

Análisis de Redes Sociales

CORE Lab + ESGE-EPG

Contents

Introducción	7
Patrocinio	8
Misión educativa	8
Editores	8
I ¿Qué es el ARSo?	11
1 ARSo: Una (Breve) Introducción	13
1.1 Conceptos Erróneos y Diferencias	14
1.2 Suposiciones	16
1.3 Términos y Conceptos Básicos	22
1.4 Resumen y Conclusiones	28
II ARSo: Tecnicas	29
2 Trabajando con Gephi	31
2.1 Introducción	31
2.2 Gephi	31
2.3 Gephi Para Principiantes	32
3 Tipos de Datos de Redes Sociales	49
3.1 Datos de Modo-Uno	49
3.2 Datos de Modo-Dos	53
3.3 Datos de Atributos	56
4 Grabación y Recopilación	59
4.1 Matrices	59
4.2 Lista de Aristas	62
4.3 Lista de Nodos	64
4.4 Ejercicio Práctico	65

5 Derivando Datos Modo-Uno de Modo-Dos	83
5.1 Modo-Dos, ¿Qué es?	83
5.2 Proyección	83
5.3 Ejercicio Práctico	88
6 Múltiples Relaciones	97
6.1 Matrices y Listas de Aristas	97
6.2 Combinar	99
6.3 Ejercicio Práctico	104
7 Extraer y Simplificar Redes	111
7.1 Extraer	111
7.2 Simplificar	113
7.3 Ejercicio Práctico	113
III ARSo: Medidas	131
8 Topología	133
8.1 Medidas Básicas de Topología	134
8.2 Interconexión	139
8.3 Centralización	141
8.4 Ejercicio Práctico	145
9 Subgrupos Cohesivos	153
9.1 Componentes	153
9.2 K-cores	155
9.3 Girvan-Newman	157
9.4 Louvain	161
9.5 Ejercicio Práctico	162
10 Centralidad	171
10.1 Datos No Dirigidos: Centralidad y Poder	171
10.2 Datos Dirigidos: Centralidad y Prestigio	182
10.3 Ejercicio Práctico	187
11 Puentes y Corredores	193
11.1 Puntos de Corte	194
11.2 Bi-componentes y Puentes	195
11.3 Intermediación de borde	196
11.4 Agujeros Estructurales, Restricción y Autonomía	198
11.5 Ejercicio Práctico	203
IV Aplicaciones	209
12 Redes Oscuras en el Perú	211

CONTENTS	5
12.1 Análisis de Redes Oscuras	211
12.2 ¿Por qué el Análisis de Redes Oscuras?	213
12.3 Conclusiones	218
13 Un Proceso para la Investigación de Redes Sociales	219
13.1 Desarrollar un Proyecto Investigativo de Redes Sociales	220
14 Ejemplos Prácticos	227
Informe de la Red Criminal en la República de Panamá	227
A Glosario de Términos	231
B Libro de Códigos	239
B.1 Libro de Códigos de Muestra	239
B.2 Libro de Códigos de Noordin Top	242
V Referencias	253
Referencias	255

Introducción



El objetivo de este manual es enseñarle los conceptos básicos del análisis de redes sociales (ARSo) y cómo estos métodos pueden ser utilizados para comprender dichas redes, asimismo crear estrategias para empoderar o socavarlas. Este proyecto cubre los puntos de vista de los autores sobre la introducción al estudio de las redes oscuras como elementos de análisis clave para todo analista de inteligencia y elementos operativos de fuerzas especiales. De tal manera que, los ejemplos prácticos y el marco teórico en que están integrados giran alrededor de las redes oscuras, definidas como redes encubiertas (o clandestinas) e ilegales (Raab & Milward, 2003). Este es un término que puede incluir redes terroristas, criminales e insurgentes. El término “oscuro” busca capturar el hecho de que las redes oscuras son aquellas que, por definición, intentan permanecer encubiertas tras ocultar sus actividades de las autoridades (Cunningham, Everton, & Murphy, 2016; Everton, 2012a).

Este manual se divide en cuatro secciones. La primera consiste de una introducción a ARSo, que aborda una breve reseña histórica, las diferencias entre este y otros campos de estudio, y los términos, conceptos y supuestos básicos que forman la base de este campo. La siguiente sección cambia el enfoque de los conceptos a la práctica. Es decir, se centra en la recopilación, codificación y manipulación de los datos necesarios para el análisis. En la tercera sección, los autores presentan varias “familias” de métricas, comúnmente utilizadas en el análisis exploratorio de redes. Finalmente, la última sección destaca ejemplos del uso de los métodos presentados en este documento por analistas de inteligencia y equipos de Asuntos Civiles y Operaciones Psicológicas.

El contenido de este manual fue diseñado conjuntamente entre el Ambiente Común de Investigación Operativa (CORE Lab, por sus siglas en inglés) de la Escuela Naval de Post Grado (NPS, por sus siglas en inglés) y la Escuela

Superior de Guerra del Ejército, Escuela de Post Grado (ESGE-EPG), Lima, Perú.

Patrocinio

El desarrollo de este manual fue posible gracias al patrocinio del Programa de Becas de Defensa Regional (RDFP, por sus siglas en inglés) de la Agencia de Cooperación para la Seguridad de la Defensa (DSCA, por sus siglas en inglés) con el propósito de apoyar el plan de estudios de ESGE-EPG.



Misión educativa

El material contenido en este manual se realizó con el propósito de revisión y discusión entre colegas para su posterior aplicación y no necesariamente refleja las opiniones del Departamento de Marina (en inglés, “Department of the Navy”) o del Departamento de Defensa (en inglés, “Department of Defense”) de los Estados Unidos y la Escuela Superior de Guerra - Escuela de Post Grado del Ejército de Perú.

Editores

Christopher J. Callaghan es un investigador asociado en el CORE Lab en el Departamento de Análisis de Defensa en NPS. Su trabajo aprovecha el análisis de datos de código abierto para comprender y modelar una variedad de problemas de seguridad nacional e internacional. Antes de unirse al NPS, se desempeñó como investigador asociado en el Instituto Middlebury de Estudios Internacionales (MIIS, por sus siglas en inglés) en Monterey California, donde trabajó en múltiples esfuerzos relacionados con la “ciudadanía” local e internacional.

Sean F. Everton es un Profesor en el Departamento de Análisis de Defensa y Codirector del CORE Lab de NPS. Ha publicado en las áreas de análisis de redes sociales, sociología de la religión, sociología económica y sociología política y se especializa en el uso del análisis de redes sociales para rastrear e

interrumpir redes oscuras (por ejemplo, redes criminales y terroristas). Su primer libro, *Disrupting Dark Networks*, fue publicado por Cambridge University Press en 2012, y su segundo libro, *Understanding Dark Networks* fue publicado por Rowman y Littlefield en 2016. Su libro más reciente , *Networks and Religion*, que explora la interacción de las redes y la religión, fue publicado por Cambridge University Press en 2018.

Gral. de Brig. EP (R) Augusto Manuel E. Álvarez Torres ostenta los títulos académicos de Maestro en Seguridad, Defensa y Desarrollo Nacional y en Ciencias Militares con mención en Planeamiento Estratégico y Toma de Decisiones, es egresado de la Escuela de Post Grado de Ciencias Políticas de las Universidad Ricardo Pala, ha seguido cursos de Terrorismo e Insurgencia, Crimen Organizado, Políticas y Estrategias de Seguridad, Estrategia y Seguridad Internacional, Relaciones Político, Civil, Militar y Liderazgo Estratégico, Estrategia de Seguridad y Defensa frente a las Amenazas Complejas en la UND/CHDS-EE. UU. Los cursos de Análisis de redes Nivel I y II, en la Escuela Naval de Post Grado en Monterey, California, EE. UU. Actualmente se desempeña como Catedrático en el Centro de Altos Estudios Nacionales (CAEN), Profesor Principal de la Catedra de Historia Militar y Profesor Adjunto en la Universidad Conjunta de Operaciones Especiales en Florida, EE. UU.

Part I

¿Qué es el ARSo?

Chapter 1

ARSo: Una (Breve) Introducción

Georg Simmel (Simmel, 1950) es visto generalmente como el precursor intelectual del análisis de redes sociales (ARSo). Argumentó que para comprender el comportamiento social, debemos estudiar los patrones de interacción, y sus estudios ofrecieron conocimientos penetrantes, por ejemplo, sobre la naturaleza de las sociedades secretas, la dinámica de las diádicas y tríadas, y cómo la creciente complejidad social ha llevado a mayores niveles de individualismo. Los estudios de Simmel continúan influyendo la disciplina actualmente, sin embargo, podemos rastrear el surgimiento del análisis de redes sociales a dos corrientes principales de pensamiento: el trabajo de los psicólogos y los antropólogos sociales. La primera corriente enfatizó cómo los patrones organizados dan forma a la forma en que vemos e interpretamos el mundo, mientras que la segunda se centró en la relación entre los patrones y la estructura sociales. Ambas ayudaron a sentar la base para el desarrollo formal de ARSo en Harvard en las décadas de 1960 y 1970, dirigido por el sociólogo Harrison White (White, 2008) y sus estudiantes (por ejemplo, Ronald Breiger, Kathleen Carley, Bonnie Erickson, Ivan Chase, Mark Granovetter, Michael Schwartz y Barry Wellman). White, quien obtuvo su doctorado en física teórica, argumentó que la sociología, a pesar de su pretensión de estudiar los fenómenos sociales, estaba en deuda con formas individualistas de análisis basadas en las características agregadas de los individuos, a menudo con la ayuda del análisis estadístico de los datos de encuestas. En respuesta, desarrolló un enfoque basado en estudios centrados en las relaciones sociales y los patrones que surgieron de ellas.

Los esfuerzos de White no ocurrieron en el vacío. Otras tradiciones teóricas también informaron el desarrollo del ARSo, como la teoría de grafos, la teoría del intercambio y la investigación sobre el reclutamiento de individuos para movimientos religiosos y sociales. En los últimos años, los economistas se

han interesado cada vez más en las redes sociales, al igual que los físicos y otros científicos. Esto ha llevado a un mayor interés en el ARSo, atrayendo a investigadores de diversas disciplinas y generando varios estudios altamente creativos.

¿Qué es el análisis de redes sociales? En pocas palabras, es una colección de teorías y métodos que asume que el comportamiento de los actores se ve afectado por sus vínculos con los demás y las redes en las que están integrados. En lugar de asumir que los actores no son afectados por quienes los rodean, ARSo asume que los patrones de interacción afectan lo que dicen, hacen y creen los actores. Aunque algunas interacciones son al azar, muchas no lo son. Es decir, los actores tienden a interactuar con otros similares, y la interacción repetida puede conducir al surgimiento de la formación social a nivel micro (por ejemplo, individual), meso (por ejemplo, grupo) y macro (por ejemplo, instituciones, naciones) que pueden ser la unidad de análisis. El ARSo se diferencia de otros enfoques cuantitativos tradicionales, que tienden a centrarse en los atributos de los actores (por ejemplo, género, raza, educación), pues el ARSo se centra en cómo los patrones de interacción afectan el comportamiento. Debido a que los atributos no cambian en los contextos sociales, no pueden explicar ninguna variación en el comportamiento de un actor, lo que sugiere que los patrones de interacción son tan (o quizás más) importantes para comprender el comportamiento que los atributos:

Una mujer que tiene un trabajo servil que requiere poca iniciativa en una oficina puede ser una líder dinámica de una asociación de vecinos y una participante asertiva de la asociación de padres de familia. Esas diferencias de comportamiento son difíciles de reconciliar con atributos inmutables de género, edad y estatus, pero son comprensibles al reconocer que las relaciones estructurales de las personas pueden variar notablemente en los contextos sociales – (Knock & Yang, 2007, p. 5).

Consecuentemente, un objetivo principal del ARSo ha sido desarrollar métricas para ayudar a los analistas a comprender mejor las características estructurales de una red en particular. Esta metodología se ha utilizado con éxito para explicar variedades de comportamiento principalmente porque obliga a los investigadores a “pensar en términos de restricciones y opciones inherentes a la forma en que se organizan las relaciones sociales” (Raab & Milward, 2003). El resto de este capítulo presenta los términos, conceptos y suposiciones básicas de ARSo. Sin embargo, antes de hacerlo, señalamos brevemente algunos de los conceptos erróneos que rodean al ARSo.

1.1 Conceptos Erróneos y Diferencias

El término análisis de redes sociales a veces se confunde con medios sociales y análisis de medios sociales. Es decir, cuando las personas escuchan que van a

aprender el análisis de redes sociales, a menudo piensan que aprenderán a extraer datos de medio sociales como Facebook y Twitter. ARSo es una colección de teorías y métodos que se han desarrollado para comprender la estructura de las redes sociales. Por el contrario, los medios sociales constan de plataformas y contenido generado por el usuario que puede incluir texto, imágenes, videos, conexiones entre usuarios y enlaces a sitios web. El hecho de que el contenido de los medios sociales es a menudo relacional puede aumentar la confusión. Para clarificar, los analistas pueden extraer datos de los medios sociales y, posteriormente, utilizar ARSo para comprender ciertos aspectos de estas redes sociales (por ejemplo, seguidores en una red de Twitter de usuario a usuario). Sin embargo, es importante hacer la diferenciación.

La forma en que algunos usan el término “red” también puede resultar confusa. Algunos lo usan para referirse a tipos de organizaciones descentralizadas, informales y/u orgánicas (Burns & Stalker, 1961; Podolny & Page, 1998; Powell, 1990; Powell & Smith-Doerr, 1994; Ronfeldt & Arquilla, 2001). Esta distinción puede ser útil, pero dentro del mundo del análisis de redes sociales, todas las organizaciones son redes. Algunas pueden ser más jerárquicas que otras, pero siguen siendo redes (Nohria, 1992). Es por ello que los analistas de redes sociales han desarrollado algoritmos que miden el grado en que una red en particular es jerárquica o no (ver, por ejemplo, Davis (1979); Nooy, Mrvar, & Batagelj (2005); Krackhardt (1994)).

Finalmente, el ARSo a veces se confunde con el análisis de enlaces, una metodología relacionada pero distinta que examina los patrones relacionales de varios objetos. Una diferencia fundamental entre los dos es que, si bien los diagramas de análisis de enlaces a menudo incluyen diferentes tipos de objetos (por ejemplo, individuos, automóviles, teléfonos celulares, etc.) y los vínculos entre ellos, los diagramas de redes sociales solo incluyen vínculos entre tipos similares de objetos. Tomemos, por ejemplo, un diagrama de análisis de vínculos en el que dos individuos (A y B) tienen vínculos con otros cinco objetos, pero los objetos con los que tienen vínculos difieren entre sí (Figura 1.1). En este ejemplo, la persona A está vinculada a la persona B, así como a una bomba, un teléfono celular, una casa y un automóvil, mientras que la persona B está vinculada a cuatro personas (A, C, D y E) y un teléfono celular. Si bien ambos tienen cinco vínculos (que es la definición de grado de centralidad; consulte la discusión sobre centralidad a continuación), no podemos comparar de manera significativa el número de vínculos de estos dos individuos porque son de diferentes tipos de objetos. Sería como comparar manzanas y naranjas.

En el ARSo, los actores tienen vínculos con objetos similares, lo que hace significativa la comparación directa del número de vínculos. Esto se ilustra en la Figura 1.2 donde en el panel de la izquierda, los individuos A y B tienen cada uno cinco vínculos con otros cinco individuos, mientras que en el derecho, tienen cinco vínculos con cinco escuelas diferentes. En ambos casos, los lazos de A son comparables a los de B porque son del mismo tipo de objeto. Aunque los analistas de redes sociales están interesados en más que el recuento de los

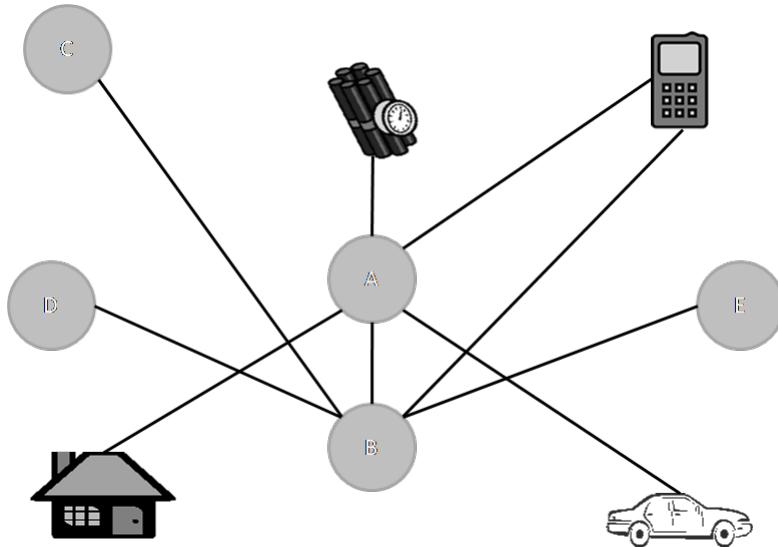


Figure 1.1: Ejemplo de un análisis de enlaces

lazos de los actores, otros algoritmos de análisis de redes sociales generalmente asumen que los lazos se dan entre tipos similares de objetos.

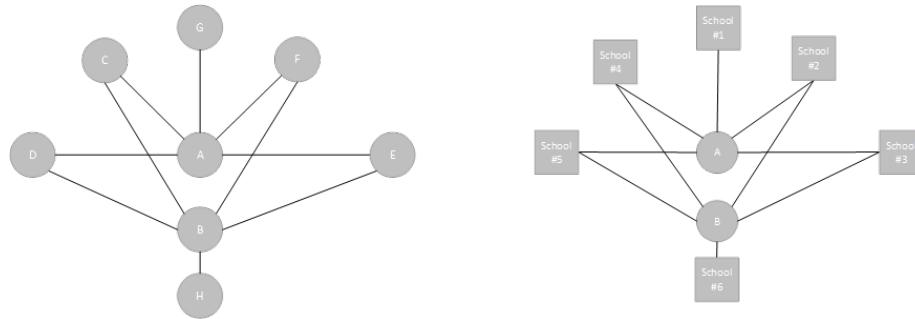


Figure 1.2: Ejemplo de diagramas de red

1.2 Suposiciones

La mayoría de las teorías y métodos de ARSo se basan en un conjunto común de suposiciones (Azarian, 2005; Knoke & Yang, 2007; Wasserman & Faust, 1994):

- Los actores y sus acciones relacionadas son interdependientes, en lugar de independientes, con otros actores
- Los vínculos entre actores son conductos para la transferencia o el flujo de diversos tipos de bienes o recursos materiales y/o no materiales (por

ejemplo, fondos, creencias, confianza, etc.)

- Las estructuras sociales son vistas en términos de patrones durables de vínculos entre actores
- Las interacciones repetidas entre actores dan lugar a formaciones sociales que cobran vida propia, siguen su propia lógica y no pueden reducirse a sus partes constituyentes aunque sigan dependiendo de estas
- La ubicación de los actores en la estructura social (es decir, su ubicación estructural) impacta sus creencias, normas y comportamiento observado
- Las redes sociales son entidades dinámicas que cambian a medida que los actores, subgrupos y vínculos entre actores entran, se forman, salen o se eliminan de la red

1.2.1 Interdependencia de los Actores

El ARSo asume que los actores no toman decisiones de forma autónoma, sino que están sustancialmente influenciados por el comportamiento y las decisiones de otros actores. Esta suposición difiere de la perspectiva de elección racional “pura” en la que los actores actúan independientemente de aquellos con quienes interactúan. Estudios como los experimentos de conformidad social de Solomon Asch (1951), los experimentos de obediencia a la autoridad de Stanley Milgram (1974) y el experimento de la prisión de Philip Zimbardo (1972) destacan cómo las redes en las que están inmersos los actores pueden influir en su comportamiento. Estos estudios sugieren que, en lugar de actuar independientemente de quienes los rodean, las personas hacen lo contrario. Los sujetos en el experimento de Asch optaron por estar de acuerdo con la multitud frente a la presión de los compañeros, incluso cuando la respuesta correcta era obvia. Los sujetos de Milgram tomaron decisiones en presencia de autoridad que probablemente no habrían hecho en un contexto diferente. Y en solo unos días, los guardias de Zimbardo adquirieron un sorprendente nivel de brutalidad, que presagió las actividades que tomaron lugar en la prisión de Abu Ghraib y el debacle que consiguió treinta años después.

Los individuos no son los únicos actores influenciados por otros actores. Por ejemplo, John Meyer, Woody Powell y Paul DiMaggio (y numerosos colegas) han demostrado que los grupos, corporaciones y estados-nación no tienen más probabilidades de actuar de forma autónoma que los individuos (ver, por ejemplo, DiMaggio & Powell (1983); Meyer, Boli, Thomas, & Ramirez (1997); Meyer & Rowan (1977)). Cuanto más interactúan entre sí, más se parecen con el tiempo. Lo hacen no necesariamente por razones instrumentales, sino para mantener su legitimidad a los ojos de otros actores similares.

Cuando una práctica o estructura organizacional se entiende comúnmente como una característica definitoria de una organización ‘legítima’ de cierto tipo, las élites de la organización se sienten presionados para instituir esa práctica o estructura. Si hay una norma cultural que dice: “Para que una organización sea una buena organización, debe tener la característica X,” las organizaciones se sienten presionadas

para instituir la característica X – (Chavez, 1961, pp. 32–33).

1.2.2 Lazos como Conductos

Otra suposición es que los lazos (es decir, las relaciones) entre actores funcionan como conductos para el flujo de varios tipos de “bienes” materiales y no materiales, como información, sentimientos, recursos financieros, normas, enfermedades, opiniones y confianza. Quizás, el ejemplo más conocido de esto es el estudio de Granovetter (1973) sobre cómo las personas encuentran sus trabajos actuales. Granovetter descubrió que era mucho más probable que las personas utilizaran contactos personales en comparación a otros medios cuando se trataba de encontrar trabajo. De aquellos que encontraron su trabajo a través de contactos personales, la mayoría de esos contactos eran débiles (es decir, conocidos) en lugar de lazos fuertes (es decir, amigos cercanos). ¿Por qué? Porque es menos probable que nuestros lazos débiles se involucren socialmente entre sí que nuestros lazos fuertes. Por ello, el conjunto de personas que componen nuestra red de conocidos tiende a ser relativamente escaso, mientras que el grupo de personas que componen nuestra red de amigos cercanos tiende a ser denso. Considere el patrón de lazos sociales sugerido por este argumento (Figura 1.3). La mayoría de las personas tendrán algunos amigos cercanos, la mayoría de los cuales se conocen entre sí. Probablemente también tengan varios conocidos, pocos de los cuales se conocen entre sí. Sin embargo, es probable que cada uno de estos conocidos tenga amigos cercanos propios, por lo que también es probable que estén integrados en sus propias redes altamente interconectadas, pero separadas de nuestro individuo original. Según Granovetter, los lazos débiles son importantes en términos de la estructura general de una red porque forman los puentes cruciales que unen a grupos de personas densamente unidas. De hecho, sin ellos, estos grupos no estarían conectados en absoluto.

Esto llevó a Granovetter a argumentar que “cualquier cosa que se difunda,” ya sea información laboral, influencia, recursos, confianza, etc., “llegará a un mayor número de personas y viajará una mayor distancia [...] cuando pasa por lazos débiles en lugar de fuertes” (Granovetter, 1973, p. 1366). Debido a esto, es más probable que los actores con pocos lazos débiles “se limiten a las noticias provinciales y las opiniones de sus amigos cercanos” (Granovetter, 1983, p. 2002). Además, los grupos o comunidades cuyos miembros carecen de vínculos débiles encontrarán más dificultades para movilizarse por una causa. Por ejemplo, al comparar dos intentos de organización comunitaria, uno que tuvo éxito y otro que no, Granovetter (1973, pp. 1373–1376) argumentó que una diferencia crítica entre los dos era que el que no tuvo éxito carecía de los lazos débiles que necesitaba para unir sus diversos círculos sociales.

Esto no significa que los lazos fuertes tienen poco o ningún valor. Es más probable que los sentimientos de confianza y solidaridad se compartan entre lazos fuertes que entre los débiles. Así, mientras que los vínculos débiles brindan a las personas acceso a información y recursos más allá de sus círculos sociales inmediatos, los vínculos fuertes suelen ser fuentes de apoyo en tiempos de incertidumbre. De

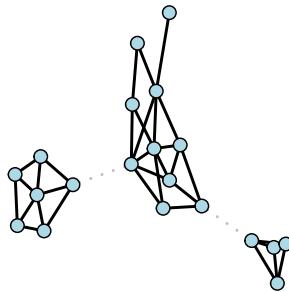


Figure 1.3: Lazos fuertes y débiles

hecho, es más probable que las personas con numerosos vínculos fuertes sean más felices y disfruten de una mejor salud.

1.2.3 Estructura Social y Formaciones Sociales Emergentes

Los científicos sociales se refieren con frecuencia al concepto de estructura social. Con este término, generalmente tienen en mente los patrones durables de comportamiento y las relaciones dentro de los sistemas sociales (por ejemplo, los roles) o las instituciones y normas sociales que se han integrado en los sistemas sociales de tal manera que dan forma al comportamiento. Las estructuras sociales “restringen quién está presente, dónde se encuentran, qué pueden hacer y cómo se relacionan entre sí. Esta estructura es tan real como los edificios que ocupan las personas” (Turner, 2006, p. 88). Sin embargo, es notoriamente difícil de captar la estructura social empíricamente, por lo que muchos encuentran atractivo el ARSo. Esta metodología conceptualiza las estructuras sociales en términos de patrones perdurables de vínculos entre actores. Por tanto, proporciona un método para estudiar de forma sistemática y empírica las causas y consecuencias de la estructura social.

La mayoría de los analistas de redes sociales también sostienen que las interacciones repetidas entre actores (por ejemplo, individuos) pueden dar lugar a formaciones sociales emergentes (por ejemplo, grupos, organizaciones formales, naciones) que siguen su propia lógica y no pueden reducirse ni explicarse por sus partes constituyentes, incluso aunque siguen dependiendo de esas partes (White, 2008). Los ejemplos de emergencia abundan en el mundo físico, como

la combinación de hidrógeno (H) y oxígeno (O) en agua (H_2O). Como el agua no puede reducirse o explicarse completamente por sus átomos subyacentes, las formaciones sociales emergentes son productos de interacciones entre actores que no pueden explicarse por completo únicamente por las características de esos actores (Smith, 2010, p. 27). En cambio, toman vida propia, lo que significa que los actores de todos los tipos y niveles ya sean individuos, grupos, organizaciones o naciones, pueden ser examinados utilizando el ARSo sin la necesidad de explorar los componentes subyacentes.

1.2.4 Ubicación Estructural: Creencias, Normas, Intenciones, Comportamiento e Identidad

Las formaciones sociales emergentes pueden ejercer lo que se llama causalidad descendente, razón por la cual el ARSo asume que las actitudes, creencias, intenciones, comportamiento e incluso sus identidades de los actores están determinadas principalmente por su ubicación en la estructura social. Por ejemplo, la investigación ha encontrado repetidamente que las personas ubicadas estructuralmente (es decir, socialmente) cerca de un movimiento particular, religioso o de otro tipo, tienen una mayor probabilidad de unirse a ese movimiento que aquellas que no lo están (Lofland & Stark, 1965; McAdam, 1986 ; Snow, Zurcher, & Ekland-olson, 1980; Stark & Bainbridge, 1980). En otras palabras, las personas tienden a no unirse a grupos al azar. En cambio, es más probable que se unan a los grupos en los que conocen a alguien.

1.2.5 Redes sociales dinámicas

Finalmente, el análisis de redes sociales asume que las redes son dinámicas. Siempre están cambiando a medida que los actores entran y salen de la red y a medida que se forman y se disuelven los lazos. Pueden crecer o encogerse de tamaño, volverse más o menos fragmentados o aumentar o disminuir su nivel de centralización. Los grupos de actores pueden agruparse o disolverse iterativamente, y los actores centrales pueden convertirse en actores periféricos (y viceversa). Para complicar las cosas, los actores también se mueven de una ubicación geográfica a otra, lo que podría afectar las operaciones de una red. Históricamente, los datos de redes longitudinales han sido difíciles de obtener y los métodos para examinarlos no han sido desarrollados. De hecho, el texto clásico del ARSo de Wasserman y Faust hace poca mención a las redes longitudinales. Afortunadamente, eso está cambiando. Gran parte del análisis hasta la fecha ha sido principalmente descriptivo, pero en los últimos años los analistas de redes sociales han desarrollado numerosos enfoques para descubrir patrones y procesos a partir de datos longitudinales (Doreian & Stockman, 1997; McCulloh & Carley, 2011; Nooy, 2011; Snijders, Bunt, & Steglich, 2010)

1.2.6 Resumen: Redes Sociales, Agencia Humana y Cultura

La prioridad que da el ARSo a las relaciones entre los actores sobre los atributos individuales conduce inevitablemente a preguntas sobre la agencia humana y la cultura. Específicamente, ¿el ARSo deja espacio para el libre albedrío de los actores y los efectos de la cultura? La respuesta corta es, sí y no. Algunas corrientes académicas dan espacio la cultura y la agencia de los actores, otras no. (Emirbayer & Goodwin, 1994, pp. 1424–1436) han identificado tres paradigmas de redes sociales a los que se adhieren la mayoría de los analistas de redes sociales: determinismo estructural, instrumentalismo estructural y construcciónismo estructural. El determinismo estructural ignora por completo (o descarta) el posible papel causal que juegan las creencias, valores y compromisos de los actores en términos de procesos sociales y cambio histórico. Los primeros escritos de Harrison White (Boorman & White, 1976; White, Boorman, & Breiger, 1976) reflejan este punto de vista, al igual que los del primer teórico de redes, Bruce Mayhew (1980, 1981). Mayhew, de hecho, creía que la conciencia humana era irrelevante para comprender el mundo social.

Por el contrario, el instrumentalismo estructural deja espacio para la acción humana, pero lo enmarca únicamente en términos de elección racional, acción instrumental y maximización de la utilidad. El análisis de Roger Gould de la Comuna de París (1991, 1993) y el trabajo de Nan Lin sobre capital social (Lin, 2001) son ejemplos. Dado que existe una amplia evidencia que respalda la premisa de que los actores responden a los incentivos (ver, por ejemplo, Becker (1976); Iannaccone (1995); Levitt & Dubner (2005)), este enfoque es más comprensivo que el anterior. Sin embargo, esta teoría no es libre críticas.

El construcciónismo estructural, como el instrumentalismo estructural, se toma en serio el papel de la agencia humana. Sin embargo, ve a los actores motivados por preocupaciones adicionales, tales como normas, valores, compromisos culturales e identidades colectivas e individuales. Por ejemplo, David Knoke argumenta, “una comprensión sofisticada de... acción requiere combinar restricciones culturales, racionales y estructurales en especificaciones complejas para problemas sustantivos dados (Knoke, 1990, p. 19). Los estudios de Doug McAdam sobre Freedom Summer (McAdam, 1986, 1988a, 1988b) ilustran este enfoque, al igual que los escritos posteriores de (White, 1992, 2008), que enfatizan la importancia de las narrativas, historias y redes de significado.

Si bien las metodologías de redes sociales tienden a ser agnósticas en cuestiones de estructura, cultura y agencia humana, a menudo ayuda hacer distinciones analíticas entre las influencias culturales, los compromisos normativos y las preocupaciones instrumentales de los actores. Considere, por ejemplo, los problemas relacionados con la recopilación de datos de redes de parentesco. Lo que constituye una red de parentesco en Occidente puede diferir considerablemente de lo que constituye una en algunas culturas de Oriente Medio y Asia. Por lo tanto, si un analista codificara una red de parentesco de Oriente Medio guiada

por supuestos occidentales, los datos resultantes podrían resultar completamente inútiles. Además, dado que los vínculos instrumentales pueden tener efectos diferentes a los vínculos culturales, tiene sentido metodológico explicarlos por separado.

1.3 Términos y Conceptos Básicos

1.3.1 Actores y Enlaces

Como se señaló anteriormente, los actores pueden tener diversas formas y tamaños (por ejemplo, individuos, grupos, organizaciones, denominaciones y naciones). Lo mismo puede decirse de los lazos. Pueden definirse en términos de tipo, fuerza y direccionalidad. Los tipos de vínculos incluyen, pero no se limitan a, vínculos de sentimiento (por ejemplo, amistad, conocido, agrado y desagrado), biología (por ejemplo, cónyuge, hermano, primo), afiliación compartida (por ejemplo, miembros de la misma iglesia, mezquita, sinagoga u organización ecuménica), comunicación (por ejemplo, cadenas de oración, llamadas telefónicas, correo electrónico, tweets), etc. Como se señaló previamente, los lazos pueden variar de fuertes a débiles. Sin embargo, la distinción entre un lazo fuerte y uno débil no siempre es obvio, ni tampoco lo es la distinción entre un lazo débil y los numerosos encuentros aleatorios y usualmente no repetidos que ocurren diariamente (Azarian, 2005, p. 37). Los lazos también pueden tener direccionalidad y capturar el flujo de recursos de un actor a otro, la comunicación entre actores (por ejemplo, el actor A puede enviar correos electrónicos repetidamente al actor B, pero B no corresponde), qué actores buscan consejo de otros actores (Krackhardt, 1992), y así sucesivamente. Finalmente, la mayoría de los actores están conectados a otros por múltiples vínculos; de hecho, uno de los desafíos del ARSo es determinar qué vínculos deben modelarse.

1.3.2 Redes Sociales

Una red social es simplemente “un conjunto finito o conjuntos de actores” ligados entre sí (Wasserman & Faust, 1994, p. 20). La Figura 1.4 muestra una red social hipotética donde los círculos representan actores y las líneas representan vínculos o relaciones. Como ilustra, los actores a menudo se agrupan en subgrupos relativamente distintos. Además, algunos están incrustados en el centro de estos subgrupos, mientras que otros se encuentran más en la periferia, a veces sirven como puentes entre subgrupos.

1.3.3 Camino (y Distancia de Caminos)

Un camino se define como una caminata (es decir, una secuencia de actores y vínculos) en la que ningún actor entre el primer y el último actor de la caminata ocurre más de una vez, mientras que la distancia de la ruta entre dos actores es el número de pasos entre los dos. Por ejemplo, En la Figura 1.4 puede trazar un

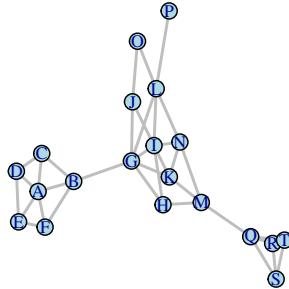


Figure 1.4: Red social hipotética

camino desde los actores R, S, y T al actor M a través del actor Q. La distancia entre los actores es dos (es decir, dos pasos).

Es bastante común que haya numerosos caminos entre actores. La ruta más corta entre dos actores se llama geodésica y la geodésica más larga de una red se considera el diámetro de la red. Se considera que los actores que pueden llegar entre sí a través de una ruta están en el mismo componente. Las redes pueden tener varios componentes y, en tales casos, se considera que están desconectados. El componente más pequeño que puede tener una red es un solo actor sin vínculos con otros actores, que se denominan aislados.

1.3.4 Topología

La topología de una red puede variar. Por ejemplo, los analistas pueden medir el tamaño de una red (la cantidad de actores en una red), su distancia promedio (la longitud promedio de las geodésicas entre todos los actores en una red) y su diámetro. Dos de las medidas topográficas más comunes son la densidad y la centralización. La densidad captura la interconexión de una red y es igual a la relación entre vínculos reales y vínculos posibles. Desafortunadamente, la medida formal de densidad está inversamente relacionada con el tamaño de la red (es decir, si todo lo demás es igual, la densidad de las redes más grandes tiende a ser menor que la densidad de las redes más pequeñas). Por lo tanto, los analistas de redes sociales a menudo recurren a medidas alternativas para capturar la interconexión de una red, como el grado de centralidad promedio, el coeficiente de agrupamiento y la cohesión/fragmentación. La centralización

mide el grado en que una red está centralizada en torno a uno o un puñado de actores. Existen múltiples medidas de centralización, y varían en términos de qué medida de centralidad se utiliza. La métrica de centralización estándar calcula una puntuación de varianza basada en la diferencia entre la puntuación de centralidad de cada actor y la puntuación de centralidad más alta de la red. Una medida alternativa calcula una puntuación de varianza basada en la diferencia entre la puntuación de centralidad de cada actor y la puntuación de centralidad promedio en la red. Para ambas, cuanto mayor es el nivel de varianza, mayor es el grado de centralización.

Las redes densas suelen estar formadas por numerosos vínculos fuertes y muy pocos débiles, mientras que las redes dispersas son todo lo contrario. Las primeras a veces se denominan redes provinciales, mientras que las segundas se denominan cosmopolitas, y la evidencia sugiere que las redes con una densidad entre los dos extremos son más eficaces o resistentes (Pescolindo & Georgianna, 1989; Uzzi, 1996 ; Uzzi & Spiro, 2005). De manera similar, los estudios académicos indican que las redes ni demasiado centralizadas o descentralizadas son más efectivas que las que se encuentran en los extremos. ¿Por qué? Por un lado, un alto nivel de centralización puede proporcionar a las organizaciones un comando y control sustanciales, lo que las hace lentas para adaptarse a un entorno cambiante. Por otro lado, las organizaciones descentralizadas pueden cambiar de rumbo rápidamente, pero pueden carecer del comando y el control necesarios para que todos estén en la misma página.

1.3.5 Subgrupos (Cluster, subredes)

Un enfoque significativo del SNA es identificar subgrupos de actores “entre los cuales hay vínculos relativamente fuertes, directos, intensos y / o positivos” (Wasserman & Faust, 1994, p. 249). Al igual que la topografía de la red, existen numerosos métodos para identificar grupos de actores. Sin embargo, todos asumen que los vínculos entre los miembros del subgrupo deberían ser más numerosos que los vínculos entre los miembros de otros grupos. La Figura 1.5 presenta un grafo donde usamos lo que se conoce como un algoritmo de detección de comunidad para identificar subgrupos dentro de la red. Aunque, aquí, probablemente no necesitábamos un algoritmo para identificar estos subgrupos, a menudo no es tan sencillo, por lo que los algoritmos de agrupación en clústeres pueden ser muy valiosos. Los algoritmos más comunes son componentes (débiles y fuertes), camarillas, núcleos k y algoritmos de detección de comunidades.

1.3.6 Centralidad

La mayoría de las redes sociales contienen personas u organizaciones que son más centrales que otras. Debido a esto, a menudo disfrutan de un mejor acceso a los recursos y están en mejores condiciones para difundir información. Los analistas de redes sociales han desarrollado varias medidas de centralidad, cada una basada en diferentes supuestos de lo que significa que un actor sea central.

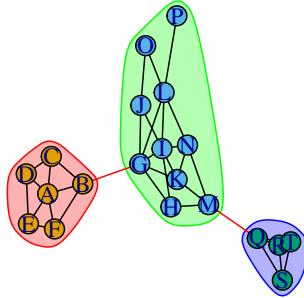


Figure 1.5: Red hipotética con comunidades demarcadas

Cuatro de las medidas más comunes son:

- Grado de centralidad, que es un recuento del número de vínculos de un actor.
- Centralidad de cercanía, que mide, en promedio, qué tan cerca (en términos de distancia del camino) cada actor está de todos los demás actores en una red.
- Centralidad de intermediación, que mide hasta qué punto cada actor se encuentra en el camino más corto entre todos los demás actores de una red.
- Centralidad del eigenvector, que asume que los vínculos con actores muy centrales son más importantes que los vínculos con actores periféricos, por lo que pondera los vínculos de un actor por la centralidad de actores adyacentes.

1.3.7 Puentes y Corredores

Los puentes son lazos que cubren las brechas en una red social, mientras que los corredores son actores que se sientan a un lado de esos puentes. En la Figura 1.4 anterior, los bordes entre el actor Q y M o G y B podrían considerarse puentes, mientras que los cuatro actores serían considerados intermediarios. De hecho, si removemos uno o múltiples de los intermediarios de la red, esta se fragmentaría en redes separadas. En lugar de adivinar qué actores son intermediarios y qué vínculos son puentes, se han desarrollado numerosos algoritmos para detectarlos. En la Figura 1.6, el tamaño del nodo refleja la medida de huecos estructurales de

Ron Burt (1992), quizás la medida de intermediación más conocida, y el grosor de los lazos refleja la centralidad de intermediación de los bordes, que algunos utilizan para identificar puentes entre subgrupos. Cuanto más grande es el nodo, más potencial de intermediación posee, mientras que cuanto más grueso es el lazo, más probable es que sea un puente. Aquí podemos ver que los actores G y M tienen un potencial de intermediación considerable. El vínculo entre los actores B y G parece más crucial que entre M y Q. El actor L también está en una posición de intermediación, principalmente debido a sus vínculos con el actor P.

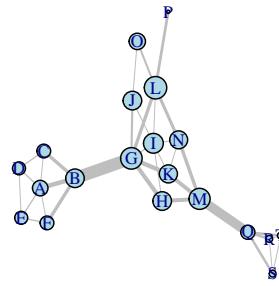


Figure 1.6: Red hipotética, bordes ponderados por intermediación y nodos dimensionados por autonomía.

1.3.8 Funciones y Posiciones

Los analistas de redes sociales generalmente analizan los datos de la red de una de dos maneras:

1. Un enfoque de conectividad social o relacional o;
2. Un enfoque posicional.

El primero se centra en los vínculos directos e indirectos entre actores y busca explicar el comportamiento y los procesos sociales a la luz de esos vínculos. De manera opuesta, el segundo enfoque tiene como objetivo identificar actores estructuralmente equivalentes, es decir, actores que pueden o no tener vínculos entre sí pero que ocupan posiciones similares dentro de una red (por ejemplo, comandantes de diferentes batallones). Este enfoque asume que es probable que los actores estructuralmente equivalentes se comporten de manera similar,

independientemente de si existe un vínculo entre ellos. Los analistas han desarrollado varios algoritmos para identificar actores estructurales equivalentes (por ejemplo, equivalencia estructural, equivalencia automórfica, equivalencia regular), y existen múltiples variaciones de estos.

1.3.9 Análisis Exploratorio y Explicativo de Redes Sociales

Podemos clasificar los métodos de redes sociales en dos amplias categorías: exploratorias y explicativas. Los métodos exploratorios pueden ayudar a los analistas a identificar actores centrales, resaltar aspectos sobre la estructura de una red, ubicar subgrupos, iluminar intermediarios y puentes, y encontrar actores estructuralmente equivalentes. Aunque estos métodos son bastante útiles, solo nos ayudan a describir los datos de la red. No nos permiten identificar las causas subyacentes de los patrones y tendencias observadas. Por el contrario, los métodos explicativos se centran más en verificar los hallazgos que surgen de los análisis exploratorios y cualitativos. Una forma de pensar lo es que los análisis exploratorios pueden ayudarnos a desarrollar hipótesis, mientras que los enfoques explicativos, como QAP, ERGM y SAOM, pueden ayudarnos a probarlos.

Debido a que los datos de las redes sociales rara vez son aleatorios y “violan” (por definición) el supuesto de independencia observacional, los analistas de redes sociales han desarrollado varios modelos que toman en cuenta las características únicas de los datos de redes sociales. Los tres más comunes son el procedimiento de asignación cuadrática (QAP en inglés), los modelos de gráficos aleatorios exponenciales (ERGM en inglés) y los modelos estocásticos orientados al actor (SAOM en inglés). QAP es similar a los modelos de regresión estándar, excepto que utiliza un método análogo al bootstrapping para estimar los errores estándar y la significancia estadística. Los ERGM y SAOM son únicos en el sentido de que examinan cómo una red social observada es una función de patrones subyacentes de vínculos (por ejemplo, diádicas, triángulos abiertos y cerrados, etc.). Los ERGM se utilizan generalmente para una red social estática, mientras que los SAOM están diseñados para datos longitudinales.

1.3.10 Atributos

Si bien el análisis de redes sociales se centra principalmente en el patrón de vínculos entre actores, la mayoría de los analistas de redes sociales no ignoran por completo los datos de atributos. Si los actores de una red son individuos, entonces los datos de atributos incluyen características como género, raza, etnia, años de educación, nivel de ingresos, edad, etc. Si son organizaciones, las variables de atributo pueden ser aquellas que indiquen el tamaño (por ejemplo, número de empleados), edad, tipo, etc. Y si los actores son países, las variables de atributo incluirían medidas como el PIB per cápita o el tamaño de la población.

A veces, el límite entre atributos y afiliaciones puede ser algo difuso. Como regla general, algo es una afiliación si la participación compartida de dos actores

indica una relación (por ejemplo, co-membresía en una organización voluntaria). Aún así, una afiliación también puede funcionar como un atributo.

1.4 Resumen y Conclusiones

Este capítulo ha proporcionado una descripción general del ARSO, cómo se diferencia de otros enfoques analíticos, los términos y conceptos básicos que emplea y los supuestos fundamentales que subyacen en gran parte del trabajo de los analistas de redes. Este capítulo también exploró los problemas relacionados con la cultura y la agencia humana y el grado en que influyen en el comportamiento humano.

Part II

ARSo: Tecnicas

Chapter 2

Trabajando con Gephi

2.1 Introducción

[El ARSo] es una colección de teorías y métodos que suponen que el comportamiento de los actores (por ejemplo, individuos, grupos u organizaciones) se ve profundamente afectado por sus vínculos con otros y las redes en las que están incrustados – (Everton, 2012a)

Como puede ver en el pasaje y en el capítulo previo, ARSo se basa en algoritmos gráficos y matemáticos complejos. Esto no quiere decir solamente “software,” puesto que las semillas de este campo de estudio empezaron a madurar mucho antes del desarrollo de computadoras modernas.

Hoy en día existen múltiples paquetes de software desarrollados para asistir a los analistas de ARSo en sus iniciativas laborales o académicas. Por consiguiente, sería ilógico presentarle métodos para trabajar con datos de redes sociales “a mano.” Cabe recalcar que el enfoque de este manual no es darle una presentación completa de las herramientas de software disponibles en este campo, para ello consulte (Scott & Carrington, 2011, Capítulo 38) o (Cunningham, Everton, & Murphy, 2016, Apéndice 3). Nuestro objetivo es mucho más limitado, presentarle métodos para grabar, manipular y visualizar datos de redes y asimismo presentarle la aplicación de herramienta de análisis en la plataforma de software Gephi.

2.2 Gephi

De acuerdo con los autores de Gephi, la plataforma es “una herramienta para analistas de datos y científicos interesados en explorar y comprender gráficos … [Donde] el usuario interactúa con la representación, manipula las estructuras, formas y colores para revelar patrones ocultos. El objetivo es ayudar a los

analistas de datos a formular hipótesis, descubrir patrones intuitivamente, aislar singularidades o fallas de estructura durante la recopilación de datos” (Bastian, Heymann, & Jacomy, 2009). Para nuestros propósitos utilizaremos esta interfaz para complementar el material y facilitar nuestro análisis exploratorio. Antes de comenzar los siguientes capítulos, descargue gratuitamente la herramienta desde el sitio web de Gephi.

Como todas las herramientas de software, esta tiene ventajas y desventajas. Es importante explicar por qué hemos elegido esta herramienta sobre algunas de las demás:

1. Es una herramienta de fuente abierta, por consiguiente, podemos descargar y utilizarla libre de costo.
2. Incluye una gran variedad de herramientas de visualización, algoritmos gráficos y matemáticos complejos, capacidades para ingerir y manipular varios formatos de datos y es extensible utilizando docenas de complementos creados por la comunidad abierta de usuarios que amplían las funcionalidades del software.
3. Existe una gran comunidad de usuarios y desarrolladores quienes activamente contribuyen al crecimiento de la plataforma, manuales de usuario, tutoriales y publican artículos utilizando esta herramienta (por ejemplo, Chaudhary, Mittal, & Arora (2019)).
4. La interfaz de Gephi se puede configurar en español.

En las siguientes secciones continuaremos introduciendo este programa, sus interfaces, características, capacidades de visualización, fortalezas, debilidades, etc.

2.3 Gephi Para Principiantes

Las siguientes subsecciones le brindaran una breve introducción a la plataforma Gephi y sus tres secciones principales. Esta no es la única vez que cubriremos algunos de los conceptos básicos de Gephi; como tal, esta sección está diseñada para ser una descripción general de alto nivel.

2.3.1 Antes de Empezar

Primero, descargue e instale el software en su computadora. Como se había mencionado previamente el software es gratuito y distribuido abiertamente. Gephi es una aplicación multiplataforma desarrollada en Java, por lo que requiere una instalación previa de Java, la cual puede obtener libremente en línea. Para computadoras Mac OS X, Java se incluye con la aplicación, por lo que no es necesario instalarlo por separado. En Windows y Linux, el sistema debe estar equipado con Java.

Si bien no hay requisitos mínimos de hardware, el rendimiento es una función de

la memoria. Además, se requiere una tarjeta gráfica compatible para ejecutar Gephi.

2.3.2 Inicie Gephi

Empiece por localizar y abrir el programa. Lo primero que vera es una ventanilla de Bienvenida, vea la Figura 2.1.

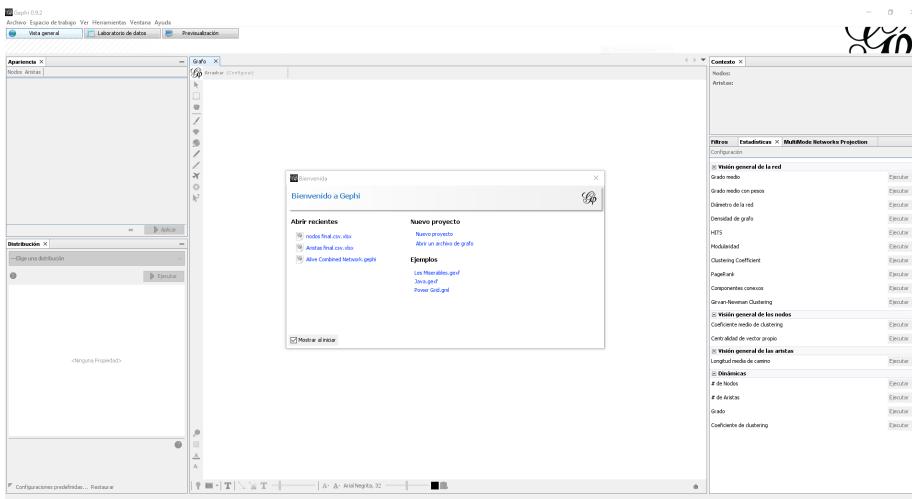


Figure 2.1: Menú de Bienvenida en Gephi

Para esta breve instrucción al software, vamos a usar uno de los conjuntos de datos proporcionados por el software, *Les Misrables.gexf* (Knuth, 1993). Localice este en la parte inferior derecha de la ventanilla de bienvenida. Al seleccionar esta opción aparecerá una nueva ventanilla, subsecuentemente explicaremos más acerca de las opciones en este espacio, por ahora seleccione el botón *Aceptar*. ¡Listo! Con el conjunto de datos cargado (Figura Figura 2.2), podemos empezar a explorar Gephi.

2.3.3 Vista General

Esta es la ventana que utilizará más a menudo, ver Figura 2.3. Aquí podrá tomar diferentes acciones para manipular la apariencia sus gráficos, filtrar información y calcular medidas de red, arista o nodo.

La característica más obvia de esta ventana es el gráfico prominente en el centro; sin embargo, hay varias ventanillas relevantes; aquí, las cubriremos en orden alfabético.

