuos. Narayanan y Shmatikov mostraron que, al cruzar dos bases de datos con estructuras de red superpuestas (por ejemplo, redes sociales de distintos proveedores), es posible reidentificar usuarios con alta precisión (Narayanan and Shmatikov, 2009; Backstrom et al., 2007). La “firma estructural” de un nodo (por ejemplo, su patrón de grados y conexiones) puede ser única, lo que plantea dilemas respecto al uso de datos anonimizados. Además, saber qué nodo es quién permite interferir con algoritmos de recomendación, publicidad dirigida y decisiones automatizadas, generando brechas de equidad y sesgos (Borgatti and Molina, 2003).

13.2. Privacidad, homofilia y filtrado

13.2.1. Homofilia y fuga de información

La homofilia —tendencia de individuos con atributos similares a conectarse— es un rasgo común de muchas redes sociales (McPherson et al., 2001). Este fenómeno da lugar a la *fuga de información*: conocer los atributos de un subconjunto de nodos permite inferir los atributos de sus vecinos con elevada probabilidad. Por ejemplo, si la mayoría de los amigos de un nodo tienen una orientación política concreta, la probabilidad de que dicho nodo comparta esa orientación es alta. De hecho, trabajos recientes han mostrado que, incluso con un porcentaje reducido de etiquetas observadas, se pueden predecir atributos sensibles (identidad étnica, orientación

sexual, estado de salud) de manera alarmantemente precisa (Narayanan and Shmatikov, 2009). Esta posibilidad pone en riesgo la privacidad individual y pone de manifiesto la necesidad de mecanismos que limiten la inferencia involuntaria de atributos.

13.2.2. Ataques de inferencia y modelos de filtrado

Se han desarrollado múltiples técnicas para inferir atributos faltantes en redes heterogéneas. Entre ellas, el *label propagation* (difusión de etiquetas) utiliza la topología para propagar las etiquetas conocidas; algoritmos de agregación de vecinos (un caso simple de GNN) extienden estos conceptos incorporando aprendizaje profundo (Hamilton, 2020). En el contexto de sistemas de recomendación, los modelos de filtrado colaborativo hacen predicciones basadas en patrones de comportamiento de usuarios similares. Sin embargo, al explotar homofilia y correlaciones, estos modelos pueden exponer información sensible de forma indirecta. Además, los propios algoritmos de recomendación pueden amplificar sesgos presentes en los datos (p. ej. invisibilizar a ciertos grupos o reforzar segregaciones existentes) (Barocas and Selbst, 2016).

13.3. Buenas prácticas y principios éticos

13.3.1. Anonimización y minimización de daños

Antes de compartir o analizar una red, se deben aplicar técnicas de anonimización. Las estrategias tradicionales incluyen eliminar identificadores directos y aplicar permutaciones, perturbaciones o agregaciones de la estructura (Hay et al., 2008). Sin embargo, el riesgo de reidentificación permanece alto. Una línea más reciente es la *diferencial privacidad* en grafos: un mecanismo M satisface ϵ -diferencial privacidad de aristas si, para cualquier par de grafos que difieren en una arista, la probabilidad de resultados de M difieren a lo sumo por un factor e^ϵ (Kasiviswanathan et al., 2013; Zhou et al., 2016). Existen variantes de privacidad de nodos, más difíciles de alcanzar pero que protegen la existencia de nodos completos.

13.3.2. Consentimiento informado y transparencia

El consentimiento debe ser específico, informado y revocable. Informar a los participantes de cómo se usarán sus datos, qué riesgos implica su recolección y con quién se compartirán es una obligación ética y legal. Además, cuando los datos provienen de terceros (por ejemplo, crawlers o API públicas), se debe respetar la licencia y las expectativas de privacidad de los usuarios (Tufekci, 2014). La transparencia también se extiende a la publicación de resultados: no se deben publicar redes sensibles o sub-redes que permitan reidentificar a individuos, especialmente en contextos de salud, sexualidad o persecución política.