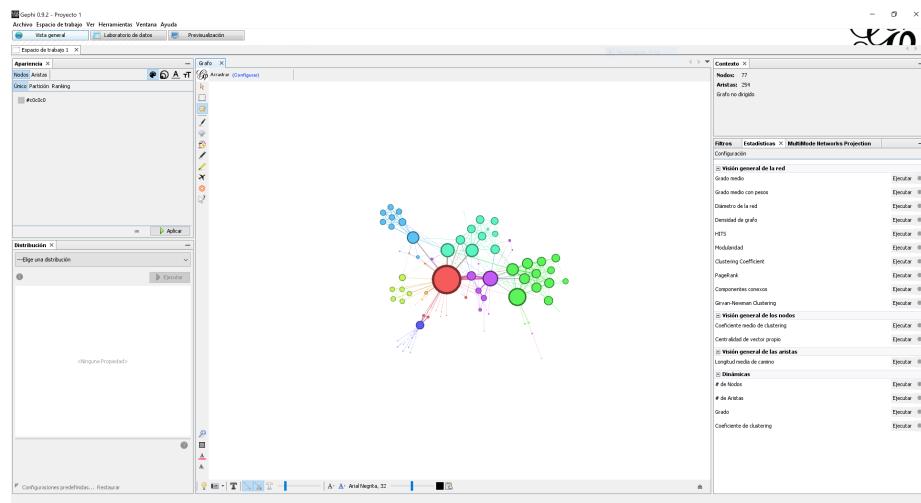


Figure 2.2: Datos Les Misrables Cargados

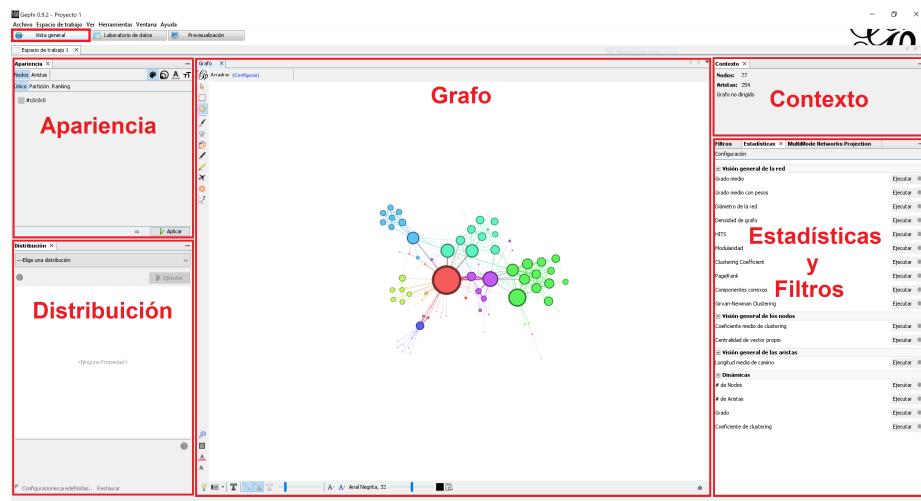


Figure 2.3: Gephi: Vista General

2.3.3.1 Apariencia

El cuadro de apariencia le permite cambiar las características visuales de los nodos, aristas y etiquetas (Ognyanova, 2016). Las opciones se bifurcan para trabajar en la apariencia de nodos y aristas, accesibles a través de las pestañas correspondientes (*Nodos* y *Aristas*). Para editar nodos o aristas se le presenta una serie de opciones en el lado derecho de este cuadro, divididas de tal manera (ver Figura 2.4):

- *Nodos*:
 1. *Color*: Modificar el color de los nodos
 2. *Tamaño*: Modificar el tamaño de cada nodo
 3. *Color de etiqueta*: Modificar el color de las etiquetas de los nodos
 4. *Tamaño de etiqueta*: Modificar el color de las etiquetas de los nodos
- *Aristas*:
 5. *Color*: Modificar el color de las aristas
 6. *Color de etiqueta*: Modificar el color de las etiquetas de las aristas
 7. *Tamaño de etiqueta*: Modificar el tamaño de las etiquetas de las aristas

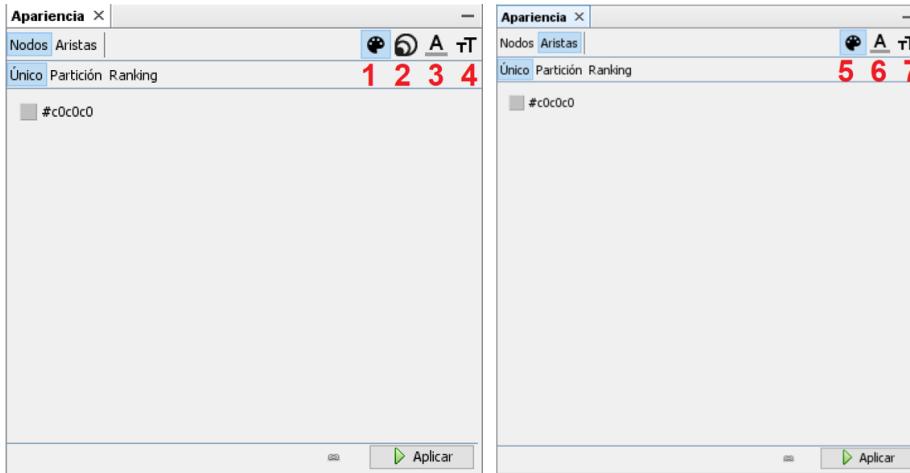


Figure 2.4: Gephi: Apariencia

Cada sección contendrá de dos a tres pestañas: *Único*, *Partición* y *Ranking*. Veamos como funcionan utilizando el software:

*Nodos > Color > Único > Clic sostenido sobre el cuadro de color > Seleccionar color > Aplicar
 Aristas > Color > Único > Clic sostenido sobre el cuadro de color > Seleccionar color > Aplicar*

Único cambia los elementos de manera uniforme por color o tamaño seleccionado. Por consiguiente, para modificar el color de los nodos en gráfico, proceda a seleccionar *Nodos > Color > Único > Clic sostenido sobre el cuadro de color > Seleccionar color > Aplicar*. Si desea modificar el color de las aristas para que reflejen un valor único, seleccione *Aristas > Color > Único > Clic sostenido sobre el cuadro de color > Seleccionar color > Aplicar*. En la Figura 2.5 los nodos han sido coloreados de negro (#000000) y las aristas de gris (#adadad). Note que Gephi define colores utilizando una combinación hexadecimial de números y letras definida por su mezcla de rojo, verde y azul.

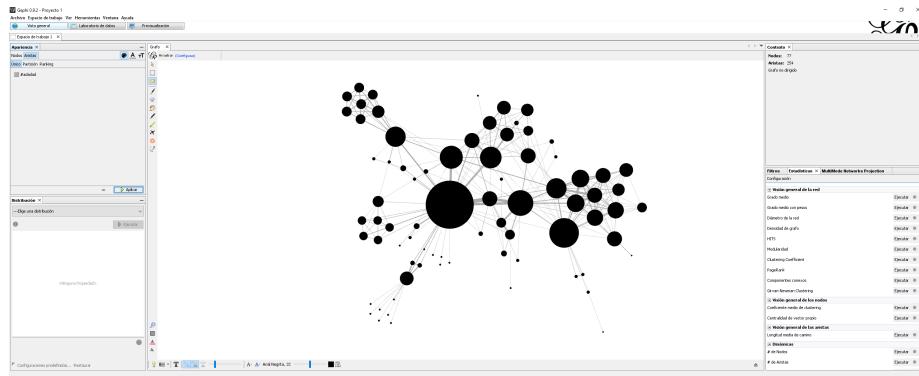


Figure 2.5: Nodos y Aristas con Colores Únicos

Nodos > Tamaño > Único > 20 > Aplicar

Para modificar el tamaño de los nodos, igualmente usaremos la opción *Único*. Sin embargo, en esta secuencia modificaremos los pasos de la siguiente manera: *Nodos > Tamaño > Único > 20 > Aplicar*. Como verá en la Figura 2.6 todos los nodos en el gráfico aparecen con el mismo tamaño.

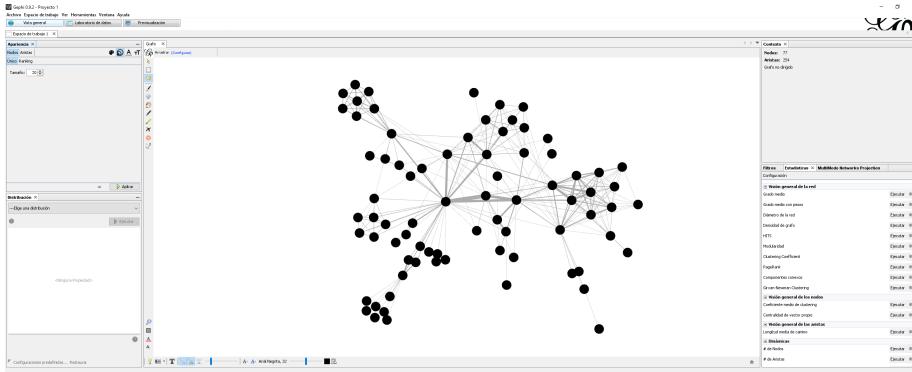


Figure 2.6: Nodos con Tamaño Único

Nodos > Color > Partición > -Escoge un atributo > Modularity Class > Aplicar

La *Partición* cambia todos los elementos visibles utilizando valores de atributos categóricos. Por ejemplo, el conjunto de datos que seleccionamos tiene un atributo categórico, **modularidad**, que representa el subgrupo al que pertenece cada nodo. Usando este atributo podemos cambiar el color de los nodos seleccionando *Nodos > Color > Partición > -Escoge un atributo > Modularity Class > Aplicar*. Una vez hecho, los colores de los nodos deberían permitirle entender qué nodos pertenecen a las misma comunidades. De manera similar, podríamos intentar cambiar los colores de las aristas basándonos en un atributo categórico; sin embargo, el conjunto de datos de ejemplo no incluye dicho tipo de variable.

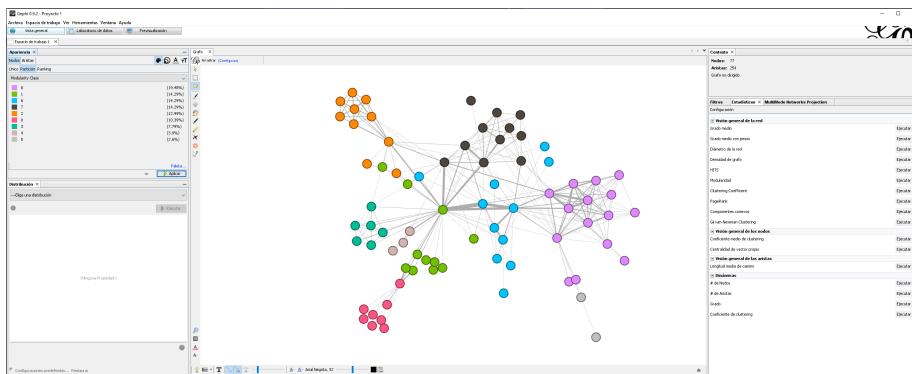


Figure 2.7: Nodos Coloreados por Atributos Categóricos.

- Nodos* > El *Ranking* cambia todos los elementos visibles utilizando valores de atributos continuos. Por ejemplo, el conjunto de datos *Tamaño* que seleccionamos tiene un atributo continuo, *grado*, que representa el numero de enlaces que tiene cada nodo. Usando *Ranking* > este atributo podemos cambiar el tamaño de los nodos seleccionando *Nodos > Tamaño > Ranking > -Escoge un atributo > Grado > Aplicar*. Ahora deberá ver que algunos nodos aparecen mucho más grandes que otros, esto se debe a que algunos tienen más conexiones. Más adelante en el manual ampliaremos el significado de esta medida y presentaremos otras. Quizás se este preguntando, ¿es posible cambiar el grosor de las aristas en función de una variable continua? La respuesta es sí, pero hacerlo requiere un menú diferente, por lo que abordaremos esta acción más adelante.

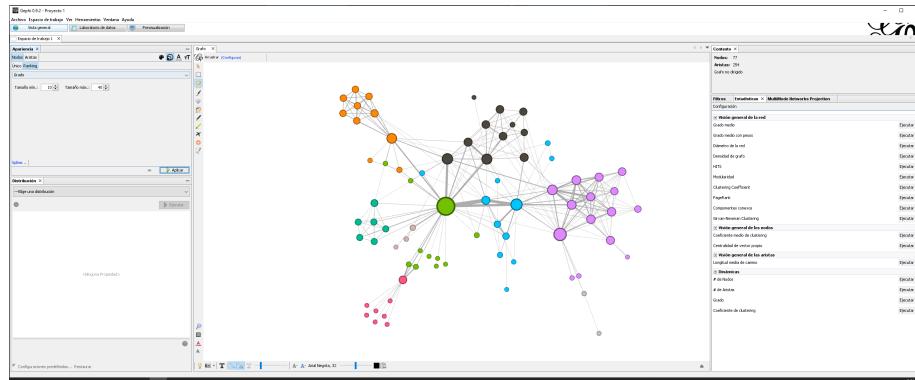


Figure 2.8: Nodos Dimensionados por Variable Continua.

2.3.3.2 Contexto

Esta sección proporciona información sobre las propiedades básicas del gráfico; específicamente, el número de nodos y aristas y si el gráfico es dirigido o no.

2.3.3.3 Distribución

ARSo depende en gran medida de la exploración de datos a través de buena visualización. Esta es una de las características clave de Gephi, contiene una buena cantidad de algoritmos de diseño de gráficos y los ejecuta rápido (todo dependiendo de la potencia de su procesador). La ventana *Distribución* le permite al usuario seleccionar entre una serie de algoritmos de diseño de gráficos y modificar los parámetros preestablecidos según sea necesario. En esta sección exploraremos algunos de los algoritmos gráficos más notables incluidos con la instalación base de Gephi.

Elije una distribución
 > *Force Atlas*
 > *Ejecutar*

Force Atlas y *Force Atlas 2* son algoritmos de diseño dirigidos por fuerza, producidos por los desarrolladores de Gephi, que permiten visualizar efectivamente gráficos de tamaño pequeño a mediano (Jacomy, 2011). Estos son algoritmos continuos, es decir que le permiten al analista editar al gráfico mientras el algoritmo dibuja. Además de *Ejecutar* el algoritmo hay que seleccionar *Parar* cuando uno este satisfecho con el diseño. Veamos como funciona este algoritmo, para ello en la sección de *Distribución* seleccione *Elije una distribución > Force Atlas > Ejecutar*. Después de un par de minutos, recuerde en dar clic sobre el botón *Parar*. Como puede ver, el algoritmo a contraído los nodos. Esta es una característica de los algoritmos dirigidos por fuerza donde los nodos vinculados se atraen entre sí y los nodos no vinculados se repelen.

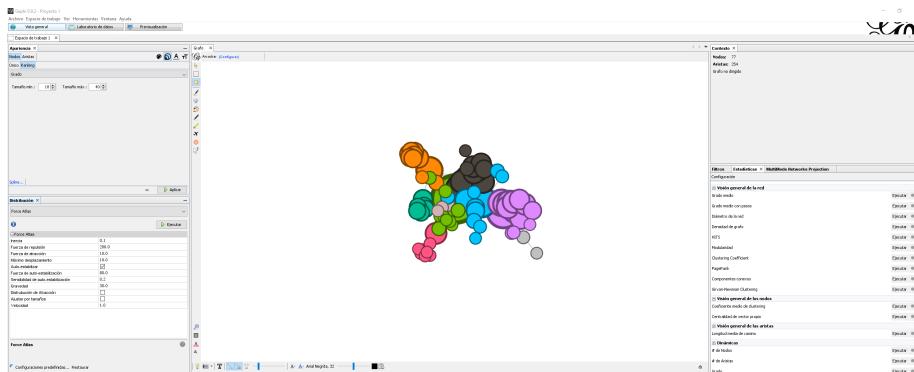


Figure 2.9: Gráfico dibujado con Force Atlas

Elije una distribución
 > *Expansión*
 > *Ejecutar*
Elije una distribución
 > *Contracción*
 > *Ejecutar*

Uno de los algoritmos de diseño más útiles es *Expansión*, el cual aumenta el tamaño de la red pero mantiene la ubicación relativa de los nodos constante. El polo opuesto de esta opción es *Contracción*, que permite compactar la red manteniendo la ubicación relativa de los nodos. Ambas opciones son ejecutadas de manera instantánea, por consiguiente, no hay que detener el algoritmo.

Por ejemplo, la Figura 2.10 es una versión expandida de a Figura 2.9. Podría decirse que es más fácil observar los patrones en este gráfico editado. Para expandir el gráfico seleccione *Elije una distribución > Expansión > Ejecutar*. De clic en *Ejecutar* un par de veces, si sobrepasa el punto deseado seleccione *Elije una distribución > Contracción > Ejecutar* para contraer el gráfico.

Por ahora, solo cubriremos estas tres opciones de diseño. Más adelante utilizaremos algoritmos de diseño de manera orgánica. Le recomendamos que juegue con

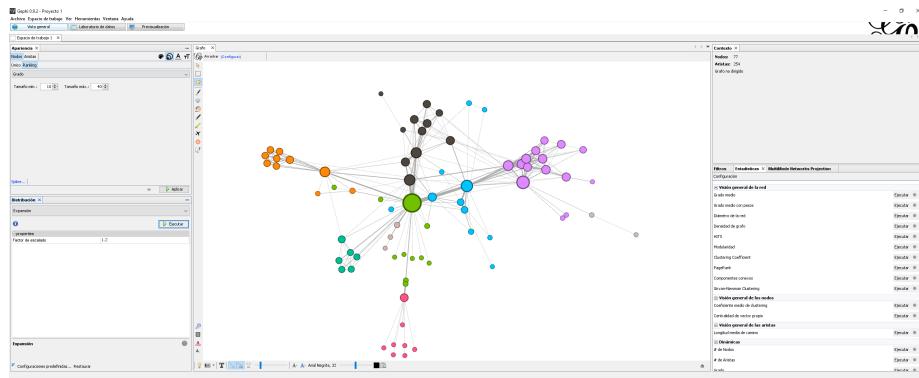


Figure 2.10: Gráfico dibujado con Force Atlas y Expandido

otras opciones, eche un vistazo a cómo se ven los gráficos resultantes, pero tenga en cuenta que algunos algoritmos no se detendrán por sí mismos, por lo que es posible que deba detenerlos manualmente.

2.3.3.4 Estadísticas

Además de utilizar visualizaciones, los analistas de ARSo dependen a menudo de medidas estadísticas utilizadas para comprender las características estructurales de la red a fondo. Gephi proporciona una cantidad limitada, pero suficiente de medidas estadísticas al descargar el software. La pestaña de *Estadísticas* sirve como ubicación principal para encontrar las medidas. Localícela a mano derecha del *Grafo*, es posible que tenga que seleccionar la pestaña ya que comparte el mismo espacio con *Filtros*.

Visión general de la red > Grado medio > Ejecutar Para ejecutar una medida, seleccione el botón *Ejecutar* a un lado de dicha opción. En capítulos posteriores explicaremos la gran mayoría de las medidas presentes, por ahora, visitemos la medida de grado. Esta se refiere al número de aristas adjuntas al nodo. Para calcular esta medida, seleccione *Visión general de la red > Grado medio > Ejecutar*.

Esta selección debe generar un informe HTML que muestre la distribución de grados del gráfico, consulte la Figura 2.11. Además de este reporte, esta acción agregará valores de centralidad de grado a cada nodo en la tabla de aristas situada en el *Laboratorio de Datos*. Proceda a cerrar este cuadro de diálogo, si desea volver a abrirlo repita los pasos previos.

2.3.3.5 Filtros

A menudo tendremos que modificar nuestros datos en Gephi, dentro de la *Vista General* encontrara la pestaña de *Filtros* a un costado del *Grafo*. Esta contiene una variedad de parámetros para filtrar los datos presentados con base

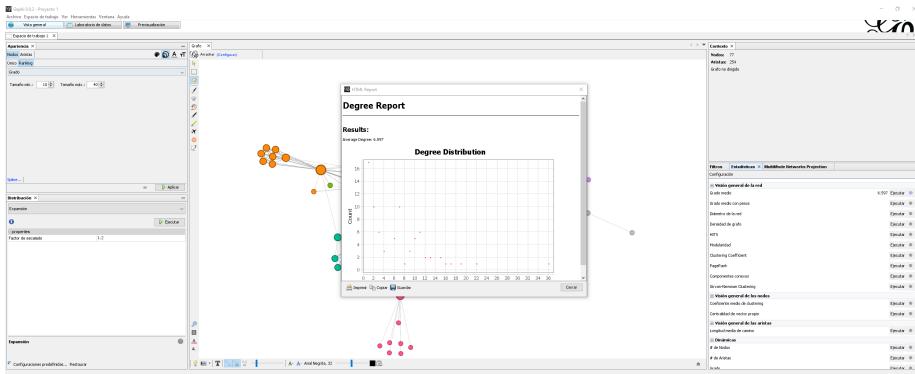


Figure 2.11: Ejecutando Estadística de Grado

a características (variables continuas o categóricas) de los nodos o las aristas. Puede aplicar uno o varios filtros a un gráfico para descubrir características ocultas o refinar la cantidad de información que se muestra en un grafo.

Arrastrar a Consultas > Rango (Grado) Configuración = 2-36 > Filtrar

Para filtrar el grafo con base al número de enlaces, por ejemplo, remover los nodos pendientes, vaya a *Biblioteca* de filtros localice *Atributos > Rango > Grado*, arrastre el filtro de rango al cuadro de la sección de *Consultas*. Al agregar el filtro en *Consultas* notará que en la parte inferior del cuadro aparecerá un rango de 1 a 36. Modifique el rango para que el número mínimo de enlaces sea 2 y su rango lea de 2 a 36. Es decir, filtraremos los nodos de manera que solo aquellos con 2 o más enlaces permanezcan en la red. Por último, de clic en el botón *Filtrar*. Note los cambios en las ventanillas de *Grafo* y *Contexto*, deberá ver un menor número de nodos y aristas.
 Si desea remover el filtro, presione el botón *Auto aplicar* en la parte inferior derecha (que solía leer *Filtrar*). Al hacer esto, sus datos volverán a aparecer.

2.3.3.6 Grafo

La última sección que cubriremos en esta introducción a la *Vista General* es el *Grafo* y algunas de las funciones asociadas con esta ventana. En el lado izquierdo e inferior del panel, encontrará las herramientas básicas para interactuar manualmente con el gráfico y editar algunos de los atributos visuales. En esta breve introducción, los definiremos de manera general, pero no entraremos en demasiados detalles, ya que los usaremos más adelante en el manual según sea necesario.

A continuación se presenta una lista de nombres de comandos y funciones en orden secuencial, ver Figura 2.13:

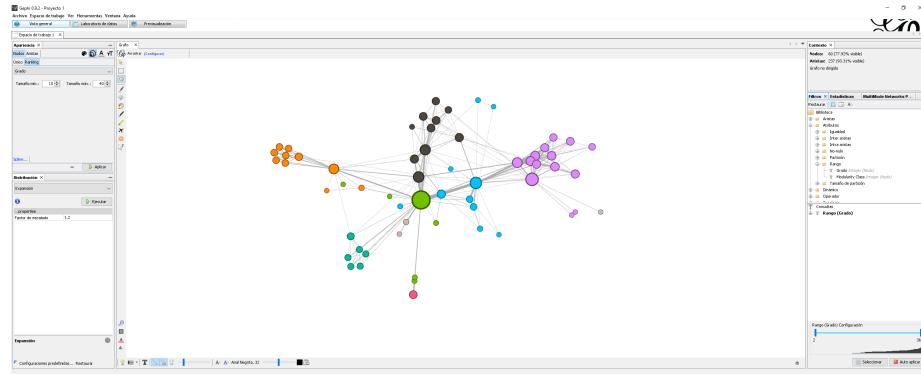


Figure 2.12: Filtrando por Grado

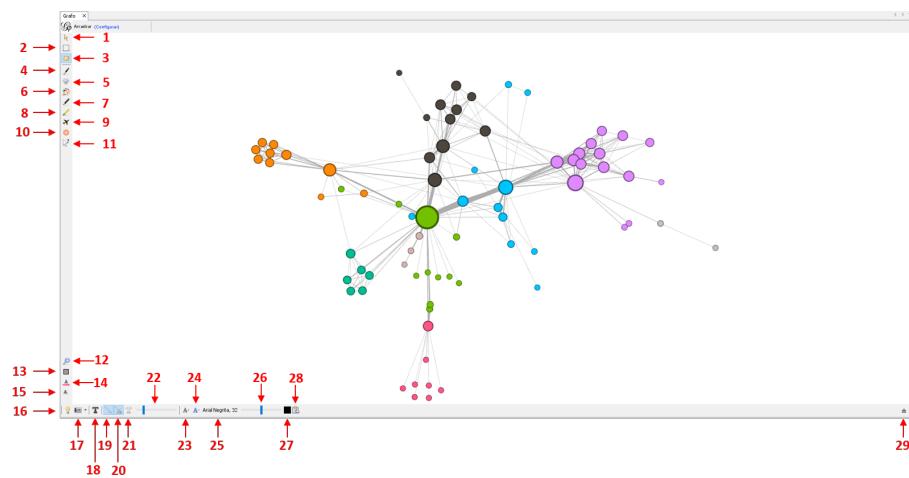


Figure 2.13: Ventanilla de Grafo

1. *Selección directa* – elegir un nodo a la vez
2. *Selección rectangular* – elegir uno o múltiples nodos en un área rectangular
3. *Desplazar* – moverse sin seleccionar
4. *Pincel* – colorea los nodos presionando el botón derecho del ratón
5. *Dar tamaño* – cambia el tamaño de los nodos presionando el botón derecho y arrastrando el ratón verticalmente
6. *Brocha* – como el pincel, da color a los nodos, pero difunde el color a los vecinos
7. *Lápiz de aristas* – añade una nueva arista al grafo
8. *Camino más corto* – muestra el camino más corto entre dos nodos seleccionados
9. *Mapa de calor* – establece la intensidad del color con base a distancia de los vecinos
10. *Editor* – editar atributos del nodo
11. *Centrar el grafo* – retorna el grafo al centro
12. *Reestablecer colores* - devolver los colores a la configuración predeterminada
13. *Reestablecer colores de etiquetas* - devolver el color de las etiquetas a la configuración predeterminada
14. *Reestablecer visibilidad de las etiquetas* - reestablecer la visibilidad de etiquetas
15. *Color de fondo* – cambiar el color de fondo
16. *Captura de pantalla* – capturar y guardar el gráfico en pantalla
17. *Mostrar etiquetas de los nodos* – mostrar las etiquetas correspondientes de cada nodo
18. *Mostrar aristas* – encender o apagar las aristas
19. *Las aristas tienen el color del nodo origen*
20. *Mostrar etiquetas de las aristas*
21. *Escala del peso de las aristas* – incrementar o disminuir el grosor de las aristas
22. *Modo de tamaño* – seleccionar modo de tamaño para las etiquetas
23. *Modo de color* – seleccionar modo de color para las etiquetas
24. *Fuente* – seleccionar fuente para las etiquetas
25. *Escala de la fuente* – incrementar o reducir el tamaño de las etiquetas
26. *Color por defecto* – elegir el color de las etiquetas
27. *Atributos* – configurar los atributos de las etiquetas para nodos o aristas
28. *Expandir* – añadir opciones para editar la apariencia *Global*, de *Aristas*, o *Etiquetas*.

2.3.4 Laboratorio de Datos

El *Laboratorio de datos* contiene tablas de los datos de nodos (Figura 2.14 izquierda) y aristas (Figura 2.14 derecha). Estas tablas permiten a los analistas interactuar con la información y editar valores directamente o mediante la serie de comandos en la parte inferior de la pantalla (Figura 2.14 abajo). Esta es una descripción muy simplificada de las capacidades de esta sección, nos centraremos en las características clave de alto nivel.

Figure 2.14: Laboratorio de datos, a la izquierda la Tabla de Nodos y a la derecha la Tabla de Aristas

2.3.4.1 Tabla de Datos: Nodos

La tabla de nodos contiene un identificador único para cada nodo en el gráfico (generalmente ubicado en la columna más a la izquierda nombrada **Id**). Además, note la columna **Label** directamente a la derecha, esta es digna de mención ya que sirve como receptáculo para las etiquetas de los nodos, es decir, los valores en esta columna aparecerán como etiqueta de nodos en el gráfico cuando se especifique.

	Id	Label	Interval	ModuloEnClas	Sols
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					

Figure 2.15: Laboratorio de datos, Tabla de Nodos

A la medida que se agregan atributos de nodo al gráfico, en esta tabla deberán aparecer nuevas columnas con estos valores. De manera similar, dentro de Gephi, los analistas a menudo ejecutan diferentes métricas a nivel nodo, estos valores se agregarán a cada registro (fila) en una nueva variable (columna).

2.3.4.2 Tabla de Datos: Aristas

Al igual que la tabla de nodos, la tabla de aristas incluye información clave del gráfico de red; específicamente, los datos relacionales. Esta información está codificada en siete columnas definidas a continuación:

- **Origen y Destino:** Todas las relaciones en un grafo vinculan un nodo inicial y un nodo final. Por ejemplo, si el nodo A está vinculado al nodo B, dibujaríamos un vínculo entre estos dos. Gephi codifica esto en dos columnas **Origen** y **Destino**, por lo que si A y B están conectados, pondríamos A en la columna **Origen** y B en la columna **Destino** o viceversa.
- **Tipo:** Denota el tipo de relación entre los nodos en las columnas **Origen** y **Destino**. Los valores pueden ser **Dirigida** o **No dirigida**. Si el tipo es **Dirigida** la relación es un arco o arista, de lo contrario, si es tipo es **No Dirigida** el vínculo es un borde.
- **Id:** Un identificador único para el vínculo. Usualmente asignado por Gephi al importar datos.
- **Label:** Etiqueta del vínculo.
- **Interval:** Para vínculos en una secuencia, esta es la columna donde encontrará los valores cronológicos.
- **Weight:** Denota el peso para vínculos ponderados.

Origen	Destino	Tipo	Id	Label	Interval	Filtre	Weight	Origen
1	destino	No dirigida	1				1,5	
2	0	No dirigida	2				8,5	
3	0	No dirigida	3				10,0	
4	2	No dirigida	4				8,5	
5	0	No dirigida	5				1,5	
6	0	No dirigida	6				1,5	
7	0	No dirigida	7				1,5	
8	0	No dirigida	8				1,5	
9	0	No dirigida	9				1,5	
10	0	No dirigida	10				1,5	
11	0	No dirigida	11				1,5	
12	0	No dirigida	12				1,5	
13	0	No dirigida	13				1,5	
14	0	No dirigida	14				1,5	
15	0	No dirigida	15				1,5	
16	0	No dirigida	16				1,5	
17	0	No dirigida	17				1,5	
18	0	No dirigida	18				1,5	
19	0	No dirigida	19				1,5	
20	0	No dirigida	20				1,5	
21	0	No dirigida	21				1,5	
22	0	No dirigida	22				1,5	
23	0	No dirigida	23				1,5	
24	0	No dirigida	24				1,5	
25	0	No dirigida	25				1,5	
26	0	No dirigida	26				1,5	
27	0	No dirigida	27				1,5	
28	0	No dirigida	28				1,5	
29	0	No dirigida	29				1,5	
30	0	No dirigida	30				1,5	
31	0	No dirigida	31				1,5	
32	0	No dirigida	32				5,0	
33	0	No dirigida	33				5,0	
34	0	No dirigida	34				5,0	
35	0	No dirigida	35				5,0	
36	0	No dirigida	36				5,0	
37	0	No dirigida	37				5,0	
38	0	No dirigida	38				4,0	
39	0	No dirigida	39				4,0	
40	0	No dirigida	40				4,0	
41	0	No dirigida	41				4,0	
42	0	No dirigida	42				4,0	
43	0	No dirigida	43				4,0	
44	0	No dirigida	44				4,0	
45	0	No dirigida	45				4,0	
46	0	No dirigida	46				4,0	

Figure 2.16: Laboratorio de datos, Tabla de Aristas

2.3.4.3 Menú Inferior

Esta última sección permite a los usuarios agregar, borrar o eliminar columnas, fusionar valores de varias columnas, copiar datos de una columna a otra, entre otras opciones. A lo largo de este manual, ampliaremos el uso de estos comandos según sea necesario.



Figure 2.17: Laboratorio de datos, Opciones

2.3.5 Previsualización

Previsualización es donde se refina la visualización para fines de exportación. Esta sección es diferente de la *Vista General* en que uno no puede interactuar dinámicamente con los datos a través de filtros, colores o diseño. El valor agregado de esta sección es enfocar la atención del analista en generar gráficos para publicación y por ello las opciones son limitadas.

Aquí nos centraremos en 4 secciones clave (indicadas en la Figura 2.18):

- Configuración de Previsualización (1) - Proporciona un conjunto de configuraciones de gráficos preestablecidos y permite al usuario guardar los ajustes preestablecidos para uso futuro.
- Parámetros (2) – Ajusta a la configuración de visualización como:
 - Color de nodo, tamaño, bordes, etiquetas, etc.
 - Modifica los bordes, ajusta el ancho, el color, la curvatura, las etiquetas, las flechas, etc.
- Opciones Adicionales (3) – Ajusta las propiedades para la previsualización, exporta el gráfico y aplica las ediciones de la sección *Parámetros*.
- Ventanilla de Previsualización (4) – Proporciona una vista del gráfico terminado.

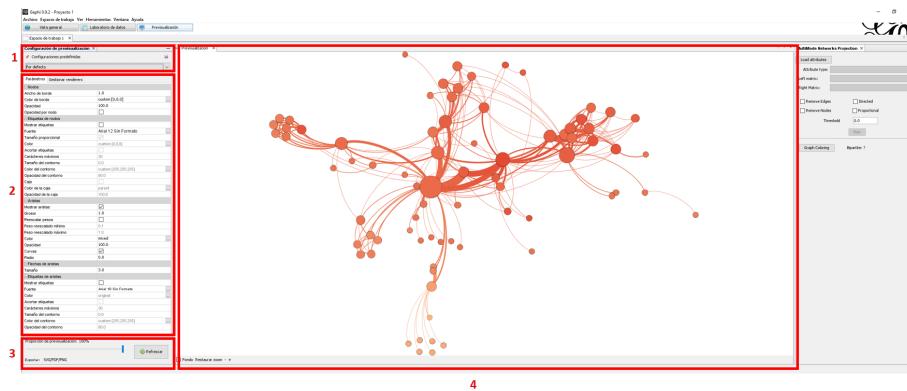


Figure 2.18: Previsualización

2.3.6 Errores Comunes

Es normal que los usuarios de este software se encuentren con problemas conocidos. Esta sección incluye una lista breve (y creciente) de problemas notificados que pueden crear un poco de confusión. Sin embargo, Gephi es un software de código abierto, por lo que puede encontrar errores desconocidos. Cuando esto suceda, infórmelo a los desarrolladores después de leer las reglas de la comunidad e investigar el problema.

Esta es una lista de errores conocidos:

- Intenta abrir Gephi pero se encuentra con el siguiente error: “Error: Cannot find Java 1.8 or higher”:
 1. Primero cheque que ha instalado Java 1.8 o posterior. Si no lo ha hecho, ese es el primer problema.
 2. Si ha instalado Java, es probable (pero no necesario) que este viendo el error en una computadora con Windows. Vaya a la carpeta de Gephi (en Windows cheque `C:\Program Files (x86)\Gephi-0.9.1\bin\`), donde vera dos versiones probablemente tituladas `gephi.exe` y `gephi64.exe`. Intente abrir una a la vez. Una deberá funcionar, cree un atajo a este programa.
- Ha desaparecido el gráfico/ventanilla/pestaña: En Gephi hay una gran cantidad de pequeñas ventanas y pestañas y todo se puede mover, minimizar y cerrar. Si por accidente cierra una ventana, vaya a el menú Ventana, de clic an la opción *Restablecer Ventanas* que reestablecerá todas las pestañas de manera predeterminada.

Chapter 3

Tipos de Datos de Redes Sociales

Tradicionalmente los analistas de redes sociales trabajan con tres tipos de datos: modo-uno, modo-dos y atributos (Everton, 2012a). En esta sección exploraremos estos tipos de datos en más detalle. El objetivo es definir y compararlos antes de empezar a recopilar y manipular estos datos.

3.1 Datos de Modo-Uno

Estos son aquellos que consisten en un solo conjunto de actores o nodos. Por ejemplo, en una red de amistad las personas están unidas a través de lazos amistosos, una red de negocios une diferentes empresas por medio de transacciones financieras, o una red diplomática une naciones con base a lazos diplomáticos. Los tres ejemplos propuestos representan relaciones de modo-uno puesto que los actores que forman parte de estas relaciones constan de un conjunto único de actores.

Sin embargo, no todas las relaciones de modo-uno son equivalentes. En las siguientes secciones diferenciaremos entre tipos de relaciones modo-uno con base a la simetría o asimetría de la relación.

3.1.1 Modo-uno: Simétricos

Las **relaciones simétricas de modo-uno** son aquellas que constan de un solo tipo de actor y donde no tomamos en cuenta la dirección de los enlaces. Por ejemplo, asumimos que la Figura 3.1 representa relaciones familiares entre miembros de esta red. En este caso, la mutualidad entre actores es asumida y por consiguiente vemos un solo borde enlazando los nodos. Es decir, si asumimos

que la persona A es familiar de la persona B, sería imposible que la persona B no sea familiar de la persona A.

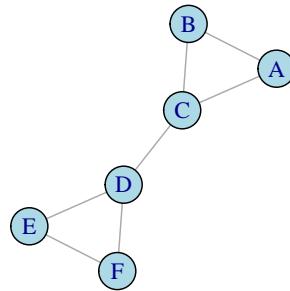


Figure 3.1: Gráfico no dirigido

Everton (2012a) nos presenta ejemplos de vínculos modo-uno simétricos utilizados para mapear la red terrorista Jemaah Islamiyah en el primer apéndice de *Disrupting Dark Network*:

Vínculo	Tipo de Actor	Definición
Amistad	Persona	Definido como vínculos de afecto y estima entre dos personas. Los lazos de amistad no son únicamente basados en reuniones y/o vínculos escolares.

Vínculo	Tipo de Actor	Definición
Compañeros de clase	Persona	Definido como individuos que reciben educación formal, sirven como empleados (maestro, miembro administrativo, etc.) y/o participan en instrucción educativa [...] simultáneamente en la misma institución.
Parentesco familiar	Persona	Definido como cualquier conexión familiar, como un hermano, un cuñado, un sobrino, etc. El parentesco incluye matrimonios actuales y matrimonios pasados debido a divorcios y/o muertes.

Los vínculos modo-uno simétricos se utilizan a menudo cuando es difícil o imposible determinar qué actor inició o envió el enlace y quién lo recibió (Cunningham, Everton, & Murphy, 2016). En los tres ejemplos previos podemos asumir mutualidad entre actores.

Ya que hemos definido este tipo de datos, es tiempo de ver como codificarlos. Los datos de modo-uno se pueden representar en un tipo de sociomatriz llamada matriz de adyacencia. La matriz (3.1.1) es una representación de los vínculos familiares de la Figura 3.1. El primer paso en generar una matriz de adyacencia es listar todos los nodos relevantes, {A, B, C, D, E, F}, como filas y columnas en la matriz. Representamos la presencia o ausencia de un vínculo con un valor, Se usan 1s para indicar que un nodo tiene una relación con otro. De lo contrario, se coloca un 0. Además, eliminamos la diagonal, ya que no es relevante que un nodo tenga un vínculo simétrico consigo mismo.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
<i>A</i>	—	1	1	0	0	0
<i>B</i>	1	—	1	0	0	0
<i>C</i>	1	1	—	1	0	0
<i>D</i>	0	0	1	—	1	1
<i>E</i>	0	0	0	1	—	1
<i>F</i>	0	0	0	1	1	—

Se estará preguntando, ¿cómo leemos la matriz? Centrémonos en un nodo a la vez. Por ejemplo, el nodo A, la primera letra de la columna más a la izquierda, está conectado con los nodos B y C, pero no con los demás {D, E, F}. Recuerde que la relación en cuestión es simétrica. Como tal, los dos lados de la matriz se reflejarán entre sí. Esto quiere decir que B y C también están vinculados a A.

3.1.2 Modo-uno: Asimétricos

Estos también son relaciones entre un solo conjunto de actores; sin embargo, las **redes asimétricas de modo-uno** generalmente se refieren a bordes que van de un nodo a otro, pero que no siempre son recíprocos (Borgatti, Everett, & Johnson, 2018; Everton, 2012a). Por ejemplo, considere la Figura 3.2. Inicialmente notará que los vínculos tienen dirección, representada por una flecha. Observe que algunos de los vínculos no son recíprocos (por ejemplo, A envía un enlace no recíproco a C), mientras que otros son correspondidos (por ejemplo, A envía un enlace a B y este lo retorna).

Cunningham, Everton, & Murphy (2016) sugieren que en situaciones donde es posible distinguir la dirección del enlace, es preferible construir matrices asimétricas de modo uno. Por ejemplo, la matriz (3.1.2) representa los datos de la Figura 3.2. Como con toda sociomatríg, comenzamos por listar todos los nodos posibles como filas y columnas {A, B, C, D, E, F}. Igualmente, la leemos de izquierda a derecha, comenzando con el nodo A, vemos que este envía lazos a los actores C y B, pero no a los demás {D, E, F}. En este caso, el actor A recibe un enlace de B pero no de C. Ahora, fíjese que ambos lados de la diagonal no son recíprocos, es decir contienen diferentes valores. Más aún, algunos nodos no envían ningún enlace solo reciben (como el actor D).

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
<i>A</i>	—	1	1	0	0	0
<i>B</i>	1	—	1	0	0	0
<i>C</i>	0	0	—	1	0	0
<i>D</i>	0	0	0	—	0	1
<i>E</i>	0	0	0	1	—	1
<i>F</i>	0	0	0	0	0	—

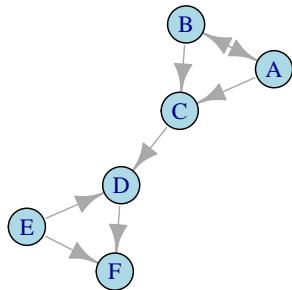


Figure 3.2: Gráfico dirigido

Para ver ejemplos de datos de red asimétricos de modo-uno, podemos volver a Everton (2012a):

Vínculo	Tipo de Actor	Definición
Comunicación	Persona	La comunicación interna se define como la transmisión de mensajes entre individuos y/o grupos dentro de la red a través de algún tipo de medio.

3.2 Datos de Modo-Dos

Las relaciones **modo-dos** son aquellas, que a diferencia de las modo-uno, consisten de vínculos entre dos clases de actores. Para entender datos de modo-dos, utilicemos la premisa de uno de los conjuntos de datos modo-dos más populares, Davis's Southern Women (Davis, Gardner, & Gardner, 1941). En este ejemplo tenemos dos conjuntos de actores. El primero representa personas {A, B, C, D, E, F}, el segundo representa eventos (E) sociales {E1, E2, E3}. En esta red modo-dos, no tenemos vínculos entre personas, sino vínculos de asistencia entre una persona y un evento. Podemos codificar estos datos utilizando una sociomatrix, listando todos los nodos {A, B, C, D, E, F, E1, E2, E3} en las filas

y columnas, ver figura (3.2). La asistencia se representa de manera binaria con un 1, mientras que la ausencia con un 0.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>E1</i>	<i>E2</i>	<i>E3</i>
<i>A</i>	—	0	0	0	0	0	1	0	0
<i>B</i>	0	—	0	0	0	0	1	0	0
<i>C</i>	0	0	—	0	0	0	1	1	0
<i>D</i>	0	0	0	—	0	0	0	1	1
<i>E</i>	0	0	0	0	—	0	0	0	1
<i>F</i>	0	0	0	0	0	—	0	0	1
<i>E1</i>	1	1	1	0	0	0	—	0	0
<i>E2</i>	0	0	1	1	0	0	0	—	0
<i>E3</i>	0	0	0	1	1	1	0	0	—

¿Qué patrones puede ver en la matriz previa? Primero, note la simetría de la matriz, es decir que los datos son simétricos. Al sumar las filas o columnas, notamos que el evento mas asistido es el primer evento (E1) con tres personas. Adicionalmente, la persona con mayor asistencia a eventos es el actor D (3 eventos). La Figura 3.3 es un sociograma de estos datos, en este las personas y eventos han sido dibujados con dos diferentes tipos de íconos para diferenciar entre las clases de nodos, sin embargo, esta diferenciación es puramente cosmética.

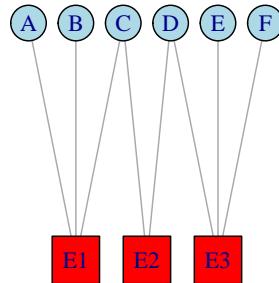


Figure 3.3: Gráfico Bipartita

Probablemente el patrón más importante que podemos observar en la matriz y el sociograma de modo-dos es que los nodos solo están conectados *entre*

clases. En este ejemplo, las personas solo pueden estar vinculadas con eventos y viceversa. Esta es la regla más importante de las redes modo-dos, solo se permiten conexiones entre nodos de diferentes conjuntos o particiones. Es por ello, que a menudo verán el término **bipartita** que refiere a redes modo-dos.

Una suposición clave que subyace al uso de redes de modo-dos es que las personas que tienden a unirse o participar en una organización común y/o eventos, a menudo tienen una razón compartida para hacerlo, como ideología, interés, etc. (Everton, 2012a). Como tal, es más probable que dos miembros de un grupo interactúen entre sí y compartan una conexión que dos personas al azar en una población. Más adelante generaremos datos modo-uno empezando con modo-dos. Por el momento, es importante entender por qué utilizamos este tipo de información en nuestro análisis.

Una vez más, veamos algunos ejemplos de redes modo-dos utilizadas para mapear Jemaah Islamiyah en Everton (2012a):

Vínculo	Tipo(s) de Actor(es)	Definición
Afiliación educativa	Persona y Organización	Las relaciones educativas se definen como escuelas donde las personas reciben educación formal, sirven como empleados (maestro, miembro administrativo, etc.) o participan en instrucción educativa o religiosa adicional en la institución.

Vínculo	Tipo(s) de Actor(es)	Definición
Afiliación Organizacional	Persona y Evento	Una organización terrorista/insurgente se define como un sistema administrativo y funcional, cuyo principal objetivo común es la conducta operativa de actividades terroristas/insurgentes, que consiste en miembros voluntariamente afiliados.
Eventos de Formación	Persona y Evento	Participación en cualquier actividad específicamente designada que enseñe el conocimiento, las habilidades y las competencias del terrorismo y la insurgencia. La capacitación no incluye la participación en un acto patrocinado por terroristas [...].

3.3 Datos de Atributos

Aunque no son el enfoque principal de ARSo, los datos de atributo se utilizan a menudo para acompañar el análisis de la red. Como se discutió anteriormente, a nivel de nodo, los atributos son características no relacionales de los nodos en un gráfico (Cunningham, Everton, & Murphy, 2016). Estos se recopilan por separado de las relaciones. Por ejemplo, considere las dos matrices (3.3) y (3.3).

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>		<i>Género</i>	<i>Edad</i>
<i>A</i>	—	1	1	0	0	0	<i>A</i>	<i>Masculino</i>	20
<i>B</i>	1	—	1	0	0	0	<i>B</i>	<i>Femenino</i>	20
<i>C</i>	1	1	—	1	0	0	<i>C</i>	<i>Femenino</i>	20
<i>D</i>	0	0	1	—	1	1	<i>D</i>	<i>Masculino</i>	29
<i>E</i>	0	0	0	1	—	1	<i>E</i>	<i>Masculino</i>	20
<i>F</i>	0	0	0	1	1	—	<i>F</i>	<i>Masculino</i>	25

La matriz de la izquierda es una sociomatrix simétrica de modo-uno, que contiene todas las relaciones entre el conjunto de nodos $\{A, B, C, D, E, F\}$. La matriz de la derecha contiene los atributos para el conjunto de nodos, específicamente, género y edad. La Figura 3.4 combina estos datos, donde los nodos están coloreados por el atributo categórico (género) y dimensionados por las variables continuas (edad).

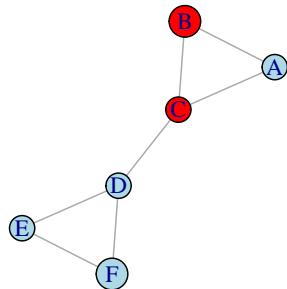


Figure 3.4: Gráfico no dirigido, con atributos

Al recopilar atributos, solo es necesario tener una entrada por nodo. Además, si se trabaja con redes multimodales (por ejemplo, bipartitas), se pueden recopilar atributos para todas las categorías de nodos.

Chapter 4

Grabación y Recopilación

Hasta este punto hemos descrito los datos de una red. Por lo que en esta sección, explicaremos cómo recopilar datos e importarlos a Gephi. Estas estrategias igualmente funcionaran con otras plataformas de software. Comenzaremos describiendo sociomatrizes. Después, veremos otra manera de codificar relaciones como listas de aristas. Luego, exploramos cómo codificar características no relacionales en listas de nodos. Finalmente, tendrá la oportunidad de poner en práctica este conocimiento a través de un ejercicio práctico en Gephi.

4.1 Matrices

4.1.1 ¿Qué son?

Comenzamos con la recopilación de datos de red en una matriz. Como señala Hanneman, la forma más común de matriz en ARSo es una matriz cuadrada con número de filas y columnas como actores en nuestro conjunto de datos, en este modelo los valores en las celdas de la matriz registran información sobre los vínculos entre cada par de actores (Hanneman & Riddle, 2005). Por ejemplo, la matriz (4.1.1) representa la forma más simple y común de codificar datos relacionales en forma de matriz binaria. En este caso, vemos un 1 si existe un vínculo, de lo contrario vemos un 0.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
<i>A</i>	—	1	1	0	0	0
<i>B</i>	1	—	1	0	0	0
<i>C</i>	1	1	—	1	0	0
<i>D</i>	0	0	1	—	1	1
<i>E</i>	0	0	0	1	—	1
<i>F</i>	0	0	0	1	1	—

Como se ha de imaginar, las matrices no siempre son binarias. De hecho, es apropiado y común el recopilar datos de magnitud para un vínculo. Por ejemplo, se puede medir la fuerza de una relación a nivel ordinal o de intervalo. Igualmente, se pueden recaudar datos relacionales con números positivos y negativos, donde los números positivos corresponden a relaciones positivas, el cero es un punto medio neutral y las relaciones negativas tienen valores menores a cero.

Recuerde que como habíamos visto previamente, los datos codificados en matrices pueden ser simétricos o asimétricos, dirigidos o no dirigidos. Por consiguiente, tradicionalmente es de buena educación ser explícito al mencionar el tipo de matriz que se está trabajando; por ejemplo, “matriz simétrica binaria modo-uno” o “matriz asimétrica ponderada modo-uno.”

En esta sección, nos enfocamos en matrices de relaciones modo-uno por su simplicidad, pero cabe recalcar que las matrices pueden ser de relación modo-dos. La mayoría de los analistas de redes trabajan con relaciones de modo-uno, sin importar si estas han sido derivadas de una relación modo-dos. Es por ello que retornaremos a las matrices modo-dos cuando toquemos el tema de derivar relaciones modo-uno de modo-dos.

4.1.2 ¿Por qué usarlas o no?

Las matrices tienen múltiples ventajas en ARSo. Primero, son simples de entender. Sin embargo, la razón clave por la cual se utilizan habitualmente es porque se pueden realizar cálculos matemáticos con estas para resumir la información de la red e ilustrar patrones relacionales (Borgatti, Everett, & Johnson, 2018).

Al trabajar con múltiples matrices de datos modo-uno con las mismas dimensiones, podemos producir una tercera matriz con el mismo número de dimensiones utilizando el producto Hadamard (operación por elementos). Por ejemplo, tomemos dos matrices con el mismo número de dimensiones una representando amistades y la otra con vínculos de negocios. Con estas podemos producir una tercera relación con las mismas dimensiones la cual nombremos “amigos en negocios,” la ecuación (4.1.2) ilustra esta operación de matrices por elemento.

$$\begin{array}{ccc}
 & A & B & C \\
 A & 0 & 1 & 0 & \circ & A & 0 & 1 & 1 \\
 B & 1 & 0 & 1 & & B & 1 & 0 & 0 \\
 C & 0 & 1 & 0 & & C & 1 & 0 & 0
 \end{array}
 = \begin{array}{cccc}
 & A & B & C \\
 A & 0x0 & 1x1 & 0x1 \\
 B & 1x1 & 0x0 & 0x0 \\
 C & 0x1 & 1x0 & 0x0
 \end{array}
 = \begin{array}{ccc}
 & A & B & C \\
 A & 0 & 1 & 0 \\
 B & 1 & 0 & 0 \\
 C & 0 & 0 & 0
 \end{array}$$

Adicionalmente, podemos multiplicar matrices “conformables”; es decir, el

número de filas en la primera matriz es igual a el número de columnas en la segunda. Las matrices cuadradas de modo-uno son conformables puesto que tienen el mismo numero de filas y columnas, por lo cual las podemos multiplicar (Hanneman & Riddle, 2005). A diferencia de el producto Hadamard, en la multiplicación de matrices el orden de estas impacta los resultados. Supongamos que trabajamos con dos matrices de modo-uno (A y B), si multiplico AxB el producto no es equivalente a BxA ($AxB \neq BxA$). La ecuación (4.1.2) ilustra el proceso de multiplicar matrices AxB . Empezamos con la parte superior izquierda de la primera matriz y multiplicamos cada valor en la primera fila por los valores de la primera columna de la segunda matriz. Los elementos de una fila de la primera matriz y de una columna de la segunda se multiplican y se suman para producir un solo producto.

$$\begin{array}{ccc} A & B & C \\ \begin{matrix} A & 0 & 1 & 0 \\ B & 1 & 0 & 1 \\ C & 0 & 1 & 0 \end{matrix} & \times & \begin{matrix} A & B & C \\ B & 1 & 0 & 0 \\ C & 1 & 0 & 0 \end{matrix} \\ & & = \begin{matrix} A & B & C \\ B & 1 & 1 & 1 \\ C & 1 & 0 & 0 \end{matrix} \end{array}$$

$$\begin{array}{ccc} A & B & C \\ \begin{matrix} A & (0x0) + (1x1) + (0x1) \\ B & (1x0) + (0x1) + (1x1) \\ C & (0x0) + (1x1) + (0x1) \end{matrix} & \begin{matrix} B & (0x1) + (1x0) + (0x0) \\ (1x1) + (0x0) + (1x0) \\ (0x1) + (1x0) + (0x0) \end{matrix} & \begin{matrix} C & (0x1) + (1x0) + (0x0) \\ (1x1) + (0x0) + (1x0) \\ (0x1) + (1x0) + (0x0) \end{matrix} \end{array}$$

El proceso de multiplicar BxA es el mismo, pero los resultados son diferentes (4.1.2).

$$\begin{array}{ccc} A & B & C \\ \begin{matrix} A & 0 & 1 & 1 \\ B & 1 & 0 & 0 \\ C & 1 & 0 & 0 \end{matrix} & \times & \begin{matrix} A & B & C \\ B & 1 & 0 & 1 \\ C & 0 & 1 & 0 \end{matrix} \\ & & = \begin{matrix} A & B & C \\ B & 0 & 1 & 0 \\ C & 0 & 1 & 0 \end{matrix} \end{array}$$

$$\begin{array}{ccc} A & B & C \\ \begin{matrix} A & (0x0) + (1x1) + (1x0) \\ B & (1x0) + (0x1) + (0x0) \\ C & (1x0) + (0x1) + (0x0) \end{matrix} & \begin{matrix} B & (0x1) + (1x0) + (1x1) \\ (1x1) + (0x0) + (0x1) \\ (1x1) + (0x0) + (0x1) \end{matrix} & \begin{matrix} C & (0x0) + (1x1) + (1x0) \\ (1x0) + (0x1) + (0x0) \\ (1x0) + (0x1) + (0x0) \end{matrix} \end{array}$$

En términos prácticos, Borgatti, Everett, & Johnson (2018) presentan el siguiente ejemplo, A es una matriz de adyacencia que representa la ‘relación de amistad’ y B es la relación de ‘jefe de.’ Entonces si tengo una relación AxB con un nodo, este es el jefe de al menos uno de mis amigos. En contraste, si tengo una relación de BxA con un nodo, entonces este es amigo de mi jefe (Borgatti, Everett, & Johnson, 2018). Como puede ver, estas son relaciones muy diferentes producidas al multiplicar las mismas matrices.

Ya que hemos recalcado las ventajas de utilizar matrices, es esencial entender algunas de las desventajas clave:

- Cada matriz representa un solo tipo de relación. Supongamos que el análisis requiere combinar relaciones familiares y financieras. Para esto, necesitaríamos recopilar dos matrices individuales antes de manipularlas. Además, el formato de matriz no permite el agregar metadatos sobre las relaciones, por ejemplo: procedencia de un dato, fecha en que ha ocurrido, entre otros.
- El mantener una o múltiples matrices requiere mantener todas las células. Por ejemplo, en una red de 100 nodos, el analista debe mantener 10,000 células (100×100) aun cuando la mayoría de estas estén vacías (sin relación).
- Esta estructura no es la manera más eficiente de guardar datos en una computadora. Específicamente, cada 0 que representa la ausencia de una relación tiene un costo en forma de memoria.

Como puede ver, las matrices (como todas las estructuras de recopilación de datos) tienen ventajas y desventajas. El álgebra de matrices es la base de la mayoría de los algoritmos utilizados por analistas de redes, por consiguiente, es esencial entender estas estructuras. Sin embargo, las desventajas mencionadas previamente son la razón por la cual los analistas de redes tienen a mantener sus datos relacionales en otros formatos más eficientes y permisibles. En la siguiente sección cubriremos uno de estos formatos llamado lista de aristas.

4.2 Lista de Aristas

4.2.1 ¿Qué son?

Son una de las estructuras más económicas en la recopilación de datos. Las listas de aristas solo consideran los vínculos existentes y excluyen lo demás. Por ejemplo, la matriz (4.2.1) puede ser colapsada como lista de aristas (4.2.1).

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>Origen</i>	<i>Destino</i>
<i>A</i>	—	1	1	0	0	0	<i>A</i>	<i>B</i>
<i>B</i>	1	—	1	0	0	0	<i>A</i>	<i>C</i>
<i>C</i>	1	1	—	1	0	0	<i>B</i>	<i>A</i>
<i>D</i>	0	0	1	—	1	1	<i>B</i>	<i>C</i>
<i>E</i>	0	0	0	1	—	1	<i>C</i>	<i>A</i>
<i>F</i>	0	0	0	1	1	—	<i>C</i>	<i>B</i>
							<i>C</i>	<i>D</i>
							<i>D</i>	<i>C</i>
							<i>D</i>	<i>E</i>
							<i>D</i>	<i>F</i>
							<i>E</i>	<i>D</i>
							<i>E</i>	<i>F</i>
							<i>F</i>	<i>D</i>
							<i>F</i>	<i>E</i>

Esencialmente, una lista de aristas es un rectángulo con un mínimo de dos columnas, cada fila representa una combinación de nodos vinculados. Cunningham, Everton, & Murphy (2016) mencionan que aunque las listas de aristas pueden ser dirigidas o no dirigidas, es decir, contienen arcos o bordes, las columnas denotan un actor que “envía” el enlace y otro que “recibe” el enlace. En Gephi, el software denota estas posiciones con nombres obligatorios; es decir, que en términos de Gephi, la columna de “envía” siempre aparecerá nombrada “Origen” (o “Source” en Inglés). Similarmente, la columna “recibe” siempre aparecerá como “Destino” (o “Target” en Inglés).

Tradicionalmente, los analistas de ARSo tienden a mantener sus listas de aristas en software como Microsoft Excel. Sin embargo, es posible integrar tablas de bases de datos (por ejemplo, SQL) como precursores.

4.2.2 ¿Por qué usarlas o no?

Las listas de aristas son comúnmente utilizadas por su simplicidad, eficiencia y versatilidad. Son eficientes pues no incluyen espacios vacíos para vínculos faltantes. Asimismo, el mantenerlas es simple pues cada fila representa una relación. Adicionalmente son versátiles pues proveen flexibilidad en el número y tipo de características de un vínculo que se pueden incluir. Como se ha mencionado previamente, el formato básico de una lista de aristas incluye dos columnas, sin embargo, no hay límite a la cantidad de variables que podemos agregar para describir cada vínculo, por ejemplo: dirigido/no dirigido, fecha, intensidad, tipo, peso, fuente de procedencia, etc.

Las Tablas 4.1, 4.2 y 4.3 son ejemplos de listas de aristas. La primera lista de aristas, Tabla 4.1, contiene los elementos mínimos requeridos; es decir, dos columnas nombradas que representan los dos vértices conectados por un vínculo relacional. Cada fila representa una relación, sin embargo, no hay manera de

Table 4.1: Lista de Aristas Básica

Origen	Destino
A	B
B	C

Table 4.2: Lista de Aristas Expandida

Origen	Destino	Relación	Tipo	Fecha	Peso
A	B	Comunicación	No dirigido	2020-11-03	5
B	C	Financiera	Dirigido	2020-11-03	10

diferenciar si estas relaciones son dirigidas/no dirigidas o si ambos vínculos representan el mismo tipo de relaciones (por ejemplo, familiares, financieras, románticas, etc.). Es por ello que la segunda lista, Tabla 4.2, expande sobre la primera al agregar un par de variables. En este ejemplo, las nuevas columnas permiten diferenciar el tipo de relación e incluso agregan información temporal y pesos a cada enlace. La última lista, Tabla 4.3, continúa expandiendo la lista de aristas previa ya que presenta relaciones multimodo. Es decir, algunas relaciones son modo-uno, mientras que otras son modo-dos.

4.3 Lista de Nodos

La lista de nodos, a diferencia de la lista de aristas, contiene los datos no relacionales de cada nodo en un sociograma. Por consiguiente, ambas listas sean mantenidas de manera independiente.

Es decir, en una lista de aristas es probable que un nodo aparezca de manera repetida, puesto que es común que cada actor en una red este atado a múltiples otros. En contraste, en una lista de nodos es recomendable el tener a un solo registro por actor. Por ejemplo, en la lista de aristas (4.3) vemos cuatro nodos {A, B, C, D} conectados por tres diferentes tipos de relaciones. La lista de nodos correspondiente (4.3) incluye información de estos nodos como género y edad, estas son características no relacionales de aquellos nodos en la red. Aunque podemos repetir el registro de cada actor, no es necesario normalmente

Table 4.3: Lista de Aristas Expandida y Multimodo

Origen	Destino	Relación	Tipo	Fecha	Peso
A	B	Comunicación	No dirigido	2020-11-03	5
B	C	Financiera	Dirigido	2020-11-03	10
A	1	Afiliación	No dirigido	2020-11-03	7

el software solo toma en cuenta un registro, normalmente el último en la lista.

<i>Origen</i>	<i>Destino</i>	<i>Relación</i>	<i>Id</i>	<i>Género</i>	<i>Edad</i>
A	B	Amistad	A	Masculino	20
A	C	Familia	B	Femenino	20
A	D	Finanzas	C	Femenino	20
			D	Masculino	29

Por su parte, Gephi es capaz de ingerir listas de nodos. Para ello, requiere que esta tabla contenga el nombre de columna *Id*. Esta variable representa la columna que el software tomará en cuenta como fuente de identificadores únicos para los nodos en una red. Por ejemplo, en la lista de nodos (4.3) vemos entradas correspondientes a los identificadores de cada nodo en la lista de aristas.

4.4 Ejercicio Práctico

Este ejercicio tiene dos propósitos:

1. Proporcionarle práctica en recopilar datos
2. Importar y manipular datos de modo-uno a Gephi

Las redes de modo-uno consiste en un conjunto único de actores, estos pueden ser personas, grupos, organizaciones, corporaciones, naciones, etc. Los lazos entre estos pueden ser lazos de amistad o parentesco, transacciones materiales (por ejemplo, transacciones comerciales), redes de comunicación (por ejemplo, élenvio o la recepción de mensajes), etc. En este laboratorio, utilizaremos los datos recopilados por Stuart Koschade de la célula terrorista Jemaah Islamiyah (JI) responsable del atentado terrorista en Bali del 2002. Todos los datos requeridos en este ejercicio deberán ser primero generados por el usuario y luego descargados.

El segundo propósito de esta actividad es presentarle algunas de las funciones de la plataforma Gephi. Específicamente, nos centraremos en visualizar redes. Estos gráficos son importantes porque pueden ayudarnos a rastrear y detectar patrones de relaciones en una red social determinada. Sin embargo, puesto que el ojo humano puede ser engañado fácilmente, los analistas de redes sociales confían en algoritmos de dibujo de gráficos que colocan sistemáticamente a los actores en el “espacio social.”

4.4.1 Preparar y Visualizar Redes de Modo-Uno en Gephi

Localice el artículo de Stuart Koschade, aquí. La única versión disponible es en inglés, por consiguiente, no es requerido que lea el texto. Utilizando Excel (u otro programa con hojas de cálculo) registre los datos de redes sociales que figuran en la Tabla 2 en el artículo de Koschade. Su matriz debe ser similar a la Figura 4.1 a continuación.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Muklas	Amrozi	Imron		Samudra	Dulmatin	Idris	Mubarok	Azahari	Ghoni	Arnasan	Rauf	Octavia	Hidayat	Junaedi	Patek	Feri	Sarijo
2	Muklas	0	2	2	1	1	5	0	1	1	0	0	0	0	0	1	0	1
3	Amrozi	2	0	0	2	0	4	5	0	0	0	0	0	0	0	0	0	0
4	Imron	2	0	0	3	5	3	0	5	5	0	0	0	0	0	5	1	5
5	Samudra	1	2	3	0	2	5	2	2	2	2	2	2	2	2	2	0	2
6	Dulmatin	1	0	5	2	0	2	0	5	5	0	0	0	0	0	5	1	5
7	Idris	5	4	3	5	2	0	2	2	2	0	0	0	0	0	0	2	0
8	Mubarok	0	5	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0
9	Azahari	1	0	5	2	5	2	0	0	5	0	0	0	0	0	2	1	2
10	Ghoni	1	0	5	2	5	2	0	5	0	0	0	0	0	0	5	1	5
11	Arnasan	0	0	0	2	0	0	0	0	0	0	2	2	2	2	0	0	0
12	Rauf	0	0	0	2	0	0	0	0	0	2	0	2	2	2	0	0	0
13	Octavia	0	0	0	2	0	0	0	0	0	2	2	0	2	2	0	0	0
14	Hidayat	0	0	0	2	0	0	0	0	0	2	2	2	0	2	0	0	0
15	Junaedi	0	0	0	2	0	0	0	0	0	2	2	2	2	0	0	0	0
16	Patek	1	0	5	2	5	2	0	2	5	0	0	0	0	0	0	1	5
17	Feri	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0	1
18	Sarijo	0	0	5	2	5	2	0	2	5	0	0	0	0	0	5	1	0
19																		

Figure 4.1: Hoja de cálculo de Excel que muestra la matriz de datos en Koschade

Vista general
> Nuevo Proyecto

Archivo >
Abrir...

1. Abra Gephi y seleccione la opción *Nuevo proyecto* en la pantalla de “Bienvenido a Gephi.” Asegúrese de estar en la ventana “Vista general” en la esquina superior izquierda de la interfaz. Una vez allí, su interfaz deberá parecerse a la Figura 4.2.
2. Primero debemos importar los datos de recopilados en el primer paso a Gephi. Para hacer esto, use el comando *Archivo > Abrir...*. Esto abrirá un cuadro de diálogo (no mostrado) desde el cual puede ubicar y seleccionar el archivo de Excel que creó anteriormente. Una vez que haya hecho esto, haga clic en *Abrir* y aparecerá un nuevo cuadro de diálogo, similar a la Figura 4.3. Tenga en cuenta que, en la esquina superior derecha, tiene la opción de importa la hoja de cálculo en varios formatos. Como registramos los datos en matriz, utilizaremos la opción *Matriz*. Haga clic en *Siguiente*, en el siguiente cuadro de diálogo, haga clic en *Terminar* (puesto que no tenemos información de tiempo asociada con estos datos).

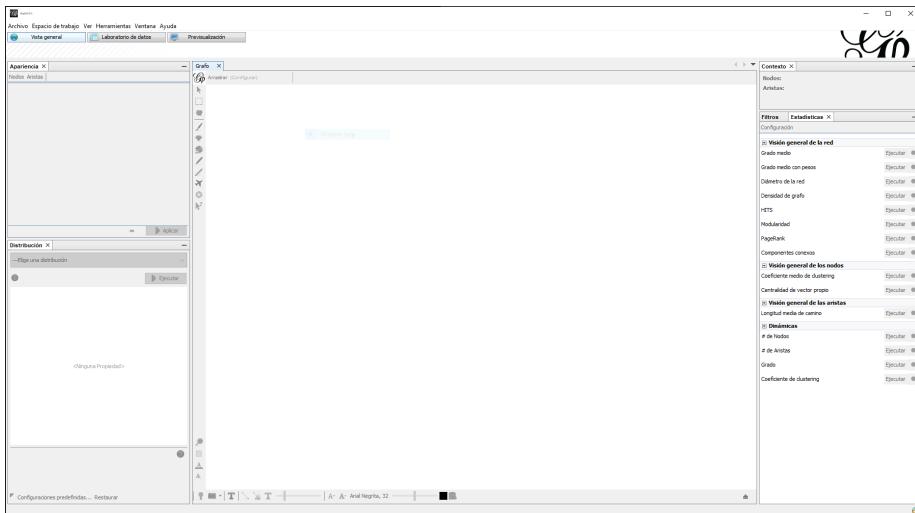


Figure 4.2: Ventana 'Vista general' en Gephi

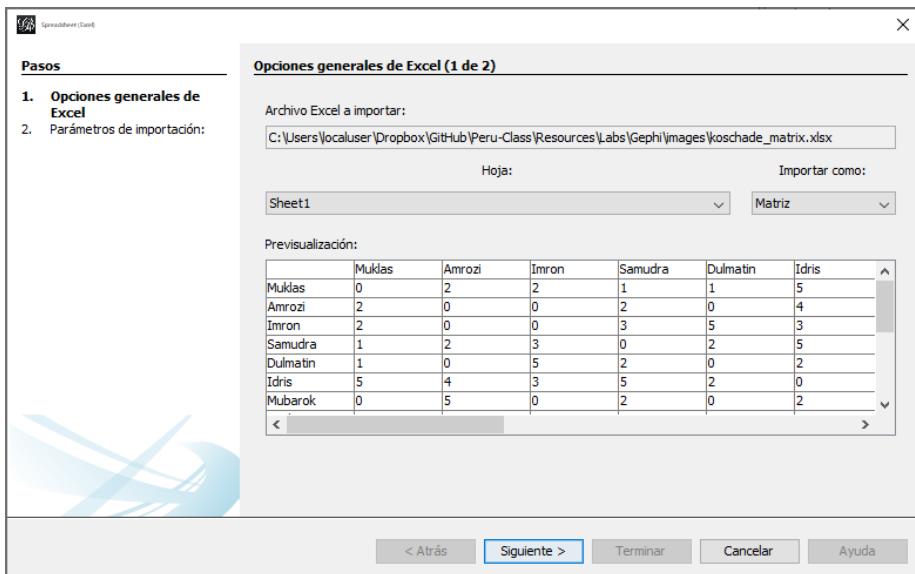


Figure 4.3: Ventana de Importación en Gephi

Tipo de grafo
 > No dirigido
Mas opciones...
 > Estrategias para combinar aristas
 > Mínimo

3. Gephi produce un *Informe de importación* (ver Figura 4.4), que proporciona varios tipos de información, por ejemplo, si ocurrió algún error durante la importación, numero de nodos (actores) en la red, la cantidad de bordes (aristas), etc. Ahora, necesitamos decirle a Gephi que esta es una red no dirigida utilizando el menú desplegable *Tipo de grafico* (el tipo predeterminado es *Mixto*). A continuación, haga clic en el enlace *Mas opciones...*, y luego en el menú desplegable *Estrategia para combinar aristas* seleccione *Mínimo*. Este último paso asegura que los pesos de los bordes no se dupliquen. Ahora, haga clic en *Aceptar*.

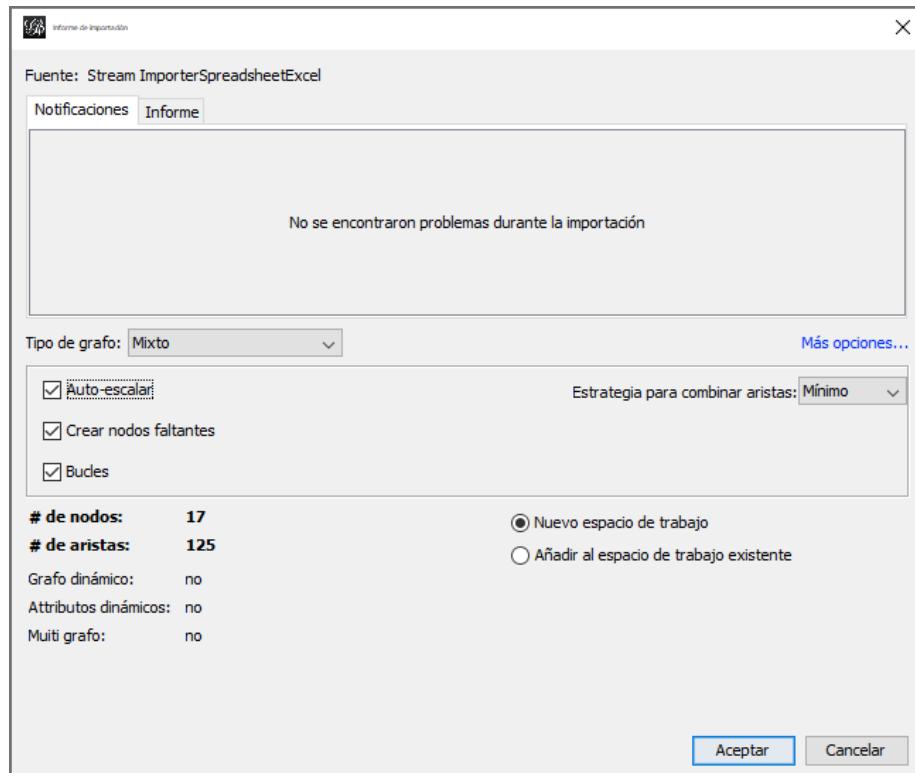


Figure 4.4: Reporte de Importación Gephi

Apariencia
> Nodos >
Único

4. Su red ahora deberá aparecer en pantalla, pero es probable que los nodos aparezcan de manera desordenada. Comencemos cambiando el color de los nodos. En la esquina superior izquierda de la interfaz vera un pestana *Apariencia* (consulte la Figura 4.5). Allí, seleccione la pestaña *Nodos* y el ícono de paleta de color (que es el ícono más a la izquierda). Debajo de la pestaña *Nodos* verá un rango de opciones, si no está seleccionada la pestaña *Único*, haga clic en esta. Debajo de esta opción, note un cuadro gris (a la izquierda de #c0c0c0); de clic y manténgalo presionado para seleccionar un nuevo color, luego seleccione *Aplicar*.

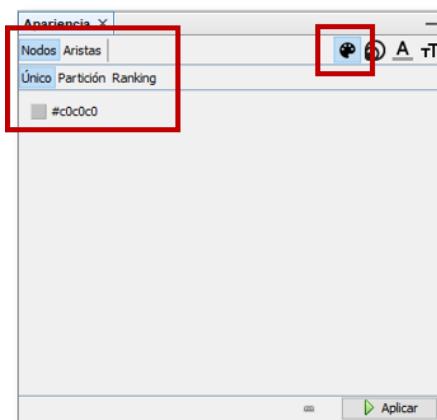


Figure 4.5: Pestaña de Apariencia

Distribución
>Fruchter-
man Reingold,
Force Atlas,
Expansion,
Yifan Hu

5. Justo debajo de la pestaña *Apariencia* encontraraá otra de nombre *Distribución* (consulte la Figura 4.6). Gephi ofrece varias opciones de distribución de nodos, spara más información sobre estas consulte la página de tutoriales de Gephi. En esta actividad solo exploraremos algunas de estas estrategias de distribución. Utilizando el menú desplegable, primero seleccione el algoritmo de nombre *Fruchterman Reigold* y haga clic en el botón *Ejecutar*. Dependiendo de la red, es posible que este algoritmo no sea el adecuado. Pruebe otros, por ejemplo *Force Atlas*. Al principio parecerá que los nodos han colapsado unos sobre otros, pero esto es engañoso. Para ver porque, elija la opción *Expansion* y haga clic repetidamente en el botón *Ejecutar* hasta que el gráfico sea los suficientemente grande para poder distinguir las conexiones entre nodos. Finalmente. Pruebe el algoritmo *Yifan Hu*, el cual, al menos con la red de Koschade, produce el gráfico más adecuado. Es decir, los subgrupos son fáciles de detectar y los nodos no se encuentran apilados unos sobre los otros.

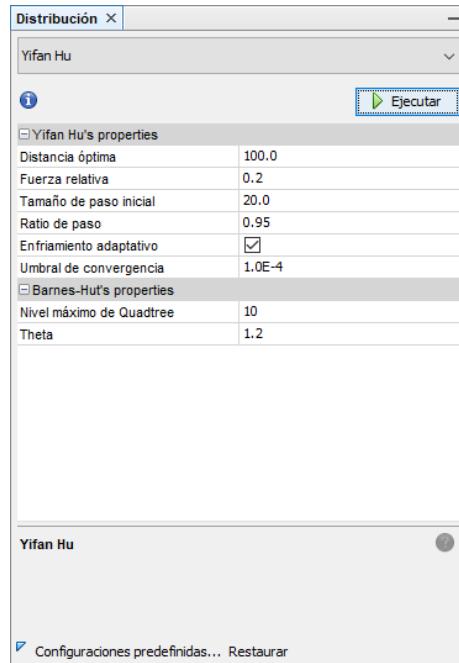


Figure 4.6: Pestaña de Diseño

Grafo

6. Estas manipulaciones deben crear un gráfico similar a la Figura 4.7 (similar pues usted ha elegido el color de los nodos). Ahora, note lo diferentes botones al costado de la ventana del *Grafo* (ver Figura 4.7), estos nos permiten hacer algunos ajustes rápidos al grafo. Por ejemplo, el botón 1 centra el gráfico, el 2 activa y desactiva las etiquetas, o puede controlar el tamaño de las etiquetas con la barra deslizante demarcada con el número 5. El botón 3 alterna el color de los nodos. Por último, note la flecha pequeña en la esquina inferior derecha (número 6), en el siguiente paso y otros laboratorios utilizaremos esta opción.

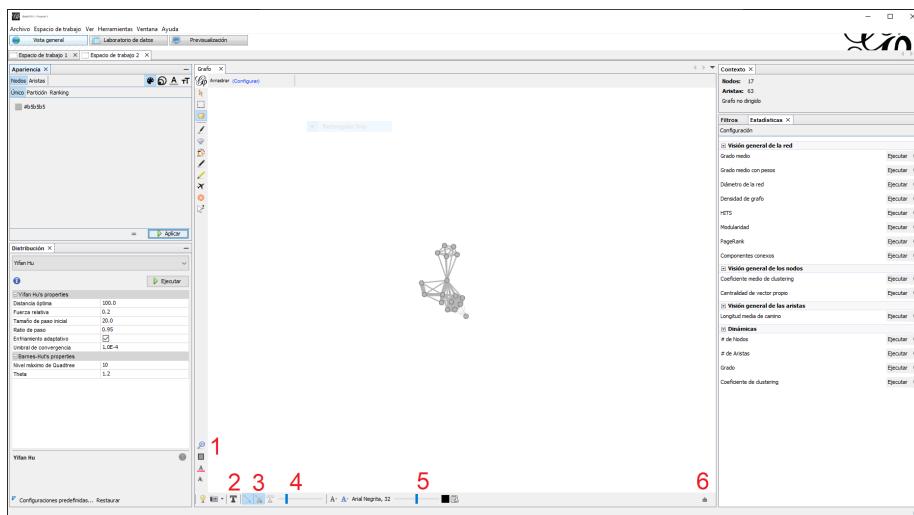


Figure 4.7: Botones en Gephi

7. Al darle clic a la pequeña flecha (número 6 en la Figura 4.7), se debe levantar un panel con controles adicionales para manipular la red. El menú que debe es el de controles *Globales* (Figura 4.8). En este, quizás el control más útil es la barra deslizante de zoom (etiqueta 2), que controla el tamaño del gráfico. En la imagen de abajo he aumentado el tamaño de la red y he centrado la imagen presionando el botón derecho del ratón y arrastrando el gráfico hacia el centro de la ventana. Además de modificar la dimensión del gráfico, puede cambiar el fondo dando clic en el cuadro en la esquina inferior izquierda (etiqueta 1).

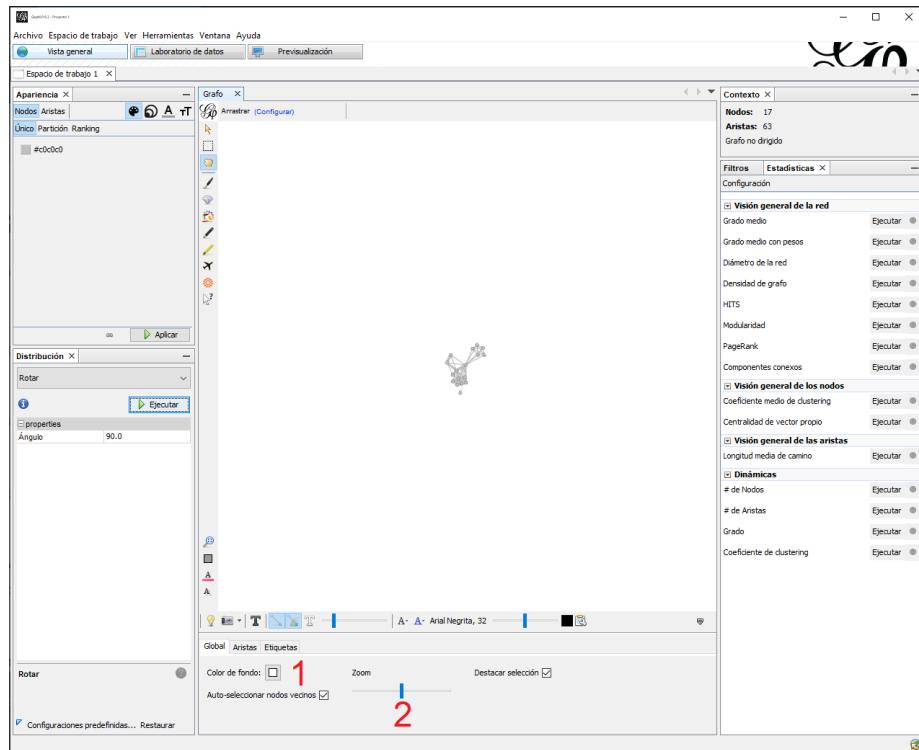


Figure 4.8: Pestaña de Controles Globales en Gephi

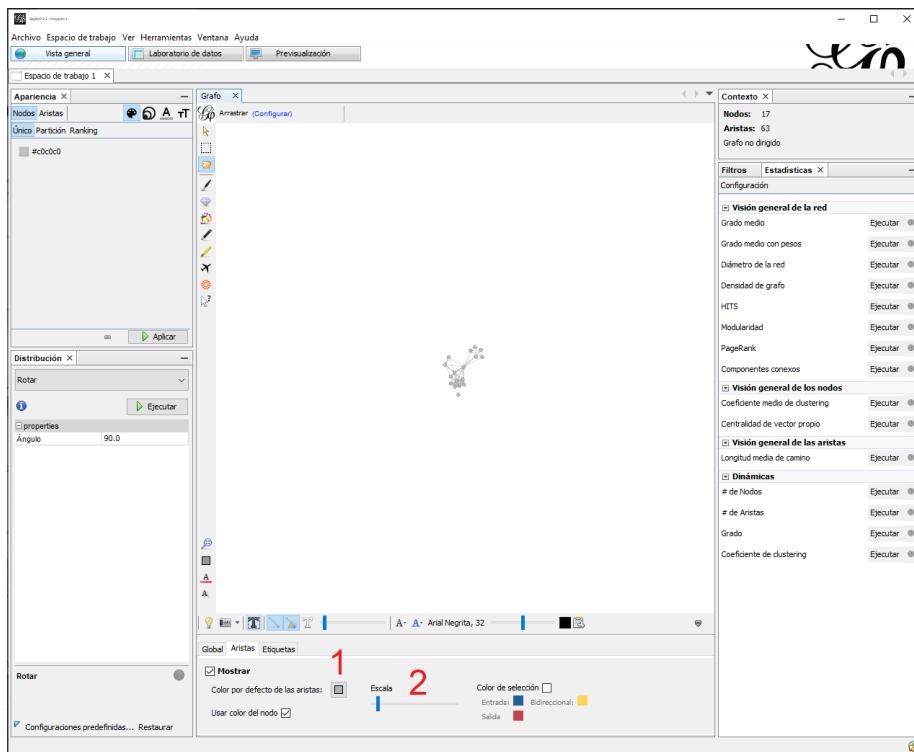


Figure 4.9: Pestaña de Control de Aristas

8. En la Figura 4.9 he seleccionado el panel de control de aristas, proceda haciendo lo mismo. Ahora, cambie el color de las aristas a gris (vea etiqueta 1 en la Figura 4.9) y use la barra deslizante (etiqueta 2) para reducir el ancho de cada línea. . Finalmente, seleccione la pestaña de etiquetas en el panel de control. Aquí, haga clic en el cuadro a un lado de *Nodos* (etiqueta 1 en la Figura 4.10) para activar las etiquetas de los nodos en el gráfico. Como verá estas serán inicialmente un poco grandes. Para ajustar estas cambie el tamaño estas utilizando la barra deslizante (etiqueta 3 en Figura 4.10). En Gephi siempre hay un par de maneras de hacer lo mismo. Por ejemplo, si selecciona el botón a un lado de “Fuentes” (etiqueta 2) note que puede cambiar múltiples propiedades de las etiquetas de cada nodo. El análisis de redes depende de gráficos claros y concisos. Tómese un par de minutos para modificar estas etiquetas a su gusto y cuando termine cierre el panel dando clic a la flecha (etiqueta 4) a mano derecha (ahora apuntando hacia abajo).
9. Terminemos de editar el gráfico. Para esto, seleccione la pestaña de *Distribución* y seleccione la opción *Ajuste de etiquetas* y haga clic en *Ejecutar*. Este algoritmo de diseño moverá los nodos ligeramente para que las etiquetas no se superpongan. Ahora, seleccione la opción *Noverlap*, y nuevamente de clic en *Ejecutar*. Verá que nuevamente los nodos deberán moverse, creando una distancia entre ellos para que sean más fáciles de discernir. Ahora, observe el grafo. ¿Surge patrones? ¿Hay algún grupo distinto? Si es así, ¿alguna persona actúa como intermediario entre los distintos grupos? ¿Alguno de los individuos parece estar más aislado socialmente que otros? ¿Cómo se comparan los dibujos con la Figura 1 en el artículo de Koschade?

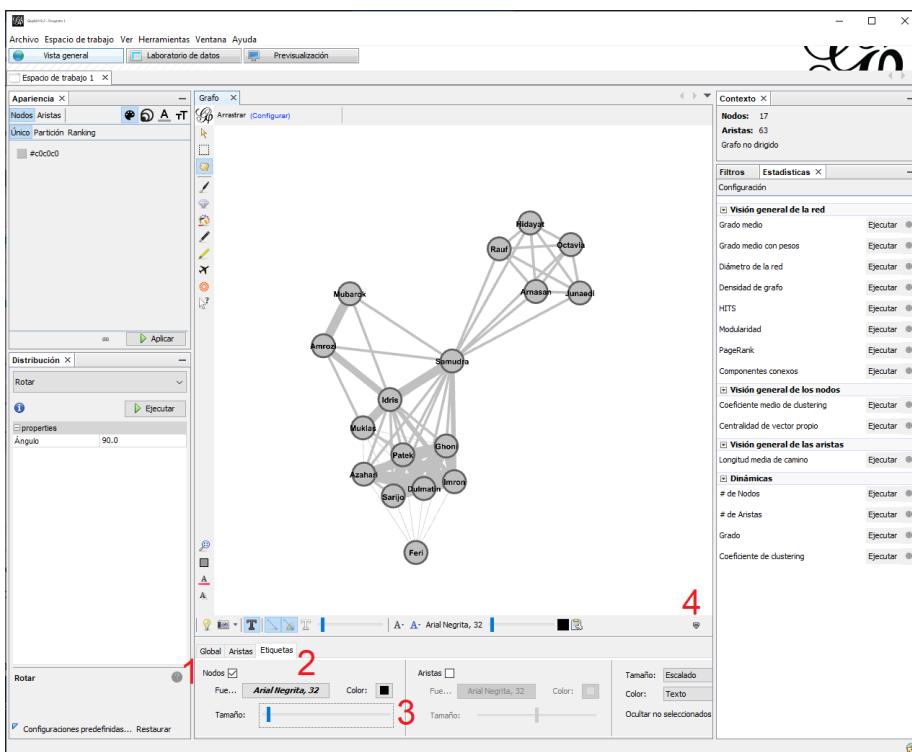


Figure 4.10: Panel de Control de Etiquetas en Gephi

10. Ahora, desplace el cursor sobre uno de los nodos, al hacer esto Gephi debe resaltar un nodo a la vez y los enlaces a otros nodos. En la Figura 4.11, presentamos el nodo de Patek resaltado, este tiene vínculos con varias personas dentro de su propia esfera de influencia, además algunos de estos nodos vecinos están conectados entre ellos. Sin embargo, note que los vecinos de Patek no están enlazados con Samudra. Utilizando esta herramienta (resaltar nodos) en su propio gráfico, **¿saltan patrones que no haya notado antes?** Finalmente, tome una captura de pantalla de su grafico dando clic en el ícono de la cámara en la esquina inferior de la ventanilla del gráfico.

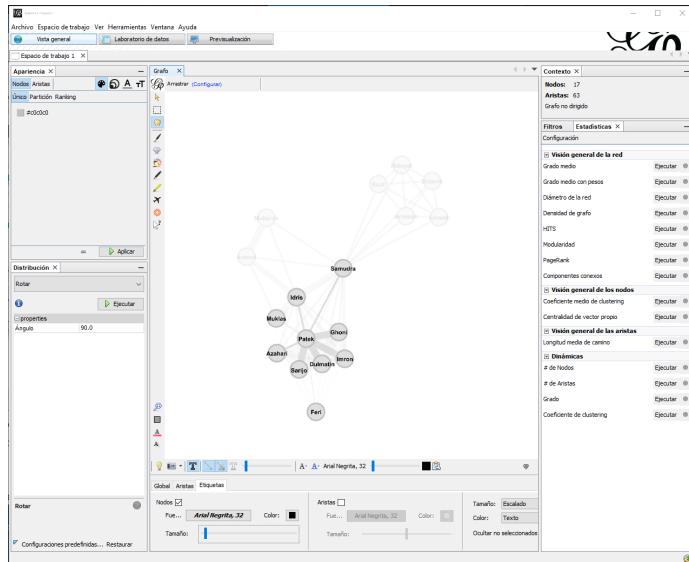


Figure 4.11: Red con Patek Resaltado

Laboratorio de datos
Tabla de datos
> Nodos

11. Echemos un vistazo a los datos de la red. Para esto, seleccione el botón *Laboratorio de datos*. Notará que en la esquina superior izquierda hay una pestaña *Tabla de datos* con dos subpestañas , *Nodos* y *Aristas*. Haga clic primero en la pestaña *Nodos* y deberá ver una lista de los identificadores (*Id*) y etiquetas de los nodos (actores) en la red.

Origen	Destino	Tipo	ID	Label	Interval	Weight
Mulás	Ameril	No dirigida	0			2,0
Mulás	Itron	No dirigida	1			2,0
Mulás	Semuda	No dirigida	2			1,0
Mulás	Duhetón	No dirigida	3			1,0
Mulás	Idris	No dirigida	4			5,0
Mulás	Azharai	No dirigida	5			1,0
Mulás	Dion	No dirigida	6			1,0
Mulás	Patek	No dirigida	7			1,0
Mulás	Sarja	No dirigida	8			1,0
Ameril	Semuda	No dirigida	9			2,0
Ameril	Itron	No dirigida	10			2,0
Ameril	Idris	No dirigida	11			4,0
Ameril	Hubarak	No dirigida	12			5,0
Itron	Semuda	No dirigida	13			3,0
Itron	Duhetón	No dirigida	14			3,0
Itron	Idris	No dirigida	15			5,0
Itron	Azharai	No dirigida	16			3,0
Itron	Dion	No dirigida	17			5,0
Itron	Patek	No dirigida	18			5,0
Itron	Feri	No dirigida	19			1,0
Itron	Sarja	No dirigida	20			5,0
Semuda	Duhetón	No dirigida	21			2,0
Semuda	Itron	No dirigida	22			5,0
Semuda	Azharai	No dirigida	23			2,0
Semuda	Hubarak	No dirigida	24			2,0
Semuda	Azharai	No dirigida	25			2,0
Semuda	Ghorai	No dirigida	26			2,0
Semuda	Ameril	No dirigida	27			2,0
Semuda	Rauf	No dirigida	28			2,0
Semuda	Octavia	No dirigida	29			2,0
Semuda	Hilwatt	No dirigida	30			2,0
Semuda	Imerai	No dirigida	31			2,0
Semuda	Patek	No dirigida	32			2,0
Semuda	Sarja	No dirigida	33			2,0
Semuda	Azharai	No dirigida	34			2,0
Semuda	Dion	No dirigida	35			2,0
Semuda	Ghorai	No dirigida	36			2,0
Duhetón	Imerai	No dirigida	37			2,0
Duhetón	Azharai	No dirigida	38			5,0
Duhetón	Ghorai	No dirigida	39			5,0
Duhetón	Patek	No dirigida	40			5,0
Duhetón	Feri	No dirigida	41			1,0
Duhetón	Sarja	No dirigida	42			5,0
Sarja	Azharai	No dirigida	43			2,0
Sarja	Dion	No dirigida	44			1,0
Sarja	Ghorai	No dirigida	45			5,0
Sarja	Hubarak	No dirigida	46			2,0
Sarja	Azharai	No dirigida	47			2,0

Figure 4.12: Lista de Aristas en la Table de Datos

*Laboratorio de datos
Tabla de datos
> Aristas*

12. Ahora, haga clic en la pestaña *Aristas* y obtendrá una lista de todos los vínculos entre todos los actores (Figura 4.12). Tenga en cuenta que esta información no se presenta en formato de matriz. En cambio, este formato llamaremos lista de bordes o lista de aristas. Estas enumeran un nodo de *Origen* (actor) y un nodo de *Destino* (actor), independientemente de si el enlace es dirigido o no.

4.4.2 Guardar Gráfico de Red en Gephi

Previsualización

1. Ahora, veamos como guardar nuestro gráfico de red. Para esto haga clic en el botón *Previsualización* para cambiar de espacio en Gephi. Al principio aparecerá una ventana vacía, así que haga clic en el botón *Refrescar* en la parte inferior de la ventana. Esto resultará en un gráfico de red similar a la Figura 4.13.

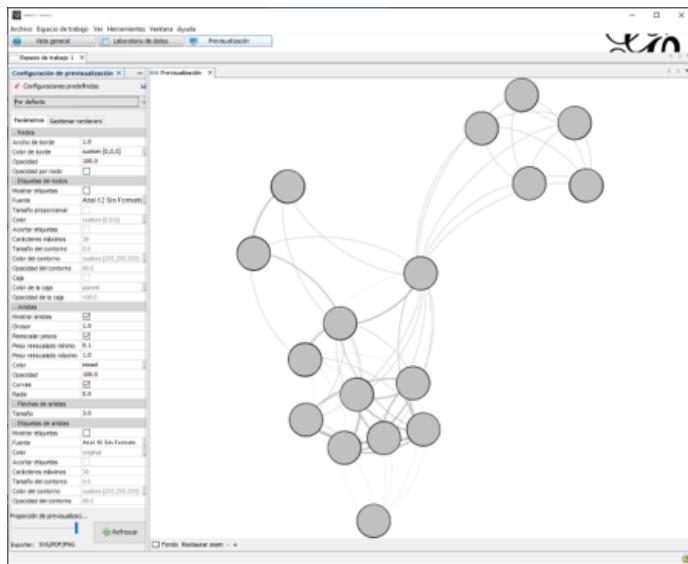


Figure 4.13: Ventana de Previsualización

*Etiquetas
de
nodos
>
Mostrar
etiquetas,
Tamaño pro-
porcional,
Fuente*

*Aristas
>
Mostrar aris-
tas, Reescalar
pesos*

2. Definitivamente debemos hacer algunos ajustes en el panel izquierdo de la ventana. Comencemos seleccionando primero la opción *Mostrar etiquetas* y deseleccionar la opción *Tamaño proporcional* (esta reducirá el tamaño de las etiquetas) en la subsección *Etiquetas de nodo*. Una vez más haga clic en *Refrescar* para ver el efecto de los cambios. Si cree que las etiquetas siguen siendo demasiado grandes, haga clic en la opción *Fuente* (en la misma subsección) y podrá reducir el tamaño de la fuente. Nuevamente, deberá hacer clic en *Refrescar* para ver el efecto de los cambios.
3. Ahora cambiemos las propiedades de las aristas. Haga clic en la opción *Mostrar aristas* (si no está autoseleccionada) en la subsección de *Aristas*. Además, seleccione la opción *Reescalar pesos*, esto reducirá el grosor de los enlaces. Tenga en cuenta que justo debajo de la opción *Reescalar pesos* se encuentran los valores predeterminados para la escalación de pesos. Si desea, puede modificar estos valores para el mínimo sea 0.2 y el máximo 2.0. Una vez más, seleccione *Refrescar*.

Aristas >
Color

4. Por último, haga clic en el cuadro de opción de *Color* en la subsección de *Aristas*. Esto abrirá un cuadro de diálogo que permite elegir el color de los bordes. Elija la opción de color *Especifico* que abrirá una paleta de colores de la cual usted puede elegir su propio color al gusto. Una vez que haya seleccionado un color, no olvide darle clic a *Refrescar*.

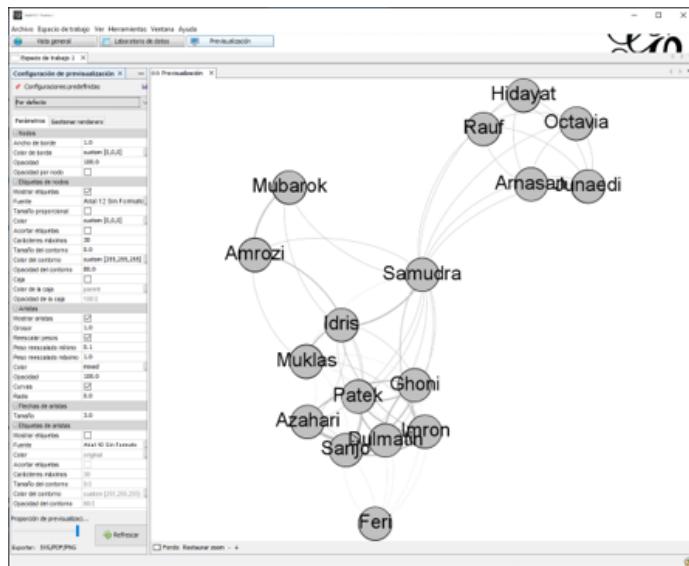


Figure 4.14: Ventana de Previsualización con Ajustes

Archivo >
Guardar como

5. Ahora, guardemos nuestro gráfico. Gephi ofrece tres opciones de formato: SVG, PDF, y PNG. En la ventana de *Previsualización* de clic en el botón *SVG/PDF/PNG* a un lado de la etiqueta *Exportar* en la parte inferior izquierda de la ventana. Por ahora, exporte su grafico en PNG. Para hacer esto, elija *Archivos PNG (.png)* en el menú desplegable. Adicionalmente, haga clic en el botón de “Opciones” en la parte inferior derecha. Esta opción abre otro cuadro de diálogo que permite aumentar el número de pixeles y la calidad de la imagen en el archivo producido. Seleccione *OK* sin cambiar las dimensiones de la imagen, cambie el nombre el archivo a producir, y guárdelo en su escritorio. Antes de pasar al siguiente paso guarde su trabajo con el comando *Archivo > Guardar como*.

4.4.3 Importación de Listas de Aristas Modo-Uno

Vista General
Archivo >
Abrir

1. Ahora, veamos como importar una lista de aristas. Hemos preparado una para este ejercicio, descárguelo `koschade_bali_edges.csv` desde aquí. Ábrala en Excel y examínela. Para importar vamos a repetir algunos pasos anteriores, use el comando *Archivo > Abrir* para abrir un cuadro de dialogo desde el cual puede importar el archivo de csv. Haga clic en *Abrir* y aparecerá un nuevo cuadro de diálogo, similar a la Figura 4.15. Tenga en cuenta que, una vez más, tenemos la opción de importar varias estructuras de datos, pero Gephi ha auto detectado que el archivo es una lista de aristas, de clic en *Siguiente*. En la siguiente venta, Gephi ofrece opciones para indicar que tipo de datos se incluirán, por ejemplo, la columna **Peso**. Ya que los valores predeterminados son adecuados, haga clic en *Finalizar*.

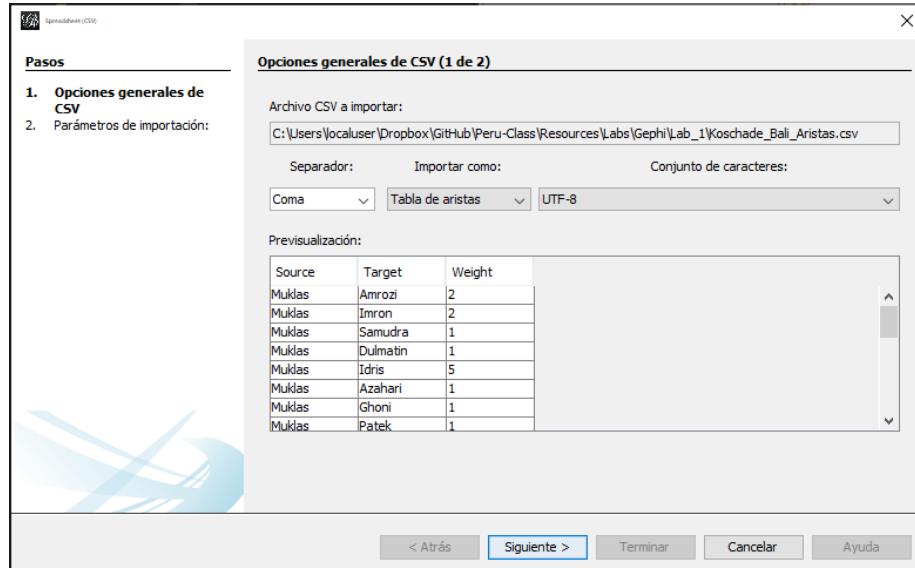


Figure 4.15: Cuadro de Diálogo de Importación

Laboratorio de datos

Tabla de datos > Nodos

2. Ahora, deberá ver un cuadro de informe de importación de Gephi. Dígale a Gephi que la red no es dirigida, seleccione la estrategia *Mínima* de combinación de aristas y haga clic en *Aceptar*. Ahora que hemos terminado de importar, cambie la ventana a *Laboratorio de datos* y haga clic en *Tabla de datos > Nodos*. Tenga en cuenta que no hasta ahora no hay etiquetas adjuntas a los nodos, solo identificadores (*Id*).

Copiar datos a otra columna

> Id

Label > Ok

3. Para agregar etiquetas a estos nodos vamos a copiar los identificadores de cada nodo a la columna de etiquetas (*Label*). Seleccione la función *Copiar datos a otra columna* (botón ubicado en la parte inferior de la ventana) y seleccione *Id*. En el cuadro de diálogo que aparecerá, seleccione *Label* y haga clic en *Ok*.

Vista general

Archivo >

Abrir

4. Ahora, importemos los datos de atributos, específicamente *roles*. Seleccione el comando *Archivo > Abrir* y lo calice y descargue el archivo *koschade_bali_nodes.csv* aquí. En el cuadro de diálogo, indique que es una *Tabla de nodos* en la opción *Importar como:* y de clic en *Siguiente*. Acepte los valores predeterminados de Gephi y de clic en *Finalizar*. Acepte los valores predeterminados, excepto que esta vez en *Informe de importación* cambie la opción de *Nuevo espacio de trabajo* a *Anadir al espacio de trabajo existente* y de clic en *Aceptar*.

Laboratorio de datos

5. Los nuevos atributos deben aparecer en la bajo la pestaña de *Nodos* en el *Laboratorio de datos*. Allí deberá ver una nueva columna de *Role*, que contiene una serie de números indicativos de los diferentes roles de los actores en la red:
 - 1 = Equipo de Comando
 - 2 = Fabricantes de Bombas
 - 3 = Asistente Operacionales
 - 4 = Equipo Lima
 - 5 = Bombarderos Suicidas

*Vista general
Apariencia
Nodos > Partición > Color
> -Escoge un atributo >
role*

6. Cambie a la ventana de *Vista general* y modifique el gráfico como lo hizo anteriormente. Esta vez coloree los nodos según los nodos de los actores. Para hacer esto, primero seleccione *Nodos > Partición > Color* en la ventanilla de *Apariencia*. Utilizando el menú desplegable *Elija un atributo*, seleccione *rol* y haga clic en *Aplicar* (consulte la Figura 4.16). De forma predeterminada, el color de los arcos reflejara el color del actor de origen. Noten que puede cambiar el color de todas las aristas en el menú de *Aristas* a un lado de *Nodos*.

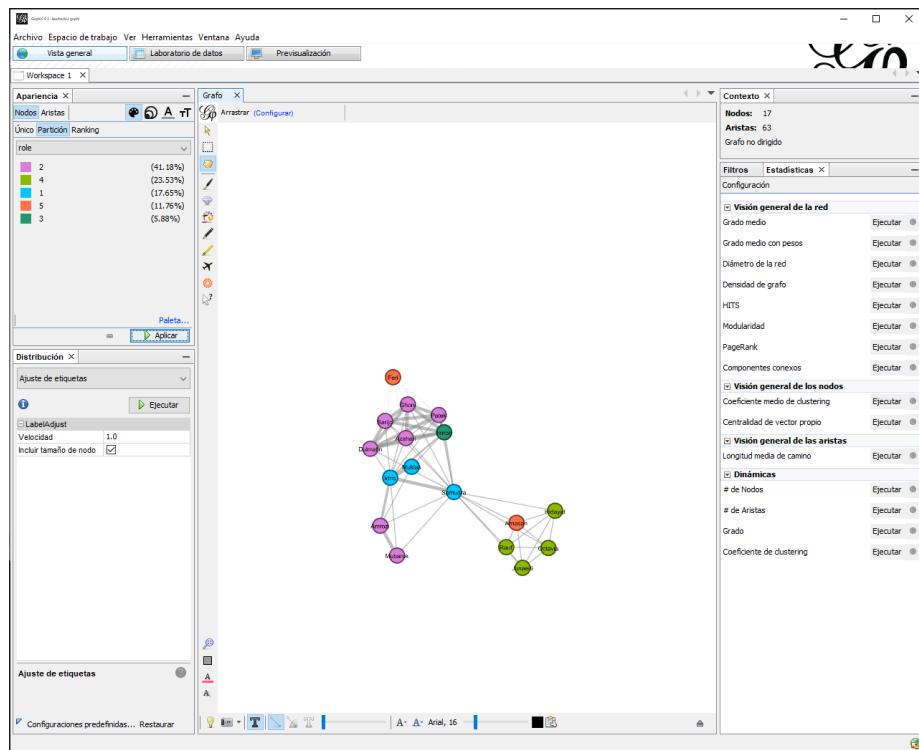


Figure 4.16: Red Koschade, Color de Nodos = Rojo

Chapter 5

Derivando Datos Modo-Uno de Modo-Dos

En este capítulo volveremos a ver datos de modo-dos, donde el objetivo principal es pasar más allí de la definición y explorar como estos sirven de precursores a relaciones de modo-uno. A lo largo de este, retornaremos a la multiplicación de matrices, si no se siente cómodo con este proceso favor de repasar el capítulo previo. Concluiremos con un ejercicio práctico en Gephi.

5.1 Modo-Dos, ¿Qué es?

Para recapitular, las redes de modo-dos son aquellas que constan de dos grupos de nodos y donde los enlaces solo existen entre nodos de diferentes clases. A veces se les conoce como redes de afiliación, redes de membresía o incidencia (Everton, 2012a). Imagínese una red modo-dos que conecta personas a eventos, por ejemplo, pandilleros conectados con pandillas u otros grupos ilícitos. Bajo la definición previa, los vínculos solo conectan pandilleros a organizaciones como en la Figura 5.1.

5.2 Proyección

El encontrar datos de modo-dos en ARSo es común. Tradicionalmente el objetivo del análisis es mapear el espacio social entre miembros de una sola clase de nodo. En la figura previa, un objetivo sería derivar las conexiones entre personas con base a sus membresías a organizaciones. Similarmente, es posible derivar relaciones entre organizaciones para entender las dinámicas organizacionales. El derivar relaciones de modo-uno empezando con datos de modo-dos es una de las herramientas más comunes y relevantes utilizadas por analistas de ARSo. Este proceso se conoce como **proyección**.

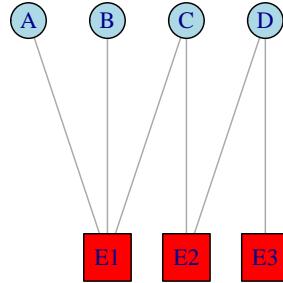


Figure 5.1: Gráfico Bipartita

Para lograr esto, retornaremos a la multiplicación de matrices. Primero, examine la matriz de adyacencia (5.2) con los datos de la Figura 5.1. Note que los vínculos solo conectan personan con eventos, por consiguiente, asumimos que lo que tenemos es una red bipartita.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E1</i>	<i>E2</i>	<i>E3</i>
<i>A</i>	—	0	0	0	1	0	0
<i>B</i>	0	—	0	0	1	0	0
<i>C</i>	0	0	—	0	1	1	0
<i>D</i>	0	0	0	—	0	1	1
<i>E1</i>	1	1	1	0	—	0	0
<i>E2</i>	0	0	1	1	0	—	0
<i>E3</i>	0	0	0	1	0	0	—

Otra forma de representar los datos de una red bipartita, además de una lista de arista, es una matriz de incidencia. En esta, vemos las que las filas y las columnas contienen las diferentes clases. Por ejemplo, en la matriz de incidencia (5.2) encontramos los eventos como columnas y las personas como filas.