13.3.3. Equidad algorítmica y mitigación de sesgos

Los algoritmos de recomendación y predicción basados en redes pueden perpetuar o amplificar sesgos. Por ejemplo, un modelo de recomendación de empleo podría favorecer sistemáticamente a candidatos de determinadas comunidades debido a la homofilia en los datos. La literatura en *fairness* y justicia algorítmica propone métricas y métodos para auditar y corregir estos sesgos (Barocas and Selbst, 2016; Chouldechova, 2017). En grafos, aún es un campo emergente: se están desarrollando GNN diseñadas para minimizar dependencia con atributos sensibles, mediante regularizadores adversarios o variantes de debiasing (Bose et al., 2019).

13.3.4. Directrices institucionales y deontología

Sociedades científicas, editoriales y agencias reguladoras han desarrollado códigos de conducta para la investigación en redes. Por ejemplo, la Asociación Internacional de Investigación de Redes (INSNA) y la American Sociological Association recomiendan principios de anonimización, minimización de daño y revisión ética. Además, la normativa europea (GDPR) y regulaciones locales definen límites y obligaciones para el procesamiento de datos personales, incluidos los de redes sociales. Es crucial que los investigadores sigan estos lineamientos y consulten a comités de ética cuando el estudio implique potenciales riesgos para los individuos.

13.4. Ética del aprendizaje automático en grafos

Los métodos de aprendizaje en redes, particularmente las GNN y los embeddings, amplifican riesgos éticos existentes. Las representaciones aprendidas pueden codificar atributos sensibles incluso cuando éstos no se observan directamente, debido a homofilia estructural y message passing. Esto implica que la anonimización clásica resulta insuficiente cuando los modelos explotan dependencias relacionales.

Además, los algoritmos de clasificación y recomendación en grafos pueden reforzar sesgos estructurales, amplificando desigualdades preexistentes. La falta de interpretabilidad en modelos profundos dificulta auditar estas decisiones. Por ello, la ética del aprendizaje en redes requiere considerar explícitamente privacidad relacional, equidad estructural y responsabilidad algorítmica desde el diseño del modelo.

13.5. Conclusiones

La ética y la privacidad en el análisis de redes no son cuestiones accesorias: la naturaleza relacional de los datos crea riesgos y dilemas que no aparecen en otros contextos. La capacidad de inferir información sensible de la estructura, la vulnerabilidad a reidentificación y el poder de alterar la visibilidad y oportunidades de grupos de personas hacen que el investigador deba ser especialmente cauto. La adopción de técnicas de anonimización robusta, diferencial privacidad, políticas de consentimiento informado y auditorías de sesgos son pasos fundamentales para mitigar riesgos. Finalmente, el rápido avance de los algoritmos de aprendizaje sobre grafos y la creciente disponibilidad de datos exigen marcos normativos y éticos que se actualicen constantemente para proteger a los individuos y colectivos implicados.

Bibliografía

- Backstrom, L., Dwork, C., and Kleinberg, J. (2007). Wherefore art thou r3579x? anonymized social networks, hidden patterns, and structural re-identification. In *Proceedings of the 16th International Conference on World Wide Web*, pages 181–190.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Barabási, A.-L. and Pósfai, M. (2016). *Network Science*. Cambridge University Press, Cambridge.
- Barocas, S. and Selbst, A. D. (2016). Big data’s disparate impact. In *California Law Review*, volume 104, pages 671–732. Discusses algorithmic discrimination and fairness, applicable to graph-based recommender systems.
- Barrat, A., Barthélemy, M., and Vespignani, A. (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press, Cambridge.
- Bartolucci, S., Caccioli, F., Caravelli, F., and Vivo, P. (2024). Distribution of centrality measures on undirected random networks via the cavity method. *Proceedings of the National Academy of Sciences of the United States of America*, 121(40):e2403682121.
- Bendahman, N. and Lotfi, D. (2024). Unveiling influence in networks: A novel centrality metric and comparative analysis through graph-based models. *Entropy*, 26(6):486.
- Bo, D., Wang, X., Yu, A., Gong, Y., Shi, C., Guo, J., and Liang, Y. (2021). Beyond homophily in graph neural networks: Current limitations and effective designs. *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*.
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27(1):55–71.
- Borgatti, S. P. and Molina, J. L. (2003). Ethical and strategic issues in organizational social network analysis. *Journal of Applied Behavioral Science*, 39(3):337–349.
- Bose, A. J., Hamilton, W. L., and Compton, R. (2019). Compositional fairness constraints for graph embedding. In *Proceedings of the 36th International Conference on Machine Learning*, pages 715–724. Proposes fairness constraints for embeddings in graphs.
- Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2):153–163.
- Clauset, A. (2022). Ethics and the science of networks. Notas de clase. Discusses power dynamics and ethical dilemmas in network measurement and analysis.
- Cohen, R., Havlin, S., and ben Avraham, D. (2003). Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, 91(24):247901.