	<i>E1</i>	<i>E2</i>	<i>E3</i>
<i>A</i>	1	0	0
<i>B</i>	1	0	0
<i>C</i>	1	1	0
<i>D</i>	0	1	1

Para multiplicar matrices estas tienen que ser conformables, es decir, es decir el número de filas en la primera matriz es igual a el número de columnas en la segunda. Es por ello que para multiplicar la matriz de incidencia (5.2) consigo misma es necesario transponerla. Supongamos que la matriz de incidencia es M , en esta situación el producto de la multiplicación de matrices sería $M \times M^T$, como se ve representado en la ecuación (5.2).

$$\begin{array}{cc} & \begin{matrix} E1 & E2 & E3 \end{matrix} \\ \begin{matrix} A \\ B \\ C \\ D \end{matrix} & \times \begin{matrix} A & B & C & D \\ E1 & 1 & 1 & 1 & 0 \\ E2 & 0 & 0 & 1 & 1 \\ E3 & 0 & 0 & 0 & 1 \end{matrix} \end{array}$$

El proceso de multiplicar dos matrices conformables es el mismo que hemos visto anteriormente. Empezando en el lado superior izquierdo de la primera matriz, multiplicamos cada célula de la primera matriz por los valores en cada célula de la primera columna de la segunda matriz, sumamos los resultados y proseguimos con cada fila de la primera matriz multiplicada por cada columna de la segunda (Hanneman & Riddle, 2005). La ecuación (5.2) incluye los resultados de dicho proceso.

$$\begin{array}{cccc} A & B & C & D \\ A (1x1)+(0x0)+(0x0) & (1x1)+(0x0)+(0x0) & (1x1)+(0x1)+(0x0) & (1x0)+(0x1)+(0x1) \\ B (1x1)+(0x0)+(0x0) & (1x1)+(0x0)+(0x0) & (1x1)+(0x1)+(0x0) & (1x0)+(0x1)+(0x1) \\ C (1x1)+(1x0)+(0x0) & (1x1)+(1x0)+(0x0) & (1x1)+(1x1)+(0x0) & (1x0)+(1x1)+(0x1) \\ D (0x1)+(1x0)+(1x0) & (0x1)+(1x0)+(1x0) & (0x1)+(1x1)+(1x0) & (0x0)+(1x1)+(1x1) \end{array} =$$

$$\begin{array}{ccccc} & A & B & C & D \\ & 1 & 1 & 1 & 0 \\ A & 1 & 1 & 1 & 0 \\ B & 1 & 1 & 2 & 1 \\ C & 0 & 0 & 1 & 2 \\ D & & & & \end{array}$$

Después de remover la diagonal, vínculos con uno mismo, el resultado es una matriz de adyacencia entre personas (5.2). Esta es simétrica y por lo tanto usualmente tratamos estos datos como no dirigidos puesto que es imposible determinar que actor inicio el enlace y quien lo recibe. La Figura 5.2 representa estos datos relacionales. Note que esta representación de la red solo contiene datos de modo-uno, por consiguiente, es posible “medir” el espacio social entre los actores. El ejemplo más simple de ello es la centralidad de grado, la cual es un conteo del número de vínculos de un actor y tiende a representar actividad, es decir, mientras más vínculos más activo es un nodo. En la Figura 5.2 el actor C tiene una centralidad de grado equivalente a 3, es decir que es el nodo más activo en participar en eventos con otros nodos.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>A</i>	0	1	1	0
<i>B</i>	1	0	1	0
<i>C</i>	1	1	0	1
<i>D</i>	0	0	1	0

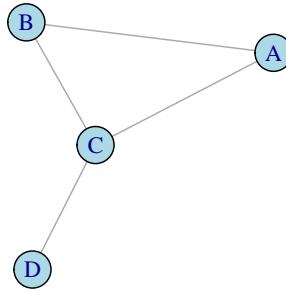


Figure 5.2: Red proyectada, persona-a-persona

Recuerde que el orden de las matrices en una multiplicación impacta el resultado. Por ejemplo, si la primera matriz en la multiplicación es la transpuesta M^T , el producto de dicha operación sería $M^T x M$, ilustrado en la ecuación (5.2).

$$\begin{array}{cccc|ccc}
 & A & B & C & D & E1 & E2 & E3 \\
 E1 & 1 & 1 & 1 & 0 & A & 1 & 0 & 0 \\
 E2 & 0 & 0 & 1 & 1 & B & 1 & 0 & 0 \\
 E3 & 0 & 0 & 0 & 1 & C & 1 & 1 & 0 \\
 & & & & & D & 0 & 1 & 1
 \end{array}$$

Una vez más, la multiplicación produce una matriz de adyacencia, pero en este caso los vínculos conectan eventos a otros eventos, como vemos en la matriz producto de la ecuación (5.2).

$$\begin{array}{ccc}
 E1 & E2 & E3 \\
 E1 (1x1)+(1x1)+(1x1)+(0x0) (1x0)+(1x0)+(1x1)+(0x1) (1x0)+(1x0)+(1x0)+(0x1) = \\
 E2 (0x1)+(0x1)+(1x1)+(1x0) (0x0)+(0x0)+(1x1)+(1x1) (0x0)+(0x0)+(1x0)+(1x1) \\
 E3 (0x1)+(0x1)+(0x1)+(1x0) (0x0)+(0x0)+(0x1)+(1x1) (0x0)+(0x0)+(0x0)+(1x1)
 \end{array}$$

	$E1$	$E2$	$E3$
$E1$	3	1	0
$E2$	1	2	1
$E3$	0	1	1

Después de remover la diagonal, la matriz puede ser dibujar como grafo (ver Figura 5.3). En este gráfico el evento 2 (E2) tiene un grado de centralidad equivalente a dos, por lo cual asumimos que es el evento más popular o asistido por más personas que han asistido a otros eventos en la red.

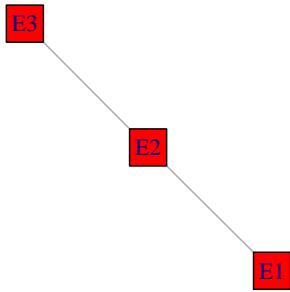


Figure 5.3: Red proyectada, evento-a-evento

Como puede ver, ambos productos de la proyección de una red bipartita pueden ser analizados por si mismos. En los siguientes capítulos de este manual presentamos una variedad de algoritmos analíticos diseñados para redes de modo-uno. Por consiguiente, antes de aplicar estas herramientas matemáticas, es crucial que el analista transforme sus datos a modo-uno.

Es probable que se esté preguntando, “¿tendré que multiplicar matrices a mano para proyectar mis datos de modo-dos a modo-uno?” La respuesta es, “no.” Gephi, y otras herramientas de software tienen la habilidad de proyectar redes modo-dos. Es por ello que a continuación tendrá la oportunidad de trabajar con redes de modo-dos a través de un ejercicio práctico en Gephi. Sin embargo, el contenido de esta sección le debe haber presentado suficiente información para entender las mecánicas del proceso que implementará el software.

5.3 Ejercicio Práctico

Los datos que usaremos en este ejercicio se les conoce como los datos de Davis: Southern Club Women. Davis y sus colegas registraron la asistencia de 18 mujeres a 14 eventos sociales diferentes. El resultado es una matriz de persona por evento, por lo tanto, cada celda (x_{ij}) es igual a 1 si la persona (i) asistió al evento social (j), y 0 en el caso contrario.

5.3.1 Antes de empezar

Herramientas
> Plugins

- Para este laboratorio requerirá el módulo (plugin) complementario para procesar datos: *MultimodeNetworks Transformation*. Para comprobar si lo ha instalado previamente, cheque utilizando los comandos *Herramientas* > *Plugins* que abrirá un cuadro de diálogo (Figura 5.4). Allí, haga clic en la pestaña *Instalados*, y si ve el módulo en esta lista, puede dar clic en *Cerrar*.

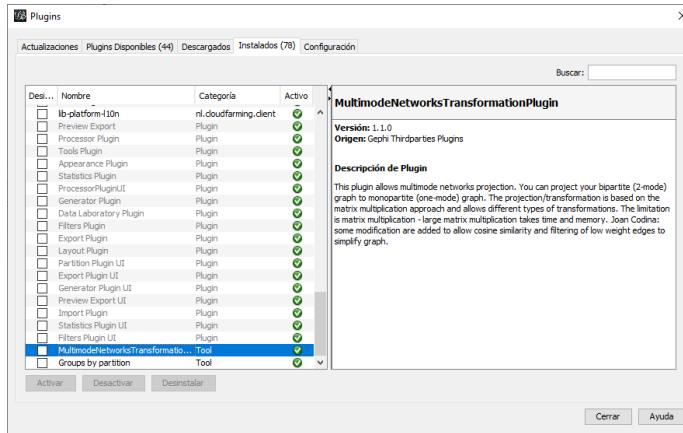


Figure 5.4: Plugins instalados

- Si no lo ha instalado previamente, haga clic en la pestaña *Plugins disponibles*, desplácese hacia abajo y localice el módulo requerido (Figura 5.5). Marque la casilla a la izquierda de la opción y haga clic en *Instalar*. Siga las instrucciones. Para que estos cambios tomen efecto deberá reiniciar Gephi.

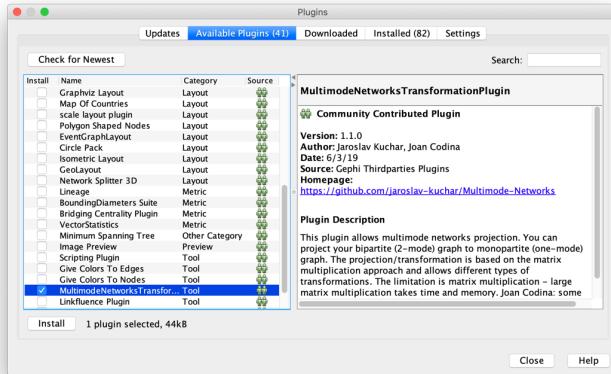


Figure 5.5: Plugins disponibles

Ventana >
Multimode
Projections

3. Después de reiniciar Gephi, asegúrese que el nuevo módulo este instalado correctamente. Para activarlo, use los comandos *Ventana > Multimode Projections*. Esta acción abrirá una nueva ventanilla (Figura 5.6). Ahora está listo para comenzar con el resto del ejercicio.

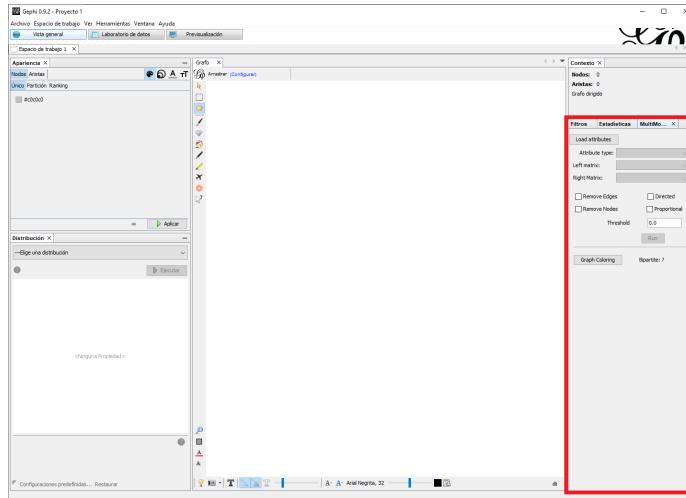


Figure 5.6: Pestaña MultiMode Projection

5.3.2 Importar Redes de Modo-dos en Gephi Utilizando Listas de Aristas

Ventana >
Multimode Projections

- Para importar datos de red modo-dos en formato de lista de bordes (o lista de aristas), usaremos dos conjuntos de datos. El primero, una lista de nodos que contiene una columna con la identificación del nodo (columna requerida, ID), etiqueta y tipo de nodo. El segundo conjunto es una lista de aristas que contiene los vínculos entre nodos. Antes de importar los datos, descargue los archivos `davis_edges.csv` y `davis_nodes.csv`.

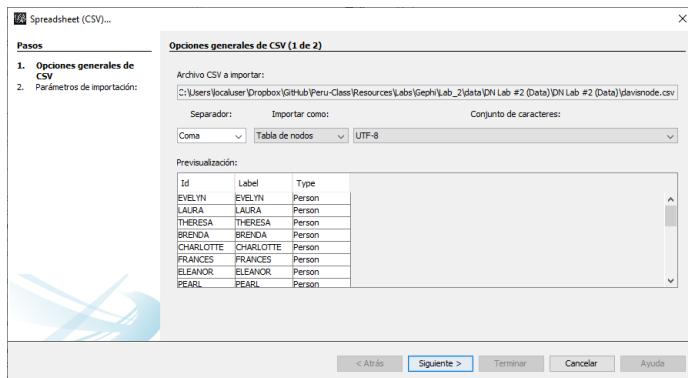


Figure 5.7: Cuadro de Diálogo: Importar Hoja de Cálculo

Archivo >
Abrir

- Primero importaremos la lista de nodos. Utilice los comandos *Archivo* > *Abrir* y luego localice y abra `davis_nodes.csv`. En el cuadro de diálogo (Figura 5.7) dígale a Gephi que el archivo es una tabla de nodos y asegure que el separador entre valores sea una coma. Haga clic en *Siguiente* y luego en el siguiente cuadro de diálogo seleccione *Terminar*. Como siempre, esto abrirá el informe de importación (Figura 5.8). Por favor indique que el tipo de grafo es *No dirigido*, la estrategia para combinar las aristas es *Mínimo*, y que desea agregar estos datos en un nuevo espacio de trabajo. Ahora, haga clic en *Aceptar*.
- Observe los datos que ha importado en la tabla de nodos en la ventana *Laboratorio de datos*. Si cambia a la ventana *Vista general* verá nodos, pero no aristas (lo cual es correcto puesto que no hemos importado los enlaces).

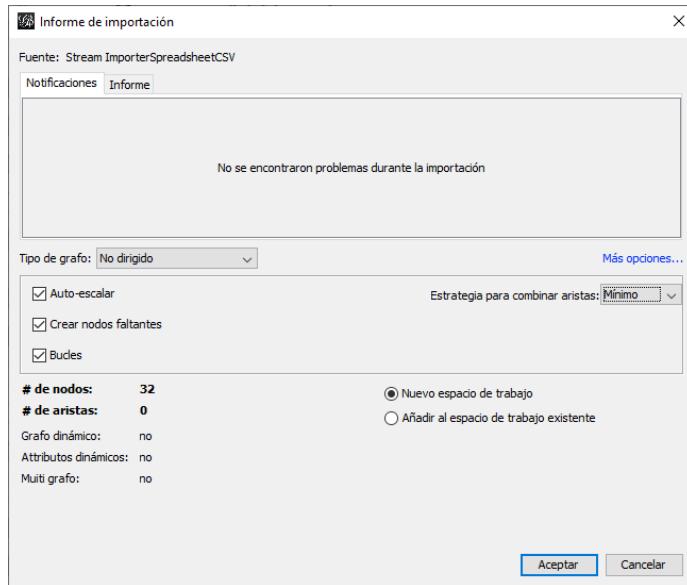


Figure 5.8: Informe de Importación para Lista de Nodos

Archivo
 Abrir

>

4. Ahora, importe la lista de aristas (`davisedge.csv`) utilizando los comandos *Archivo > Abrir*. Esta vez, importe la lista de aristas, asegurándose que Gephi sepa que se está importando este tipo de archivo (*Tabla de Aristas*). Luego haga clic en *Siguiente* y en el siguiente cuadro de diálogo seleccione *Terminar*. En el informe de importación indique que el tipo de grafo es *No dirigido*, la estrategia para combinar las aristas es *Mínimo*, y que desea añadir al espacio de trabajo existente. Finalice dando clic en *Aceptar* (Figura 5.9).

5. Puesto que la tabla de nodos incluye un atributo (`type`) que indica si un nodo es `persona` o `evento`, no necesitamos la herramienta *MultiMode Networks Projection* que descargamos para diferenciar entre clases de nodos.

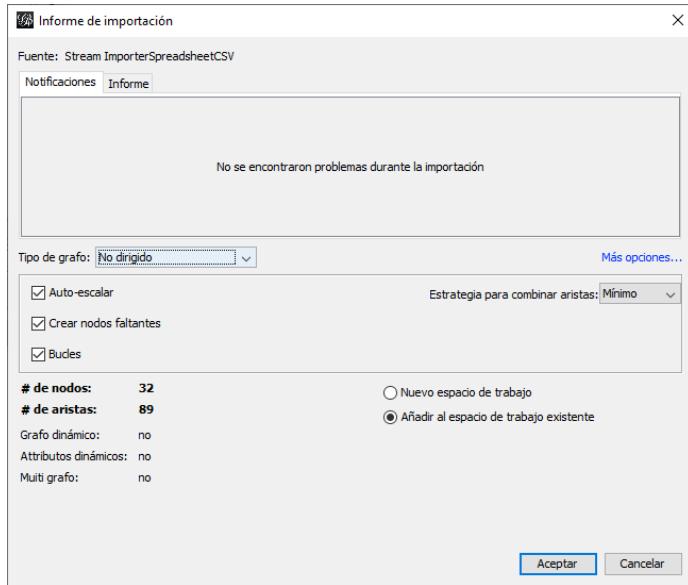


Figure 5.9: Informe de Importación para Lista de Aristas

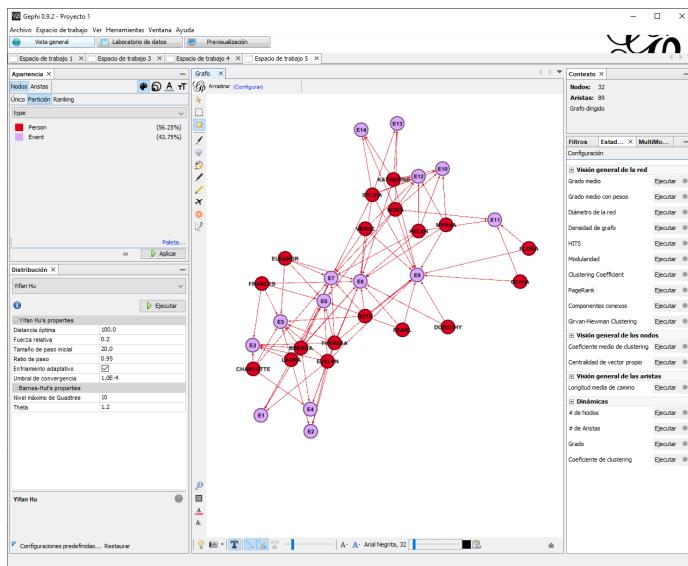


Figure 5.10: Datos Bipartitas Desde Lista de Aristas

Vista general

Apariencia > Nodos > Color > Partición > -Escoge un attribute > type

6. Cambie a la ventana *Vista general* y, en la pestaña *Apariencia*, seleccione *Nodos > Color > Partición* y, a continuación, con el menú desplegable *-Elija un atributo*, seleccione *type*. Antes de presionar *Aplicar*, haga clic en el enlace *Paleta...* justo arriba del botón *Aplicar* y elija la paleta roja/azul. Ahora presione *Aplicar*. Una vez que haya terminado, guarde sus datos, pero no cierre Gephi. En la siguiente parte, exploraremos cómo “plegar” una red de modo-dos para crear dos redes de modo-uno.

5.3.3 Plegar Redes de Modo-dos en Gephi

Vista general

MulitMode Networks Projection > Load Attributes > Attribute types > type

1. En la ventana *Descripción general*, seleccione la pestaña *MultiMode Networks Projection* queque instalamos previamente (si no lo puede localizar seleccione el menú de *Ventana* y de clic en *MultiMode Projections*). De clic en el botón *Load attributes* que cargará los atributos de los nodos. Ahora, en el menú desplegable *Attribute type* seleccione la opción *type*. Al plegar queremos crear dos redes de modo-uno, la primera entre personas y la segunda entre eventos.

Left matrix > Person-Event Right matrix > Event-Person Run

2. Comencemos creando la red “persona a persona” (mujeres conectadas solo a mujeres). Para hacer esto, modifique los menús desplegables *Left matrix* y *Right matrix* de la siguiente manera, luego de clic en *Run*.

Filtros > Atributos > Partición > type Consultas > Person > Filtrar

3. Ahora agregaremos un filtro para extraer las relaciones de entre los nodos **person**. Seleccione la pestaña *Filtros*, abra la carpeta *Atributos* y localice la opción *type* dentro de la sección *Partición*. Arrastre el filtro de *type* a la sección de *Consultas* (Figura 5.11). En respuesta a esta acción vera una nueva ventanilla al fondo de este panel (titulada *Partición (type) Configuración*), seleccione la opción *Person* y de clic en el botón *Filtrar* (Figura 5.11). Ahora exporte este gráfico (y los datos filtrados) a un nuevo espacio de trabajo dando clic en el botón *Exportar el grafo filtrado en un nuevo espacio de trabajo* (el botón resaltado por una caja roja en la Figura 5.11).

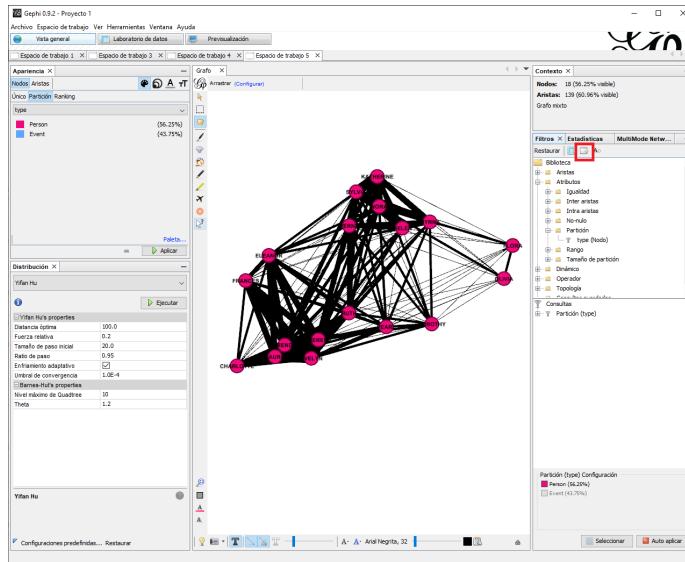


Figure 5.11: Red Bipartita Filtrada, Partición = Person

Apariencia
 > Nodos >
Tamaño >
Unique

MultiMode
Networks
Projection
Right matrix
 > *Event-*
Person Left
matrix >
Person-Event
Run

4. En el espacio de trabajo recién creado, modifique la visualización de la red usando las herramientas de diseño en Gephi. La Figura 5.12 es un ejemplo (diseño, etiquetas y el tamaño de nodos editados). Observando esta red conteste las siguientes interrogantes. **¿Qué mujeres parecen estar socialmente cerca entre sí? ¿Algunas son más centrales? ¿Qué mujeres parecen ser las más distantes del grupo?**

5. Regrese al espacio previo de trabajo donde generó la red de modo-uno. En la ventanilla de *Filtros* seleccione *Auto aplicar* (al hacer esto el botón cambiará y dirá *Fitrar*). Regrese a la pestaña *MultiMode Networks Projection*, modifique los menús desplegables *Left matrix* y *Right matrix* de la siguiente manera, luego de clic en *Run*.

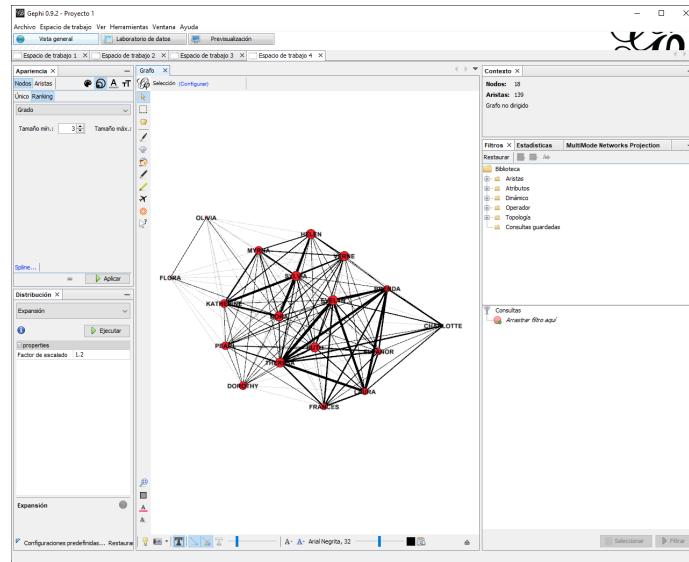


Figure 5.12: Enlaces de Persona-a-Persona

Filtros >
Atributos >
Partición >
type
Consultas
> *Event* >
Filtrar

Filtros >
Atributos >
Partición >
type
Consultas
> *Event* >
Filtrar

6. Una vez más vera nuevos enlaces en el gráfico. Ahora agregaremos un filtro para extraer las relaciones de entre los nodos *Evento*. Seleccione la pestaña *Filtros*, abra la carpeta *Atributos* y localice la opción “*type*” dentro de la sección *Partición*. Arrastre el filtro de *type* a la sección de *Consultas*. En respuesta a esta acción vera una nueva ventanilla al fondo de este panel (titulada *Partición (type) Configuración*), seleccione la opción *Event* y de clic en el botón *Filtrar*. Ahora exporte este gráfico (y los datos filtrados) a un nuevo espacio de trabajo dando clic en el botón *Exportar el grafo filtrado en un nevo espacio de trabajo*.
7. En el nuevo espacio de trabajo vera una red de modo uno que contiene relaciones de evento-a-evento. Como hicimos con la red de persona-a-persona modifique el gráfico con las herramientas de diseño en Gephi (Figura 5.13). Mirando esta red, ¿Qué eventos parecen socialmente cercanos entre sí? ¿Alguno parece más central, más popular que los demás? ¿Alguno parece ser menos popular?

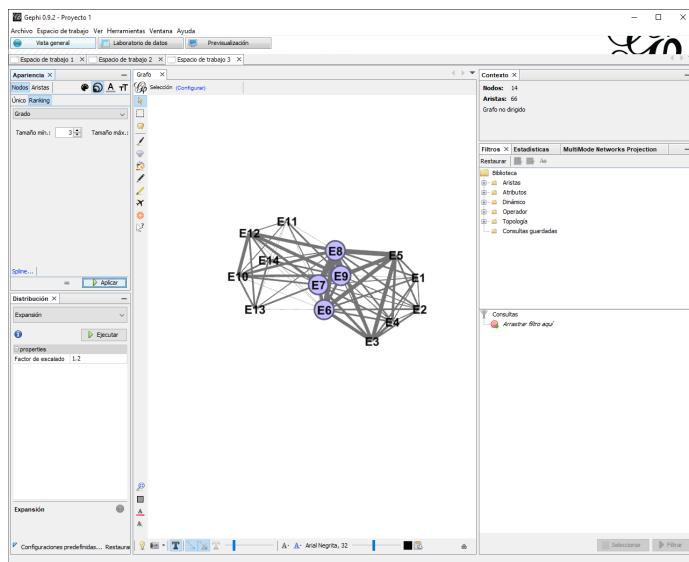


Figure 5.13: Enlaces de Evento-a-Evento

Chapter 6

Múltiples Relaciones

Hasta este punto, solo hemos trabajado con un tipo (modo-uno o modo-dos) de relaciones a la vez. Sin embargo, en el mundo real los actores incrustados en redes sociales se encuentran entrelazados los unos con otros por múltiples relaciones (Hanneman & Riddle, 2005). Visualice su red social, es decir, los círculos sociales en los que usted se encuentra involucrado. ¿Qué relaciones ha considerado? Es probable que incluya vínculos familiares, amistosos, profesionales, membresía ligas de deportes o clubes religiosos. El punto clave es que las redes sociales constan de múltiples tipos de relaciones que unen a los actores.

Por consiguiente, los analistas de redes trabajan con múltiples relaciones para poder abarcar la gama de interacciones en un sistema social. En este capítulo, cubriéremos algunas técnicas utilizadas al trabajar con múltiples tipos de relaciones, tomando pauta para recalcar las ventajas y desventajas de diferentes enfoques. Concluiremos con un ejercicio práctico en Gephi que le dará la oportunidad de utilizar estos conceptos.

6.1 Matrices y Listas de Aristas

Empecemos por considerar como trabajar con múltiples relaciones utilizando matrices. Las matrices subsecuentes ((6.1) y (6.1)) representan dos relaciones diferentes, la matriz izquierda contiene relaciones familiares y la matriz derecha financieras. Por separado, cuentan una historia diferente de los miembros de la red.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
<i>A</i>	0	1	1	0	0	0		0	0	0	0	0	0
<i>B</i>	1	0	1	0	0	0		0	0	0	0	0	0
<i>C</i>	1	1	0	0	0	0		0	0	0	1	0	0
<i>D</i>	0	0	0	0	1	1		0	0	1	0	0	0
<i>E</i>	0	0	0	1	0	1		0	0	0	0	0	0
<i>F</i>	0	0	0	1	1	0		0	0	0	0	0	0

El objetivo es analizar la combinación de estas relaciones de manera que podamos identificar los límites de la red. La Figura 6.1 es una representación de los datos descritos previamente en formato de sociomatrizes. El grafo incluye dos triadas cerradas que consisten de miembros familiares, los cuales se unen por una relación financiera.

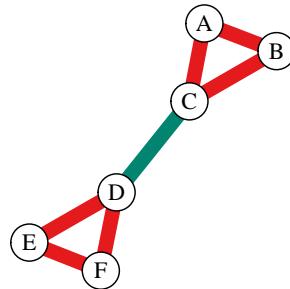


Figure 6.1: Múltiples Relaciones

Como hemos mencionado previamente, el recopilar información en sociomatrizes es laborioso e ineficiente. Por consiguiente, al recopilar múltiples tipos de relaciones es sugerible utilizar una lista de aristas, el formato básico de dos columnas puede ser expandido para incluir una columna clasificando el tipo de vínculo. Recuerde que en una lista de aristas cada fila representa un vínculo, por lo tanto, es posible apilar relaciones de la siguiente manera:

El incorporar múltiples relaciones en una sola lista con una variable categórica, permite la importación de datos de manera efectiva. Sin embargo, es común encontrar analistas que mantienen sus relaciones en múltiples archivos por

Origen	Destino	Relación
A	B	Familiar
A	C	Familiar
B	C	Familiar
D	E	Familiar
D	F	Familiar
E	F	Familiar
C	D	Financiera

propósitos de manipulación de datos.

6.2 Combinar

El combinar datos relacionales en ARSo se puede lograr mediante diferentes métodos. Cunningham, Everton, & Murphy (2016) proponen tres técnicas: apilar, agregar y recombinar. En esta sección discutiremos las tres.

6.2.1 Apilar

El apilar es una técnica comúnmente utilizada por analistas para mantener y utilizar múltiples relaciones de manera independiente (Cunningham, Everton, & Murphy, 2016). La mantener e incorporar estos archivos varía con base a la herramienta de software. Por ejemplo, las matrices subsecuentes ((6.2.1) y (6.2.1)) contienen relaciones de manera independiente, al apilarlas podemos crear un gráfico (ver Figura 6.2) que incorpora ambas sin tener que unir estos datos; por consiguiente, todos los datos se mantienen íntegros.

	A	B	C	D	E	F		A	B	C	D	E	F
A	0	1	1	0	0	0		0	0	0	0	0	0
B	1	0	1	0	0	0		0	0	0	0	1	0
C	1	1	0	0	0	0		0	0	0	1	0	0
D	0	0	0	0	1	1		0	0	1	0	0	0
E	0	0	0	1	0	1		0	1	0	0	0	0
F	0	0	0	1	1	0		0	0	0	0	0	0

El objetivo de apilar es el poder agregar o remover datos al análisis con base en las necesidades del analista. Algunas plataformas de software tratan datos apilados como separados y por consiguiente calculan las medidas de cada matriz de manera independiente. Por ejemplo, el contar vínculos de las matrices apiladas previamente ((6.2.1) y (6.2.1)) produce dos valores, 6 y 2. Es por ello, que usualmente el apilar relaciones solo se utiliza como una herramienta visual.

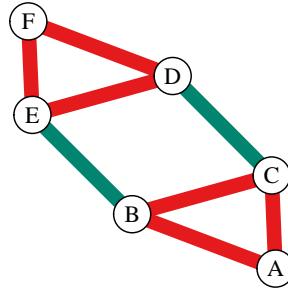


Figure 6.2: Relaciones Apiladas

6.2.2 Agregar

Como lo sugiere el nombre, agregar es una técnica que consiste en crear una nueva red de un conjunto de relaciones; usualmente, sumando las relaciones en un nuevo producto (Cunningham, Everton, & Murphy, 2016). Nuevamente, el proceso de agregar relaciones varía con base al programa de software. Sin embargo, podemos visualizar el proceso utilizando una sumación de matrices. Por ejemplo, las matrices (6.2.2) y (6.2.2) representan dos tipos de relaciones asimétricas. Note que ambas matrices constan del mismo número de filas y columnas, lo cual es requerido para sumar matrices.

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>A</i>	0	1	1	0
<i>B</i>	1	0	1	0
<i>C</i>	1	1	0	0
<i>D</i>	0	0	0	0
<i>E</i>	0	0	0	1
<i>F</i>	0	0	0	1

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>A</i>	0	0	0	0
<i>B</i>	0	0	0	0
<i>C</i>	0	0	0	1
<i>D</i>	0	0	1	0
<i>E</i>	0	1	0	0
<i>F</i>	0	0	0	0

El proceso en este ejemplo es simple, se suman las entradas correspondientes (ver (6.2.2)) y genera una nueva matriz producto. Aquí el producto de la suma se representa en la Figura 6.3. Sin embargo, diferentes herramientas de software incluyen modificaciones de esta operación como el tomar un promedio de valores correspondientes, o solo incluir el valor mínimo o máximo, etc.

$$\begin{array}{rrrr} & A & B & C & D \\ A & 0 & 1 & 1 & 0 \\ B & 1 & 0 & 1 & 0 \\ C & 1 & 0 & 0 & 0 \\ D & 0 & 0 & 1 & 0 \end{array}
 +
 \begin{array}{rrrr} & A & B & C & D \\ A & 0 & 1 & 0 & 0 \\ B & 0 & 0 & 0 & 0 \\ C & 0 & 0 & 0 & 1 \\ D & 0 & 0 & 1 & 0 \end{array}
 =
 \begin{array}{rrrr} & A & B & C & D \\ A & 0 & 2 & 1 & 0 \\ B & 1 & 0 & 1 & 0 \\ C & 1 & 0 & 0 & 2 \\ D & 0 & 0 & 1 & 0 \end{array}$$

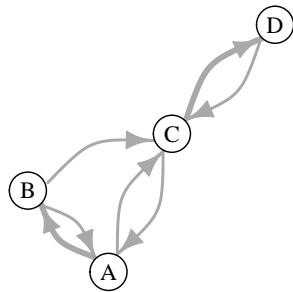


Figure 6.3: Relaciones Agregadas

El agregar redes tiene ventajas y desventajas. Permite consolidar datos antes de ejecutar medidas. Sin embargo, al consolidarlos perdemos información acerca del tipo de vínculo. Por ejemplo, en la matriz producto anterior vemos algunos valores de “2,” los cuales corresponden a la presencia de vínculos precursores en ambas matrices. Sin embargo, para los vínculos con un valor de “1” es imposible determinar a qué categoría o matriz precursora pertenece dicho vínculo sin evaluar las entradas correspondientes. Por consiguiente, los analistas de redes tienden a agregar datos y retener los conjuntos precursores.

6.2.3 Recombinar

Ocasionalmente al trabajar con datos de redes encontraremos situaciones en las cuales es necesario evaluar si la presencia de un enlace es suficiente evidencia para determinar la conexión entre dos actores. Particularmente al trabajar con redes oscuras donde las actividades se mantienen encubiertas, el triangular una relación a menudo requiere utilizar umbrales o filtros para confirmar una conexión entre actores (Cunningham, Everton, & Murphy, 2016).

Cunningham, Everton, & Murphy (2016) da un ejemplo práctico de recombinación redes. Supongamos que tenemos dos redes modo-dos. La primera (ver (6.2.3)) vincula a personas con ubicaciones. La segunda (ver (6.2.3)) vincula personas con roles operacionales.

El

	<i>Lima</i>	<i>Cusco</i>	<i>Arequipa</i>		<i>Finanza</i>	<i>Logística</i>	<i>Operativo</i>	
<i>A</i>	0	0	1		<i>A</i>	1	0	0
<i>B</i>	0	0	1		<i>B</i>	1	0	0
<i>C</i>	1	0	0		<i>C</i>	1	0	0
<i>D</i>	0	1	0		<i>D</i>	0	1	0
<i>E</i>	0	1	0		<i>E</i>	0	1	0
<i>F</i>	0	1	0		<i>F</i>	0	0	1
<i>G</i>	0	1	0		<i>G</i>	0	0	1
<i>H</i>	0	1	0		<i>H</i>	1	0	0
<i>I</i>	1	0	0		<i>I</i>	0	1	0

vincular a todas las personas en una ubicación no siempre es adecuado, por ejemplo, en una ciudad de millones de habitantes es probable que dos personas que residen en dicha ciudad no se conozcan. Similarmente, el hecho de que dos personas comparten el mismo rol operacional en una organización no significa que estas estén enlazadas. Sin embargo, podemos recombinar nuestros datos para generar una sociomatriz donde actores en la misma ubicación y con el mismo rol operacional comparten un vínculo mutuo. Para ello, utilizamos los siguientes pasos:

$$\begin{array}{ccc}
 & \begin{matrix} \textit{Lima} & \textit{Cusco} & \textit{Arequipa} \end{matrix} & \\
 \begin{matrix} \textit{A} \\ \textit{B} \\ \textit{C} \\ \textit{D} \\ \textit{E} \\ \textit{F} \\ \textit{G} \\ \textit{H} \\ \textit{I} \end{matrix} & \times & \begin{matrix} \textit{Lima} & \begin{matrix} \textit{A} & \textit{B} & \textit{C} & \textit{D} & \textit{E} & \textit{F} & \textit{G} & \textit{H} & \textit{I} \end{matrix} \\ \textit{Cusco} & \begin{matrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{matrix} \\ \textit{Arequipa} & \begin{matrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \end{matrix}
 \end{array}$$

○

$$\begin{array}{ccc}
 & \begin{matrix} \textit{Finanza} & \textit{Logística} & \textit{Operativo} \end{matrix} & \\
 \begin{matrix} \textit{A} \\ \textit{B} \\ \textit{C} \\ \textit{D} \\ \textit{E} \\ \textit{F} \\ \textit{G} \\ \textit{H} \\ \textit{I} \end{matrix} & \times & \begin{matrix} \textit{Finanza} & \begin{matrix} \textit{A} & \textit{B} & \textit{C} & \textit{D} & \textit{E} & \textit{F} & \textit{G} & \textit{H} & \textit{I} \end{matrix} \\ \textit{Logística} & \begin{matrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{matrix} \\ \textit{Operativo} & \begin{matrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \end{matrix} \end{matrix}
 \end{array}$$

El proceso es simple, primero cada matriz de incidencia ((6.2.3) y (6.2.3)) para recombinar se debe multiplicar por su matriz equivalente transpuesta, como se presento en la sección de Proyección. Puesto que el objetivo es recombinar los enlaces, el siguiente paso es multiplicar las matrices resultantes utilizando el

producto Hadamard (operación por elementos) que vimos previamente. Esta última operación genera una sociomatriz modo-uno donde los nodos se encuentran enlazados si viven en la misma ubicación y comparten el mismo rol operacional (ver (6.2.3) y 6.4).

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>	<i>I</i>
<i>A</i>	—	1	0	0	0	0	0	0	0
<i>B</i>	1	—	0	0	0	0	0	0	0
<i>C</i>	0	0	—	0	0	0	0	0	0
<i>D</i>	0	0	0	—	1	0	0	0	0
<i>E</i>	0	0	0	1	—	0	0	0	0
<i>F</i>	0	0	0	0	0	—	1	0	0
<i>G</i>	0	0	0	0	0	1	—	0	0
<i>H</i>	0	0	0	0	0	0	0	—	0
<i>I</i>	0	0	0	0	0	0	0	0	—



Figure 6.4: Relaciones Recombinar

Además de multiplicaciones de matriz, existen múltiples otras estrategias para recombinar relaciones. Sin embargo, esta discusión será continuada más adelante. Por ahora, solo nos enfocaremos en familiarizarnos con el concepto de recombinación.

6.3 Ejercicio Práctico

Los datos que utilizaremos son los datos de la red del Monasterio Sampson recopilados por Samuel Sampson, quien registró las interacciones sociales entre un grupo de hombres preparándose para unirse a una orden monástica. Grabó cuatro tipos de “lazos” entre estos actores: gusto (en tres períodos de tiempo diferentes) y disgusto (un período de tiempo); alta y baja estima; influencia positiva y negativa; alabanza y culpa. Es decir, este conjunto de datos tiene nominaciones positivas y negativas. Cada novicio clasificó sus tres preferencias para cada tipo de enlace, donde 3 indica su primera opción, 2 la segunda y 1 la tercera. Por ello, los vínculos son ponderados.

Durante el período de observación de Sampson, ocurrió una “crisis en el claustro” en respuesta a algunos de los cambios propuestos por el Concilio Vaticano II. Esto condujo a la expulsión de cuatro novicios y la partida voluntaria de varios otros. Según sus observaciones, Sampson dividió (es decir, ordenó y clasificó) a los novatos en cuatro grupos: (1) los jóvenes turcos, (2) la oposición leal, (3) los marginados y (4) los neutrales. Los jóvenes turcos llegaron más tarde y cuestionaron las prácticas del monasterio, a diferencia de la oposición leal que defendieron estas prácticas. Los marginados eran novatos que no fueron aceptados en el grupo principal, mientras que los neutrales fueron aquellos que no tomaron partido en el debate sobre las prácticas del monasterio. La mayor parte de la oposición leal había asistido a un seminario, “Cloisterville,” antes de su llegada al monasterio.

6.3.1 Importar múltiples redes a Gephi

Vista general
Archivo >
Abrir Importar como >
Tabla de nodos
Separador >
Coma

1. Para “apilar” redes en Gephi, cargaremos redes por separado. Por esto, encontramos nuestros datos de red en formato de tablas de nodos y de aristas. Comience importando la tabla de nodos de Sampson, que puede descargar aquí. Para hacerlo, use el comando *Archivo > Abrir*. Asegúrese de que Gephi reconozca que la hoja de cálculo como una tabla de nodos y separada por comas. Dado que esta hoja de cálculo solo consta de nodos y sus atributos, en el informe de importación no tiene que preocuparse por las opciones. Sin embargo, debemos ser un poco más selectivos cuando importamos las listas de borde. Después de importar la tabla de nodos, examine brevemente los datos en la ventana *Laboratorio de datos*.

Vista general

Archivo >
 Abrir Importar como >
 Tabla de aristas Separador > Coma

2. Use el comando *Archivo > Abrir* para importar el archivo **Liking 1.csv**, que puede descargar aquí . En el primer cuadro de diálogo (Figura 6.5), asegúrese de que Gephi sepa que es una tabla de aristas y que los datos están separados por comas. Haga clic en *Siguiente* y en el siguiente cuadro de diálogo (no se muestra) acepte los valores predeterminados de Gephi y haga clic en *Terminar*. En el *Informe de importación* (Figura 6.6), indique que los datos no son dirigidos, no desea mezclar los bordes, pero sí desea agregar los datos al espacio de trabajo existente. Luego haga clic en *Aceptar*.

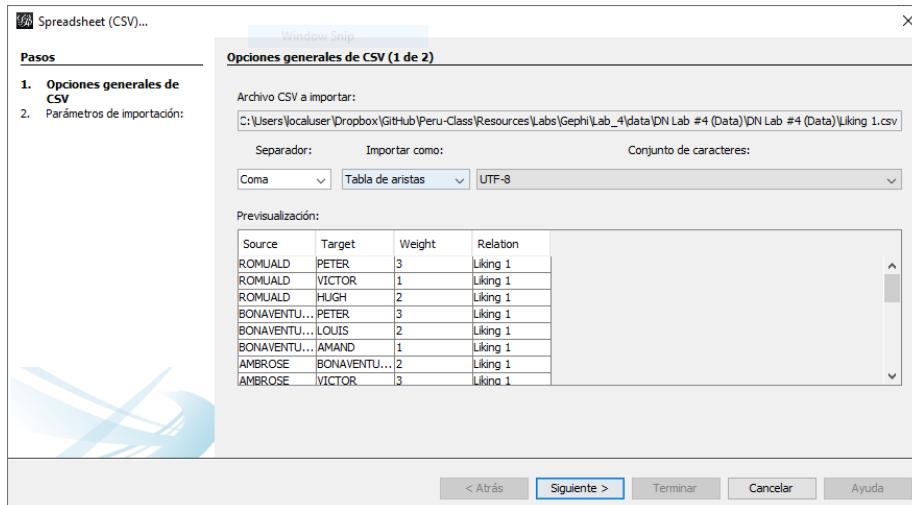


Figure 6.5: Importar lista de aristas

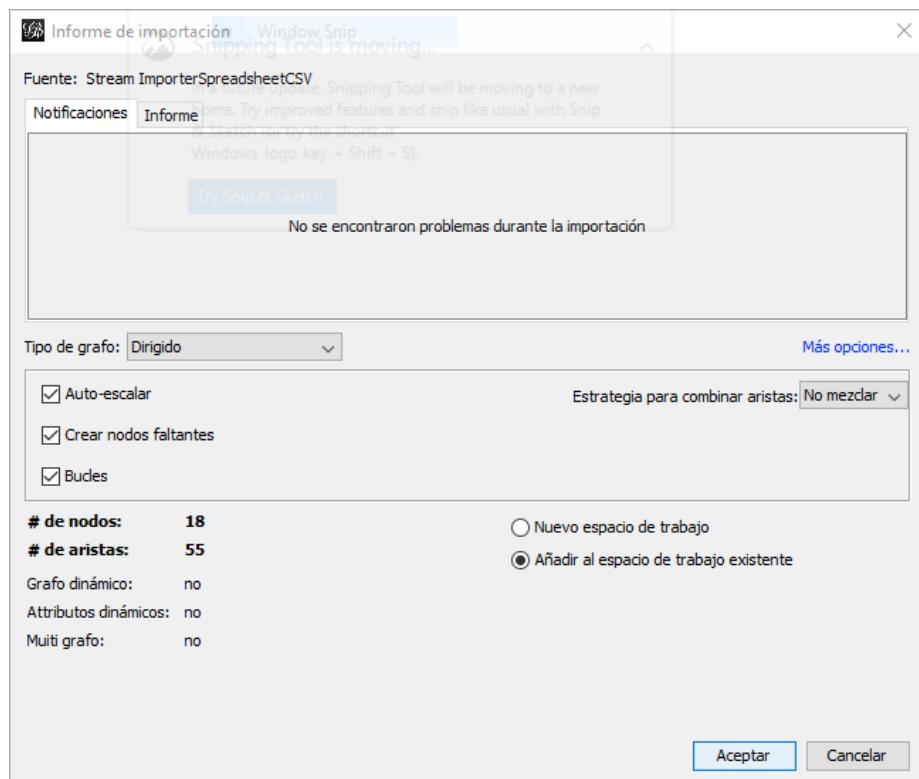


Figure 6.6: Reporte de Importación

3. Repita el proceso para importar las nueve listas de aristas restantes (descarguelas individualmente: `blame.csv`, `dislike.csv`, `highEsteem.csv`, `liking_2.csv`, `linking_3.csv`, `lowEsteem.csv`, `negativeInfluence.csv`, `positiveInfluence.csv`, y `praise.csv`). Tenga cuidado al importar cada archivo, fíjese en importar cada relación al espacio de trabajo existente. Note que en el *Informe de importación* para cada lista de aristas, Gephi mantendrá seleccionadas la mayoría de las opciones anteriores, pero no *Nuevo espacio de trabajo*. Cada vez que importe un nuevo conjunto de vínculos, el gráfico parecerá cada vez más denso. Esto se debe a que cada vez que importa una nueva relación, agrega vínculos a la red. Si selecciona la *Tabla de bordes* en la ventana *Laboratorio de datos* y luego se desplaza hacia abajo, verá que cada conjunto de vínculos (relaciones) se apilan unos sobre otros.

Vista general
Appearance >
Edges Color > Partición
-Escoge un atributo >
relation >
Apply
Paleta... >
Generar

4. En la ventana *Vista general*, ubique la pestaña *Apariencia* y luego haga clic en la pestaña *Aristas*, luego en el ícono de la paleta de colores, luego en *Partición* y luego con el menú desplegable *-Escoge un atributo*, seleccione *relation* (en español: relación) y haga clic en *Aplicar*. Recuerde que el número predeterminado de colores de partición de Gephi es ocho, pero aquí tenemos 10 tipos de vínculos (relaciones). Por lo tanto, haga clic en el enlace *Paleta ...*, y en el cuadro de diálogo, seleccione *Generar*, al igual que nosotros hizo anteriormente, cambie el número de colores a 10 y genere un nuevo conjunto de colores. Una vez más, haga clic en *Aplicar*. El gráfico de red probablemente se verá un poco colorido (Figura 6.7). Esto se debe a que cada conjunto de bordes está coloreado por una sola clasificación.
5. Ahora podríamos aplicar uno de los algoritmos de diseño para hacer el gráfico más atractivo, pero por la combinación de vínculos positivos y negativos, el gráfico podría ser engañoso. **¿Por qué?** En las siguientes secciones extraeremos diferentes conjuntos de relaciones y las examinaremos por separado. Antes de esto, guarde su trabajo en un archivo de Gephi.

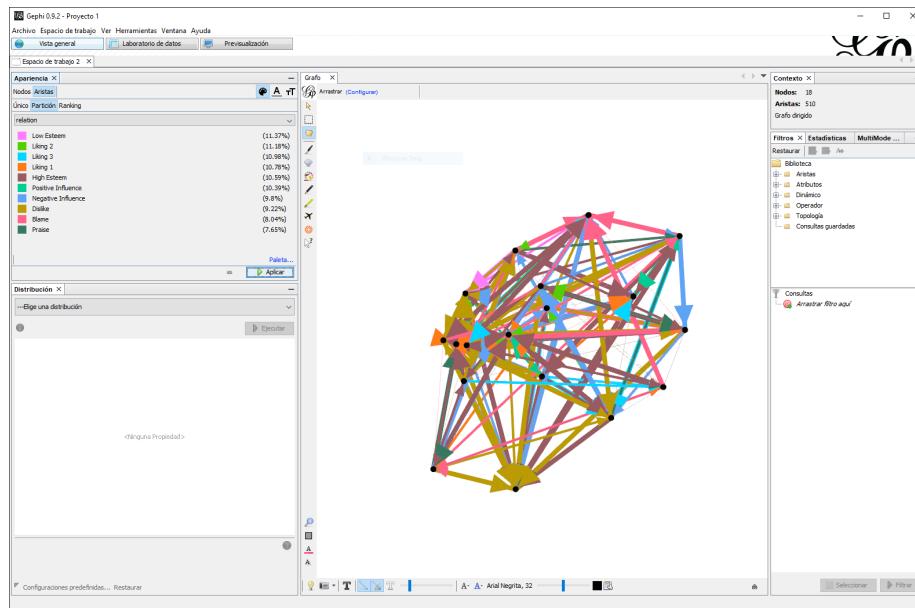


Figure 6.7: Red apilada de datos Monasterio de Sampson

6.3.2 Extrayendo y visualizando múltiples lazos en Gephi

Vista general
Filtros >
Atributos >
Partición
Consultas >
Partición >
Filtrar > *Exportar gráfico filtrado a un nuevo espacio de trabajo*

1. Extraer múltiples relaciones en Gephi es similar a extraer un subconjunto de red basado en atributos de nodo, que hemos hecho previamente. En esta sección, separaremos los enlaces positivos de la red apilada. En la pestaña *Filtros* a la derecha, primero seleccione *Atributos*, luego *Partición*. Arrastre la partición *relation* a la sección *Consultas*. En el cuadro debajo de la sección Consultas, seleccione **Liking 3** (Me gusta 3), **Positive Influence** (Influencia positiva), **High Esteem** (Alta estima) y **Praise** (Elogio). Luego haga clic en *Filtrar* y use el botón *Exportar gráfico filtrado a un nuevo espacio de trabajo*.

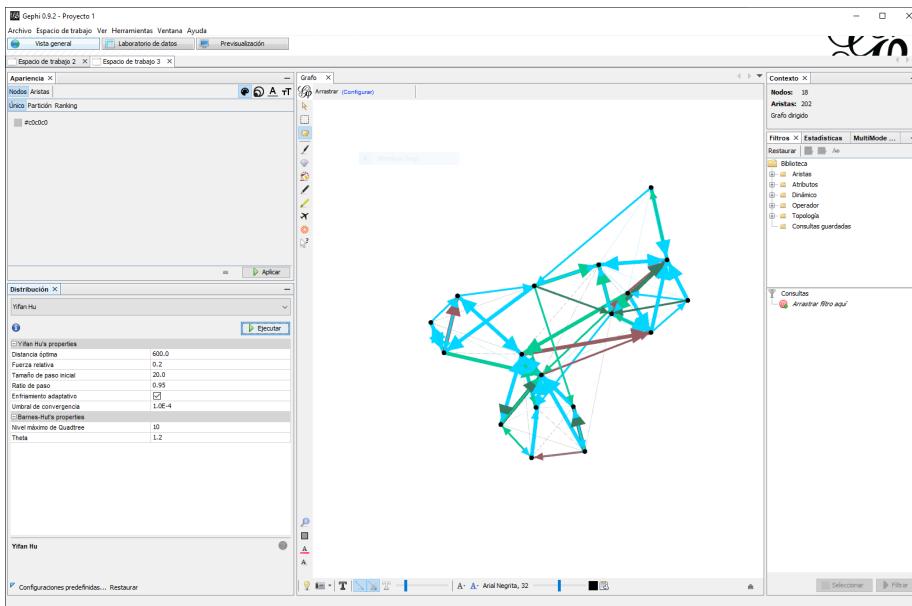


Figure 6.8: Enlaces positivos de la red Monasterio de Sampson

Vista general
Espacio de Trabajo >
Renombrar
Diseño >
Yifan Hu

2. En el nuevo espacio de trabajo, que puede renombrar **Vínculos positivos**, aplique el algoritmo de diseño Yifan Hu. Los nodos probablemente estarán demasiado apagados y será difícil distinguir cualquier patrón. Para rectificar esto, en el cuadro debajo del menú desplegable del algoritmo de diseño, cambie la *Distancia óptima* de 100 a 600 y haga clic en *Ejecutar*. Usando algunas de las otras técnicas discutidas en ejercicios prácticos anteriores, termine de modificar la red de lazos positivos en la red del Monasterio Sampson. La Figura 6.8 es un ejemplo. Como antes, el color de los lazos refleja varios tipos de relaciones. Si no le gusta el color que heredaron cuando exportó el gráfico a este espacio de trabajo, puede volver a la pestaña *Apariencia*, hacer clic en *Aristas*, luego en el ícono de la paleta de colores, luego en la opción *Particiones*, y luego con el menú desplegable *Elegir un atributo*, seleccione *relation* y haga clic en *Aplicar*. Exporte su grafico de relaciones positivas con etiquetas. **¿Qué patrones observa? ¿Qué nodos son centrales? ¿Qué significa ser central en esta red? (Tome en cuenta los tipos de enlaces).**

Chapter 7

Extraer y Simplificar Redes

En este último capítulo de técnicas de manipulación de datos, abarcaremos los temas de simplificación y extracción de subredes. Everton (2012a) menciona que a menudo al trabajar con redes sociales grandes es difícil apreciar los patrones de interacciones por el volumen de información; por esto, los analistas utilizan un conjunto de técnicas para encoger la red o extraer subconjuntos de datos con características relevantes. El objetivo de este capítulo es presentarle herramientas comúnmente utilizadas para reimaginar redes con propósito de localizar patrones emergentes. Para mantener claridad, en esta sección tocaremos cada técnica de manera individual. Antes concluir, tendrá la oportunidad de practicar el uso de estas herramientas en Gephi.

7.1 Extraer

Extraer es una técnica utilizada para aislar porciones de una red con base a un conjunto de características. Estas pueden ser características de los nodos o las aristas. Por ejemplo, en una red de pandilleros tenemos miembros asociado por diferentes tipos de enlaces, algunos de los miembros se encuentran libres y vivos, vivos y en búsqueda, y encarcelados/detenidos. La Figura 7.1 es una representación gráfica de este ejemplo. En ella los nodos libres y vivos están coloreados de negro, los vivos y en búsqueda en gris, y los encarcelados/detenidos en blanco. Por su parte los enlaces rojos son familiares, los verdes financieros, rosados amistosos y azul son de superior-subordinado.

Podemos enfocarnos en las características de los nodos y extraer una subred con base a estos atributos. Por ejemplo, si el enfoque del análisis es localizar a los nodos vivos y libres más activos, es importante extraer de la red solo nodos libres o en búsqueda. La Figura 7.2 sólo incluye nodos vivos y libres, por ello se han removido seis nodos del conjunto de datos original.

Similarmente, podemos extraer subredes con base en los patrones o tipos de

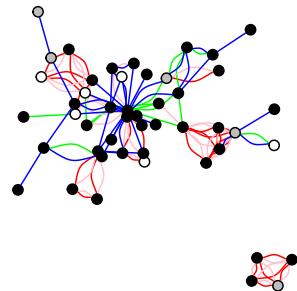


Figure 7.1: Red de Pandilleros

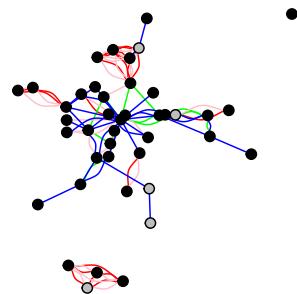


Figure 7.2: Red de Pandilleros Vivos y Libres

relaciones. Supongamos que el objetivo es aislar nodos enlazados únicamente por vínculos financieros, la Figura 7.3 representa la subred de nodos exclusivamente conectados por vínculos de finanzas.

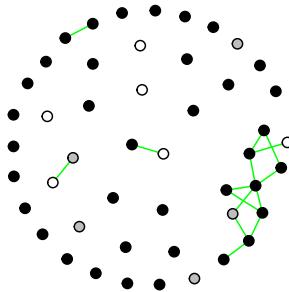


Figure 7.3: Red de Pandilleros con Lazos Financieros

7.2 Simplificar

Simplificar, tambien llamado colapsar, es una estrategia que permite el aglomerar múltiples nodos con un atributo compartido a uno solo nodo. Como sugiere Everton (2012a), esta estrategia es valiosa si el objetivo del análisis es examinar los patrones de enlaces entre tipos de nodos y no entre nodos individuales.

Una vez más, retornemos a la red de pandilleros. Esta vez, los nodos han sido coloreados con base a la afiliación de cada miembro a pandillas, Figura 7.4.

Utilizando estas categorías, podemos colapsar la red. Por ejemplo, la Figura 7.5 es una representación de la red donde cada nodo representa una categoría. Cada vínculo entre estas categorías simboliza la presencia de una relación entre dos personas con afiliaciones a dicho grupo.

7.3 Ejercicio Práctico

En este ejercicio, veremos técnicas para combinar datos relacionales con datos de atributos. Estos son útiles para resaltar diferentes aspectos de la red o para simplificar las redes. En este ejercicio utilizaremos un subconjunto de los datos principales de Noordin: “Red Operativa,” junto con los atributos asociados.

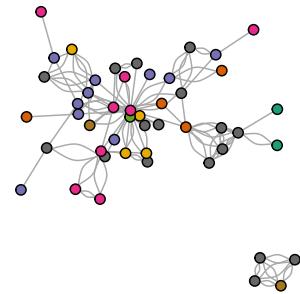


Figure 7.4: Red de Pandilleros con Nodos Coloreados por Pandilla

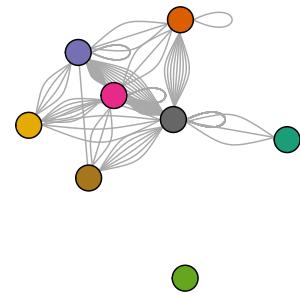


Figure 7.5: Pandilleros, Collapsados por Grupo Ilícito

7.3.1 Antes de empezar

Para poder manipular los datos adecuadamente es necesario descargar e instalar un módulo adicional de software, *Grupos by partition* ó *Grupos por partición*. Haga esto siguiendo los mismos pasos que utilizamos en para descargar el modulo *MultimodeNetworks Transformation* en ejercicio práctico en el capitulo Derivando Datos Modo-Uno de Modo-Dos.

Recuerde que deberá reiniciar Gephi.

7.3.2 Importar redes y atributos en Gephi

Vista general
Archivo > Abrir Importar como > Tabla de nodos

1. Comenzaremos importando la red operacional junto con sus atributos en Gephi. Primero, abra el archivo **Attributes (Nodes Table).csv** utilizando el comando *Archivo > Abrir* en Gephi. Este archivo contiene una lista de los nodos de red y sus atributos correspondientes. Gephi debe detectar el tipo de archivo que está importando, pero mantenga en mente que a veces no funciona correctamente. En el cuadro de diálogo de importación (Figura 7.6), asegúrese de indicar que el *Separador* sean comas y de importar el archivo como una *Tabla de Nodos*. Luego haga clic en *Siguiente*.
2. En la siguiente pantalla de importación (Figura 7.7) vera que se ha detectado una serie de atributos: *Education Level*, *Contact with People*, *Current Status (ICG Article)*, *Military Training*, *Nationality*, *Noordin's Network*, *Primary Group Affiliation*, y *Role*, junto con la identificación y etiqueta de cada nodo (actor) en la red. Gephi probablemente importará de manera predeterminada todos los atributos como *String* (caracteres de texto). Por lo tanto, usando los menus desplegables debajo de cada atributo, cámbielo de *String* a *Integer* (números enteros). Luego, de clic en *Terminar*.
3. A continuación, debería ver el *Informe de importación* de Gephi (Figure (Figura 7.8)). Aquí, indique que desea agregar los datos a un nuevo espacio de trabajo. No cambie ningún otro valor predeterminado de importación (como esta tabla no incluye aristas, ninguna de las otras opciones afectará el resultado). Haga clic en *Aceptar*.

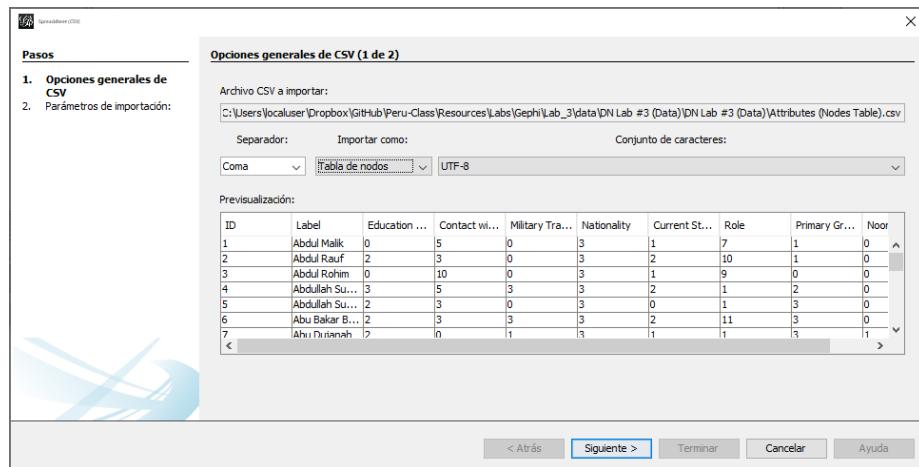


Figure 7.6: Cuadro de dialogo para importar un archivo CSV (1 de 2)

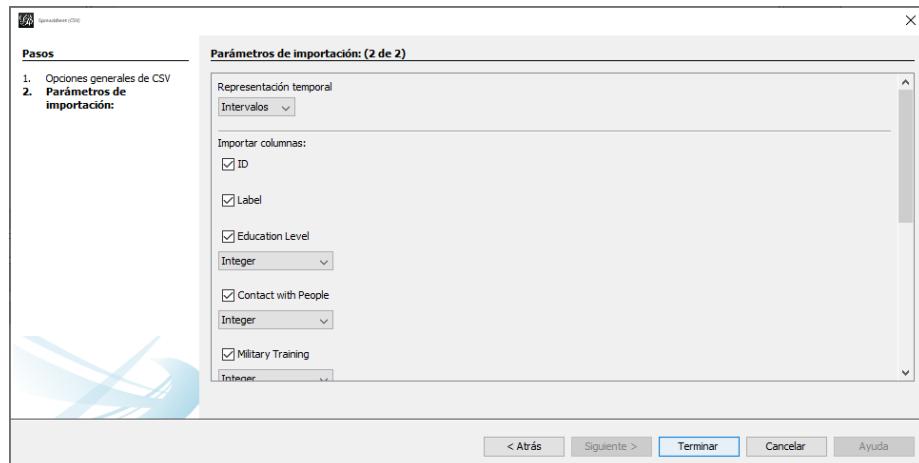


Figure 7.7: Cuadro de dialogo para importar un archivo CSV (2 de 2)

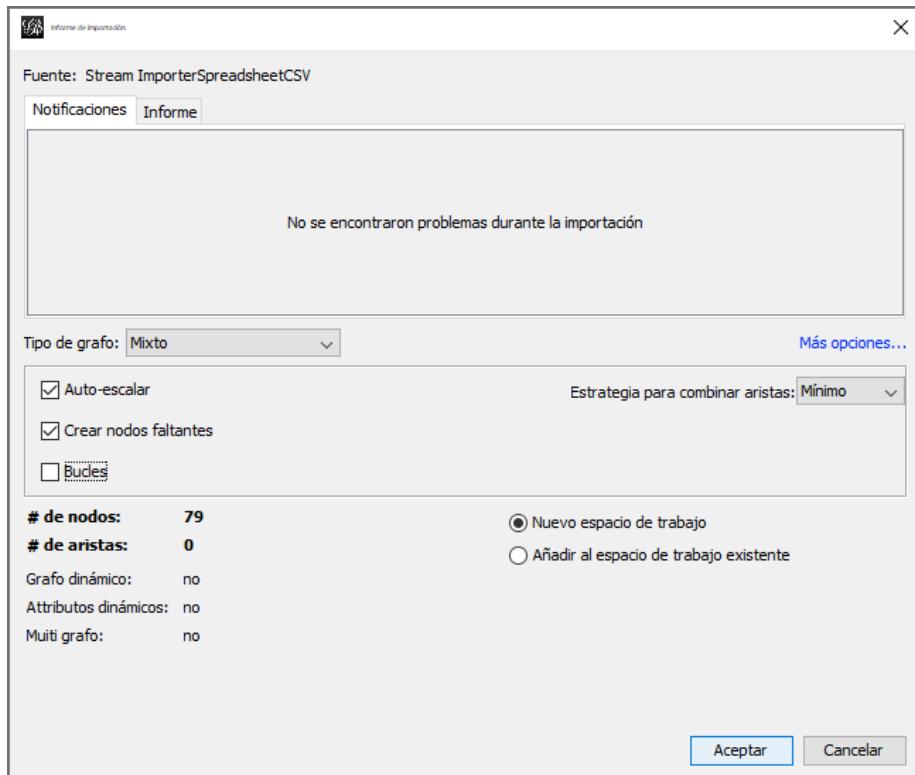


Figure 7.8: Informe de importación de Gephi

Vista general
Archivo >
Abrir Im-
portar como
> Tabla de
aristas

4. Ahora importemos la lista de aristas correspondiente, *Operational (Edge List).csv*, usando el comando *Archivo > Abrir*. Gephi debe detectar el formato del archivo correctamente, pero asegúrese que los separadores sean comas y que el archivo de importación sea reconocido como *Tabla de aristas* (Figura 7.9). Luego haga clic en *Siguiente*. En el siguiente cuadro de diálogo, acepte todos los valores predeterminados de Gephi y haga clic en *Terminar*.

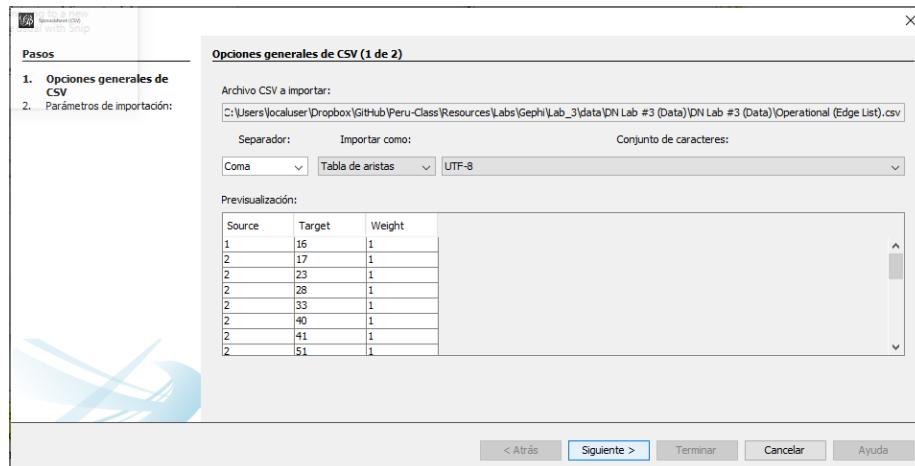


Figure 7.9: Cuadro de diálogo para importar un archivo CSV

5. Una vez más, deberá ver el informe de importación de Gephi (Figure 7.10). Aunque, es importante que le indique a Gephi que el grafo no es dirigido, que desea utilizar la estrategia para combinar aristas *Mínimo* y que desea agregar los datos al lugar de trabajo existente. Note que el informe indica que solo hay 68 nodos, sin embargo, el reporte de importación de los atributos (Figure 7.8) indica que hay 79. ¿Por qué? La respuesta es que la red contiene 11 nodos aislados, estos son, individuos que no tienen enlaces a otros en esta red. Haga clic en *Aceptar*. Antes de comenzar, guarde sus datos como archivo de Gephi.

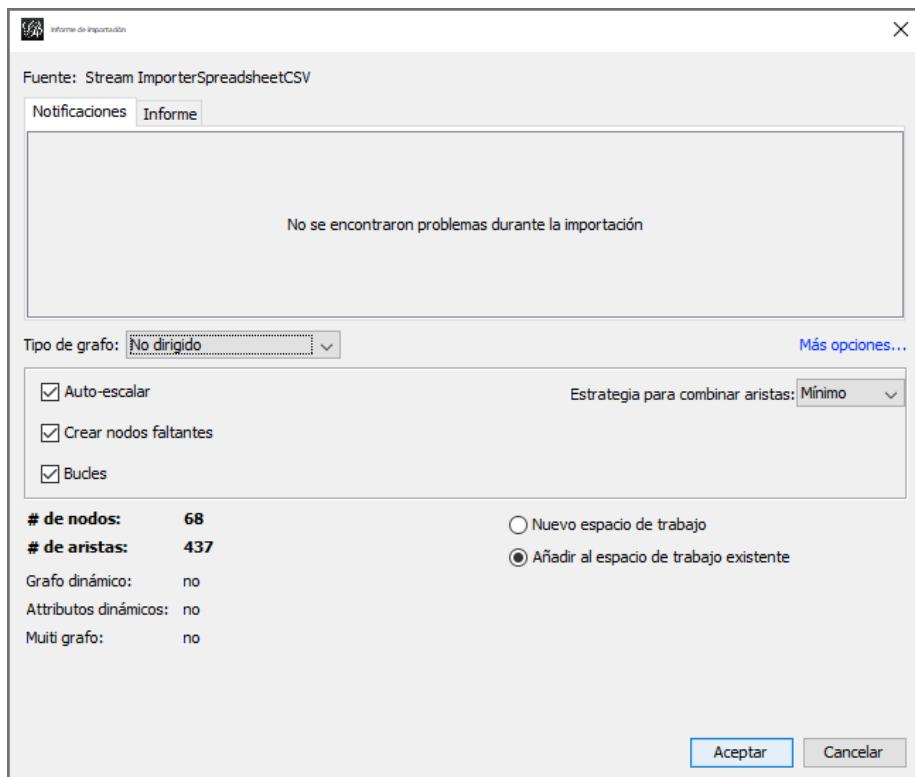


Figure 7.10: Reporte de importación

7.3.3 Uso de atributos de nodo en visualizaciones de red

Vista general
Apariencia
> Nodos >
Partición >
-Escoge un
atributo >
Current status
(icg article)

1. Ahora, veamos cómo podemos incorporar algunos de los datos de atributos de nodo en nuestras visualizaciones de red. En la pantalla *Vista general*, ubique la pestaña *Apariencia*, luego haga clic en *Nodos* y el ícono de la paleta de colores a la derecha, y finalmente en la opción *Partición*. Usando el menú desplegable de atributos, seleccione **current status (icg article)** y de clic en *Aplicar*. Los nodos ahora deben exhibir tres colores diferentes, cada uno de los cuales refleja el estado actual de un actor (es decir, si está muerto “0,” vivo y libre “1” o en la cárcel “2”; todo esto está en el libro de códigos). El color asociado con cada atributo se encuentra justo debajo del menú desplegable. Si no le gustan los colores predeterminados, puede dar clic en el cuadro junto al número o en el enlace *Paleta...*, que abrirá una ventanilla con múltiples opciones.

Vista general
Apariencia
> Nodos >
Partición >
-Escoge un
atributo >
role
Paleta... >
Generate...

2. Ahora, seleccione **role** y de clic en *Aplicar*. Este atributo contiene 12 roles, desplácese hacia el final de la lista de colores asociados con los roles, notará que Gephi no posee suficientes colores para este número de categorías. Para rectificar esto, haga clic en el enlace *Paleta...* y elija *Generar...* en el cuadro de diálogo (Figura 7.11), aumente el número de colores a 12, elija uno de los ajustes preestablecidos, haga clic en *Generar* y luego *Aceptar*. En la ventana *Descripción general*, haga clic en *Aplicar*, y ahora todos sus nodos deben tener un color individual.

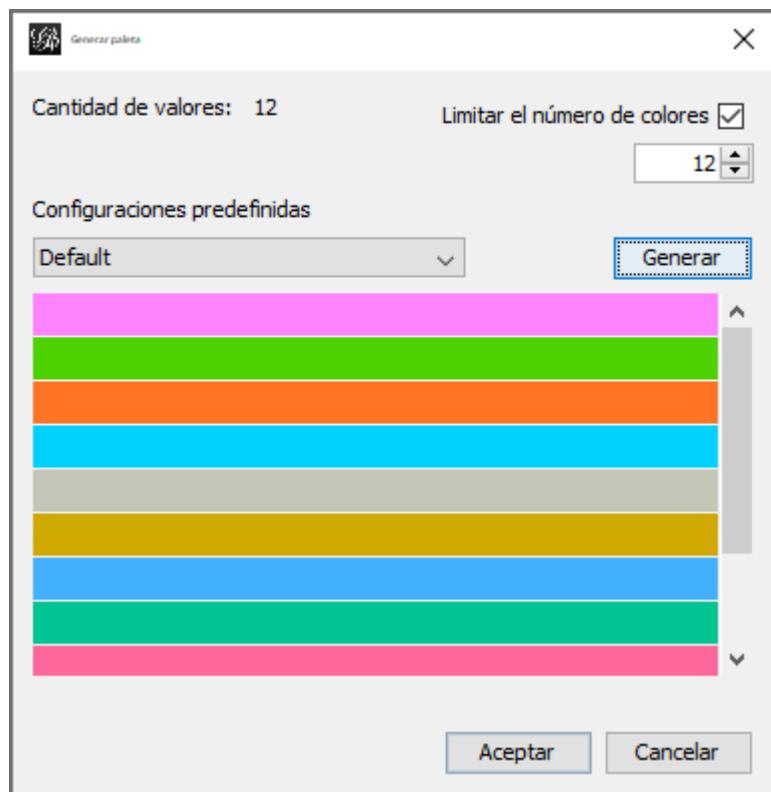


Figure 7.11: Cuadro de diálogo 'Generar paleta'

Vista general
Apariencia
> Nodos >
Tamaño >
Ranking >
Attributo >
Education
Level

3. Ahora, ajustemos el tamaño de los nodos según el nivel educativo. En la pestaña *Apariencia*, haga clic en *Nodos*, luego en el ícono de tamaño (a la derecha de la paleta de colores) y luego en *Ranking*. Usando el menú desplegable de atributos, seleccione **education level** y haga clic en *Aplicar*. El tamaño de los nodos ahora debe reflejar el nivel de educación. Los tamaños de nodo mínimo y máximo predeterminados de Gephi predeterminados constan de un rango arbitrario de valores. Si observa los valores de este atributo en la ventana *Laboratorio de datos*, verá que el nivel educativo va de 0 a 8. Para distinguir los nodos un poco mejor, ajusté el mínimo y el máximo de 3 a 27 (no usamos 0 como mínimo porque los nodos desaparecerán). Use las técnicas de visualización cubiertas en nuestros laboratorios anteriores. Este autor presenta un ejemplo en la Figura 7.12 generado utilizando el diseño *Force Atlas* porque no empuja los aislados lejos del resto de la red. **¿Detecta algún patrón relacional asociado con los atributos de los nodos (por ejemplo, los actores con niveles de educación más altos se encuentran en el centro de la red)?**

7.3.4 Simplificación (colapso / reducción) en Gephi

Vista general
Herramientas
> Generate
groups by
partition

1. Solo prosiga a completar esta sección si ha descargado el módulo requerido, *Grupos por partición*.
2. En la ventana *Descripción general*, seleccione el comando *Herramientas > Generate groups by partition*. Aparecerá un cuadro de diálogo que le preguntará si desea crear un nuevo espacio de trabajo o sobreescribir el actual. Elija *Crear*. Esto generará un nuevo espacio de trabajo donde los nodos han sido colapsados según el rol de cada nodo. Es importante tener en cuenta que este módulo contrae la red en base a la partición de nodos seleccionada en la ventana *Vista general*. Si hubiéramos querido colapsar la red basándose, por ejemplo, en el estado actual, entonces habríamos tenido que volver a colorear y revisualizar la red usando otra partición, siguiendo los comandos detallados en la sección Uso de atributos de nodo en visualizaciones de red.

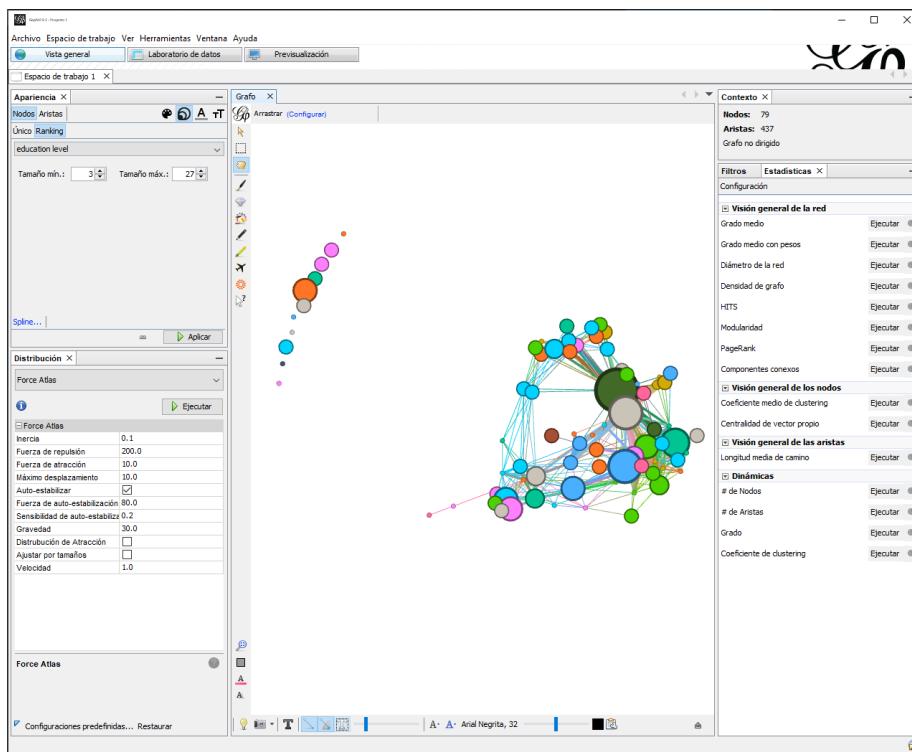


Figure 7.12: Red operacional de Noordin

3. Finalmente, necesitamos cambiar el nombre de las etiquetas para que correspondan con los diferentes roles. Cambie a la ventana *Laboratorio de datos*, seleccione la *Tabla de nodos* y verá que ahora solo hay 12 nodos correspondiendo a los grupos colapsados (por ejemplo, **Group of Jabir**). Al colapsar los nodos el nombre asociado con el grupo corresponde a uno de los actores que pertenecen a este subconjunto. Para modificar las etiquetas con los roles, vamos a tener que abrir la tabla original de atributos (**Attributes (Nodes Table).csv**) y la clave a continuación:

Código	Definición
0	Sin información/poca claridad
1	Estratega: Planificador de alto nivel
2	Fabricante de bombas: Individuo que construye bombas
3	Bombardero/combatiente: Individuo que participa en ataques con bombas o que se describe como un combatiente
4	Entrenador/instructor: Individuo que entrena o instruye a nuevos miembros
5	Terrorista suicida: Individuo que planea o ya ha realizado un ataque suicida
6	Reconocimiento y vigilancia: Participa en la vigilancia y reconocimiento
7	Reclutador: Se dedica a identificar y reclutar nuevos miembros
8	Mensajero /Intermediario: Facilita las comunicaciones entre miembros
9	Propagandista: Desarrolló campañas de información
10	Facilitador: Asiste en el funcionamiento de la red
11	Líder religioso: Proporciona formación y apoyo religioso
12	Comandante/Líder táctico: A cargo de las operaciones a nivel local/táctico

En Excel observara que el rol de Jabir es 7, y en la clave vemos que el rol 7 es el de **Recruiter** (o reclutador en español). En la columna de etiqueta (*Label*) en la tabla de nodos haga doble clic en la fila **Group by Jabir**, que abrirá un cuadro de texto. Allí, reemplace la etiqueta con el rol (**Reclutador** o **Recruiter**) y luego repita este proceso para todos los grupos restantes. Cuando haya cambiado el nombre de todas las etiquetas, la red resultante tendrá un aspecto similar a la Figura 7.13.

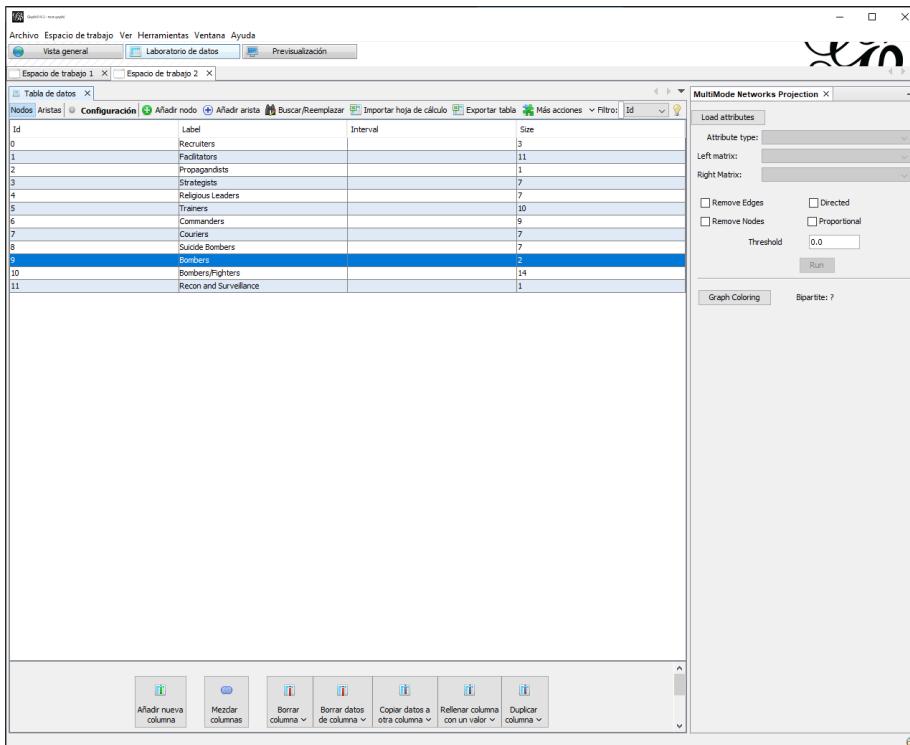


Figure 7.13: Nodos por rol

Vista General
 Nodos >
 Tamaño >
 Único

4. Ahora, cambie la Ventana *Vista general* y visualice la red. Los nodos probablemente variarán en tamaño, por lo que se recomienda modificar el tamaño utilizando los comandos *Nodos > Tamaño > Único* en esta secuencia. También use el algoritmo de diseño *Force Atlas* para editar el gráfico. Note que los pesos de las aristas reflejan en número de enlaces entre cada par de nodos, por lo que podemos reducir el grosor de estos. Además, agregue el tamaño de las etiquetas a su gusto (vea nuestra sugerencia en la Figura 7.14). Antes de pasar al siguiente paso, haga clic derecho en el nodo *propagandista* (*o propagandist*). Esta acción abrirá un cuadro de diálogo; allí, seleccione *Bloquear*. Hacer esto evita que el nodo se mueva al aplicar algoritmos de diseño adicionales.

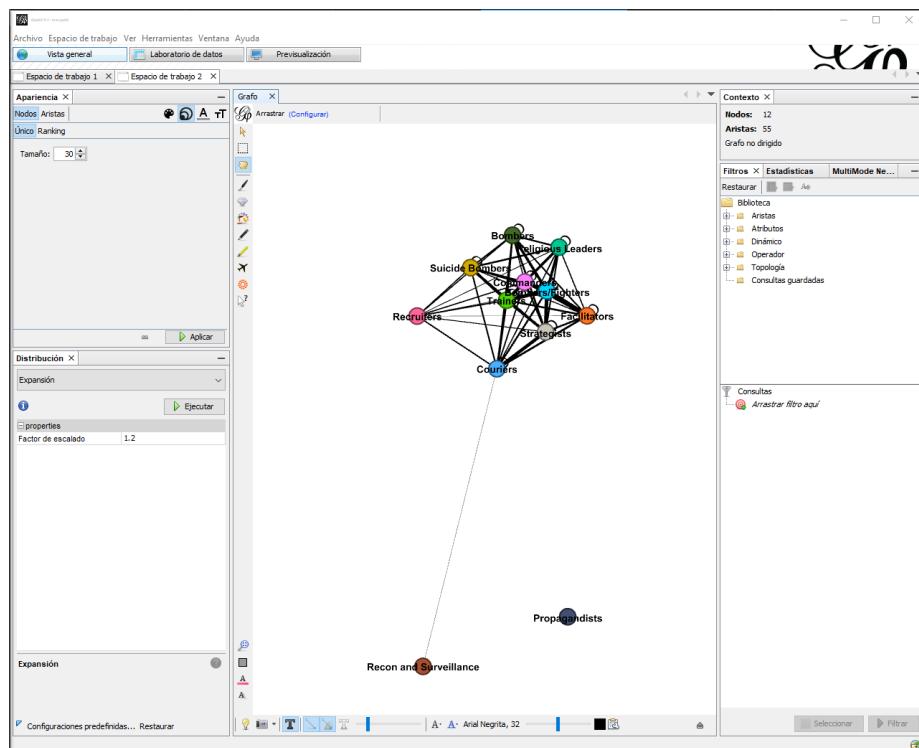


Figure 7.14: Red colapsada por atributo de rol (diseño Force Atlas)

5. Experimenté con varios algoritmos de diseño, en mi grafico (Figura 7.15) vera un sociograma modificado con el algoritmo de diseño *OpenOrd*. Mirando el gráfico (que usted mismo puede producir en Gephi), **¿Qué sugieren los patrones entre los roles?** Tenga en cuenta que algunos de los nodos tienen vínculos consigo mismos (búcles). **¿Qué crees que indican estos patrones? ¿Por qué cree que el nodo “Recon and surveillance” no se encuentra más cerca a los demás nodos?**

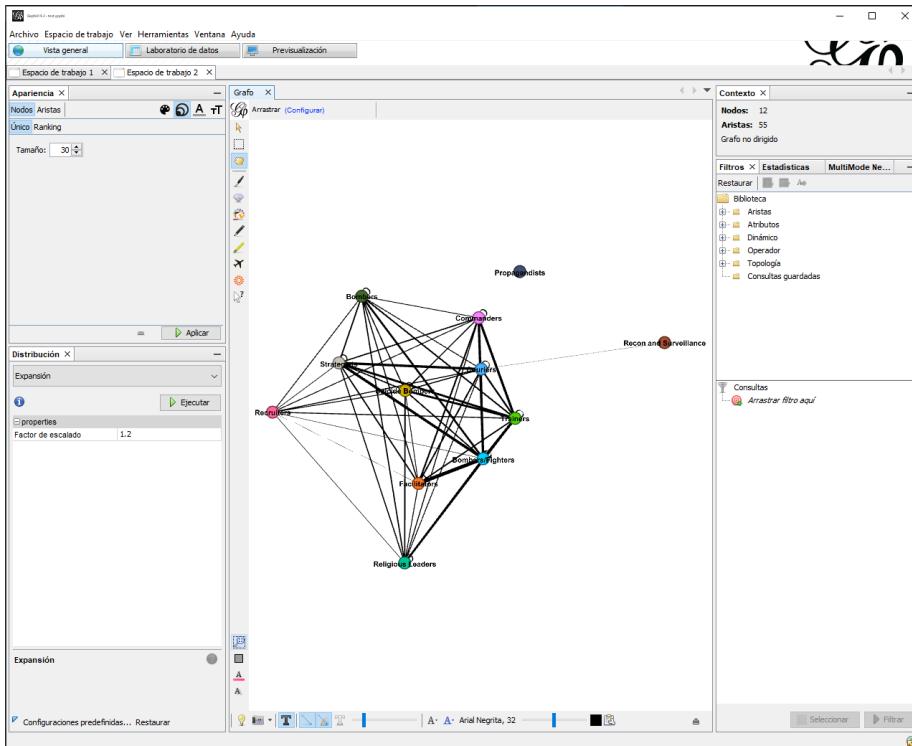


Figure 7.15: Red colapsada por el atributo de rol (diseño OpenOrd)

7.3.5 Extracción en Gephi

Vista general
Filtros >
Atributos >
Partición >
current status
Consultas >
Partición > 1
Filtrar > Exportar gráfico filtrado a un nuevo espacio de trabajo

1. Ahora, vamos a extraer un conjunto de red más grande. Para hacer esto, volvamos a la red original (sin colapsar), que probablemente se encuentra el primer espacio de trabajo. En la pestaña *Filtros* en el lado derecho de la ventana *Descripción general*, seleccione *Atributos*, luego *Partición* y luego *current status (icg articles)* a la sección *Consultas*. En el cuadro debajo de la sección *Consultas - Partición (current status (icg article)) Configuración* – deberá ver tres cuadros, etiquetados *2*, *1*, y *0*, que indican si el actor esta muerto *0*, vivo y libre *1*, o en la cárcel *2*. Aquí, extraeremos la red *Viva y libre* marcando la casilla *1*. Haga clic en *Filtrar* y luego use el botón *Exportar gráfico filtrado a un nuevo espacio de trabajo*, que utilizamos en el laboratorio anterior (cuadro rojo en la Figura 7.16).

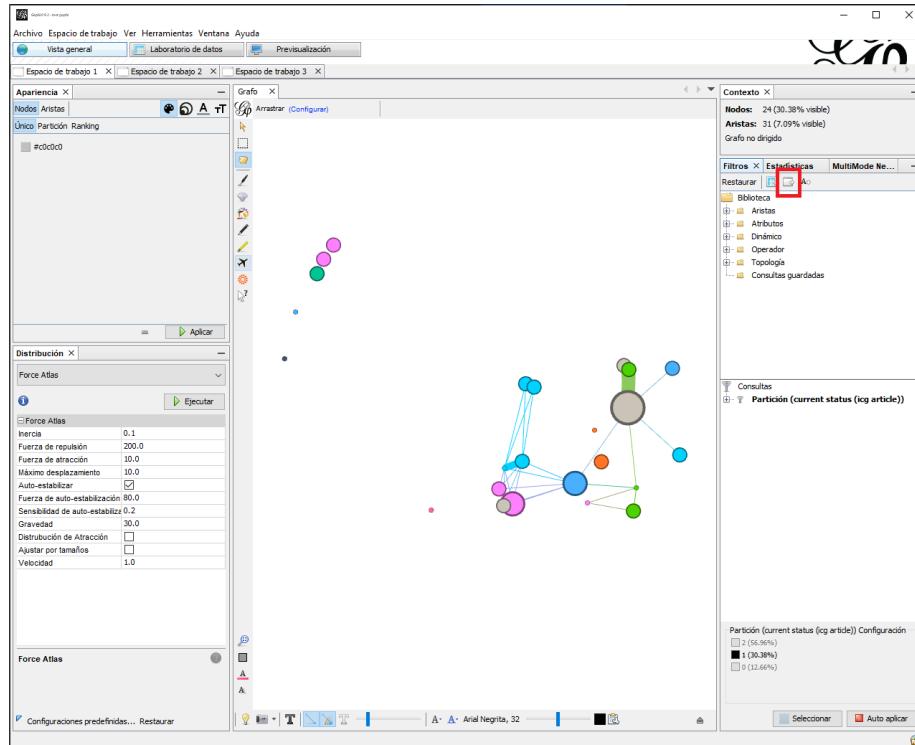


Figure 7.16: Red viva y libre filtrada

2. Cambie el nuevo espacio de trabajo y cámbiele el nombre (por ejemplo, **Red viva y libre**). Después de experimentar con algunos algoritmos de diseño, tamaño de nodo con base en el nivel educativo, tamaño de etiqueta y pesos de borde, produje el grafico en la Figura 7.17. Tenga en cuenta que vez elegí no mostrar los nodos aislados. Los colores reflejan el rol.

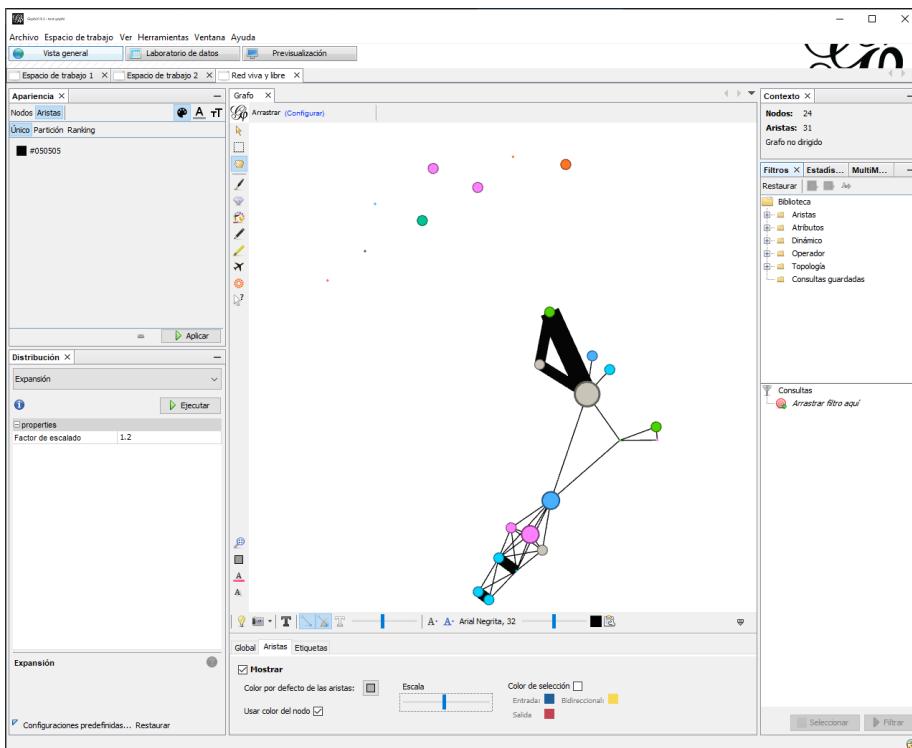


Figure 7.17: Red viva y libre (sin aislados)

Part III

ARSo: Medidas

Chapter 8

Topología

La topología consta de una familia de medidas que son calculadas sobre el conjunto completo de datos relacionales que componen una red; por consiguiente, proveen una indicación cuantitativa sobre la estructura de esta (Valente, 2010). Es relevante puesto que ciertos estudios sugieren que redes centralizadas alrededor de un conjunto de nodos centrales tienden a ser más eficientes en movilizar recursos humanos o financieros (McCormick & Owen, 2000); asimismo, esta configuración hace a la red más vulnerable a disruptión al remover un puñado de actores clave. De manera opuesta, las redes altamente descentralizadas son relativamente invulnerables a la eliminación de actores clave, pueden tener dificultades para movilizar recursos (Cunningham, Everton, & Murphy, 2016).

Observe las Figuras 8.1 y 8.2, ambas tienen la misma cantidad de nodos, pero las características estructurales son diferentes y presentan ventajas y desventajas. Por ejemplo, en la Figura 8.1 la mayoría de los nodos rodean a un actor en el centro de la red. Este tipo de configuración es comúnmente asociada con organizaciones jerárquicas con protocolos establecidos de comando y control. La ventaja de esta topología es un que un nodo central controla el flujo de recursos e información, en teoría haciendo la diseminación de estos más eficiente. Sin embargo, esta estructura tiene un defecto clave. ¿Qué pasa si removemos al actor central? La red se deshace por completo y el flujo es interrumpido de manera indeterminada.

Ahora, en la Figura 8.2 no encontramos una formación centrada alrededor de un actor clave. Al contrario, en este caso todos los nodos se encuentran conectados entre ellos. La ventaja es clara, no hay un solo punto de disruptión como en la Figura 8.1. Sin embargo, las características estructurales no sugieren un solo camino para el flujo de información o recursos, lo cual en algunas ocasiones retrasa ciertos procesos organizacionales. Es decir, si todos son líderes, nadie puede estar a cargo de dar instrucciones a los demás.

Es probable que sea mejor considerar ambas configuraciones (centralizada y

descentralizada) como polos opuestos de un continuo. Como tal, el nivel óptimo de centralización en dicho espectro cambiará con el tiempo y en función de las presiones ambientales (Cunningham, Everton, & Murphy, 2016; Everton, 2012a). Por ejemplo, en un entorno hostil, la red puede priorizar la seguridad de actores clave; así, avanzando hacia una estructura más descentralizada. Es precisamente por este dinamismo intrínseco a las redes sociales que el estudio de la estructura de red debe ir acompañada con contexto en el que se encuentra incrustada.

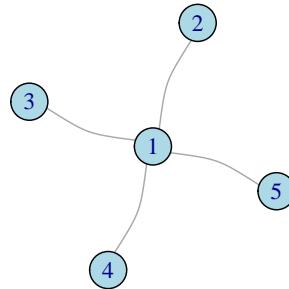


Figure 8.1: Red estrella

En este capítulo, examinaremos una serie de medidas que capturan características estructurales de la red completa. Comenzaremos con una serie de medidas básicas comúnmente utilizadas para describir una red como: tamaño, distancia y diámetro. Después, exploraremos dos dimensiones topológicas: interconexión y centralización. Esta discusión será complementada por una serie de medidas correspondientes. Finalmente, este capítulo concluirá con un ejercicio práctico en Gephi.

8.1 Medidas Básicas de Topología

Estas proveen al analista contexto sobre la red en discusión. Aquí cubriremos: tamaño, distancia promedio, y diámetro.

8.1.1 Tamaño

El tamaño se refiere a el número de actores en una red (Everton, 2012a). Es importante pues afecta otras medidas topológicas de la red. Por ello, conocer el

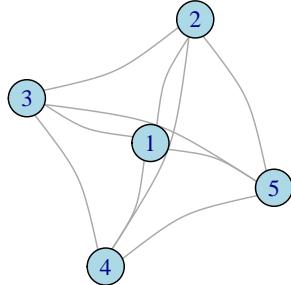


Figure 8.2: Red completa

tamaño de una red es importante al interpretar medidas basadas en este valor, como densidad o diámetro (Cunningham, Everton, & Murphy, 2016).

Comúnmente denotamos el tamaño de un a red con la letra N . Por ejemplo, en las Figuras 8.3 y 8.4 N es equivalente a 5 y 10 respectivamente.

En teoría, hay un límite práctico a la cantidad de enlaces que puede mantener cada persona en la red, por ello, el tamaño habitualmente no es una de las medidas más interesantes sino un indicador de los límites de la red (Valente, 2010). Es por ello, que el analista debe considerar como establecer los límites para reflejar de manera adecuada el tamaño real de la red (Cunningham, Everton, & Murphy, 2016).

8.1.2 Distancia Promedio

La distancia promedio se refiere a la longitud promedio de todos los caminos más cortos entre todos los actores de una red (Everton, 2012a). Para un grafo no dirigido binario G , definimos la distancia promedio como:

$$l_G = \frac{1}{N(N - 1)} \sum_{i \neq j} d(v_i, v_j)$$

Examinemos esta definición, con la Figura 8.5. Supongamos que el objetivo del actor 1 es transferir información a los otros actores en la red de manera más eficiente.

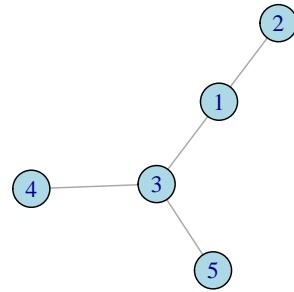


Figure 8.3: Grafo de red, tamaño = 5

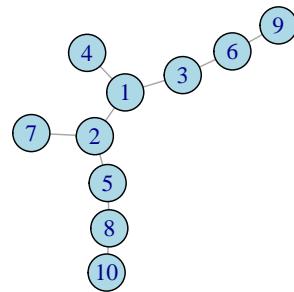


Figure 8.4: Grafo de red, tamaño = 10

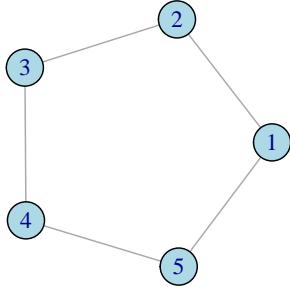


Figure 8.5: Grafo circular

Si d es la distancia entre un par de nodos, podemos expresar la distancia entre el nodo 1 los demás de la siguiente manera:

$$d(1, 2) = 1$$

$$d(1, 3) = 2$$

$$d(1, 4) = 2$$

$$d(1, 5) = 1$$

Por consiguiente, debemos calcular la distancia (d) entre cualquier nodo (i) y todos los demás en la red (j), sumar estos valores ($\sum_{i \neq j} d(v_i, v_j)$) y multiplicar este valor por la fracción del total de enlaces posibles en el gráfico ($\frac{1}{N(N-1)}$). La Figura 8.5 se vería reflejada como la siguiente operación:

$$l_G = \frac{1}{5(5-1)} \times 5(1 + 2 + 2 + 1) = 1.5$$

Al calcular la distancia promedio, podemos decir que en promedio todos los nodos en la Figura 8.5 se encuentran a 1.5 pasos entre ellos. Puesto que este es un promedio de distancias, podemos comparar este valor con el promedio de otra red.

En términos prácticos, la distancia promedio es indicativa de la velocidad con la cual la información viaja a través de una red. Es decir, en una red compacta

la información se debe difundir rápidamente; al contrario, en una red con una distancia promedio alta la difusión de información tomará lugar en un periodo de tiempo más amplio (Everton, 2012a). Esta interpretación no es ley, sino una suposición comúnmente hecha por analistas. Es importante considerar que en ciertas situaciones las redes con distancia promedio pequeñas no gozan de difusión rápida puesto que algunos actores tienden a retener información de sus vecinos (Cunningham, Everton, & Murphy, 2016). Adicionalmente, se debe recalcar que la fórmula de distancia promedio se centra en eficiencia, es decir, distancias más cortas entre nodos. Es posible que, en una red oscura donde los nodos no conozcan a sus vecinos o a los vecinos de sus vecinos, un actor envié información a través de caminos con distancias mayores al promedio. Por lo tanto, la distancia promedio solo sugiere eficiencias con base a la estructura de una red.

8.1.3 Diámetro

La distancia geodésica mayor entre todos los nodos en un grafo (Everton, 2012a). Formalmente, Wasserman & Faust (1994) definen esta medida como la máxima $d(i, j)$ o $\max d(i, j)$. Por ejemplo, podemos calcular las distancias geodésicas de la Figura 8.6:

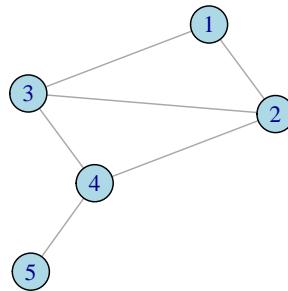


Figure 8.6: Grafo ejemplo

El diámetro del grafo es equivalente es $\max d(i, j) = d(1, 5) = 3$.

Esta medida es importante pues cuantifica cuán lejos se encuentran los nodos más distantes entre sí (Wasserman & Faust, 1994). Como con la distancia promedio, consideremos una red en la cual la información viaja a través de los vínculos

entre actores. Suponiendo que los mensajes siguen las rutas más cortas, tomará tiempo para que un mensaje recorra toda la red cuando el diámetro es mayor.

8.2 Interconexión

En una red oscura, la interconexión entre los actores es crucial para la red como organismo dinámico. Por ejemplo, puesto que la seguridad de la organización es crucial, a menudo el reclutamiento de nuevos miembros toma lugar a través de los vínculos preexistentes de miembros de la red (Cunningham, Everton, & Murphy, 2016). Es por este aislamiento que las redes oscuras enfocadas en el anonimato tienden volverse cada vez más densas con vínculos entre miembros de la red y pocos enlaces al mundo exterior. Sin embargo, el cortar vínculos con el mundo exterior conlleva un costo al limitar el acceso a recursos (por ejemplo, materiales, logísticos, financieros, etc.). Es por ello, que algunos investigadores y analistas sugieren que los vínculos a actores externos a la red oscura son esenciales para acceder recursos y sobrevivir retos (Uzzi, 1996).

Esta dinámica entre acceso a recursos y seguridad, fuerza a las organizaciones y a los actores que las conforman a balancear aislamiento (por ejemplo, interconexión alta) con acceso a recursos exteriores (por ejemplo, interconexión moderada o baja). En esta sección, nos enfocamos en las medidas comúnmente utilizadas para medir este proceso, específicamente, densidad y grado medio.

8.2.1 Densidad

La densidad (d) es el número de enlaces en un grafo dividido por el número total posible de enlaces en el mismo (Cunningham, Everton, & Murphy, 2016). En una red no dirigida, se define como:

$$d = \frac{L}{n(n - 1)/2}$$

Donde L se refiere al número actual de enlaces y n es el número de actores en la red. Para una red dirigida, no es necesario dividir el denominador entre dos, por lo que la ecuación es diferente:

$$d = \frac{L}{n(n - 1)}$$

En términos prácticos, ambas ecuaciones producen puntuaciones de densidad de 0 a 1. Es decir, una red con una densidad de 0 es una en la que no existen vínculos entre los actores. Mientras que una red con una densidad de uno es un gráfico completamente conectado en el que todos los actores están vinculados entre sí.

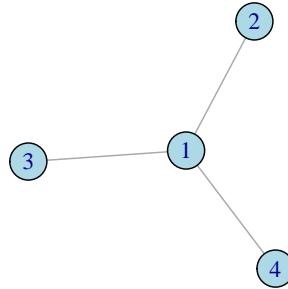


Figure 8.7: Grafo no dirigido

Pongamos esta herramienta en práctica; específicamente, con la Figura 8.7 la cual podemos representar como matriz de adyacencia (8.2.1).

$$\begin{matrix}
 & 1 & 2 & 3 & 4 \\
 1 & - & 1 & 1 & 1 \\
 2 & 1 & - & 0 & 0 \\
 3 & 1 & 0 & - & - \\
 4 & 1 & 0 & 0 & -
 \end{matrix}$$

El número actual (L) de enlaces es seis y contamos con cuatro nodos en este gráfico. Entonces, la densidad de esta red es equivalente a $d = \frac{6}{3} = 0.5$. Es decir, la interconexión de la red es 0.5 o solo la mitad de los actores están conectados.