- DeCanio, S. and Watkins, W. (2025). Inequality emerges from networks. *Qeios*. Preprint.
- Evans, T. S. and Chen, B. (2022). Linking the network centrality measures closeness and degree. *Communications Physics*, 5(1):172.
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486:75–174.
- Grover, A. and Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 855–864. ACM.
- Hamilton, W. L. (2020). *Graph Representation Learning*, volume 14 of *Synthesis Lectures on Artificial Intelligence and Machine Learning*. Morgan & Claypool Publishers.
- Hamilton, W. L., Ying, R., and Leskovec, J. (2017). Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 1024–1034.
- Hay, M., Li, G., Miklau, G., and Jensen, D. (2008). Resisting structural re-identification in anonymized social networks. In *Proceedings of the VLDB Endowment*, volume 1, pages 102–114.
- Karrer, B. and Newman, M. E. J. (2010). Message passing approach for general epidemic models. *Physical Review E*, 82(1):016101.
- Kasiviswanathan, S. P., Nissim, K., Raskhodnikova, S., and Smith, A. (2013). Analyzing graphs with node differential privacy. *The Theory of Cryptography Conference (TCC)*, 7194:457–476. Introduces node differential privacy and mechanisms for graph data.
- Kempe, D., Kleinberg, J., and Tardos, É. (2015). Maximizing the spread of influence through a social network. *Theory of Computing*, 11(4):105–147. Extended version of KDD 2003 paper.
- Kipf, T. N. and Welling, M. (2016). Variational graph auto-encoders. *NIPS Workshop on Bayesian Deep Learning*. arXiv:1611.07308.
- Kipf, T. N. and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *Proceedings of the International Conference on Learning Representations (ICLR)*. arXiv:1609.02907.
- Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., and Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature Physics*, 6:888–893.
- Li, J., Shomer, H., Mao, H., Zeng, S., Ma, Y., Shah, N., Tang, J., and Yin, D. (2023). Evaluating graph neural networks for link prediction: Current pitfalls and new benchmarking. In *Advances in Neural Information Processing Systems 36 (Datasets and Benchmarks Track)*.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444.
- Menczer, F., Fortunato, S., and Davis, C. A. (2020). *A First Course in Network Science*. Cambridge University Press, Cambridge.
- Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., and Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, 328(5980):876–878.
- Narayanan, A. and Shmatikov, V. (2009). De-anonymizing social networks. *Proceedings of the 30th IEEE Symposium on Security and Privacy*, pages 173–187.