8.2.2 Grado Medio

La medida de densidad está inversamente relacionada con el tamaño de la red, puesto que al agregar actores a la red, el número de vínculos posibles aumenta exponencialmente mientras que el número de vínculos que cada actor puede mantener se mantiene constante (Everton, 2012a). Es por ello, que a menudo las redes más grandes tendrán densidades más bajas. Con base en esta observación, los analistas de redes utilizan la medida de grado medio para comparar el nivel de interconexión entre redes de diferentes tamaños. El grado medio se define formalmente como el promedio de enlaces que tiene cada actor en la red. Si el grado de cada nodo es k , entonces:

$$\bar{k} = \frac{E}{N}$$

Donde N es el número de nodos y E es el número de vínculos en la red. En términos prácticos, esta ecuación retorna un número promedio de vínculos para todos los actores en la red. Esto nos ayuda a comparar redes de diferentes tamaños. Por ejemplo, las redes 8.8 y 8.9 tienen un grado medio similar, de 1.98 y 1.99 respectivamente. La densidad es equivalente a 0.02 y a 0.01.

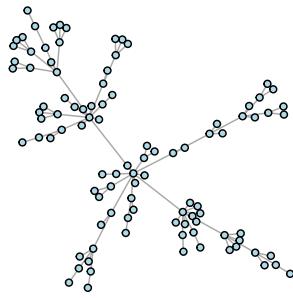


Figure 8.8: Grafo sin escala de 100 nodos

Es decir, aunque la densidad de estos gráficos es disímil podemos atribuir esta discrepancia a el número de nodos en cada red. Sin embargo, la medida más adecuada en esta situación sería el grado medio, la cual como vemos nos permite comparar el nivel de interconexión entre redes de diferente tamaño.

8.3 Centralización

Para recapitular, la eficiencia operacional de una red se ve afectada por la habilidad de esta para movilizar recursos materiales o humanos (McCormick & Owen, 2000). Algunos sugieren que esta capacidad operacional se encuentra íntimamente relacionada con la topología de la red social; específicamente, la centralización de la red (Cunningham, Everton, & Murphy, 2016; Morselli, 2009). Sin embargo, las redes altamente centralizadas no son libres de desafíos. Sino que pueden ser vulnerables a la eliminación de un puñado clave de nodos (Bakker, Raab, & Milward, 2011). Por ello, Everton (2012b) hace el argumento “demasiado

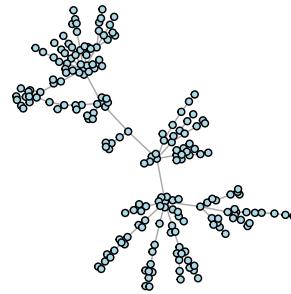


Figure 8.9: Grafo sin escala de 200 nodos

de algo bueno puede llevar a las redes a tener un rendimiento inferior . . . [desde la perspectiva de la red oscura] no deben encontrarse demasiado centralizadas o descentralizadas, mientras que desde la nuestra así es exactamente como queremos que sean” (Everton, 2012b).

Es crucial que como analistas de redes sepamos como determinar el nivel de centralización de una red. En la siguiente sección presentamos la medida de centralización, la cual es comúnmente utilizada para capturar de manera cuantitativa el nivel de centralización de una red.

8.3.1 Centralización

La centralización es el grado en que los vínculos de una red se centran en una persona o un conjunto de personas, de modo que en una red centralizada, una o varias personas ocupan puestos de poder y control en la red, mientras que las redes descentralizadas tienen estructuras de poder y control difusas (Valente, 2010).

Esta medida se basa en la centralidad de cada nodo en la red para calcular que tan variables o heterogéneas es la distribución de estos índices para cada actor (Wasserman & Faust, 1994). Podemos observar un ejemplo de red centralizada en la Figura 8.1. En el grafo estrella el actor uno (1) tiene contacto directo con todos los demás, los cuales no tienen contacto entre ellos. Utilizando el concepto de centralidad de grado, que es el número de vínculos que tiene cada nodo, exploremos como calcular la centralización de grado. Aquí utilizamos

la definición general matemática de Freeman (1979) para calcular el índice de centralización de grado a nivel grupo:

$$C = \frac{\sum[C(n^*) - C(n_i)]}{\max \sum[C(n^*) - C(n_i)]}$$

Donde $C(n_i)$ representa el índice de centralidad de un actor y $C(n^*)$ es el mayor valor del índice de centralidad en la red. Por lo tanto, $\sum[C(n^*) - C(n_i)]$ es la suma de las diferencias entre el valor máximo y los valores observados; adicionalmente, $\max \sum[C(n^*) - C(n_i)]$ es la suma teórica máxima (Wasserman & Faust, 1994). El producto de esta operación es un índice entre 0 y 1. Una red con un índice de centralización 0 sería aquella donde todos los actores tienen la misma centralidad. De manera opuesta, una red con índice de centralización 1 representa una configuración donde un solo actor domina el índice de centralidad comparado con los demás nodos en la red. Utilizando la fórmula de Freeman, calculemos la centralización de grado para la red estrella (ver Figura 8.1).

$$\begin{aligned} C_D &= \frac{\sum[C(n^*) - C(n_i)]}{\max \sum[C(n^*) - C(n_i)]} \\ C_D &= \frac{(4-4)+(4-1)+(4-1)+(4-1)+(4-1)}{12} \\ C_D &= 1.0 \end{aligned}$$

En la red estrella la centralización de grado es 1.0 (ver (8.3.1)). Por consiguiente, de manera matemática asumimos que los nodos en este gráfico convergen alrededor de un actor central.

Nodo	Grado
1	4
2	1
3	1
4	1
5	1

Veamos que sucede al introducir vínculos adicionales al grafo. En la Figura 8.10 el actor uno retiene el mismo número de vínculos que le proveen contacto directo con todos los demás, sin embargo, en esta ocasión los actores dos (2) y tres (3) tiene acceso mutuo.

Recalculemos, la centralización de la red con base a la centralidad de grado:

$$\begin{aligned} C_D &= \frac{\sum[C(n^*) - C(n_i)]}{\max \sum[C(n^*) - C(n_i)]} \\ C_D &= \frac{(4-4)+(4-2)+(4-2)+(4-1)+(4-1)}{12} \\ C_D &= 0.833 \end{aligned}$$

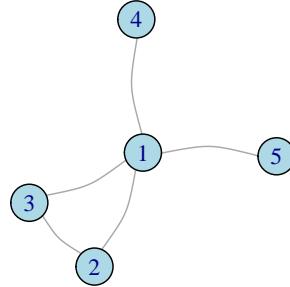


Figure 8.10: Red estrella modificada

El índice de centralización de grado es 0.833 (ver (8.3.1)). Lo cual sugiere que la mayoría de los vínculos en la red se encuentran enlazados con un actor central. Sin embargo, al agregar vínculos en la periferia de la red vemos una reducción en la homogeneidad de los valores individuales, la cual reduce el índice grupal.

<i>Nodo</i>	<i>Grado</i>
1	4
2	2
3	2
4	1
5	1

Antes de concluir esta sección, es importante recalcar dos cosas. Primero, como se ha mencionado la centralización es un índice grupal generado a base de las medidas de centralidad para cada vértice. En los siguientes capítulos veremos múltiples medidas de centralidad que pueden servir como precursores para calcular una medida de centralización. En este capítulo solamente hemos utilizado la centralidad de grado como precursor, pero es importante tener en mente que es posible generar índices de centralización con medidas de centralidad como intermediación, cercanía, o eigenvector. Segundo, Gephi no tiene la capacidad de calcular centralización, sin embargo, otras plataformas de software (como ORA, UCInet, o igraph) tienen esta capacidad y complementan las funciones de Gephi.

8.4 Ejercicio Práctico

En este ejercicio consideramos varias métricas para examinar la topología de la red, por ejemplo: tamaño de la red, densidad, grado promedio, distancia promedio, diámetro y centralización. El tamaño es igual al número de actores en una red. La densidad es igual al número de enlaces dividido por el total de lazos posibles. Sin embargo, el número de enlaces posibles en una red está inversamente relacionado con el tamaño de la red (es decir, cuanto más grande es la red, menor es la densidad); por consiguiente, los investigadores usan otras medidas como el grado promedio para comparar redes de diferente tamaño. La distancia promedio de la ruta es la longitud promedio de todas las rutas más cortas (es decir, geodésicas) entre todos los actores en una red, y el diámetro es la geodésica más larga de una red. Finalmente, la centralización mide el grado en que una red se centra en un solo actor (o grupo de actores).

Para este ejercicio utilizaremos un conjunto de datos con diez redes diferentes apiladas, que se encuentran ya cargadas en el archivo: **Noordin.gephi**. Descargue y guarde el archivo, abra el archivo siguiendo los pasos que se muestran a continuación.

8.4.1 Antes de Empezar

Para este ejercicio, deberá instalar el módulo *Clustering Coefficient* (ó coeficiente de agrupación) en Gephi. Haga esto siguiendo los mismos pasos que utilizamos en para descargar el modulo *MultimodeNetworks Transformation* en ejercicio práctico en el capitulo Derivando Datos Modo-Uno de Modo-Dos.

8.4.2 Topología de redes en Gephi

Gephi
Archivo>Abrir

Vista General
Contexto

1. Abra el archivo **Noordin.gephi**. En el primer espacio de trabajo, encontrará la red Noordin apilada. Los otros 10 espacios de trabajo son subredes por tipo de relación individual.
2. Comenzaremos por ver cómo encontrar el tamaño de una red. En la esquina superior derecha de la ventana *Descripción general*, debe ver una pestaña *Contexto* (cuadro rojo en la Figura 8.11; si no se encuentra visible, seleccione la opción *Contexto* en el menú *Ventana*). En esta ventanilla vera el tamaño (número de nodos), cuántos lazos (aristas) tiene la red/subred y si los datos son dirigidos o no. **¿Cuál es el tamaño de “Noordin (Red apilada)”** (en inglés “Noordin (Stacked Network)”? Cambie al espacio de trabajo “Red de amistad (en inglés “Friendship”). **¿Qué tamaño tiene? ¿Cuántas aristas hay?**

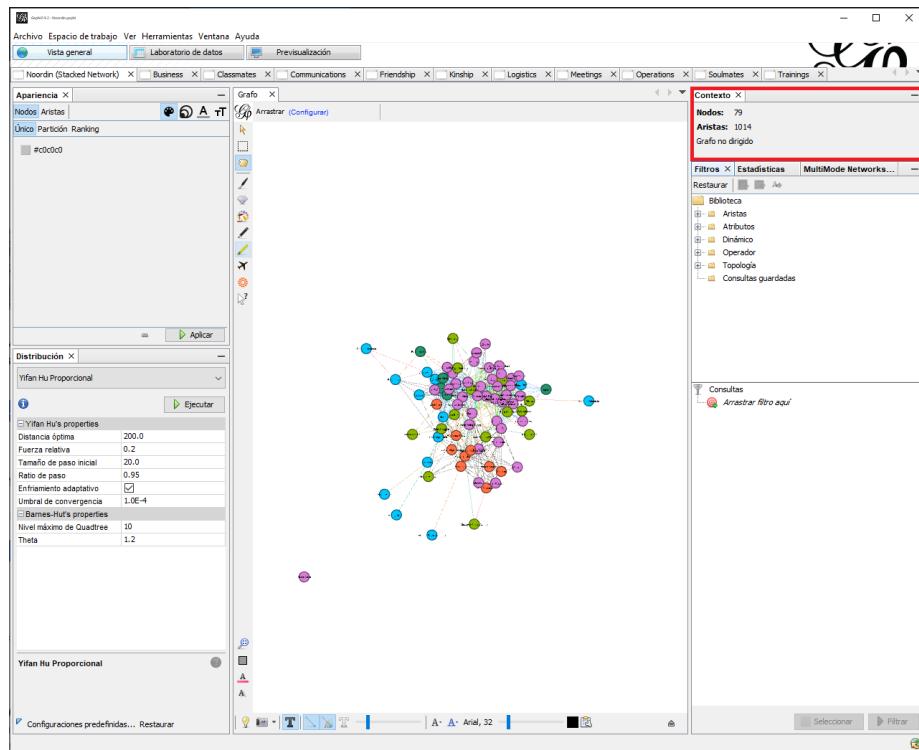


Figure 8.11: Ventana de descripción general, ventanilla de contexto

Estadísticas
 > *Densidad de grafo* >
Ejecutar

3. Calcular la densidad de la red en Gephi es sencillo. Seleccione la pestaña *Estadísticas* que se encuentra debajo de la ventana *Contexto* (consulte la Figura 8.12) y haga clic en el botón *Ejecutar* junto a la opción *Densidad de grafo*. Cuando se le solicite, trate el gráfico como no dirigido. ¿Cuál es la densidad de la red apilada? De clic en *Cerrar* para cerrar la ventanilla de reporte. Si desea volver a consultar la densidad, el resultado aparece junto a *Densidad de grafo*. Siempre puede volver a ver el reporte dando clic en el círculo con el ? a la derecha de la opción *Ejecutar*. Repita el proceso para las redes de Amistad (en inglés “Friendship”) y Comunicaciones (en inglés “Communications”). ¿Cuál es la densidad de la red de amistad? ¿Y la red de comunicaciones?

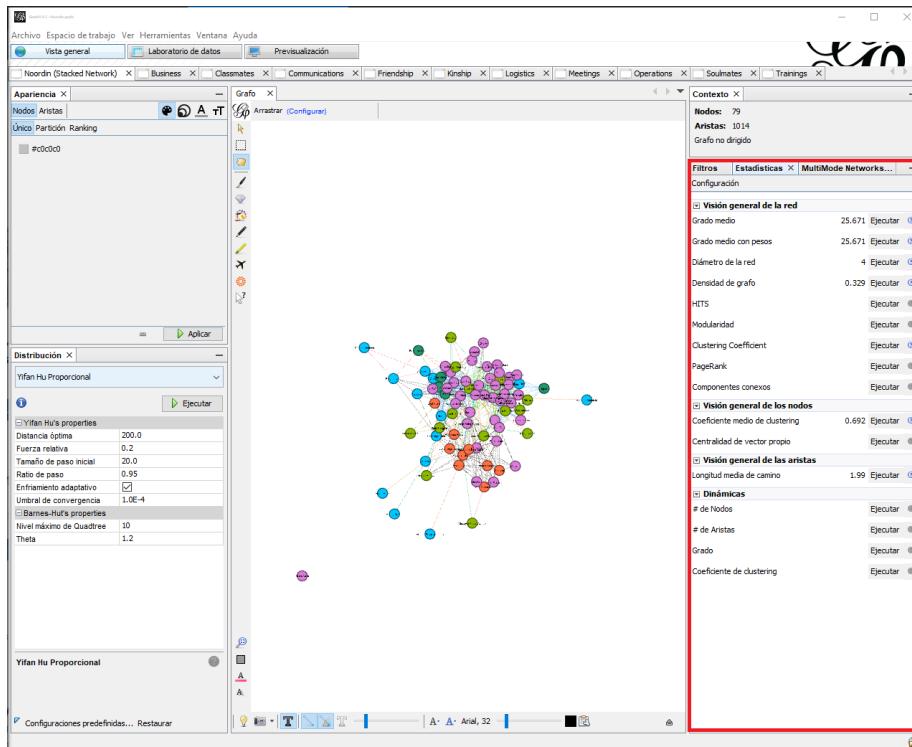


Figure 8.12: Ventana de descripción general, pestaña Estadísticas

*Estadísticas >
Grado medio >
Ejecutar*

4. La densidad de la red es igual a el número de enlaces dividido por todos los lazos posibles. Sin embargo, los puntajes de densidad tienden a disminuir a medida que las redes sociales se hacen más grandes porque el número de enlaces posibles aumenta exponencialmente con cada actor adicional, pero el número de lazos que cada actor puede mantener permanece relativamente constante. Por esto, la densidad de la red se debe usar para comparar redes del mismo tamaño (que hacemos aquí). Una medida alternativa a la densidad es el grado medio puesto que esta no es sensible al tamaño de la red. Por consiguiente, podemos utilizar esta segunda medida para comparar redes de diferentes tamaños. Podemos estimar el grado promedio de la red apilada haciendo clic en “Ejecutar” junto a “Grado medio” en la ventana de “Estadísticas.” En este caso, solo nos interesa la medida de grado promedio estándar en lugar de la medida de “Grado medio con pesos” porque no queremos incorporar los pesos de las conexiones. La puntuación aparecerá en la esquina superior izquierda de la ventana del informe (no se muestra). ¿Cuál es el grado promedio de la red apilada? El informe también produce un gráfico de “distribución de grados” que indica el número de actores en diferentes niveles de centralidad de grado. Si observa de cerca, verá que parece que hay un actor con un grado de centralidad de 130. Volveremos a este informe en un laboratorio posterior cuando veamos los puntajes de centralidad individuales. ¿Cuál es el grado medio de la red de amistad? ¿Comunicaciones? De las 10 redes individuales, ¿cuál tiene el grado medio más alto?

Estadísticas
> *Clustering*
Coefficient >
Ejecutar

5. Otra medida que podemos usar para capturar el nivel general de cohesión es el coeficiente de agrupación. Se estima identificando primero la red del ego para cada actor (es decir, los lazos de cada actor con otros actores vecinos o alteros y los lazos entre ellos), después calcular la densidad de esta misma (sin incluir el ego o los lazos del ego en el cálculo) , y luego tomando el promedio de todos los puntajes de densidad de la red del ego. Esta medida toma en cuenta el número de los vecinos de cada actor que están contados entre sí (también conocido como cierre triádico) y calcula el promedio de toda la red. Para obtener esta medida, en la pestaña *Estadísticas*, busque la sección denominada *Visión general de la red*. Haga clic en *Ejecutar* junto a *Clustering Coefficient*. Indique que desea utilizar el *Basic method* (método básico) y haga clic en *Aceptar*. El resultado aparecerá en la parte inferior de la ventana del informe después de las puntuaciones individuales, bajo *General C.* ¿Cuál es el coeficiente de agrupación de la red apilada? ¿La red de la amistad? ¿Red de comunicaciones? ¿Qué coeficiente de agrupamiento de red es más alto?

*Estadísticas >
Longitud media de camino
> Ejecutar*

6. Ahora veremos cómo obtener la distancia promedio y el diámetro en una red. La primera medida captura la distancia promedio entre todos los pares de actores (conectados) en la red, y puede indicar la velocidad con que la información puede viajar a través de una red. En otras palabras, la información se ha de difundir mas eficazmente en una red de menor distancia promedio que en una red con una distancia promedio alta. El diámetro de la red se refiere a la geodésica más larga de una red (es decir, la ruta más corta entre todos los pares de actores) y podría indicar qué tan dispersa o extendida está una red. Por lo tanto, puede capturar el grado en que una red está centralizada o descentralizada. Sin embargo, debemos tener un poco de cuidado al interpretar el diámetro porque es, en parte, una función del tamaño de la red. Por ejemplo, las redes grandes probablemente tendrán diámetros mayores. Para calcular la distancia y el diámetro promedio de la ruta de la red apilada, regrese a la pestaña *Estadísticas* y busque la sección *Visión general de las aristas*. Allí, haga clic en el botón *Ejecutar* a la derecha de *Longitud media de camino*. En el cuadro de diálogo, seleccione la opción no dirigida, no se preocupe por normalizar la centralidad en este momento (lo cubriremos en la sección de centralidad) y haga clic en *Aceptar*. Esto abrirá un informe (Figura 8.13) que proporciona mucha información, parte de la cual no nos ocuparemos en este momento. En la parte superior izquierda del informe, en *Results* (Resultados), Gephi nos dice cuál es el diámetro (*diameter*) y la distancia promedio de la ruta (*Average Path length*) para la red Noordin apilada **¿Cuál es la distancia y el diámetro promedio de las redes de comunicaciones y amistad?**

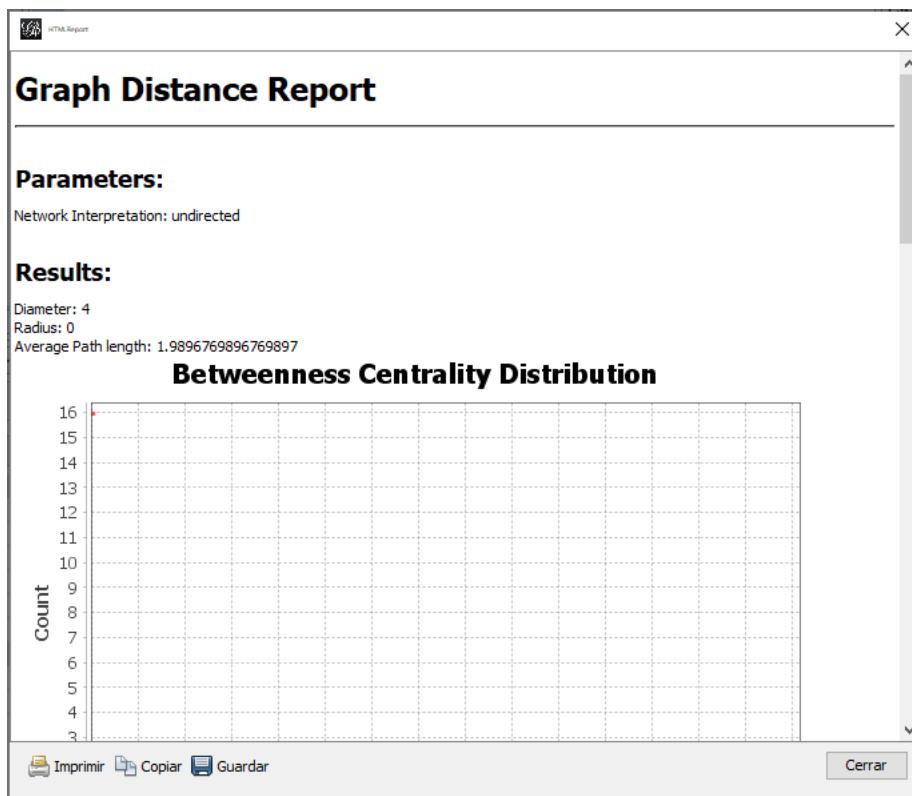


Figure 8.13: Ventana de descripción general, pestaña Estadísticas

Chapter 9

Subgrupos Cohesivos

En este capítulo, exploraremos algunos (pero no todos) métodos para hallar subgrupos cohesivos de actores dentro de una red social. Los subgrupos cohesivos son subconjuntos de actores entre los cuales existen fuertes lazos, por ello los métodos que utilizaremos formalizan la existencia de un grupo social a través de características estructurales de una red (Wasserman & Faust, 1994).

En esta breve introducción cubriremos: componentes, núcleos (k -core), grupos Girvan-Newman y de Louvain. Para poner este nuevo conocimiento en práctica, el capítulo concluye con un ejercicio práctico de como implementar estas medidas en Gephi.

9.1 Componentes

Estos son la forma más simple de subgrupos cohesivos en ARSo. Representan subredes en las cuales los miembros están conectados los unos a los otros (directa o indirectamente) pero no a miembros de otras subredes (Everton, 2012a, p. 171). Los componentes pueden ser divididos en dos campos, fuertes o débiles.

9.1.1 Componentes Débiles

Los nodos en un grafo desconectado se pueden dividir en dos o más subconjuntos de manera que no haya caminos entre los nodos en diferentes conjuntos (Wasserman & Faust, 1994), estos subconjuntos del grafo se llaman componentes débiles. Localizar componentes débiles es apropiado para datos dirigidos o no dirigidos, puesto que el objetivo es localizar subconjuntos conectados. Tome por ejemplo la Figura 9.1. En esta imagen, el color de los nodos denota el componente débil al que pertenecen. Como puede observar, encontramos tres componentes débiles en dicho gráfico, uno con cinco actores, una diada, y un aislado.

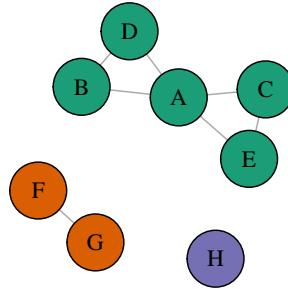


Figure 9.1: Red con múltiples componentes débiles

El detectar componentes débiles no solo sirve para colorear los grafos, de hecho, una de las aplicaciones más comunes es el aislar componentes. Por ejemplo, la densidad de el grafo en la Figura 9.1 es 0.25. Este índice sugiere una red escarza, sin embargo, de manera visual podemos ver que algunos componentes son más densos que otros. El componente principal, aquel de mayor tamaño, aparenta ser más denso. Es por ello que podemos aislar este componente principal y al ejecutar la medida de densidad vemos un valor de 0.6. Este proceso nos permite enfocarnos en diferentes subconjuntos del grafo en nuestro análisis.

9.1.2 Componentes Fuertes

Este tipo de subgrupo solo es adecuado cuando los datos son dirigidos. Easley & Kleinberg (2010) define los componentes fuertes en un grafo dirigido como subconjuntos de nodos en los cuales (A) que todos los nodos en el subconjunto pueden alcanzarse entre sí y (B) el subconjunto no es parte de un conjunto mayor con la propiedad que los nodos pueden llegar a todos los demás. Por ejemplo, en la Figura 9.2 el color de los nodos denota membresía a un subcomponente fuerte.

Observe que en esta ocasión encontramos cuatro componentes fuertes, el primero se encuentra compuesto por un actor aislado {H}, el segundo es una diada reciproca {F, G}, el tercero es un solo actor {E}, y el cuarto sería el componente fuerte principal {A, B, C, D}. Es probable que se esté preguntando, ¿Por qué E no forma parte del componente fuerte principal? La razón la podemos encontrar en la definición de Easley y Kleinberg. Note que en el componente principal todos los nodos pueden alcanzarse entre sí, es decir, si cualquier nodo en el componente

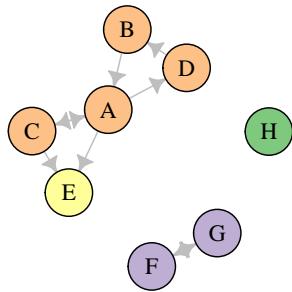


Figure 9.2: Red con múltiples componentes fuertes

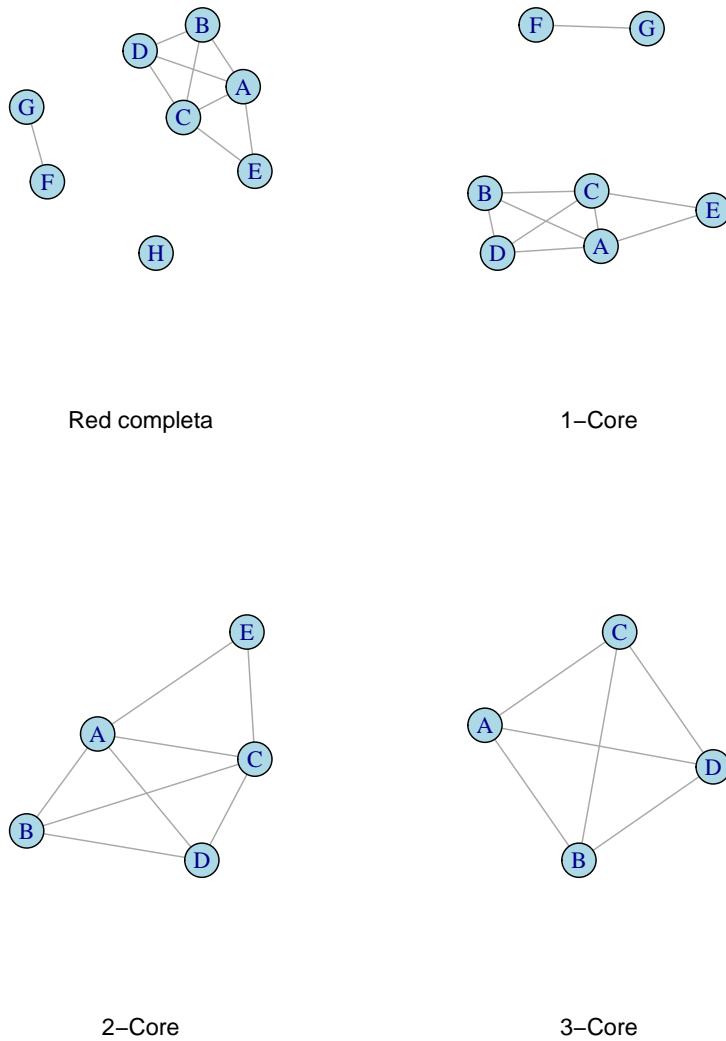
fuese a difundir alguna noticia a través de sus contactos esta información podría retornar al originador. Sin embargo, E se encuentra en un componente único puesto que no alcanza a otros nodos.

Como puede ver, los componentes fuertes son más restrictivos que los componentes débiles. Sin embargo, son útiles al trabajar con datos dirigidos puesto que permiten el análisis de subconjuntos donde todos los nodos pueden alcanzarse entre sí tomando en cuenta la direccionalidad de vínculos dirigidos y el flujo potencial de información y/o recursos.

9.2 K-cores

Otra forma de localizar subgrupos cohesivos en una red es basándose en grados nodales; específicamente, utilizando k-cores (o k-núcleos) que se definen en términos del grado mínimo. Donde k , en un subgrafo es el número mínimo de nodos adyacentes (Wasserman & Faust, 1994). Por ejemplo, la primer grafo en la Figura 9.3 representa un grafo completo, es decir, 0-core. En este, vemos nodos aislados con cero enlaces, nodos pendientes con un solo vínculo y nodos con una variedad de grados de centralidad. Un 1-core incluye todos los nodos conectados al menos a una persona (ver Figura 9.3). Similarmente, un 2-core incluye nodos conectados al menos dos otros nodos (ver Figura 9.3). De igual manera podemos definir un 3-core como nodos con tres o más vínculos (ver Figura 9.3). Al aumentar k comenzamos a localizar el núcleo de actores más interconectados de la red, de igual manera el subconjunto restante aumentará

en densidad (Valente, 2010).



2–Core

3–Core

Figure 9.3: Red completa y 1-, 2-, y 3-core

El remover nodos incrementando k es una estrategia comúnmente utilizada por analistas de redes para describir la estructura de la red. Específicamente, para aislar el núcleo de la periferia, o los actores más activos de los actores con pocas conexiones.

9.3 Girvan-Newman

El algoritmo de Girvan y Newman (2004) detecta subcomunidades enfocándose en remover vínculos intermedios entre subconjuntos de nodos con el propósito de localizar subcomponentes. Por ejemplo, en la red 9.4 (izquierda) vemos dos triadas cerradas conectadas a través de un puente. En este gráfico los vínculos han sido dibujados de manera ponderada con base en la intermediación de cada enlace. El algoritmo de Girvan y Newman comienza por calcular la intermediación de los enlaces, remueve de manera sistemática aquellos con el mayor índice de intermediación y procede a recalculársela y remover enlaces con alto valor de manera iterativa. Supongamos que ejecutamos este algoritmo por un ciclo, entonces al remover el puente entre traídas se produce dos subconjuntos, mejor descritos como componentes débiles (ver, 9.4 (derecha)).

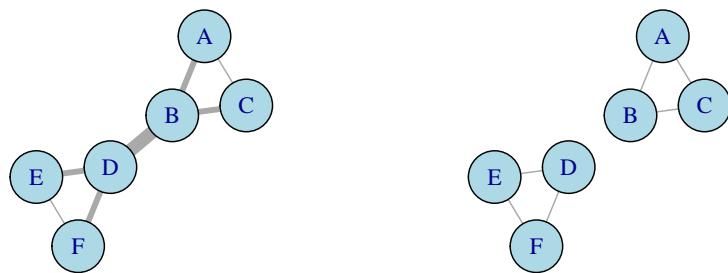


Figure 9.4: Red hipotética (izquierda) y componentes débiles (derecha)

En práctica, el algoritmo ejecuta este proceso y retorna un valor categórico correspondiente a la subcomunidad a la que pertenece cada nodo. Por lo tanto, el gráfico original no se ve afectado y el analista gana información sobre la comunidad Girvan-Newman correspondiente para cada nodo. Por ejemplo, la Figura 9.5 contiene la red previa con comunidades Girvan-Newman resaltadas.

El algoritmo de Girvan-Newman provee una forma de generar particiones de la red a grupos mutuamente exclusivos y un índice que mide la adecuación de estas particiones, de tal forma que el analista puede elegir entre estas particiones en busca de la más adecuada conforme a los datos de la red (Valente, 2010, p. 106). Normalmente, la partición de la red que genera el índice de modularidad más alto es considerado el punto óptimo (Everton, 2012a).

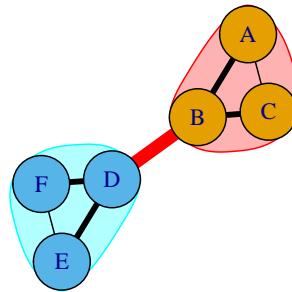


Figure 9.5: Comunidades Girvan-Newman y puentes resaltados.

9.3.1 Modularidad

La modularidad mide la fuerza de la división de una red en módulos (a veces llamados particiones, comunidades, etc.). En términos prácticos, las redes con alta modularidad tienen conexiones densas dentro de los módulos pero escasa entre estos. Se define como:

$$Q = \frac{1}{2m} \sum_{i,j} [A_{ij} - \frac{k_i k_j}{2m}] \delta(c_i, c_j)$$

Donde A_{ij} es el peso del vínculo entre los nodos i y j . k_i y k_j son los grados de cada nodo correspondiente. $\delta(c_i, c_j)$ o la delta Kronecker tiene un valor de 1.0 si ambos i y j corresponden a la misma partición, de otra manera el valor es 0.0. Por último, m corresponde al número de vínculos en la red.

El siguiente es un ejemplo de cómo calcular la modularidad de una red a mano a se base en el ejemplo de Matthew Joseph en ResearchGate y luego Abhishek Mishra. Empecemos con nuestra pequeña red 9.6 y su matriz de adyacencia (9.3.1).

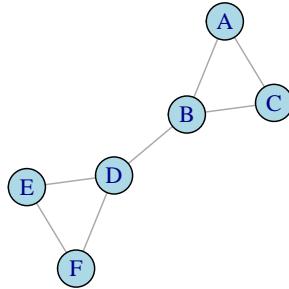


Figure 9.6: Red hipotética

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>
<i>A</i>	—	1	1	0	0	0
<i>B</i>	1	—	1	1	0	0
<i>C</i>	1	1	—	0	0	0
<i>D</i>	0	1	0	—	1	1
<i>E</i>	0	0	0	1	—	1
<i>F</i>	0	0	0	1	1	—

Como mencionamos previamente, la implementación del algoritmo Girvan-Newman comienza por remover el vínculo con la mayor intermediación, en este caso el puente entre los actores D y B. Aquí utilizaremos esta partición como punto de inicio, es decir, la primera partición contiene los nodos {A, B, C} y la segunda {D, E, F}. Igualmente, el número de enlaces en este grafo es $m = 7$ y el grado de cada nodo es equivalente a:

	<i>Grado</i>
<i>A</i>	2
<i>B</i>	3
<i>C</i>	2
<i>D</i>	3
<i>E</i>	2
<i>F</i>	2

Con esta información disponible, podemos calcular la modularidad del grafo:

$$\begin{aligned}
 Q = \frac{1}{2 \times 7} (& \\
 & \text{A a A: no connectados (0), mismo grupo (1)} \\
 & (0 - \frac{2 \times 2}{(2 \times 7)}) \times 1 + \\
 & \text{A a B: connectados (1), mismo grupo (1)} \\
 & (1 - \frac{2 \times 3}{(2 \times 7)}) \times 1 + \\
 & \text{A a C: connectados (1), mismo grupo (1)} \\
 & (1 - \frac{2 \times 2}{(2 \times 7)}) \times 1 + \\
 & \text{A a D: no connectados (0), otro grupo (0)} \\
 & (0 - \frac{2 \times 3}{(2 \times 7)}) \times 0 + \\
 & \text{A a E: no connectados (0), otro grupo (0)} \\
 & (0 - \frac{2 \times 2}{(2 \times 7)}) \times 0 + \\
 & \text{A a F: no connectados (0), otro grupo (0)} \\
 & (0 - \frac{2 \times 2}{(2 \times 7)}) \times 0 + \\
 & \text{B a B: no connectados (0), mismo grupo (1)} \\
 & (0 - \frac{3 \times 2}{(2 \times 7)}) \times 1 + \\
 & \text{B a A: connectados (1), mismo grupo (1)} \\
 & (1 - \frac{3 \times 2}{(2 \times 7)}) \times 1 + \\
 & \text{B a C: connectados (1), mismo grupo (1)} \\
 & (1 - \frac{3 \times 2}{(2 \times 7)}) \times 1 + \\
 & \text{B a D: connectados (1), otro grupo (0)} \\
 & (1 - \frac{3 \times 3}{(2 \times 7)}) \times 0 + \\
 & \text{B a E: no connectados (0), otro grupo (0)} \\
 & (0 - \frac{3 \times 2}{(2 \times 7)}) \times 0 + \\
 & \text{B a F: no connectados (0), otro grupo (0)} \\
 & (0 - \frac{3 \times 2}{(2 \times 7)}) \times 0 + \\
 & \dots)
 \end{aligned}$$

Podemos simplificar la operación previa como:

$$Q = \frac{1}{2 \times 7} (4(1) + 2(0.5)) = 0.3571$$

La modularidad con los nodos {A,B,C} y {D, E, F} en particiones separadas es igual a $Q = 0.3571$. En este ejemplo no continuaremos removiendo enlaces y asignando nuevas particiones, como lo haría el algoritmo Girvan-Newman. El objetivo de este ejercicio breve es presentarle en mayor detalle el mecanismo utilizado para calcular la modularidad de una red y determinar el número óptimo de subcomunidades.

9.4 Louvain

Publicado originalmente en 2008, el método Louvain es un método de optimización codicioso que intenta optimizar la modularidad de las particiones de la red. Esencialmente, la optimización se lleva a cabo en dos pasos. Primero, el método encuentra pequeñas comunidades locales optimizando la modularidad. Luego, agrega nodos que pertenecen a las comunidades identificadas y construye una nueva red con esos nodos colapsados. El proceso continúa hasta que se identifique la máxima modularidad y se produzca una jerarquía de comunidades (Blondel, Guillaumé, Lambiotte, & Lefebvre, 2008).

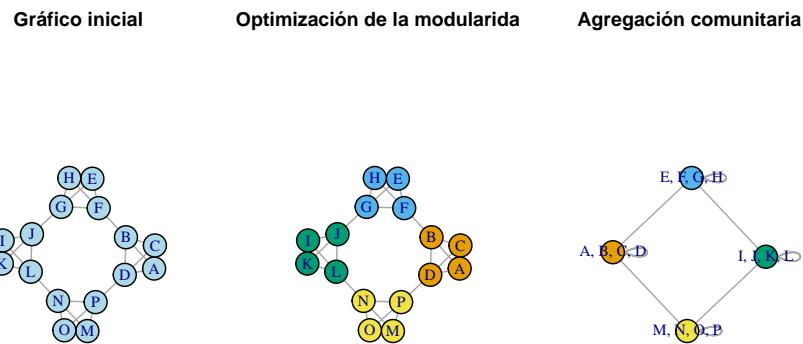


Figure 9.7: Visualización de los pasos en el algoritmo

La Figura 9.7 es una representación gráfica de una sola ronda en el proceso. Sin embargo, en términos prácticos, el proceso resulta en particiones que denotan las subcomunidades a las que pertenece cada nodo. La red izquierda de la Figura 9.8

consta de 100 nodos, a su derecha vemos las subcomunidades óptimas detectadas por el algoritmo de Louvain. Es importante recalcar que el objetivo del algoritmo es optimizar la modularidad y por consiguiente los grupos resultantes son una representación matemática de los patrones de enlaces en el gráfico. Sin embargo, este tipo de herramienta analítica es útil en localizar comunidades cohesivas donde los miembros tienen normas, valores o aptitudes similares. Otra manera de pensar en esto es que los nodos dentro de cada grupo pueden seleccionarse a sí mismos como parte del grupo o las presiones sociales del grupo pueden persuadir a los nuevos miembros para que adopten normas o valores del grupo (Valente, 2010, p. 108).

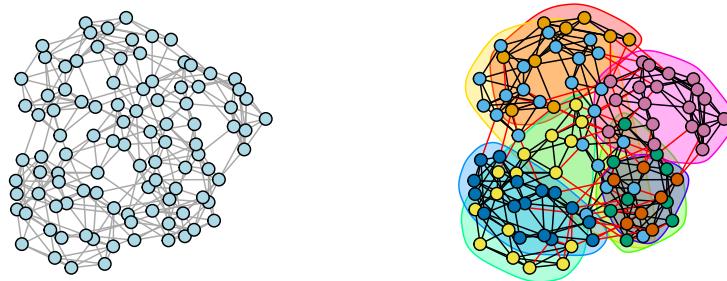


Figure 9.8: Red sin escala (izquierda) y red sin escala con particiones (derecha)

9.5 Ejercicio Práctico

Los científicos sociales generalmente asumen que “la interacción social es la base de la solidaridad, las normas compartidas, la identidad y el comportamiento colectivo, por lo que es probable que las personas que interactúan intensamente se consideren parte de un grupo social.” Por lo tanto, una herramienta principal del análisis de redes sociales es identificar grupos densos de actores “entre los cuales hay vínculos relativamente fuertes, directos, intensos y / o positivos.” Por lo general, se denominan “subgrupos cohesivos,” “subredes” o “subgrupos.” Una forma de agrupar a los actores se basa en atributos compartidos (por ejemplo, raza, género, etc.). Otro es utilizar el patrón de vínculos entre actores.

En un mundo ideal, habría un solo algoritmo para identificar subgrupos cohesivos, pero este no es un mundo ideal, por lo que los analistas de redes sociales

han desarrollado una variedad de algoritmos para identificar subredes. No los consideraremos todos aquí; simplemente hay demasiados. En cambio, nos centraremos en algunos. En Gephi consideraremos componentes, k-cores y algoritmos de detección de comunidades; y en snExplorer veremos brevemente los recuentos de componentes, k-núcleos y camarillas.

9.5.1 Antes de Empezar

Para este ejercicio vamos a necesitar el módulo *Newman-Girvan* clustering antes de comenzar. Haga esto siguiendo los mismos pasos que utilizamos en para descargar el modulo *MultimodeNetworks Transformation* en ejercicio práctico en el capítulo Derivando Datos Modo-Uno de Modo-Dos.

9.5.2 Identificación de Subgrupos en Gephi

Gephi
Archivo >
Abrir

1. Abra Gephi y lea el archivo `alive_trust_network.gephi`, que puede descargar aquí. Comenzaremos explorando componentes, estos son subgrupos de actores que están conectados (ya sea directa o indirectamente) entre sí. En las redes dirigidas, puede identificar dos tipos de componentes: fuertes y débiles. Los componentes fuertes toman en cuenta la dirección de los lazos, mientras que los componentes débiles no. Con redes no dirigidas, solo puede identificar componentes débiles.

*Vista general
Estadísticas >
Componentes
conexos >
Ejecutar*

*Laboratorio de
datos
Tabla de datos
> Nodos
Vista general
Apariencia
> Nodos
> Color >
Partición >
-Escoge un
atributo >
Component ID
> Aplicar*

2. Para identificar componentes en Gephi, haga clic en *Ejecutar* junto a la función *Componentes conexos* en la pestaña *Estadísticas*. Aparecerá un cuadro de diálogo titulado *Parámetros de Componenets Conectados* (no se muestra) que ofrece la opción de indicar si nuestros datos son *Dirigidos* o *No dirigidos*. Aquí, estamos trabajando con datos no dirigidos, así que elija esta opción y haga clic en *Aceptar*. Ahora deberá aparecer el informe *Connected Componenets Report* (Reporte de Componentes conectados) (Figura 9.9), que reporta los parámetros de la red, el número de componentes *Results* (Resultados) y un gráfico que indica el número y tamaño de cada uno de los componentes. Como puede ver, Gephi encontró 8 componentes, pero al comparar esta cifra con el gráfico de red, podemos ver que 7 de estos son nodos aislados. Igualmente podemos llegar a esta conclusión examinando el gráfico en el informe. Si observa verá que hay 7 componentes de tamaño 1 y un componente de un poco más grande que 60 (62 en realidad). Cierre la ventana del informe cuando haya terminado.
3. También puede ver los resultados navegando al *Laboratorio de datos* y buscando una nueva columna (partición) en la sección *Nodos* denominada *Component ID* (identificación del componente). Gephi genera una serie de números, comenzando con 0, que identifican los componentes respectivos de cada actor. También podemos usar esta variable categórica para visualizar la red, seleccionando primero la opción *Nodos> Color> Partición* en la pestaña *Apariencia*, y luego con el menú desplegable *-Escoge un atributo*, seleccione *Component ID* y de clic en *Aplicar*. El resultado no es muy interesante (Figura 9.10) ya que cada nodo aislado posee un color diferente, mientras que los actores conectados en el componente más grande (llamado componente principal) son todos del mismo color.

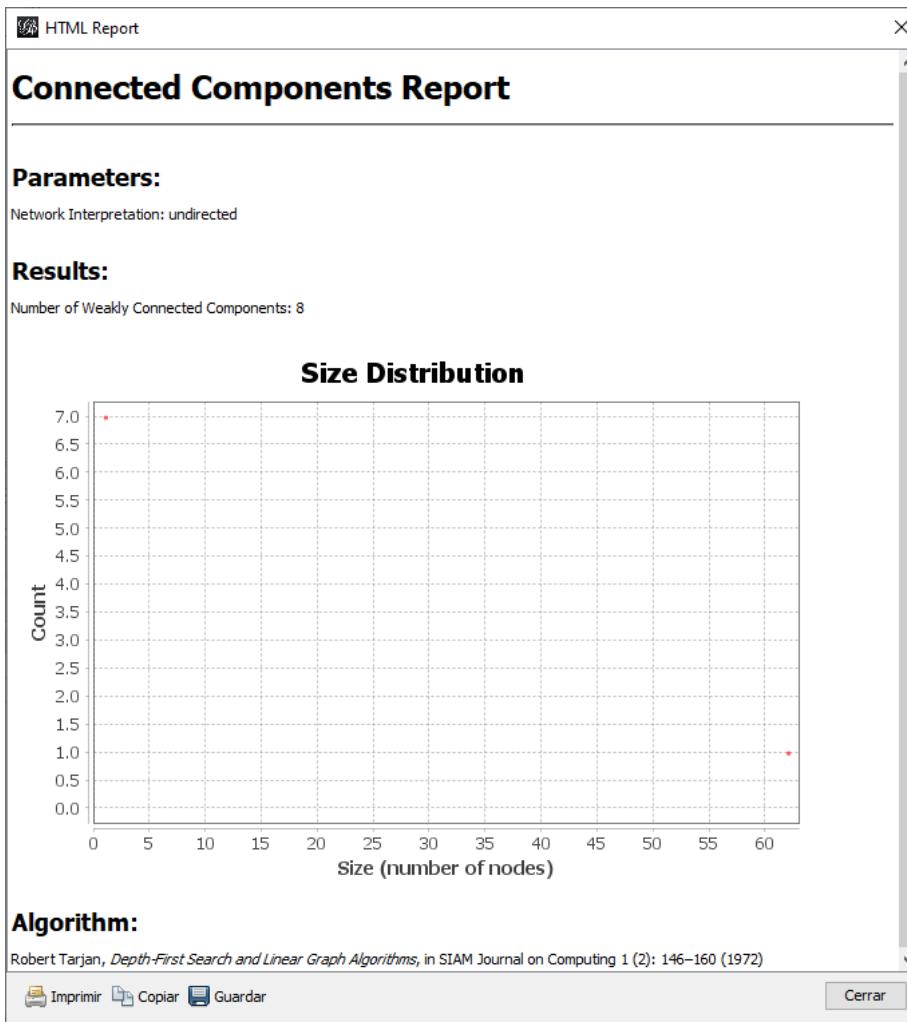


Figure 9.9: Reporte de Componentes Conectados en Gephi

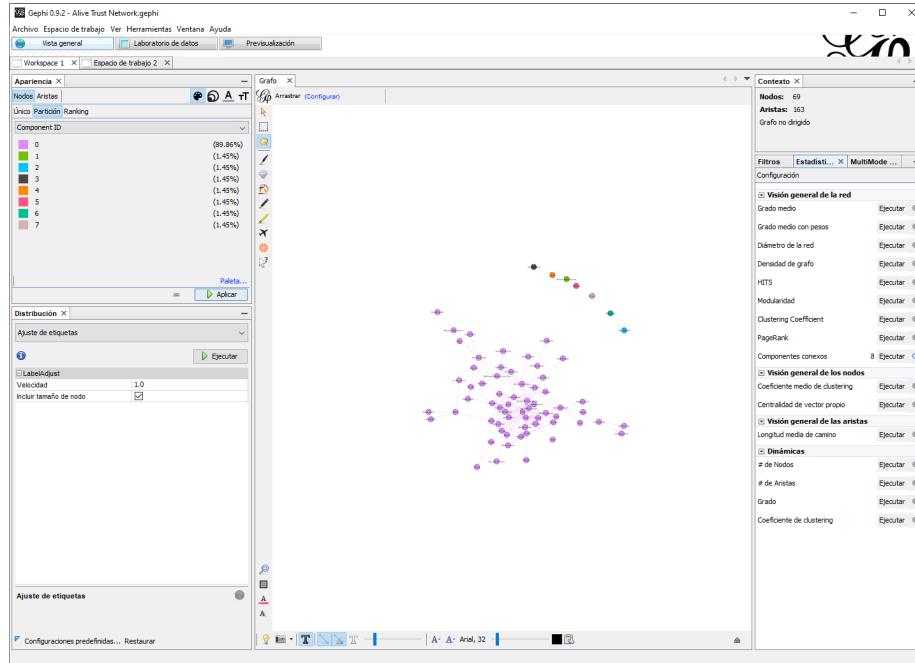


Figure 9.10: Red de confianza con actores vivos, color por componente

Vista general
Filtros >
Atributos >
Partición >
> Queries
> Partition
> Filter >
Export filtered graph to a new workspace

4. No obstante, puede que algunas ocasiones sea útil identificar el componente principal. De hecho, no es raro que los investigadores extraigan el componente principal y lo analicen por separado. Por lo tanto, exportaremos el componente principal a un espacio de trabajo separado donde lo analizaremos por separado. Para hacer esto, en la pestaña *Filtros* a la derecha de la ventana *Vista general*, primero seleccione *Atributos* y luego *Partición*. A continuación, arrastre la partición *Component ID* a *Consultas*. En las opciones abajo, haga clic en el cuadro que tiene el mayor porcentaje de nodos (probablemente 0), haga clic en *Filtrar* y luego use el botón *Exportar el grafo filtrado en un nuevo espacio de trabajo* (Figura 9.11). ¿Cuál es el tamaño del componente principal? ¿Cuál es su densidad? ¿Grado medio?

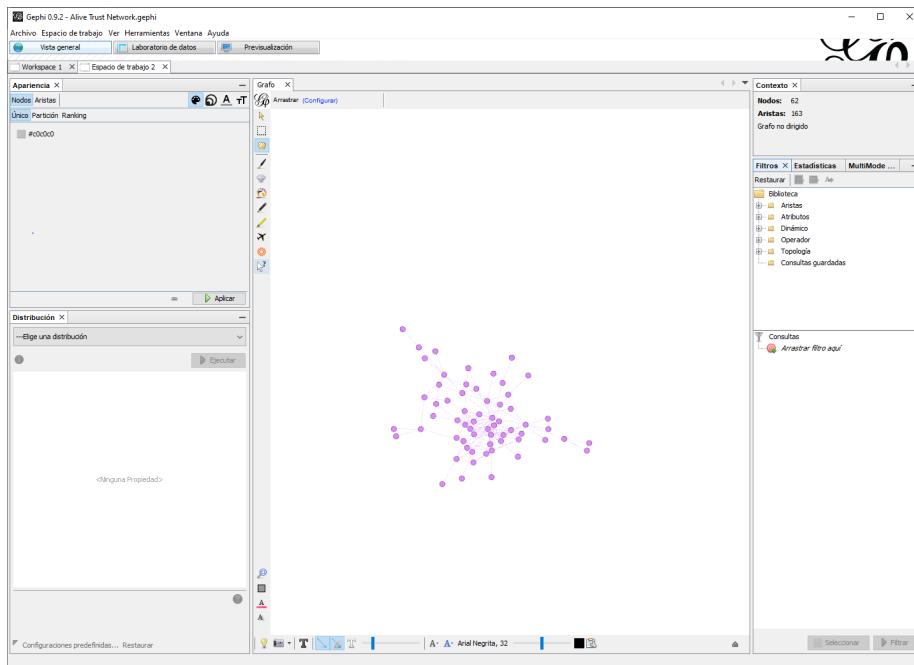


Figure 9.11: Red de confianza con actores vivos, componente principal

5. Veamos si identificar *k-cores* es una mejor estrategia para identificar subgrupos cohesivos para esta red en particular. La “k” en k-core indica el número mínimo de vínculos de cada actor dentro del núcleo (“core”); por ejemplo, un núcleo 2 incluye a todos los actores que tienen dos o más conexiones con otros actores dentro del núcleo, y un núcleo 3 incluye todos los actores que tienen tres o más vínculos con otros actores dentro de un núcleo. Es importante tener en cuenta que el k-core más alto de una red no corresponde necesariamente al puntaje de centralidad de grado más alto obtenido por un actor en la red. Más bien, es el subgrupo en el que cada actor comparte el mayor número de vínculos. Imagínese una red en la que un actor tiene seis vínculos, pero nadie más tiene más de cuatro. El k-core más alto no sería un 6-core porque todos los demás en el subgrupo también necesitarían tener seis enlaces.

Filtros >
Topología >
K-core

6. Regrese a la pestaña *Alive Trust Network*. Una vez más usaremos un filtro para encontrar k-cores. Esta vez, arrastre la opción *K-core* que se encuentra sección de *Topología* en la pestaña *Filtro a Consultas*. Hacia la parte inferior de la pestaña *Filtros*, debería ver un cuadro *K-core Configuración* que estará predeterminado en 1. Haga clic en el botón *Filtrar* y los aislados deberán desaparecer. Aumente la configuración a 2 y no debería haber ningún cambio. Eso es porque Gephi aparentemente cuenta los lazos dos veces, para reflejar las dos direcciones de un enlace aunque, en este caso, Gephi sabe (o debería saber) que se trata de una red no dirigida. Por lo tanto, el filtro de 1 y 2 capturan el primer núcleo ($k = 1$), 3 y 4 capturan el segundo núcleo ($k = 2$), y así sucesivamente. Si sigue aumentando la configuración, todos los nodos desaparecen al llegar a quince, lo que significa que el núcleo k más alto de la red es el séptimo núcleo ($14/2$). Vea la Figura 9.12. Exporte el 7-core a un nuevo espacio de trabajo. **¿Qué actores parecen ser los más centrales en la red del séptimo núcleo ($k = 7$)?**

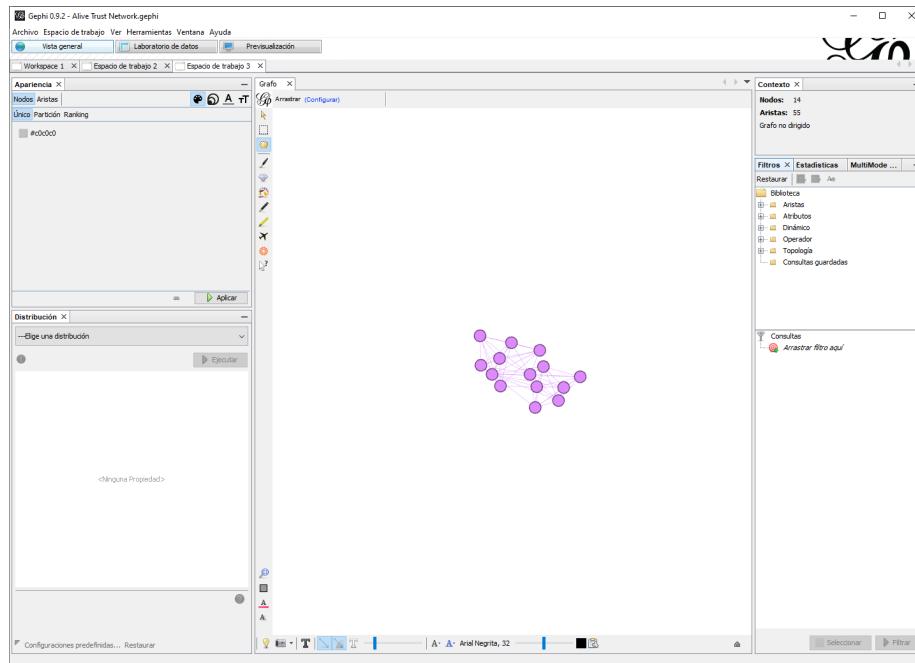


Figure 9.12: Red de confianza con actores vivos, 7-core

Vista general
Estadísticas
> *Girvan-*
Newman
Clustering >
Ejecutar

7. Ahora consideremos otro método para identificar subgrupos: algoritmos de detección de comunidades. Consideraremos dos: Girvan-Newman y Louvain. El segundo estará implementado automáticamente en Gephi, pero Girvan-Newman no, por lo que deberá descargar e instalar el módulo Newman-Girvan Clustering, si aún no lo ha hecho. En la pestaña *Estadísticas* de Alive Trust Network, haga clic en el botón *Ejecutar* a un lado de función *Girvan-Newman Clustering*. Esto abrirá un cuadro de diálogo (no se muestra). Acepte sus valores predeterminados y haga clic en "Aceptar". Esto producirá un informe (no se muestra) que indica el número de comunidades que detectó y la modularidad. **¿Cuántas comunidades encontró el algoritmo?**

Apariencia
> *Nodos*
> *Color* >
Partición >
-> *Escoge un*
atributo >
Cluter-ID >
Aplicar

8. Asimismo, coloree los nodos según la partición creada por el algoritmo (Cluter-ID). Note que el algoritmo Girvan-Newman ha asignado cada uno de los aislados a una comunidad separada. No todos los algoritmos de detección de comunidades tratan los aislamientos de esta manera, por lo que cuando compare los resultados de dos o más de ellos, tenga esto en cuenta.

Estadísticas >
Modularidad >
Ejecutar

9. Ahora, repitamos el proceso con el algoritmo de Louvain, que se implementa mediante la función *Modularidad* en la pestaña *Estadísticas*. Esto abrirá un cuadro de diálogo (no se muestra). Tenga en cuenta que incluye un parámetro de ajuste (*Resolución*), que le permite ajustar si desea obtener menos o más comunidades. Por ahora, haga clic en *Aceptar*. Esto produce un informe (no se muestra). **¿Cuántas comunidades encontró el algoritmo? ¿Qué es la modularidad? ¿Cómo se compara esto con los resultados de Girvan-Newman?**

Apariencia
> *Nodos* >
Color>*Partición*
> -> *Escoge un*
atributo
>*Modular-*
ity Class >
Aplicar

10. Coloree los nodos usando la partición creada por el algoritmo de Louvain (*Clase de modularidad*). Debido a que hay más de ~8 comunidades, deberá ajustar la paleta de colores como hicimos en laboratorios anteriores. **¿Louvain asigna a cada aislado en una comunidad separada?**

*Generate
groups
by
partition*

11. Ahora colapse la red usando la herramienta *Generate groups by partition* (Generar grupos por partición). Asegúrese de decirle a Gephi que cree un nuevo espacio de trabajo. Recuerde que los atributos son aspectos no relacionales de las redes. Pueden ser previamente dados (por ejemplo, rol, género, nacionalidad) o pueden ser generados por nuestro análisis de una red. Los algoritmos de agrupación, como los núcleos k y los de detección de comunicación, clasifican actores en varios subgrupos, y la partición de “pertenencia” del subgrupo es un atributo de actor.
12. Visualice la red reducida. **¿Aparece algún subgrupo (o subgrupos) más central que otros?**

Chapter 10

Centralidad

La centralidad es uno de los conceptos más antiguos del ARSo. Un actor central puede ser alguien que tiene numerosos vínculos con otros actores (grado), alguien que está más cerca (en términos de distancia del camino) a todos los demás (cercanía), alguien que se encuentra en el camino más corto (geodésico) entre dos actores cualesquiera (intermediación), o alguien que tiene vínculos con otros actores muy centrales (eigenvector). En algunas redes, los mismos actores obtienen una puntuación alta en las cuatro medidas, en otras no.

Por supuesto, hay más de cuatro medidas de centralidad. UCINET implementa más de 20, el programa R CINNA incluye más de 40, y David Schoch ha creado una tabla periódica interactiva de más de 100 medidas de centralidad.

Aquí discutiremos las tres centralidades más comunes, originalmente desarrolladas por Freeman (1979), es decir: grado, cercanía e intermediación. Además, otras medidas importantes serán introducidas como eigenvector (Bonacich, 1987), grados de entrada y salida (Wasserman & Faust, 1994) y centros de actividad y autoridades (Kleinberg, 1999). Para diferenciar entre medidas destinadas para datos dirigidos y no-dirigidos, dividiremos el capítulo en dos secciones. Adicionalmente, utilizaremos el marco teórico, originalmente postulado por Freeman (1979), utilizado en Cunningham, Everton, & Murphy (2016) para clasificar las medidas de centralidad en tres familias: medidas basadas en frecuencia, medidas basadas en distancia y medidas basadas en caminos. Como en todos los capítulos, este concluirá con un ejercicio práctico en Gephi.

10.1 Datos No Dirigidos: Centralidad y Poder

Aquí consideramos algunas de las medidas de centralidad apropiadas para trabajar con datos no dirigidos. Para clarificar algunas de las similitudes y diferencias entre estas medidas, serán presentadas utilizando un sistema de clasificación en Cunningham, Everton, & Murphy (2016). Específicamente, en tres secciones:

Table 10.1: Tabla de índices, grado

Nombre	Grado
A	4
B	1
C	1
D	2
E	2

- Basadas en frecuencia: la característica unificadora de estas es el enfoque en la frecuencia con la que un actor tiene vínculos con otros en la red. Por ello, estas suponen que la alta frecuencia de enlaces es correspondiente a poder o importancia en la red.
- Basadas en distancia: estas asumen que las distancias entre actores es el factor principal en determina la centralidad y poder de cada nodo.
- Basadas en caminos: estas se basan en la ubicación de un nodo en la red, principalmente si esta ubicación es una posición de corretaje o intermediación.

10.1.1 Medidas Basadas en Frecuencia

10.1.1.1 Centralidad de Grado

La centralidad de grado es la medida más simple y fácil de calcular y entender, por ello comisaremos aquí. Esta medida se define como el conteo de vínculos de un actor. En una red no dirigida, no tomamos en cuenta si estos enlaces son “de” o “hacia” el nodo. Con esto en mente, para una red binaria la centralidad de grado se define formalmente como:

$$C_i^{deg} = \sum_{j=1}^n x_{ij}$$

Donde la medida se calcula para todos los nodos en un grafo. Es decir, el grado se define como la suma de los vínculos entre un nodo (i) hacia actores adyacentes (x_{ij}). Por ejemplo, el grado de centralidad de los nodos en la Figura 10.1 aparecen en la Tabla 10.1. El nodo con mayor grado en este grafo es A puesto a que se encuentra enlazado con cuatro otros, D y E tienen un grado de centralidad de dos y finalmente C y B tienen un grado de uno.

Existen múltiples maneras de calcular esta medida de manera programática. Sin embargo, el método mas simple es multiplicar la matriz de adyacencia por un vector de unos, donde el largo del vector es correspondiente a el numero de vértices en el grafo (Meghanathan, 2016). La ecuación (10.1.1.1) ilustra como computar dicho índice.

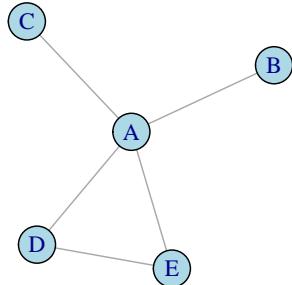


Figure 10.1: Grafo no dirigido

	A	B	C	D	E		C_i^{deg}
A	0	1	1	1	1	1	A 4
B	1	0	0	0	0	1	B 1
C	1	0	0	0	0	1	C 1
D	1	0	0	0	1	1	D 2
E	1	0	0	1	0	1	E 2

Como todas las medidas en esta sección, la manera en la que se calcula el grado de centralidad es indicativo de el objetivo de la medida y como se debe interpretar. Cunningham, Everton, & Murphy (2016) postulan que el grado mide la popularidad de un actor en términos de influencia directa entre este y otros nodos directamente adyacentes (Cunningham, Everton, & Murphy, 2016, p. 147). La idea principal aquí es que los vínculos sirven como canales para el flujo de influencia, materiales, ideas, etc. Por consiguiente, un nodo como A en la Figura 10.1 que a invertido su tiempo y recursos en enlazarse con cuatro otros nodos en una red, es posiblemente mas activo en la red que un actor como B conexo solo a un nodo.

Similarmente, los nodos con alto grado de centralidad tienden a ser más visibles en la red. Si asumimos que ciertos bienes – por ejemplo, información – tienden a difundirse en una red por medio de los enlaces, entonces podemos asumir que aquellos nodos con alto grado de centralidad estarán bien informados (Borgatti, Everett, & Johnson, 2018). Por ejemplo, en la Figura 10.1 el nodo A puede recibir información de tres ubicaciones aisladas, de los nodos B o C, y de la traída

compuesta por D y E. En contraste, B y C solo pueden acceder a información de la red a través de A. Similarmente, D y B dependen de A para acceder a recursos no redundantes en la triada.

10.1.1.2 Centralidad de Eigenvector

La centralidad de eigenvector, o centralidad de autovector (“eigen” significa “propio”), puede ser descrita de múltiples maneras. Aquí la categorizamos como una medida basada en frecuencia de vínculos similar a la centralidad de grado, pero con una variación clave, en esta medida ponderamos el grado de cada nodo por la centralidad de los nodos adyacentes (Borgatti, Everett, & Johnson, 2018). Esta medida asume que los enlaces a actores altamente centrales son más importantes que los enlaces a nodos en la periferia (Cunningham, Everton, & Murphy, 2016). En otras palabras, es preferible tener conexiones a nodos bien conectados. Por ejemplo, en la Figura 10.2 los nodos A y F tienen la misma centralidad de grado (2). Sin embargo, A se encuentra conectado a nodos quienes a su vez tienen un número de enlaces limitado. En contraste, F se encuentra conectado a G y H los cuales son adyacentes a tres otros nodos. En este ejemplo, si el analista enfoca su análisis en centralidad de grado, que mide actividad directa, A y F son equivalentes. Sin embargo, si expandimos el enfoque del análisis para considerar la influencia indirecta, es aparente que F se encuentra en una posición superior al nodo A en base al alcance de sus conexiones indirectas.

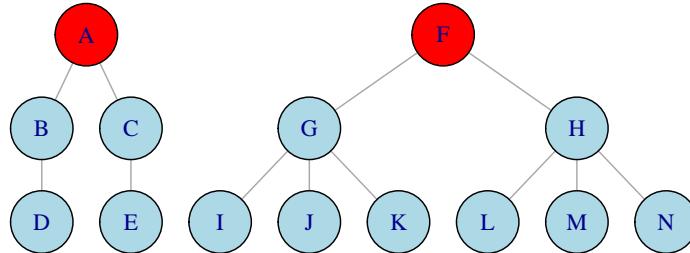


Figure 10.2: Comparación de nodos (A y F) con el mismo grado

De manera formal, Bonacich (1987) postula que la centralidad de eigenvector para cada nodo (i) corresponde a:

$$e_i = \frac{1}{\lambda} \sum_j R_{ij} e_j$$

Donde R corresponde a la matriz de adyacencia, la cual es usualmente (pero no obligatoriamente) simétrica y los elementos de la diagonal son iguales a cero. Además, λ es la constante llamada eigenvalor requerida. Por convención, el eigenvalor más grande es preferido. Para determinar el eigenvalor principal de una red podemos emplear una serie de herramientas de software (Gephi calcula este valor en segundos) o el método de las potencias (Meghanathan, 2016). Las operaciones en (10.1.1.2) representan el proceso de calcular la centralidad de eigenvector para el grafico en la Figura 10.3.

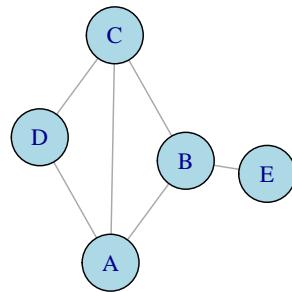


Figure 10.3: Grafo no dirigido

Primera Iteración :

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>			
<i>A</i>	0	1	1	1	0		1	3
<i>B</i>	1	0	1	0	1		1	3
<i>C</i>	1	1	0	1	0		1	3
<i>D</i>	1	0	1	0	0		1	2
<i>E</i>	0	1	0	0	0		1	1

Vector normalizado = 5.6569

Segunda Iteración :

$$\begin{array}{ccccccccc}
 & A & B & C & D & E & & & \\
 A & 0 & 1 & 1 & 1 & 0 & 0.5303 & 1.4142 & 0.6365 \\
 B & 1 & 0 & 1 & 0 & 1 & 0.5303 & 1.2374 & 0.5569 \\
 C & 1 & 1 & 0 & 1 & 0 & 0.5303 & 1.4142 & 0.6365 \\
 D & 1 & 0 & 1 & 0 & 0 & 0.3535 & 1.0606 & 0.4773 \\
 E & 0 & 1 & 0 & 0 & 0 & 0.1768 & 0.5303 & 0.2387
 \end{array}$$

$$\text{Vector normalizado} = 2.221984$$

Tercera Iteración :

$$\begin{array}{ccccccccc}
 & A & B & C & D & E & & & \\
 A & 0 & 1 & 1 & 1 & 0 & 0.6365 & 1.6707 & 0.5337 \\
 B & 1 & 0 & 1 & 0 & 1 & 0.5569 & 1.5117 & 0.4829 \\
 C & 1 & 1 & 0 & 1 & 0 & 0.6365 & 1.6707 & 0.5337 \\
 D & 1 & 0 & 1 & 0 & 0 & 0.4773 & 1.273 & 0.4067 \\
 E & 0 & 1 & 0 & 0 & 0 & 0.2387 & 0.5569 & 0.1779
 \end{array}$$

$$\text{Vector normalizado} = 3.130236$$

Cuarta Iteración :

$$\begin{array}{ccccccccc}
 & A & B & C & D & E & & & \\
 A & 0 & 1 & 1 & 1 & 0 & 0.5337 & 1.4233 & 0.5389 \\
 B & 1 & 0 & 1 & 0 & 1 & 0.4829 & 1.2454 & 0.4715 \\
 C & 1 & 1 & 0 & 1 & 0 & 0.5337 & 1.4233 & 0.5389 \\
 D & 1 & 0 & 1 & 0 & 0 & 0.4067 & 1.0675 & 0.4041 \\
 E & 0 & 1 & 0 & 0 & 0 & 0.1779 & 0.4829 & 0.1828
 \end{array}$$

$$\text{Vector normalizado} = 2.641108$$

Quinta Iteración :

$$\begin{array}{ccccccccc}
 & A & B & C & D & E & & & \\
 A & 0 & 1 & 1 & 1 & 0 & 0.5389 & 1.4146 & 0.5356 \\
 B & 1 & 0 & 1 & 0 & 1 & 0.4715 & 1.2607 & 0.4773 \\
 C & 1 & 1 & 0 & 1 & 0 & 0.5389 & 1.4146 & 0.5356 \\
 D & 1 & 0 & 1 & 0 & 0 & 0.4041 & 1.0778 & 0.4081 \\
 E & 0 & 1 & 0 & 0 & 0 & 0.1828 & 0.4715 & 0.1785
 \end{array}$$

$$\text{Vector normalizado} = 2.641162$$

Sexta Iteración :

$$\begin{array}{ccccccccc}
 & A & B & C & D & E & & & \\
 A & 0 & 1 & 1 & 1 & 0 & 0.5356 & 1.4210 & 0.5380 \\
 B & 1 & 0 & 1 & 0 & 1 & 0.4773 & 1.2497 & 0.4731 \\
 C & 1 & 1 & 0 & 1 & 0 & 0.5356 & 1.4210 & 0.5380 \\
 D & 1 & 0 & 1 & 0 & 0 & 0.4081 & 1.0712 & 0.4055 \\
 E & 0 & 1 & 0 & 0 & 0 & 0.1785 & 0.4773 & 0.1807
 \end{array}$$

$$\text{Vector normalizado} = 2.641176$$

...

Decimonovena Iteración :

$$\begin{array}{ccccc}
 \text{Nodo} & A & B & C & D & E \\
 \text{Eigenvector} & 0.537 & 0.475 & 0.537 & 0.407 & 0.179
 \end{array}$$

$$\text{Vector normalizado} = 2.641186$$

Como puede ver, el método de las potencias es un proceso iterativo. Comienza al multiplicar la matriz por un vector con valores de uno, el cual produce el grado de centralidad. Normalizamos los valores de dicho vector, en la primera ronda 5.6569, y dividimos el producto de la multiplicación de matriz por el valor normalizado, esto produce un eigenvector tentativo. La siguiente ronda comienza multiplicando la matriz de adyacencia por el eigenvector tentativo de la ronda previa. Una vez más normalizamos el vector resultante de la multiplicación y dividimos dichos valores por el valor normalizado. Esencialmente este proceso continua hasta que el producto de la normalización del vector converge (Meghanathan, 2016). En el ejemplo previo, este proceso toma 19 iteraciones para concluir. Sin embargo, alrededor de la cuarta iteración empezamos a ver un valor similar a valor convergente de en la última y penúltima ronda.

Después de tanta aritmética, se ha de estar preguntando ¿cómo interpreto los índices de centralidad resultantes? Como se mencionó previamente, esta medida de centralidad toma en cuenta los enlaces de cada nodo, así como los enlaces de sus vecinos. En términos prácticos el software habitualmente generara una serie de valores donde los índices de mas grandes representan nodos con mayor influencia indirecta. Como dice el dicho, “lo importante no es lo que sepas, sino a quien conozcas.” Especialmente la centralidad de eigenvector toma esta idea y la cuantifica en términos matemáticos.

10.1.2 Medidas Basadas en Distancia

10.1.2.1 Centralidad de Cercanía

La medida más común basada en distancia entre nodos es la centralidad de cercanía definida por Freeman (1979) como la distancia geodésica entre un nodo y los demás en una red. Se calcula sumando las distancias geodésicas entre cada nodo y los demás y después invirtiendo el valor para cambiar el índice de una

medida de distancia a cercanía (Valente, 2010, p. 85). Es común calcular la centralidad de cercanía como un índice normalizado (con valores de 0 a 1) de la siguiente manera:

$$C_i^{clo} = \frac{N - 1}{\sum_{j=1}^n d_{ij}}$$

Donde d_{ij} es la distancia geodésica entre dos nodos. Adicionalmente, el número máximo teórico de cercanía en cualquier red es $N - 1$, por lo tanto, al dividir las distancias sumadas por este numerado generamos un valor normalizado. El dividir el valor máximo por la suma de las distancias genera un índice donde los valores grandes corresponden a nodos más centrales (Borgatti, Everett, & Johnson, 2018, p. 199).

Calcular esta medida es simple, el grafo 10.4 servirá como ejemplo. Comenzaremos por encontrar los caminos geodésicos entre todos los nodos, presentado en la matriz izquierda de la ecuación (10.1.2.1). El siguiente paso consiste en sumar las distancias para cada nodo. Finalmente, dividimos el número máximo teórico de cercanía ($N - 1$) por la suma de las distancias de cada nodo. Por ejemplo, el nodo A tiene una distancia cumulativa de 15 y el valor teórico máximo de cercanía es 6. Por lo tanto, la cercanía normalizada del nodo A es equivalente a $\frac{6}{15} = 0.4$.

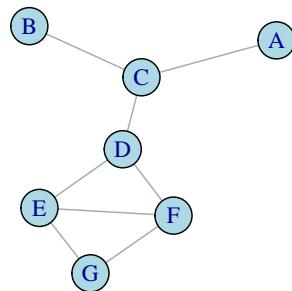


Figure 10.4: Grafo no dirigido

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>Suma</i>	<i>Cercanía Normalizada</i>
<i>A</i>	0	2	1	2	3	3	4	15	0.4
<i>B</i>	2	0	1	2	3	3	4	15	0.4
<i>C</i>	1	1	0	1	2	2	3	10	0.6
<i>D</i>	2	2	1	0	1	1	2	9	0.667
<i>E</i>	3	3	2	1	0	1	1	11	0.545
<i>F</i>	3	3	2	1	1	0	1	11	0.545
<i>G</i>	4	4	3	2	1	1	0	15	0.4

La centralidad de cercanía es una medida intuitiva. Borgatti, Everett, & Johnson (2018) sugieren que en términos de flujo de recursos los nodos más cercanos a menudo reciben materiales, recursos o información rápidamente (Borgatti, Everett, & Johnson, 2018, p. 199). Esto es significativo puesto que los recursos tienden a perder fidelidad a lo largo de distancias de camino largas. Por ejemplo, la información tiende a ser distorsionada conforme cambia de manos en una red o la transferencia de fondos financieros tiende a acumular costos con cada brinco entre nodos. Es por ello que los nodos centrales en términos de cercanía no solo tienden a recibir información, recursos o materiales de manera acelerada, sino también de alta fidelidad.

Antes de concluir la discusión de centralidad de cercanía, es importante entender una de las limitaciones clave de esta medida. Principalmente, el hecho que esta medida no debe ser aplicada al trabajar con grafos con componentes no conexos, por ejemplo, aquel con múltiples componentes. En estas situaciones Cunningham, Everton, & Murphy (2016) sugieren aislar el componente principal antes de calcular la centralidad de cercanía.

10.1.3 Medidas Basadas en Caminos

10.1.3.1 Centralidad de Intermediación

La tercera medida de Freeman (1979) que consideraremos en este capítulo mide la frecuencia con la que un nodo se encuentra en los caminos mas cortos conectando a todos los demás en la red. La centralidad de intermediación es comúnmente utilizada por los analistas de redes puesto que identifica actores en posiciones de corretaje claves. Wasserman & Faust (1994) recalca que interacciones de dos actores no adyacentes pueden depender de aquellos que se encuentran entre esos dos; es decir, los intermediarios podrían, potencialmente, tener algún control sobre las interacciones entre nodos no adyacentes, o podríamos también decir, tener más influencia interpersonal sobre otros (Wasserman & Faust, 1994).

Calcularemos centralidad de intermediación con la definición formal de Freeman (1979), quien fue uno de los primeros investigadores en cuantificar esta idea. Donde g_{jk} es el número de caminos más cortos, geodésicos, que unen a dos actores. En un grafo donde los vínculos tienen el mismo peso, es probable que existan múltiples geodésicos entre dos nodos, por ello asumimos que todos los caminos geodésicos tienen la misma probabilidad de ser elegidos, simplemente

podemos articular esta idea en términos matemáticos como $\frac{1}{g_{jk}}$. Puesto que el objetivo es identificar el índice de intermediación de un nodo (i), $g_{jk}(n_i)$ representa la porción de los caminos geodésicos que conectan a dos actores g_{jk} y que contienen al actor i . Freeman define el índice de intermediación de cualquier actor (n_i) en un grafo como la suma de todas las probabilidades estimadas para todos los pares de nodos en la red sin incluir al actor i (Wasserman & Faust, 1994):

$$C_B(n_i) = \sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}}$$

El índice alcanza su valor máximo cuando el nodo n_i se encuentra en todas las distancias geodésicas. Por lo tanto, el valor máximo es $\frac{(n-1)(n-2)}{2}$ que corresponde a el número de actores sin incluir n_1 . Podemos utilizar el valor máximo para normalizar el índice a una medida con valores entre 0 y 1, lo cual facilita la comparación con otros índices. Por lo tanto, la centralidad de intermediación centralizada se puede calcular como:

$$C'_B(n_i) = \frac{\sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}}}{\frac{(n-1)(n-2)}{2}}$$

Pongamos esta ecuación en práctica. La Figura 10.5 incluye un grafo no dirigido. ¿Qué nodo tendrá la centralidad de intermediación más alta? En este ejemplo le prestamos el cálculo de la centralidad de intermediación de los vértices en el gráfico. Para esta operación el valor máximo es 6 (o $\frac{(5-1)(5-2)}{2}$). Los índices presentados han sido normalizados.

Intermediación del nodo A : 0

Intermediación del nodo B : 0

Intermediación del nodo C : 0.083

$$\text{Par } (A, B) = 1/2$$

$$\frac{0.5}{6} = 0.083$$

Intermediación del nodo D : 0.583

$$\text{Par } (A, B) = 1/2$$

$$\text{Par } (A, E) = 1/1$$

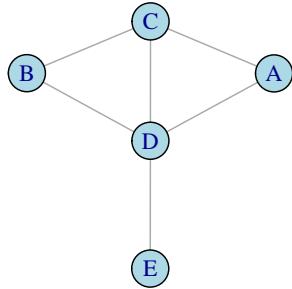


Figure 10.5: Grafo no dirigido simple.

$$Par(B, E) = 1/1$$

$$Par(C, E) = 1/1$$

$$\frac{0.5 + 1 + 1 + 1}{6} = 0.583$$

Intermediación del nodo E : 0

Note que en este índice un nodo puede tener un valor nulo, por ejemplo $\{A, B, E\}$. Esto ocurre pues estos nodos no se encuentran en ningún camino geodésico entre otros nodos. El ejemplo más sencillo es el nodo E, que no se encuentra en ningún camino entre nodos pues es un pendiente en la red. Los nodos A y B son redundantes para los demás nodos pues no caen en un camino geodésico, es decir, si E quiere navegar a C la ruta procedería de E a B y luego a C. Otros nodos sirven como posibles intermediadores, por ejemplo C. Este se encuentra en un camino entre A y B. Sin embargo, puesto que existen dos caminos geodésicos entre estos nodos, la porción de caminos entre A y B en la que encontramos a C es solo 0.5. Finalmente, algunos nodos son cruciales para el flujo de la red. Por ejemplo, el nodo D es el punto de acceso a E y se localiza en la mitad de los caminos geodésicos entre A y B.

El proceso demostrado en esta sección es largo y un poco tedioso, sin embargo, los programas de software utilizados por analistas de ARSo producen estos índices con bastante rapidez. Esperamos que pueda ver el valor de esta medida como analista de redes. Simplemente permite identificar nodos cuyo potencial como corredores es mayor en base a su ubicación estructural en una red de nodos. Antes de concluir con esta sección es importante recalcar un punto clave presentado por Cunningham, Everton, & Murphy (2016). Los autores notan que no existe ninguna garantía que dos nodos elijan tomar el camino más corto entre estos en la red (Cunningham, Everton, & Murphy, 2016). En ciertos contextos sociales, el camino mas largo puede ser preferible. Por ejemplo, una ruta larga puede ser la única ruta conocida por un par de actores y por lo tanto el camino preferido. Esto no quiere decir que la centralidad de intermediación no tiene valor para el analista de redes, esta medida es útil y comúnmente utilizada pues a menudo aporta nueva información, pero esta debe ser evaluada a par con el contexto de la red.

10.2 Datos Dirigidos: Centralidad y Prestigio

10.2.1 Medidas Basadas en Frecuencia

10.2.1.1 Centralidad de Grado: Entrada y Salida

Retornemos por un segundo a la idea de grados de centralidad, previamente definida en una red no dirigida como el conteo de vínculos de un actor. Sin embargo, en un grafo dirigido podemos expandir este concepto para captar la dirección del vínculo adyacente a cada nodo. Específicamente, podemos enfocarnos en los vínculos adyacentes hacia o desde el nodo, que de manera formal llamamos grado de entrada o de salida. Wasserman & Faust (1994) postula que es necesario considerar ambos por separado pues uno cuantifica la tendencia de los nodos a forjar vínculos mientras que el otro la tendencia de recibirlos (Wasserman & Faust, 1994, p. 151).

Considere la Figura 10.6 que incluye un grafo dirigido. La dirección de la flecha indica el nodo que recibe un enlace, del otro lado de la flecha encontramos a aquel nodo que envía dicho arco. La interpretación del grado de entrada o salida depende del contexto del grafo. Por ejemplo, supongamos que los vínculos en la Figura 10.6 representan enlaces de confianza; es decir, cada nodo nomina a otros en la red en quien confía. En esta situación interpretaríamos a un nodo con alto grado de entrada como alguien confiable y un nodo con alto grado de salida como alguien confiado.

Wasserman & Faust (1994) definen el grado de entrada hacia un nodo n_i como $d_I(n_i)$, donde el grado de entrada es igual al numero de arcos que terminan en n_i ($l_k = \langle n_j, n_i \rangle$) para todos los arcos ($l_k \in L$) y nodos ($n_j \in N$) en el grafo ($G(N, L)$). Similarmente, el grado de salida de un nodo $d_O(n_i)$ es el numero de nodos adyacentes desde n_i . Es decir, todos los enlaces que comienzan en n_i ($l_k = \langle n_i, n_j \rangle$) para todos los arcos ($l_k \in L$) y nodos ($n_j \in N$) en el grafo

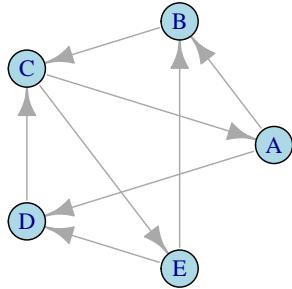


Figure 10.6: Datos dirigidos

(Wasserman & Faust, 1994). En una matriz de adyacencia, la suma de las filas para cada nodo es equivalente a el grado de salida, y los totales de las comunas es igual al grado de entrada. Formalmente:

$$d_O(n_i) = \sum_{j=1}^g x_{ij} = x_{i+},$$

y

$$d_O(n_i) = \sum_{j=1}^g x_{ij} = x_{+i}.$$

En términos prácticos el calcular estos valores es simple. La matriz (10.2.1.1) contiene los datos dirigidos. En este ejemplo, hemos agregados una fila ($d_I(n_i)$) y columna ($d_O(n_i)$) adicional con los valores de centralidad de entrada y salida. Para calcular dichos índices hemos sumado los valores en la sociomatrix. El sumar las filas, de arriba hacia abajo, resulta en el grado de entrada para cada nodo. Similarmente, el sumar las columnas, de izquierda a derecha produce el grado de salida.

	A	B	C	D	E	$d_O(n_i)$
A	0	1	0	1	0	2
B	0	0	1	0	0	1
C	1	0	0	1	0	2
D	0	0	1	0	0	1
E	0	1	0	1	0	2
$d_I(n_i)$	1	2	2	3	0	

Note que algunos nodos no reciben ningún enlace, por ejemplo, E tiene un grado de entrada de cero $d_I(n_E) = 0$. Sin embargo, esta falta de vínculos hacia E no impacta de manera directa el número de enlaces que comienzan con E, $d_O(n_E) = 2$. El punto clave en interpretar estos índices es el contextualizar las medidas, es decir, hacerse la pregunta como analista: ¿Qué significa un alto grado de entrada o salida con las relaciones presentes?

10.2.1.2 Centros de Actividad y Autoridades

El último algoritmo que exploraremos en esta sección es centros de actividad, o hubs, y autoridades, comúnmente conocido como “hubs and authorities” en inglés o simplemente HITS. Su autor, Kleinberg (1999), originalmente desarrollo esta rutina para clasificar sitios de web. Este algoritmo es similar a la centralidad de eigenvector (Cunningham, Everton, & Murphy, 2016), sin embargo, tiene diferencias cruciales. Fue diseñado para trabajar con datos dirigidos, por lo tanto, genera dos medidas un índice de centro de actividad y otro de autoridad.

Kleinberg (1999) postula que un centro de actividad sería un nodo enlazado a múltiples autoridades. Es decir, los centros de autoridad actúan como portales que dirigen el tráfico de enlaces hacia nodos eminentes en un tema. Estos últimos son las autoridades, quienes reciben el tráfico de las autoridades (Monge & Contractor, 2003). Recuerde que este algoritmo fue diseñado para categorizar sitios de web, por ello que podemos pensar en centros de actividad como sitios de web a los que acuden las personas buscando información sobre un tema, por ejemplo, Wikipedia. Las autoridades en este contexto serían los sitios de web en la sección de referencias de un tema.

Antes de proceder a calcular los índices, piense ¿En qué contextos sociales piensa usted que un analista de ARSo utilizaría el algoritmo de HITS? Es importante recalcar que, aunque el algoritmo fue diseñado para clasificar sitios de web, tiene múltiples aplicaciones en contextos sociales. Por ejemplo, en una red de comunicación dirigida entre miembros de una organización podemos aislar quienes son los centros de actividad o los nodos que representan las autoridades de la red. En este ejemplo, el localizar un centro de autoridad sería valioso pues como mencionamos estos sirven como portales comúnmente utilizados en la red para acceder a autoridades. Similarmente, dependiendo de nuestro objetivo,

un nodo en posición de autoridad se encuentra en una posición estructural prestigiosa de la red en cuestión.

En términos prácticos, el calcular los índices para cada nodo de centro de actividad y autoridad es relativamente simple. Observe la red 10.7, los centros de autoridad apuntan a las autoridades, esto sucede por la relación mutuamente reforzada donde un buen centro de actividad apunta a muchas buenas autoridades y del otro lado de la moneda una buena autoridad es aquel nodo que recibe múltiples vínculos de buenos centros de actividad (Kleinberg, 1999).

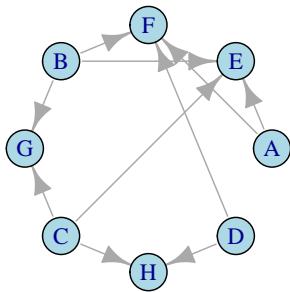


Figure 10.7: Grafo dirigido

Suponga que el nodo i tiene un índice de autoridad de a_i y de centro de actividad h_i donde i es equivalente a $1 : n$. Asuma que la autoridad y centro de actividad inicial de cada nodo i será a_i^0 y h_i^0 . El método iterativo HITS actualizara los valores iniciales con las siguientes sumas:

$$a_i^k = \sum_j h_j^{(k-1)}$$

$$h_i^k = \sum_j a_j^{(k-1)}$$

Como proceso iterativo, podemos seleccionar el número de iteraciones k , en este ejemplo utilizaremos $k = 1$. Podemos reescribir las ecuaciones previas como multiplicaciones de matriz:

$$a^k = A^T h_{(k-1)}, \quad h^k = A a_k$$

Para calcular el primer vector de autoridad a^1 , primero debemos transponer la matriz de adyacencia (A^T) y multiplicarla por el vector inicial de centros de actividad h^0 de valor uno (10.2.1.2).

	A	B	C	D	E	F	G	H		a^1
A	0	0	0	0	0	0	0	0	1	0
B	0	0	0	0	0	0	0	0	1	0
C	0	0	0	0	0	0	0	0	1	0
D	0	0	0	0	0	0	0	0	\times	$1 = 0$
E	1	1	1	0	0	0	0	0	1	3
F	1	1	0	1	0	0	0	0	1	3
G	0	1	1	0	0	0	0	0	1	2
H	0	0	1	1	0	0	0	0	1	2

Con el vector de autoridad resultante a^1 , calcularemos centros de actividad h^1 al multiplicar la matriz de adyacencia A por el vector de autoridad (10.2.1.2).

	A	B	C	D	E	F	G	H		h^1
A	0	0	0	0	1	1	0	0	0	6
B	0	0	0	0	1	1	1	0	0	8
C	0	0	0	0	1	0	1	1	0	7
D	0	0	0	0	0	1	0	1	\times	$0 = 5$
E	0	0	0	0	0	0	0	0	3	0
F	0	0	0	0	0	0	0	0	3	0
G	0	0	0	0	0	0	0	0	2	0
H	0	0	0	0	0	0	0	0	2	0

Para concluir la primera iteración $k = 1$ de este algoritmo, calculamos los eigenvectores normalizados $\|a\|$ y $\|v\|$ y dividimos cada índice por el valor normalizado. De la siguiente manera:

$$\|v\| = \sqrt{0^2 + 0^2 + 0^2 + 0^2 + 3^2 + 3^2 + 2^2 + 2^2} = 3.605551$$

$$\|u\| = \sqrt{6^2 + 8^2 + 7^2 + 5^2 + 0^2 + 0^2 + 0^2 + 0^2} = 13.19091$$

	a^1	h^1
A	$\frac{0}{3.6056} = 0$	$\frac{6}{13.1909} = 0.4549$
B	$\frac{0}{3.6056} = 0$	$\frac{8}{13.1909} = 0.6065$
C	$\frac{0}{3.6056} = 0$	$\frac{7}{13.1909} = 0.5307$
D	$\frac{0}{3.6056} = 0$	$\frac{5}{13.1909} = 0.379$
E	$\frac{3}{3.6056} = 0.832$	$\frac{0}{13.1909} = 0$
F	$\frac{3}{3.6056} = 0.832$	$\frac{0}{13.1909} = 0$
G	$\frac{2}{3.6056} = 0.5467$	$\frac{0}{13.1909} = 0$
H	$\frac{2}{3.6056} = 0.5467$	$\frac{0}{13.1909} = 0$

El siguiente ciclo $k = 2$, comenzara utilizando los valores de centros de actividad (h^1) y autoridades (a^1) del ciclo previo. El proceso repite una y otra vez dependiente en el numero de iteraciones.

Los resultados previos deben corresponder con nuestra intuición, E y F son los nodos mas autoritarios, puesto que reciben la mayoría de los enlaces de entrada. Similarmente, B y C son los centros de autoridad más activos pues envían la mayor porción de vínculos. Sin embargo, el algoritmo HITS nos permite cuantificar que nodos ocupan que posición.

10.3 Ejercicio Práctico

En este ejercicio practicaremos algunas de las medidas de centralidad presentes en Gephi. Para ello necesitara descargar el conjunto de datos titulado: `alive_combined_network.gephi`.

10.3.1 Centralidad en Gephi

Gephi Archivo
> Abrir
Vista general
Estadísticas >
Grado medio >
Ejecutar

1. Abra `alive_combined_network.gephi`, que puede descargar aquí. Los colores de los nodos de la red reflejan los subgrupos de acuerdo con el algoritmo de Louvain. Gephi implementa un puñado de medidas de centralidad. La medida más común de centralidad es el grado, que en una red no dirigida y binaria es simplemente un recuento del número de vínculos de cada actor (es decir, el número de vecinos). Los vínculos en una red son ponderados cuando los actores comparten múltiples entre sí, por un par de actores puede compartir más de un vínculo, como parentesco, religioso, escolar. Para la centralidad de grado, Gephi ofrece la opción de tomar en cuenta el peso de los vínculos o ignorarlo. Para calcular la centralidad de grado no ponderada, ubique la opción *Grado medio* en la pestaña *Estadísticas* y haga clic en *Ejecutar*. Esto genera un informe, que hemos visto antes, que le presenta la centralidad de grado medio y produce un gráfico que indica la distribución de los valores de grado.

Laboratorio de datos Tabla de datos > Nodos

- En este laboratorio nuestro interés es capturar los puntajes de centralidad de grado de cada actor. Para hacer esto, cambie a la ventana *Laboratorio de datos*, haga clic en *Nodos* en la pestaña *Tabla de datos*, y vera una columna en el extremo derecho con la etiqueta *Grado* (Figura 10.8). Puede ordenar los actores por grado (ya sea ascendente o descendente) haciendo clic en la etiqueta. Por lo tanto, con un par de clics, puede determinar qué actores son más o menos influyentes en términos de grado. Si observa la Figura 10.8, puede ver que Noordin Top ocupa el primer lugar (46), seguido de Suramto (28), Ubeid (27), Abdullah Sunata (25) e Iwan Dharmawan (25). Si desea puede exportar esta table y manipular la información en Excel.

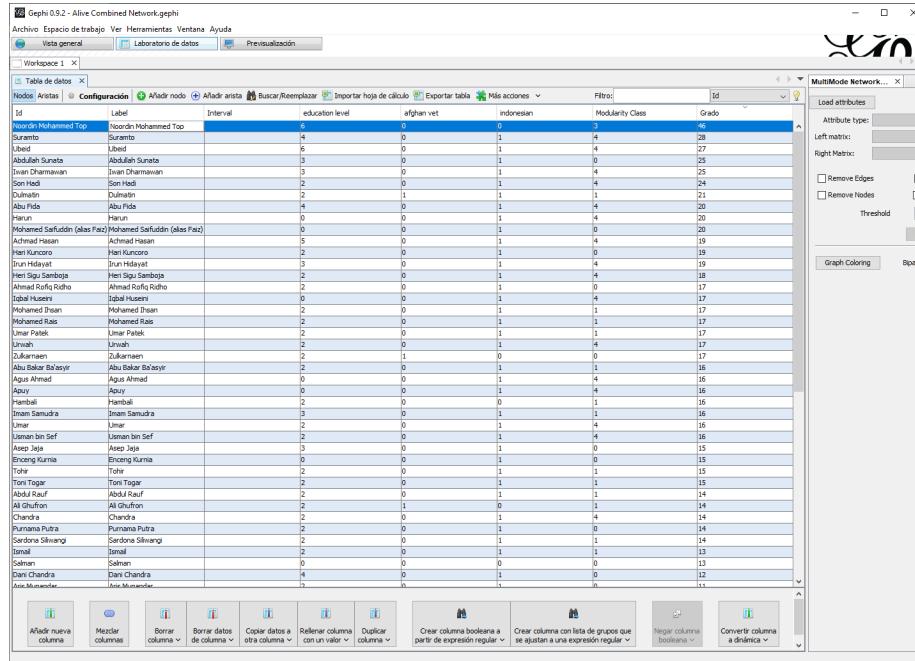


Figure 10.8: Red combinada de nodos vivos, Laboratorio de datos

Apariencia
> *Nodos*
Tamaño >
Ranking >
Aplicar

- Para visualizar la red donde el tamaño del nodo refleja el grado de centralidad, en la pestaña *Apariencia*, seleccione *Nodos*, *Tamaño* y elija *Ranking*. En el menú desplegable, elija *Grado* (no importa si elige el primero o el segundo). Haga clic en *Aplicar*. El gráfico de red resultante debe tener un aspecto similar al de la Figura 10.9.

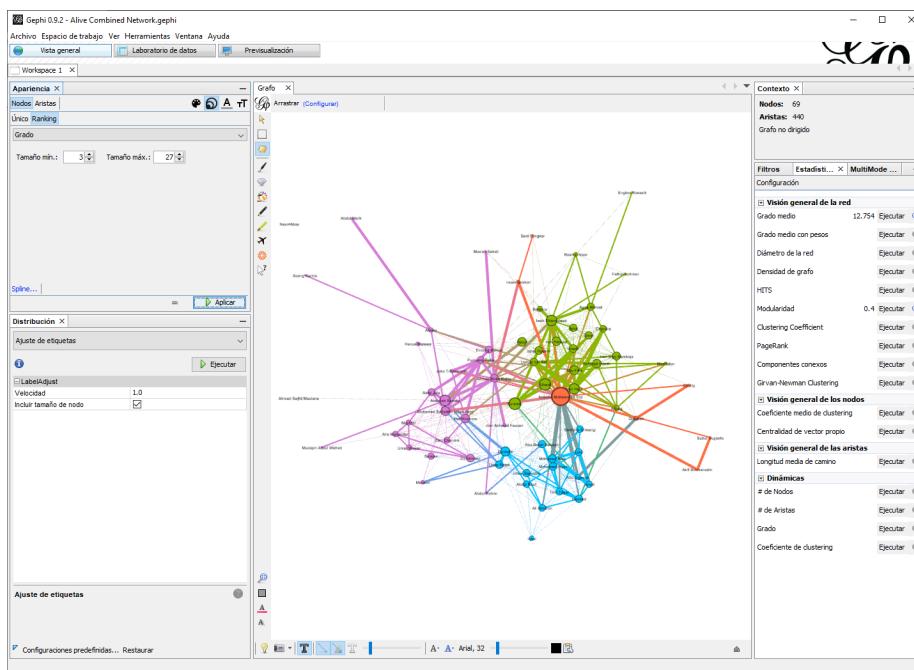


Figure 10.9: Red combinada de nodos vivos, el tamaño refleja el grado de centralidad

*Vista general
Estadísticas >
Grado medio
con pesos >
Ejecutar*

*Estadísticas >
Longitud media de camino
> Ejecutar*

*Estadísticas
> Centralidad
del vector
propio >
Ejecutar*

4. Ahora, repita el proceso excepto que esta vez use *Grado medio con pesos* en lugar de *Grado medio*. Mire los resultados en la ventana *Laboratorio de datos*. ¿Qué actor ocupa el puesto más alto? ¿Cuál es su puntuación? ¿Cómo se comparan las clasificaciones con las anteriores que no tenían en cuenta el peso del enlace?
5. Para obtener la centralidad de cercanía e intermediación en Gephi, podemos usar *Longitud media de camino* en la pestaña *Estadísticas*. Al hacer clic en *Ejecutar*, aparece un cuadro de diálogo (Figura 10.10), que presenta un par de opciones para tratar la red como dirigida o no dirigida (aquí, detecta correctamente que la red no está dirigida). Seleccione la opción para normalizar las centralidades en el rango. La normalización es buena para comparar diferentes medidas. Note que esta opción también calcula la centralidad de excentricidad, que (como indica el cuadro de diálogo) es la distancia desde un nodo en particular hasta el nodo más lejano en la red. Haga clic en *Aceptar* y Gephi producirá un informe, en el verá que Gephi calcula dos medidas de proximidad: proximidad (Freeman) y proximidad armónica. Dado que se trata de una red desconectada, utilizaremos la segunda. Como hicimos anteriormente, cambie a la ventana *Laboratorio de datos* y vea qué actores obtienen la puntuación más alta en términos de centralidad de intermediación y cercanía armónica (Harmonic Closeness Centrality).
6. Finalmente, para estimar la centralidad de eigenvector, usamos la función *Centralidad del vector propio* de Gephi. En el cuadro de diálogo, indique si la red está dirigida o no y haga clic en *Aceptar*. Al igual que con las otras funciones de centralidad, genera un informe y almacena las puntuaciones de centralidad para cada actor en la tabla de nodos.

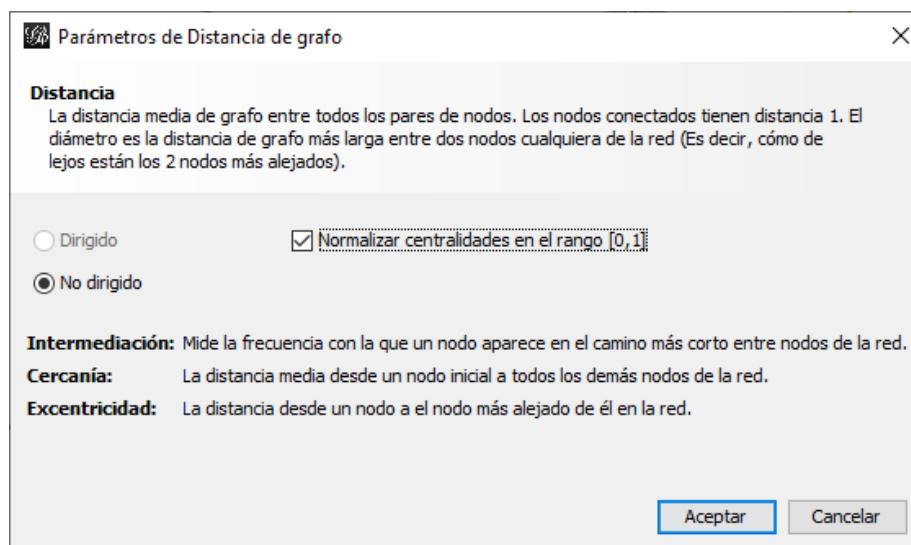


Figure 10.10: Diámetro de red/distancia promedio de camino

Chapter 11

Puentes y Corredores

En este capítulo, exploraremos una variedad de algoritmos para detectar puentes y corredores en una red: puntos y conjuntos de corte, bi-componentes y puentes, agujeros estructurales, e intermediación de borde. El tema central de este capítulo que el enfoque es localizar ubicaciones en la red valiosas en términos de flujo de recursos, materiales o información. A menudo, los analistas de redes requieren identificar un puñado de estrategias para disrupir la efectividad de una organización, aquí presentamos algunas opciones.

En general, un corredor es un actor que, en virtud de su ubicación estructural, posee el poder potencial para controlar el flujo de recursos a través de una red, mientras que un puente puede considerarse como un vínculo que conecta dos subgrupos. Un bi-componente es un componente que es inmune a la eliminación de un solo nodo, mientras que un punto de corte es un nodo que, si se elimina, desconecta la red. Un bi-componente con solo dos actores se considera un puente porque si se elimina, la red se desconecta. El concepto de agujeros estructurales fue desarrollado por Ron Burt. Se basa en la noción de Mark Granovetter de “la fuerza de los vínculos débiles,” pero en lugar de centrarse en el tipo de vínculo (es decir, fuerte frente a débil), se centra en detectar las brechas (es decir, huecos) en la estructura social. Finalmente, la intermediación de los bordes identifica los bordes (en lugar de los nodos) que se encuentran entre otros bordes de la red. Asume que los bordes que obtienen un puntaje alto en la intermediación tienen más probabilidades de ser puentes que los que no lo hacen.

Como en capítulos previos, comenzaremos por expandir a fondo los conceptos brevemente mencionados en el párrafo previo y concluiremos con un ejercicio práctico en Gephi.

11.1 Puntos de Corte

Un nodo, n_i es punto de corte si el número de componentes en el grafo inicial es menor que el número de componentes en el subgrafo que resulta al remover dicho nodo (Wasserman & Faust, 1994, p. 139). Por ejemplo, en el grafo izquierdo de la Figura 11.1 el nodo A es un punto de corte. ¿Por qué? Como dice la definición previa, al remover A se descompone el grafo en un número mayor de componentes (11.1 derecha).

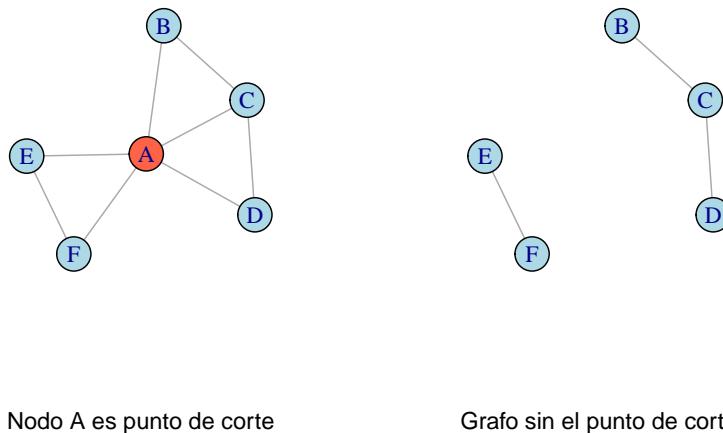
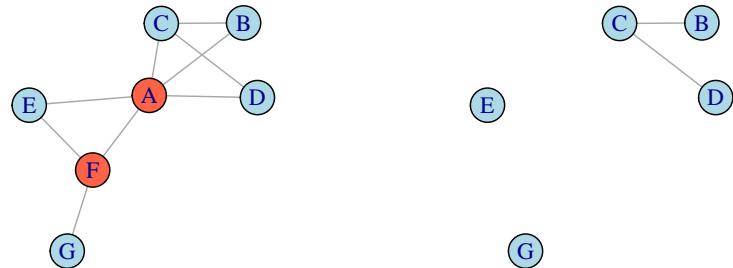


Figure 11.1: Ejemplo de punto de corte

El concepto de puntos de corte puede extenderse a un conjunto de nodos necesarios para mantener el grafo conexo, es decir un conjunto de corte (Wasserman & Faust, 1994, p. 140). De misma forma, el número de componentes en un grafo que contiene un conjunto de corte es menor que el que resulta al remover estos nodos. Por ejemplo, en la Figura 11.2 el conjunto de corte $\{A, F\}$ contiene dos nodos que al ser removidos incrementa el numero de componentes en la red.