- Newman, M. (2010). *Networks: An Introduction*. Oxford University Press, Oxford.
- Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582.
- Newman, M. E. J. (2000a). Models of the small world: A review. <https://arxiv.org/abs/cond-mat/0001118>. arXiv preprint cond-mat/0001118.
- Newman, M. E. J. (2000b). Models of the small world: A review. arXiv preprint.
- Newman, M. E. J. (2002). Spread of epidemic disease on networks. *Physical Review E*, 66(1):016128.
- Pastor-Satorras, R., Castellano, C., Mieghem, P. V., and Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, 87:925–979.
- Pastor-Satorras, R. and Vespignani, A. (2002). Immunization of complex networks. *Physical Review E*, 65(3):036104.
- Perozzi, B., Al-Rfou, R., and Skiena, S. (2014). Deepwalk: Online learning of social representations. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 701–710.
- Poursafaei, F. et al. (2022). Towards better evaluation for dynamic link prediction. In *NeurIPS Datasets and Benchmarks Workshop*.
- Rodrigues, F. A. (2018). Network centrality: an introduction. In Macau, E. E. N., editor, *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems*, volume 9 of *Nonlinear Systems and Complexity*, pages 177–196. Springer.
- Rodrigues, F. A. (2019). Network centrality: An introduction. In Macau, E. E. N., editor, *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems*, pages 177–196. Springer, Cham.
- Rossi, E., Chamberlain, B., Frasca, F., King, O., Bronstein, M., and Monti, F. (2021). Temporal graph networks for deep learning on dynamic graphs. *Proceedings of the International Conference on Machine Learning (ICML)*. arXiv:2006.10637.
- Rossi, R. A. and Ahmed, N. K. (2015). Role discovery in networks. *IEEE Transactions on Knowledge and Data Engineering*, 27(4):1112–1131.
- Shi, C., Kong, X., Kuang, Y., Wang, M., and Sun, Y. (2020). Heterogeneous graph neural network. *Proceedings of the 2019 World Wide Web Conference*, pages 1054–1064.
- Sánchez-Lengeling, B., Reif, E., Pearce, A., and Wiltchko, A. B. (2021). A gentle introduction to graph neural networks. *Distill*, 6(9):e33.
- Tang, J., Qu, M., Wang, M., Zhang, J. Y., Qin, B., and Mei, Q. (2015). Line: Large-scale information network embedding. *Proceedings of the 24th International Conference on World Wide Web*, pages 1067–1077.
- Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. *Proceedings of the International AAAI Conference on Web and Social Media*, 8:505–514. Addresses ethical issues in using public social media data.
- Van Mieghem, P., Omic, J., and Kooij, R. (2009). Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14.

- Veličković, P., Cucurull, G., Casanova, A. R., Liò, P., and Bengio, Y. (2018). Graph attention networks. *Proceedings of the International Conference on Learning Representations (ICLR)*. arXiv:1710.10903.
- Viles, W. D. and O’Malley, A. J. (2023). Constrained community detection in social networks. *New England Journal of Statistics in Data Science*, 2(3):368–379.
- Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, UK.
- Watts, D. J. and Strogatz, S. H. (1998a). Collective dynamics of small-world networks. *Nature* **393**, 440–442.
- Watts, D. J. and Strogatz, S. H. (1998b). Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442.
- Yanchenko, E. and Sengupta, S. (2023). Core-periphery structure in networks: a statistical exposition. *Statistics Surveys*, 17:42–74. arXiv preprint arXiv:2202.04455.
- Ying, R., You, J., Morris, C., Ren, X., Hamilton, W. L., and Leskovec, J. (2018). Hierarchical graph representation learning with differentiable pooling. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 4805–4815.
- Yoo, Y., Shi, X., Collins, M., and Zust, J. (2020). Graph u-nets for graph neural network based graph classification. *arXiv preprint arXiv:1905.05178*. arXiv:1905.05178.
- Zhou, B., Bu, Z., Wang, C., Gao, S., Li, K., Cao, Q., and Song, D. (2016). Edge differential privacy for social network data publishing. *IEEE Transactions on Knowledge and Data Engineering*, 28(9):2413–2423.
- Zhu, J., Zhou, Y., Ioannidis, V. N., Qian, S., Ai, W., Song, X., and Koutra, D. (2023). Pitfalls in link prediction with graph neural networks: Understanding the impact of target-link inclusion & better practices. *arXiv preprint*.