Los puntos y conjuntos de corte son cruciales para las redes donde el flujo de recursos, información o materiales son cruciales para el objetivo de la organización. Por ejemplo, en una red de comunicación la información suele viajar a través de estos nodos entre partes distantes de la red. Por consiguiente, el remover estos intermediadores resulta en una ruptura de comunicación. Sin embargo, cabe recalcar que no todos los nodos en un punto o conjunto de corte son creados iguales. Por ejemplo, en la Figura 11.2 el remover el nodo F solo desconecta al nodo G del grafo y pero crea dos componentes en el proceso. Similarmente, el remover el nodo A de manera individual genera dos componentes $\{B, C, D\}$ y $\{E, F, G\}$. Sin embargo, se puede hacer un argumento que al remover A impactamos



Nodos A y F son el conjunto de corte

Grafo sin el conjunto de corte

Figure 11.2: Ejemplo de conjunto de corte

la red de mayor manera que al remover F. Es por ello, que los analistas de redes tienden a utilizar los puntos o conjuntos de corte en juntos con otras medidas de intermediación para determinar nodos o combinaciones optimas para disrupir una red de manera eficaz.

11.2 Bi-componentes y Puentes

Un bi-componente es formalmente un componente sin punto de corte. Por lo tanto, estos son invulnerables a fragmentación en múltiples subestructuras al remover un solo punto de corte (Everton, 2012a, p. 265). Por ejemplo, en la Figura 11.3 encontramos tres bi-componentes: {A, B, C}, {D, E, F, G} y {F, H, I}. Note el traslapo entre componentes, por ejemplo, {D, E, F, G} y {F, H, I} comparten el nodo F, el cual es un punto de corte. De hecho, los nodos compartidos entre bi-componentes son siempre puntos de corte.

A menudo, cuando los analistas hablan de bi-componentes se refieren a estructuras con tres o más nodos y sin un punto de corte (Everton, 2012a, p. 265). Adicionalmente, de particular interés son los vínculos entre bi-componentes, como la relación en la Figura 11.3 entre los nodos C y D en rojo. Para los enlaces existe una noción análoga al punto de corte, conocida como puentes. De acuerdo con Wasserman & Faust (1994) un puente es “una línea que hace que el grafo que contiene esta línea tenga menos componentes que el subgrafo que resulta una vez que se quita la línea (los nodos incidentes con la línea siguen estando en el subgrafo)” (Wasserman & Faust, 1994, p. 140). Otros autores definen los

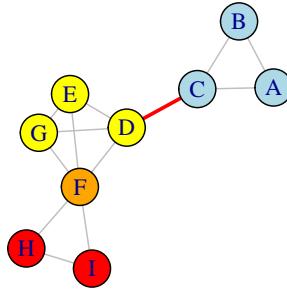


Figure 11.3: Análisis de bi-componentes

puentes como bi-componentes de dos nodos que al ser removidos desconectan la red (Everton, 2012a, p. 265). Ambas definiciones capturan un punto clave, los puentes son enlaces claves que conectan regiones de la red, sin ellos el flujo efectivo de recursos, ideas o materiales es imposible.

11.3 Intermediación de borde

La intermediación de bordes es definida de manera similar a la intermediación de nodos (ver Centralidad de Intermediación). Borgatti, Everett, & Johnson (2018) la definen como un recuento del número de veces que un enlace se encuentra en una ruta geodésica entre un par de vértices (Borgatti, Everett, & Johnson, 2018, p. 221). De manera formal podemos definir este índice con la misma notación desarrollada por Brandes (2001) y utilizada por Hashimoto y col. (2009) para describir la intermediación de bordes. Donde la intermediación de un borde e en el grafo G , se denota como $I(e)$ y mide la frecuencia en la que e aparece en el camino más corto (geodésico) entre dos distintos nodos a y b . Aquí σ_{ab} es el número de geodésicos entre a y b en un grafo G . Mientras que $\sigma_{ab}(e)$ es el número de geodésicos entre a y b que contienen e . Entonces, la intermediación de bordes sería:

$$I(e) = \sum_{a,b} \frac{\sigma_{ab}(e)}{\sigma_{ab}}$$

Donde la operación toma en cuenta todas las combinaciones distintas de nodos a y b en G . El resultado de esta operación es un índice numérico de intermediación para cada vínculo en la red.

La medida fue desarrollada por Girvan y Newman (2002) con el propósito de localizar vínculos entre comunidades. De acuerdo con los autores, si una red contiene comunidades conectadas a través de un numero escaso de vínculos intergrupales, entonces las rutas geodésicas entre estas comunidades deben atravesar estos enlaces (Girvan & Newman, 2002). Por consiguiente, estos enlaces entre comunidades deberán tener un índice alto de intermediación de borde.

Por ejemplo, en la Figura 11.4 los bordes han sido ponderados con base al índice de intermediación de cada enlace. En esta red los actores en cada subconjunto, $\{A, B, C, D\}$ y $\{E, F, G, H\}$, se encuentran densamente interconectados internamente. Adicionalmente, el subconjunto $\{E, F, G, H\}$ es adyacente a un nodo pendiente, I. El único camino entre estos subconjuntos consiste en el vínculo entre D y E. Por ello si un nodo en el subconjunto $\{A, B, C, D\}$ desea compartir información con alguien en el subconjunto $\{E, F, G, H\}$ o con I esta debe fluir a través de el vínculo con mayor intermediación, es decir el camino más corto más a menudo.

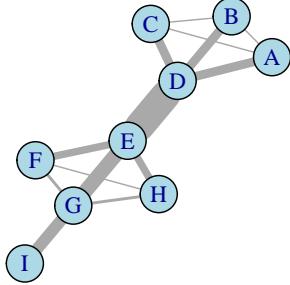


Figure 11.4: Representación esquemática de una red con dos comunidades densamente enlazadas entre si por vínculos con alta intermediación.

Recuerde que el algoritmo de intermediación optimiza el flujo entre nodos seleccionando los caminos mas cortos. Es decir que los vínculos conexos a pendientes tendrán un índice relativamente alto puesto que solo existe un camino para acceder dicho nodo. Por ejemplo, en la Figura 11.4 el vínculo entre {E, F, G, H} y I es relativamente alto, aunque solo conecta a un solo nodo al resto del grafo.

En términos prácticos, esta medida es utilizada a menudo en combinación con los puntos de corte y los puentes para iluminar caminos entre comunidades. A diferencia de los puentes, esta medida le provee al analista un índice continuo de valores. Es decir, a diferencia de una medida binaria esta presenta un rango de valores para que el analista decida que vínculos son más importantes y que diferencia cuantitativa existe entre estos.

11.4 Agujeros Estructurales, Restricción y Autonomía

La última medida de corretaje que cubriremos en este capítulo es autonomía. Sin embargo, para poder entender esta medida es importante presentar el contexto teórico y matemático presentados por el autor de este índice, Ronald Burt.

La teoría de los agujeros estructurales de Burt (1992) propone que las personas acumulan recursos sociales, o capital social, el cual pueden invertir en ciertos contextos sociales para generar una ganancia. Un agujero estructural es una ubicación dentro de la red donde los actores se encuentran desconectados. Por consiguiente, estos agujeros presentan oportunidades donde los actores en la red pueden invertir dicho capital social y cerrar estas aperturas conectando a dos o más actores adyacentes (Monge & Contractor, 2003, p. 143). Otra manera de pensar en el valor de los agujeros estructurales es que estos permiten que un actor funcione como intermediador o corredor entre otros nodos no-conexos entre sí. Por lo tanto, con base a su posición en la red el corredor tiene la habilidad de influenciar el flujo de información, es decir, puede mantener un agujero estructural abierto y continuar como intermediario o puede cerrarlo.

Por ejemplo, en la Figura 11.5 el actor A se localiza entre tres agujeros estructurales, estos son los espacios abiertos entre los nodos {B, H, F}. Estos tres agujeros ponen a A en una posición de corretaje donde esta puede coordinar el intercambio de información, materiales o recursos entre {B, H, F}. Adicionalmente, A se encuentra conectado a una variedad de círculos sociales, por lo tanto, tiene mayor acceso a recursos, información o materiales de cada grupo. En contraste, los nodos en cada comunidad periferia ($\{B, C, D\}$, $\{H, I, J\}$ y $\{F, E, G\}$) tienen menor acceso a información no redundante puesto que cada grupo se encuentra esencialmente aislado.

Para construir un índice que captura las propiedades de los agujeros estructurales,

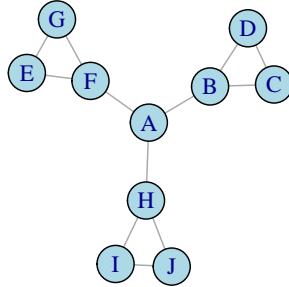


Figure 11.5: Red social hipotética donde a se encuentra entre tres agujeros estructurales

Burt utiliza como base la red del ego para cada actor, es decir, las triadas en las que se encuentra incrustado cada actor (Cunningham, Everton, & Murphy, 2016, p. 181). Para ilustrar este punto, examinemos la Figura 11.6 que contiene cuatro posibles cuatro configuraciones posibles de la misma triada de actores ($\{A, B, C\}$). La primera triada (1) es una cerrada o completa; por consiguiente, ningún nodo se localiza entre agujeros estructurales y por extensión goza de una posición de corretaje. Las otras tres triadas (2 a 4), son abiertas y cada una contiene un agujero estructural y un nodo en posición de corredor. En esta Figura el ego, actor A, a sido coloreado de rojo mientras que los alternos se encuentran en azul. Burt sostiene que los alternos no conectados ofrecen recursos diferentes al ego (Burt, 1992). Por lo tanto, desde la perspectiva del ego la configuración preferible es aquella que observamos en la segunda triada (2), donde el ego se encuentra en una posición de corretaje.

El índice desarrollado para localizar nodos con poco grado de intermediación es la restricción. Esta mide cuantos alternos al ego tienen enlaces entre sí, lo cual determina el grado en que el ego invierte tiempo y energía en alternos que se conocen entre sí (Borgatti, Everett, & Johnson, 2018). Otra manera de pensar en restricción es pensar en la triada cerrada ((1) en la Figura 11.6). Esta es la triada con la mayor restricción para el ego. Por lo tanto, las acciones y percepciones del ego son controladas por sus alternos. De manera opuesta, en una red del ego sin restricción ((2) en la Figura 11.6) los alteros no están conectados entre sí, por ello estos no pueden coludir entre si para influenciar las acciones o el acceso a recursos de ego (Valente, 2010, p. 70).

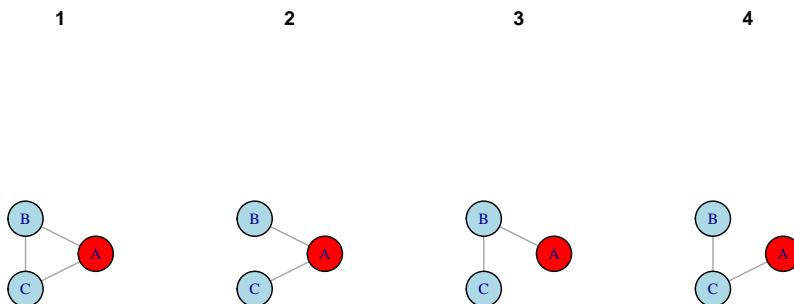


Figure 11.6: Cuatro tipos de triadas

Note que la medida de restricción no mide de manera directa el numero de agujeros estructurales. De lo contrario, un nodo con alta restricción se encontrará a menudo entre pocos agujeros estructurales. Por lo tanto, a diferencia de otros índices del vértice, actores con baja restricción son aquellos en una posición beneficiosa.

De manera formal, podemos definir la restricción de cada nodo i como:

$$C_i = (p_{ij} + \sum p_{iq}p_{qj})^2, q \neq i, j$$

Donde p_{ij} representa la porción de los enlaces de i compartidos con el alterno j . Además, p_{iq} representa la proporción de los enlaces de i compartidos con el alterno q , la cual se multiplica por p_{qj} , que representa la proporción de los vínculos de q compartidos con el alterno j .

Pongamos en práctica esta formula con un ejemplo sencillo, proporcionado por Ruqin Ren aquí. En la Figura 11.7 vemos una red de 5 nodos. Como práctica, calculemos la restricción de el nodo A a mano.

Primero, note que A se encuentra conectado a dos nodos adyacentes ($\{B, C\}$), por ello, calcularemos la restricción para entre el ego A y cada nodo adyacente. Utilizaremos la centralidad grado (ver (11.4)) para determinar la porción de vínculos entre dos pares de nodos. Por ejemplo, la porción de vínculos entre el ego (A) y cada alterno (p_{ij}) es 0.5 puesto que A solo tiene un vínculo con B y C respectivamente, pero dos enlaces en total ($\frac{1}{2}$).

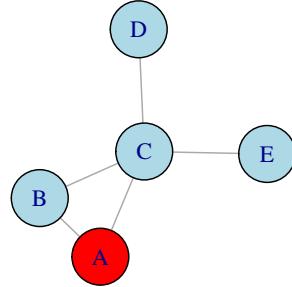


Figure 11.7: Red pequeña

	Grado
A	2
B	2
C	4
D	1
E	1

Podemos calcular los valores restantes de utilizando los indices en centralidad de grado de la siguiente manera:

Origen (i)	Destino (j)	Camino (ij)	Camino (iq)	Camino (qj)
A	B	1	1	1
	C	1	1	1
		Porción (ij)	Porción (iq)	Porción (qj)
A	B	$\frac{1}{2} = 0.5$	$\frac{1}{2} = 0.5$	$\frac{1}{4} = 0.25$
	C	$\frac{1}{2} = 0.5$	$\frac{1}{2} = 0.5$	$\frac{1}{2} = 0.5$

Con estos valores, podemos calcular la restricción del nodo A:

$$C_A = C_{AB} + C_{AC} = (0.5 + (0.5 \times 0.25))^2 + (0.5 + (0.5 \times 0.5))^2 = 0.953$$

En términos prácticos, el software calculará este índice de manera eficaz. Sin embargo, es importante entender un poco sobre las mecánicas de este algoritmo y los índices resultantes. El valor máximo de restricción es 1.125, por lo tanto, en el ejemplo previo suponemos que el nodo A tiene una restricción alta y por lo tanto pocas opciones de corretaje.

Cunningham, Everton, & Murphy (2016) proponen invertir el índice de restricción para generar un valor de “autonomía,” es decir, el antónimo matemático. La autonomía se calcula restando el índice de restricción para cada nodo del valor máximo teórico de restricción (1.125). El valor resultante indica un valor potencial de corretaje, donde puntajes altos reflejan un alto potencial. Por ejemplo, el índice de restricción para el nodo A es 0.953, por lo tanto, su autonomía sería equivalente a 0.172. En otras palabras, A no se encuentra en una posición de corretaje. Contrastemos esta observación con la autonomía del nodo C, el cual tiene un índice de restricción de 0.406. Por lo tanto, su autonomía es equivalente a 0.718. Como notara en la Figura 11.7, el nodo C se encuentra entre múltiples agujeros estructurales y su autonomía indica un potencial de corretaje es moderadamente alto.

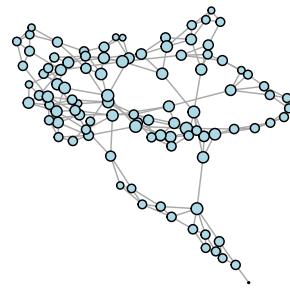


Figure 11.8: Red con nodos dimensionados por autonomía

Puesto que la idea de autonomía es una medida más intuitiva que la restricción, muchos analistas tienden a utilizarla a menudo para diferenciar de manera visual o cuantitativa los nodos con un potencial de corretaje alto (vea Figura 11.8). Sin embargo, ambos índices son valiosos en el contexto adecuado.

11.5 Ejercicio Práctico

Desafortunadamente, por el momento Gephi solo calcula una de estas métricas: centralidad de puente, y deberá descargar el complemento Bridging Centrality para calcularlo.

Para este ejercicio usaremos la red de comunicación de la planta de procesamiento de madera donde los trabajadores rechazaron un nuevo paquete de compensación y finalmente se declararon en huelga. Por ello, la gerencia contrató a un consultor externo para analizar la estructura de comunicación del empleado porque consideró que los negociadores sindicales no estaban comunicando eficazmente la información sobre el paquete a todos los empleados. El consultor externo pidió a todos los empleados que indicaran, en una escala de 5 puntos, la frecuencia con la que hablaban de la huelga con cada uno de sus colegas, desde ‘casi nunca’ (menos de una vez por semana) a ‘muy a menudo’ (varias veces por día). El consultor utilizó 3 como valor de corte para identificar un vínculo entre dos empleados. Si al menos una de las dos personas indicaron que hablaron sobre el trabajo con una frecuencia de tres o más, se incluyó un vínculo entre ellos en la red.

La red se encuentra clasificada en grupos por edad e idioma: empleados hispanohablantes de 30 años o menos, empleados angloparlantes de 37 años o menos y empleados angloparlantes de 38 años o más. Todos los vínculos entre grupos tienen antecedentes especiales. Entre los hispanos, Alejandro es el más competente en inglés y Bob habla algo de español, por lo que existe un vínculo entre ellos. Bob le debe a Norm por conseguir su trabajo y Ozzie es el padre de Karl, lo que ayuda a explicar sus lazos. Sam y Wendle son los negociadores sindicales y fueron los encargados de explicar el nuevo programa propuesto por los gerentes.

11.5.1 Puentes y Corredores en Gephi

Gephi
Archivo >
Abrir

1. En Gephi, primero importe el archivo **strike.net** (los archivos Pajek (.net) son un formato común de herramientas para ARSo). Asegúrese de indicar que el archivo no está dirigido y de elegir *Mínimo* como estrategia de combinación de bordes. A continuación, importe el archivo de atributos **brokerage_scores.csv**, este debe leerse como una tabla de nodos separados por comas. En el segundo de los dos cuadros de diálogo de importación (Figura 11.9). Finalmente, en el informe de importación de Gephi, indique que desea agregar estos datos al espacio de trabajo existente y haga clic en *Aceptar*.

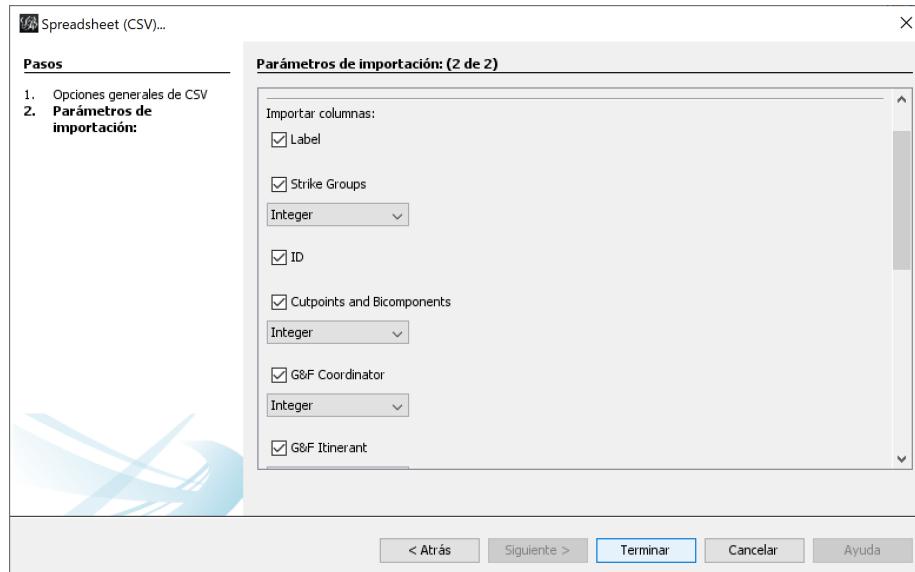


Figure 11.9: Cuadro de diálogo Importación de Gephi

Laboratorio de datos
Tabla de Datos
> Nodos

2. Verifique la *Tabla de nodos* en la ventana *Laboratorio de datos* para verificar que los atributos se hayan importado correctamente. Regrese a la ventana *Descripción general* y visualice la red donde el color del nodo refleja el grupo en el que se clasificaron (empleados hispanohablantes de 30 años o menos, empleados angloparlantes de 37 años o menos y empleados angloparlantes de 38 años o más). La Figura 11.10 presenta un ejemplo.
3. En la red, el vínculo entre Alejandro y Bob es el único canal de intercambio de información entre los empleados hispanos y todos los demás empleados. Eliminarlo aislaría a los empleados hispanos de la información que circula entre los demás empleados. Dicho de otra manera, cuando quitas el lazo entre Alejandro y Bob, desconectas a los trabajadores hispanos del resto de la red de comunicación. Tenga en cuenta que eliminar el vínculo entre Frank y Gill también desconecta la red: Frank se convierte en un aislado.

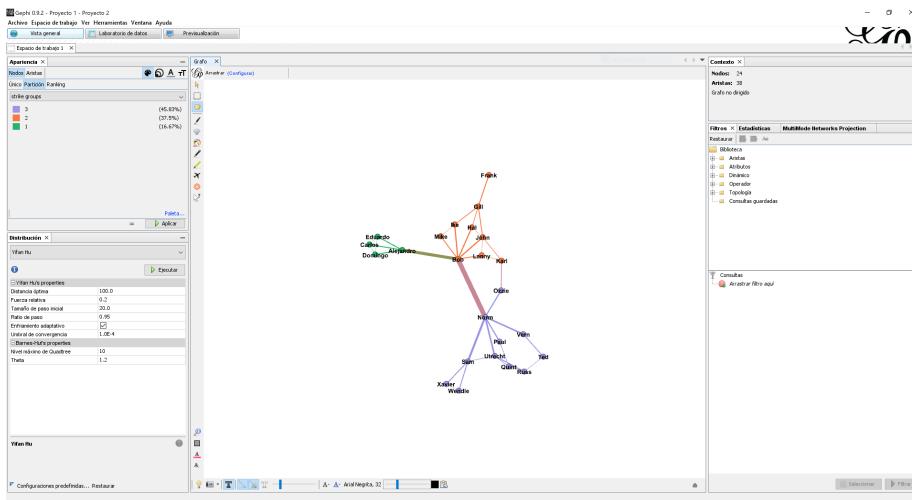


Figure 11.10: Red Huelga, Color = Strike Groups, Peso del borde = Intermediación de Borde

4. La eliminación de un actor puede tener el mismo efecto que la eliminación de un enlace. Si Bob fuera removido, todos sus vínculos desaparecerían, incluido el puente con Alejandro. Bob se conoce como punto de corte, punto de articulación o amplificador de límites (según el programa que utilice) porque eliminarlo desconecta la red. Los puntos de corte son cruciales para el flujo de información en una red. Son “cuellos de botella” en el sentido de que controlan el flujo de recursos de una parte de una red a otra. Formalmente, los puntos de corte son actores que pertenecen a dos o más bi-componentes, que son subredes que son invulnerables a la eliminación de un solo actor. En un bi-componente, ningún actor puede controlar completamente el flujo de información entre otros actores porque siempre hay un camino alternativo que los recursos pueden seguir.

Vista General
Apariencia >
Nodos > Color
> Partición >
Aplicar

Estadísticas >
Grado medio >
Ejecutar
Apariencia
> Nodos >
Tamaño >
Ranking >
Aplicar

5. Actualmente, Gephi no puede identificar bi-componentes y puntos de corte, por ello hemos importado los resultados de otros programas a Gephi. Coloree los nodos para reflejar los bi-componentes y los puntos de corte en la red Strike seleccionando la partición *cutpoints and bicomponents*. Ahora, los puntos de corte (Alejandro, Bob, Norm, Sam y Gill) deben ser todos del mismo color, mientras que los colores del resto de los actores deben reflejar sus respectivos bi-componentes. Puede comprobar por sí mismo que la eliminación de un punto de corte desconectará la red.
6. A continuación, en la pestaña *Estadísticas* en la ventana *Visión general de la red*, Ejecute la función *Bridging Centrality*. Esto genera un informe, que puede cerrar. También calcula tres métricas: centralidad de intermediación, coeficiente de vinculación (agrupamiento) y centralidad de vinculación, siendo esta última el producto de las dos primeras. Primero, ajuste el tamaño de los nodos en términos de centralidad de intermediación. **¿Existe alguna asociación entre la centralidad de intermediación y qué nodos son puntos de corte?** A continuación, ajuste el tamaño de los nodos en términos de coeficiente de puenteo (agrupamiento). **¿Qué sucedió?** Por último, ajuste el tamaño de los nodos en términos de centralidad de puenteo (Figura 11.11). **¿Se compara esta medida con qué nodos son puntos de corte? ¿Cómo se compara con la centralidad de intermediación?** Recuerde, puede verificar las puntuaciones de centralidad de intermediación y puente en la ventana “Laboratorio de datos.”
7. Finalmente, ajuste el tamaño de los nodos en términos de puntajes de corretaje totales (no dirigidos) (*undirected g&f total*) (Figura 11.12) y luego la medida de autonomía de Burt (*burt's autonomy*) (Figura 11.13). **¿Cómo se comparan estas visualizaciones con las anteriores? ¿Alguno parece más útil que otros?**

11.5. EJERCICIO PRÁCTICO

207

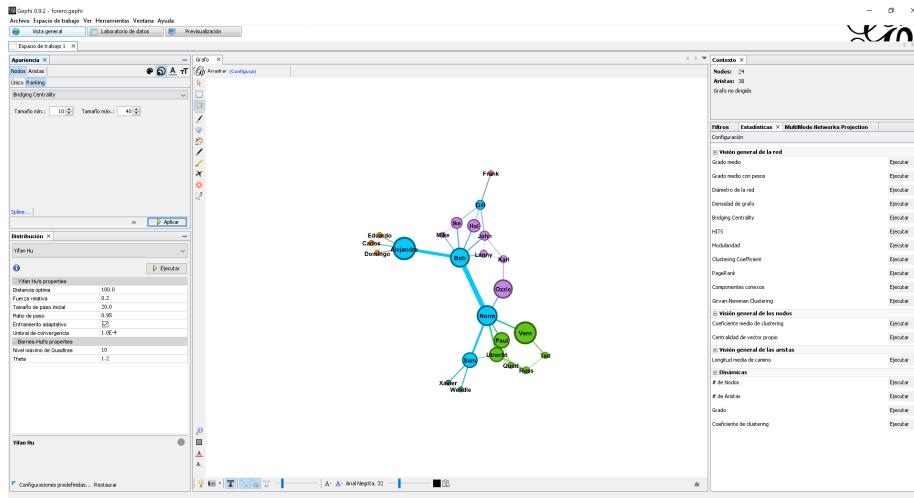


Figure 11.11: Red Huelga: puntos de corte, bi-componentes y centralidad de puente

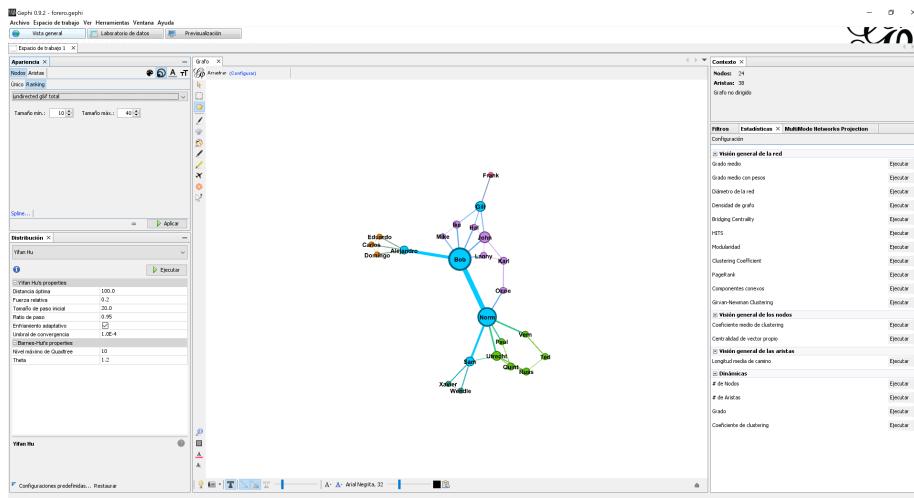


Figure 11.12: Red Huelga, puntos de corte, bi-componentes y corretaje

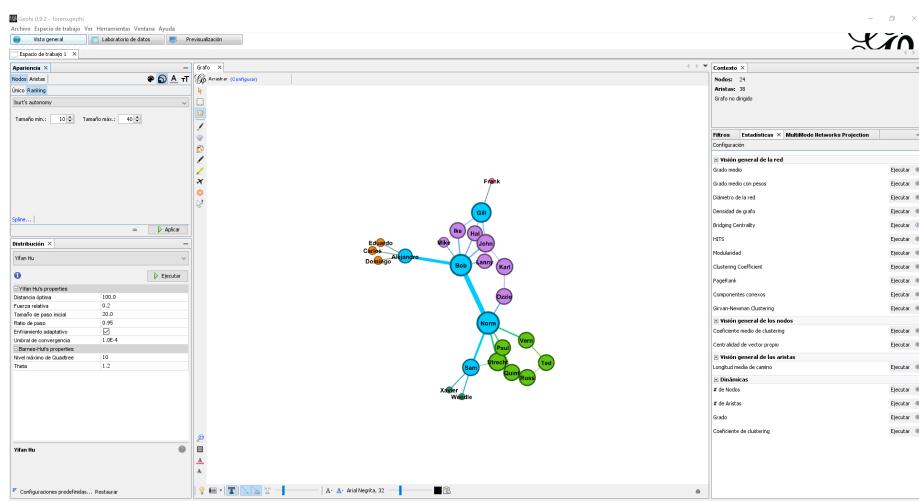


Figure 11.13: Red Huelga: puntos de corte, bi-componentes y autonomía

Part IV

Aplicaciones

Chapter 12

Redes Oscuras en el Perú

Este proyecto de tener una doctrina propia, que se adapte a la realidad del Perú es un punto de vista del autor sobre la introducción al estudio de las redes oscuras como elemento de análisis clave para todo analista de inteligencia y elementos operativos de fuerzas especiales, basados en el estudio analítico de la interacción de las redes sociales.

Este capítulo trata de definir lo que es una “red oscura,” palabras muy escuchadas en los trabajos de organización, redes y jerarquías, basada en la definición de guerra en red, de los Doctores Arquilla y Ronfeldt (2001), así como en definiciones y postulados de Everton (2012a), Cunningham y col. (2016), y de la Escuela Superior de Guerra del Ejercito del Perú, Escuela de Post Grado (ESGE/EPG).

En este capítulo único nos abocaremos, principalmente a definir lo que se denominan redes oscuras en el Perú, su organización, características y la metodología para degradarlas y neutralizarlas. Así mismo estudiar los métodos aplicados por los analistas de Inteligencia, de Fuerzas Especiales, de Asuntos civiles y Operaciones Psicológicas en la identificación de redes oscuras.

12.1 Análisis de Redes Oscuras

Cuando hablamos del análisis de redes sociales (ARSo), aplicadas al terrorismo, crimen organizado trasnacional (COT) y otros flagelos de este tipo, podemos decir que nos referimos a las redes oscuras. Las cuales podemos definir como un conjunto de entidades clandestinas, encubiertas, ilegítimas, subrepticias, delictivas o escondidas fuera de la ley que se relacionan entre sí (Cunningham, Everton, & Murphy, 2016). Por ejemplo, las organizaciones que participan en el terrorismo, la minería ilegal, el tráfico ilícito de drogas, la corrupción y otros crímenes que forman parte de lo que se denomina COT y las redes ilícitas que este desarrolla.

El análisis de redes sociales se ha utilizado en el estudio de redes oscuras de manera exitosa para comprender mejor las redes de terroristas/ insurgentes (Cunningham, Everton, & Murphy, 2016; Koschade, 2006; Krebs, 2002), distribución de narcóticos (Morselli & Petit, 2007), dinámicas entre pandillas (Andrew Papachristos & Braga, 2013), intercambio de información en grupos de hackers (Decary-Hetu, 2014), entre otras aplicaciones. En el contexto ilícito, el objetivo del analista puede ser el utilizar las herramientas empíricas del ARSo, que hemos visto a lo largo de este manual, para captar y comprender las estructuras de estos grupos con el objetivo de mejorar el entendimiento de estas y forjar mejores estrategias de disruptión. Es por ello que el estudio de redes oscuras, a través del prisma del ARSo como se ha presentado, llega a ser un método ordenado, sistemático y lógico para estructurar datos, visualizar redes, y generar hipótesis sobre el comportamiento de la red y sus miembros que se pueden llegar a sustentar por medio de herramientas y métodos adicionales.

Este proceso de entendimiento de las redes oscuras, desde nuestra perspectiva, se centra alrededor de cuantificar las relaciones entre entidades sociales para entender el contexto social en el que se sitúa un actor (Wasserman & Faust, 1994). Igualmente suponemos que el comportamiento de los actores se ve afectado por sus vínculos con los demás y las redes en las que estos están incrustados (ver Capítulo 1 para una lista completa de los supuestos en el análisis de redes). Por ejemplo, en una red terrorista, los vínculos pueden servir como conductos de información para compartir ideas, mejores tácticas criminales o terroristas, tecnología, conocimiento, acciones, etc. Al analizar la red desde la perspectiva del análisis de redes sociales oscuras, intentamos reconocer las relaciones entre entidades para plasmarlas en un mapa del sistema social y llegar a identificar :

- La macro, meso, y microestructuras de la red (ver Capítulos 8, 9 y 10 de este manual);
- Quienes forman parte de la red y donde se encuentran (periferia o centro); y
- Las distancias sociales entre nodos, las cuales proveen pistas sobre el potencial flujo de información, conocimiento u otros recursos compartidos entre nodos

Como conjunto de teorías y métodos empíricos, el estudio de redes oscuras a través del ARSo se da por medio de la recolección y uso de conjuntos de datos de relaciones entre actores (llamados nodos o vértices). Este manual incluye ejercicios analíticos utilizando la red terrorista Noordin Top (Roberts, Everton, & Cunningham, 2014) (ver Apéndice B, “Libro de Códigos”). Este conjunto de datos contiene :

- Relaciones, divididas en dos modalidades:
 - Modo Uno: relaciones entre el mismo conjunto de actores (persona a persona, organización a organización, grupo insurgente a grupo insurgente, etc.), estas pueden ser dirigidas o no dirigidas. En este conjunto de datos, los autores se centran en analizar el dominio humano. Por ello, las relaciones modo uno han sido recopiladas de

persona a persona. Por ejemplo, en la red de Noordin Top los autores incluyen lazos de comunicación, parentesco, compañerismo en clases, reclutamiento, amistad, mentorías, entre otros.

- Modo Dos: relaciones entre dos conjuntos de actores (persona a evento, persona a grupo insurgente, etc.). En la red de Noordin Top, los autores incluyen relaciones de modo dos como afiliaciones comerciales entre personas y organizaciones con fines de lucro, afiliación con una organización terrorista/insurgente, participación en entrenamientos designados para enseñar el conocimiento y habilidades terroristas entre otros. Como mencionan Cunningham y colegas, el recopilar interacciones entre miembros de una organización que intenta mantenerse oculta es difícil por naturaleza, por ello en ocasiones los analistas de redes recopilan sus datos de afiliación como modo dos, pues el registrar cuales sitios o eventos visitan miembros de una red oscura en relación con la organización puede ser más viable (Cunningham, Everton, & Murphy, 2016).
- Atributos: Datos no relacionales sobre los actores de la red, como su nivel educativo, entrenamiento militar, nacionalidad del individuo, rol, etc.

El ejemplo de la red terrorista de Noordin Top demuestra el uso de vínculos formales (personas a criminales) e informales (amistad) en el análisis de la red con el objetivo de entender como la combinación de relaciones puede facilitar o impedir, generar o degradar la creación y difusión de ideas en estas organizaciones criminales. Más aun, una de las dificultades principales de trabajar con redes oscuras es la interdependencia entre múltiples diferentes entidades ilícitas (Everton, 2012a); por ejemplo, vemos evidencia de vínculos entre Al Qaeda y estados fallidos en África Occidental, donde la compra y venta de diamantes provee los fondos necesarios para Al Qaeda y de manera recíproca este mercado ilícito facilita la transferencia de armamento a los combatientes de milicias locales (Raab & Milward, 2003). Similarmente, en Perú donde uno de los desafíos de seguridad involucra una combinación interdependiente de terrorismo, narcotráfico y otras actividades ilícitas el entender las redes que habilitan estos desafíos requiere un enfoque holístico basado en la interrelación entre actores, así como las organizaciones a las que pertenecen y como estas se interconectan entre ellas.

12.2 ¿Por qué el Análisis de Redes Oscuras?

Cabe preguntarse, ¿por qué es importante el análisis de las redes oscuras? En los conflictos modernos el dominio humano es el factor más importante que todo adversario quiere conocer, influenciar y dominar; por ello, el comprender las redes sociales del enemigo, es el elemento más crítico y desafiante en los conflictos del Siglo XXI. El analizar el dominio humano con una herramienta empírica, que facilite el estudio e investigación de las organizaciones criminales,

grupos radicales y/u organizaciones terroristas, tiene un valor incalculable para los miembros de las fuerzas armadas y seguridad nacional.

En términos prácticos, el análisis de redes oscuras nos permite de manera sistemática contestar preguntas sobre las redes ilícitas de interés. Por ejemplo, si el enfoque del analista es detectar nodos cruciales para la red. Uno identificaría a los nodos centrales los cuales pueden representar actores muy activos en la red, los cabecillas de estas organizaciones, o miembros de alto valor. Con esta información a la mano, una estrategia para neutralizar la red podría centrarse en remover dichos nodos valiosos, continuar investigándoles, etc. Sin embargo, como menciona Everton (2012a), el enfoque en identificar y remover nodos de “alto valor” suele ser una estrategia atractiva y a plazo corto puede proporcionar satisfacción; sin embargo, el remover nodos centrales en una red estructurada alrededor de un puñado de actores importantes no es lo mismo que tomar la misma acción contra una red sin jerarquía (Everton, 2012a). Por ello la estrategia del analista de redes dependerá del tipo de estructura de red.

Puesto a que las propiedades de las organizaciones terroristas o del COT varían con base a el contexto social, histórico y geográfico de cada nación, el análisis de redes es una herramienta interesante y necesaria para toda agencia de Inteligencia o Policial . Las conexiones entre nodos en una red sociales son mantenidas entre dos o más nodos, los nodos pueden ser personas, grupos, organizaciones, estados, etc. Traducido a términos relevantes al analista de Inteligencia , Fuerzas Especiales y Policiales podemos contextualizar los nodos como sujetos estratégicos.

Como ha sugerido este autor, la estrategia del analista debe basarse en el contexto de la red, así como la estructura de esta. La Figura 12.1, adaptada de una conferencia sobre Redes Sociales a cargo de la Doctora Nancy Roberts de la Escuela Naval de Postgrado, presenta la estructura de redes como polos opuestos (Roberts, 2014). Esta dicotomía se basa en el estudio de organizaciones, la cual separa las jerarquías de las redes. Asimismo, compartimenta características, asociando las redes con estructuras sociales informales, igualitarias y de gobernanza compartida; igualmente, asume que las organizaciones jerárquicas son formales, estructuradas y con gobernanza centralizada (Everton, 2012a).

Al analizar redes oscuras es importante preguntarse si esta distinción es útil. Algunos (Everton, 2012a) argumentan que la distinción es apropiada en algunos contextos, pero desde la perspectiva del análisis de redes esta división es inapropiada (Everton, 2012a). Esto es puesto que tanto en las organizaciones jerárquicas como en las redes el análisis se basa en las relaciones entre actores, el entendimiento sobre las propiedades organizacionales de la red se determina por medidas estructurales que cuantifican la jerarquía o heterarquía de dicha red. En este manual, hemos explorado las medidas estructurales de las redes en el Capítulo 8 “Topología.”

En términos prácticos, un analista de redes puede comenzar su exploración de redes investigando esta dimensión estructural. Por ejemplo, haciéndose la

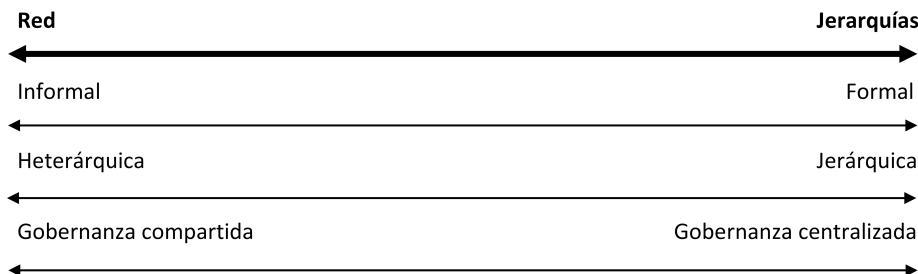


Figure 12.1: Dimensiones de Organizaciones

pregunta “El MOVADEF, ¿cómo está estructurado? ¿Cómo red o jerarquía?” es un buen comienzo en entender las características estructurales de la red con el fin de generar la estrategia que mejor se acople a la realidad. Es importante mencionar, que hay organizaciones que, sin tener una jerarquía establecida, tienen un orden informal. Para aclarar esto podemos utilizar un ejemplo presentado por el Capitán de Navío, R. Devoto (2016) en su presentación para los oficiales alumnos de la Escuela de Inteligencia del Ejército en el curso básico y superior de Intelligencia:

Dentro de Sendero Luminoso, teníamos comunistas y simpatizantes de acuerdo a su propia definición. Luego estos se subdividían en Dirigentes, Cuadros, Militantes, Combatientes y Masas, implícitamente ya adoptan una jerarquía a pesar de que siempre pregonaban que todos eran iguales

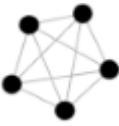
Podemos complementar las observaciones de Cap. de Navío Devoto, con la definición de la guerra en red de Arquilla y Ronfeldt (2001, página 6), que para ser exacto citare textualmente:

El termino de guerra en red hace referencia a un modo emergente de conflicto (y de delincuencia) en el ámbito social, alejado de la guerra militar tradicional, en el que los protagonistas utilizan estructuras de organización en red y doctrinas, estrategias y tecnologías en relación con aquellas, acordes con la era de la información.

Esta definición nuevamente retorna el enfoque a la estructura de la red, pero incluye las cualidades adicionales de las redes oscuras, como las doctrinas, estrategias y tecnologías que van mano a mano con la guerra en red. Nuevamente, podemos recalcar el valor del análisis sistemático de redes utilizando los métodos empíricos presentados en este manual y en la literatura del ARSo.

El ejemplo previo de Sendero Luminoso recalca una de las cualidades adicionales de la guerra en red, según Arquilla y Ronfeldt, una red se compone de muchas piezas pequeñas y dispersas, las cuales comparten conjuntos de ideas y están

interconectadas para lograr un fin común (Arquilla & Ronfeldt, 2001). Nuevamente, Arquilla y Ronfeldt presentan tres tipos de topologías de redes que podemos ver en la siguiente tabla:

Tipo	Explicación	Dibujo
Cadena	Como en una cadena de contrabando de personas, bienes, o información donde estos se mueven a lo largo de la línea de contactos y donde la comunicación de un extremo a otro viaja únicamente por nodos aledaños.	
Estrella	Donde un conjunto de actores está ligado con un nodo central o actor, para comunicarse entre sí o coordinarse deben atravesar el nodo central.	
Malla	Una estructura de red en la que todo el mundo está conectado.	

En una red oscura, la composición estructural puede contener una mezcla de estas tipologías (Figura 12.2). Por ejemplo, el núcleo de la red puede estar organizado como malla, pero en estrellas o cadenas en otras partes de la red como en los comandos operativos (Arquilla & Ronfeldt, 2001). Más aun, los actores generalmente están conectados entre sí de varias maneras, y estas conexiones pueden empujarlos en direcciones múltiples y a veces contradictorias (Cunningham, Everton, & Murphy, 2016). Por ejemplo, la mayoría de los individuos tienen lazos de intercambio, amistad y parentesco; las organizaciones criminales y terroristas generalmente, intercambian personal, armas, dinero, información y forman alianzas entre ellos.

En el Perú, el análisis de redes oscuras nos ha permitido mapear, degradar, y neutralizar redes criminales como la red criminal “Los cuellos blancos del Puerto,” la facción de Sendero Luminoso denominada Línea Oportunista de Derecha (LOD), entre otras (ver Capítulo 13, “Ejemplos Prácticos” de este

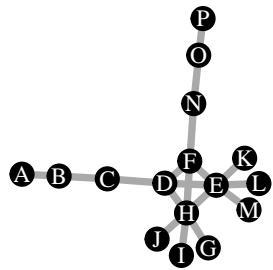


Figure 12.2: Red hipotética, mezcla de cadenas, estrellas y mallas

manual). El análisis de redes oscuras también se ha utilizado para verificar la convergencia que existe entre las organizaciones criminales y los partidos políticos, tema de alta importancia pues en los últimos años hemos presenciado en el mundo y especialmente en la región Sudamericana, el surgimiento de redes de actores, dedicados a limitar las acciones de fuerza de diversos gobiernos, así como activistas que reclaman la no contaminación del medio ambiente, la movilización en los conflictos sociales contra las asimetrías de poder, terroristas en convergencia con partidos políticos de ideología radical a fin de derrocar a un gobierno legítimamente constituido, organizándose en redes inmensas formadas por poblaciones interconectadas por múltiples canales, apoyándose en la internet, desde donde se han apoyado, coordinado y organizado (Arquilla & Ronfeldt, 2001).

Hay que tener en cuenta que la difusión del crimen organizado trasnacional constituye la amenaza más importante para la seguridad de las Américas (Garzon, 2013), y estas organizaciones se manifiestan en redes. Es por ello, que nuevamente retornamos a las sabias palabras de Arquilla y Rondefeld, quienes afirman que “se necesita una red para luchar contra las redes” (Arquilla & Ronfeldt, 2001). Esto depende, hasta cierto punto de nuestra habilidad de comprender al enemigo y desarrollar estrategias innovativas basadas en nuestro entendimiento de la red.

12.3 Conclusiones

Para concluir este breve capítulo, el autor propone algunas conclusiones basadas en la literatura del análisis de redes (dentro y fuera de este manual), la realidad en el Perú y sus observaciones como catedrático y miembro en situación de retiro de las fuerzas armadas:

1. Las redes y organizaciones pertenecen a un sistema complejo e interdependientes y hay que estudiarlas como un todo, y cuando estas son demasiadas extensas podemos dividirlas en subsistemas, pero sin perder el punto de vista sistémico. Por ejemplo, las redes terroristas pueden dividirse por su afiliación política o ideológica, sin embargo, es importante no olvidar como estas se interrelacionan.
2. Las redes sociales oscuras sean cual fuera su tipo, son difíciles de comprender porque pueden ser de cualquier estructura y pueden estar totalmente dispersas.
3. Observar las redes/organizaciones como sistemas nos ayuda a identificar a sus elementos, características y componentes. Entendiendo todos estos aspectos podemos contrarrestar a este tipo de organizaciones con estrategias que tomen en cuenta los patrones estructurales de la red, sus miembros, y los retos y oportunidades que presentan.
4. Debido a la complejidad de las redes y la dificultad en codificar interacciones sociales, es crucial que los equipos intra e interorganizacionales de seguridad nacional creen y compartan estándares de documentación, recopilación y codificación de datos de red. En este manual, hemos presentado una serie de “libros de código” (ver Apéndice B, “Libro de Códigos”) para ejemplificar el uso de esta estrategia en el análisis de redes oscuras. En el Perú, en la instrucción de ARSo hemos denominado este documento guía como “libro de términos.” El enfoque ha sido el incluir definiciones y características de los grupos en estudio adaptadas a la realidad del país. Por ejemplo, en el libro de código de Roberts, Everton y Cunningham (2014) los autores utilizan términos como “tribu” el cual debe ser adaptado a “centro poblado menor” (CPM). Además de términos y definiciones consistentes, es importante tener un entendimiento común en diferentes niveles organizacionales de los departamentos de defensa y seguridad nacional. Por ejemplo, el libro de términos creado para el estudio del MOVADEF (brazo político de Sendero Luminoso) debe ser el mismo para todos los niveles de análisis encargados de su estudio, desde el nivel táctico hasta el nivel estratégico nacional.
5. Durante la enseñanza del análisis de redes sociales, se debe preparar e instruir a los analistas de redes en diversos campos; sin embargo, el objetivo debe ser crear expertos en determinado campo. De tal manera que, el analista pueda responder preguntas de investigación enfocadas. Es un error que un analista sea “todista,” porque nos puede llevar a errores. La formación de redes entre entidades policiales y de seguridad es una estrategia para dividir y conquistar.

Chapter 13

Un Proceso para la Investigación de Redes Sociales

Después de leer y trabajar con el material de las secciones anteriores, la pregunta por contestar es ¿para qué sirve todo esto? Podemos hacer un intento de responder; diciendo que todos estos datos sirven para que el analista pueda diseñar una estrategia de interrupción o empoderamiento de la red y presentarla para la toma de decisiones.

Para apoyar dicha meta, los autores abordarán en esta sección algunos de los errores más comunes que se pueden cometer. Este manual a proporcionado datos de red para facilitar el aprendizaje, presentándole un puñado de conjuntos de datos y orientándole sobre cómo utilizarlos el software. Sin embargo, el cómo desarrollar un proyecto antes del análisis de datos ha estado mayormente ausente de la discusión. A raíz que muchos analistas aprenden ARSo usando conjuntos de datos listos para usar y con metas de aprendizaje definidas, algunos no saben cómo comenzar a diseñar un proyecto de investigación de red. Es por ello que esta sección aborda el tema de diseñar investigaciones de redes, tratando de presentarle cuestiones que deben tenerse en cuenta al construir un estudio. El lector reconocerá que muchos de los temas discutidos aquí no son específicos del análisis de redes sociales; además, esta no es la única discusión sobre cómo navegar este proceso (ver Borgatti, Everett, & Johnson (2018) o Cunningham, Everton, & Murphy (2016)). Sin embargo, el siguiente es un marco desarrollado por NPS y ESGE para guiar a nuevos analistas de redes en el proceso de construir un buen diseño de investigación.

Esta sección del manual se divide en dos partes. La primera como desarrollar un proyecto investigativo de redes sociales el cual se enfoca en cómo definir

un problema, identificar los componentes de la red y enmarcar las preguntas analíticas. La segunda parte presenta ejemplos desarrollados por estudiantes de cursos previos proporcionando información concreta de cómo los analistas han embarcado en el mapeo de redes.

13.1 Desarrollar un Proyecto Investigativo de Redes Sociales

A continuación se presenta un orden sugerido de cómo abordar un problema. Sin embargo, de acuerdo con el criterio del analista podrá utilizar los insumos de la manera que contribuya a su trabajo.

13.1.1 Pregunata(s) de Investigación

Un analista debe poder contestar brevemente ¿qué red, grupo o individuo se abordará en la investigación? y ¿Por qué? Este punto de partida puede ser por elección propia o debido a órdenes de un superior. El contestar esta breve interrogante previene que el proceso cambie y pierda su trayectoria. Cunningham, Everton, & Murphy (2016) nos proveé el siguiente ejemplo sobre la importancia de tener un punto de partida claro:

... Un comandante militar puede solicitar que los analistas examinen la “red de apoyo” de una red oscura en particular basándose en una brecha de información. El término “red de apoyo” es importante y no debe asumirse su significado. [...] ¿“Red de apoyo,” en este caso, indica activo, pasivo o alguna combinación? ¿Cómo deberían definir el apoyo pasivo frente al apoyo activo? ¿En qué momento el apoyo pasivo se convierte en apoyo activo? – (Cunningham, Everton, & Murphy, 2016, p. 37)

Como señalan los autores, definir la red es solo la mitad de la batalla. Definir claramente los conceptos y términos relevantes es crucial para comenzar a recopilar datos antes del análisis, particularmente cuando se trabaja con equipos o en un entorno con interrupciones constantes.

En este documento nos hemos centrado en explorar redes a través del prisma de estadísticas de red exploratorias o descriptivas. Las siguientes interrogantes son ejemplos de preguntas exploratorias que los analistas pueden hacer en las primeras etapas del análisis (adaptadas de Everton (2012a)):

- ¿Qué podemos aprender de la estructura de la red y cómo podemos utilizar esta información en nuestras operaciones?
 - Si la estructura es descentralizada ¿Por qué? ¿Qué ventajas tienen las redes descentralizadas? ¿Qué estrategias debemos implementar contra una red descentralizada?
 - Si la estructura es centralizada ¿Por qué? ¿Qué ventajas tienen las

13.1. DESARROLLAR UN PROYECTO INVESTIGATIVO DE REDES SOCIALES221

redes centralizadas? ¿Qué significa esto con respecto a la implementación de operaciones?

- ¿Es densa? ¿Qué ventajas tiene una red densa? ¿Qué desventajas?
- ¿Es escasa? ¿Qué ventajas tiene una red escasa? ¿Qué desventajas?
- ¿A qué tipología corresponde (por ejemplo, mundo pequeño, sin escala, aleatoria)? Con base en esta tipología, ¿qué tipo de estrategia podemos tomar para fragmentar la red?

- ¿Hay algunos vínculos que parecen ser más importantes que otros para la estructura general de la red?
- ¿Hay subredes? ¿Hay uniones entre estas que puedan explotarse?
- ¿Hay actores o vínculos claves cuyas eliminación o aislamiento hará que la red sea menos efectiva?
- ¿Hay nodos que aparentan ser más relevantes?
 - ¿Qué tipo de relevancia (intermediación, actividad, etc.)?

13.1.2 Limites

La cita anterior enfatizó la importancia de aclarar los términos centrales como un paso integral en el diseño de una investigación de redes sociales. Igual de importante es comprender los bordes del análisis. Dado que este análisis se centra en los actores y las relaciones, una consideración clave es definir qué datos se incluirán. Una vez más, Cunningham, Everton, & Murphy (2016) proporciona una excelente guía para especificar límites, por ello la lista de preguntas que sigue fueron adaptadas de su libro:

- ¿Cómo se define la población a sí misma? En términos de atributos, tipos de relaciones, participación en eventos o una combinación de estos. Por ejemplo, miembros de un cartel pueden identificarse como parte de dicha red con base en atributos característicos (tatuajes, roles, etc.), un objetivo compartido y participación en eventos.
- ¿Cómo defino la relevancia de un actor para el análisis?
 - Si el investigador se enfoca en atributos, puede preguntarse ¿son los actores relevantes porque poseen algún tipo de característica o conjunto de características, como un rol, afiliación organizacional y función logística? Por ejemplo, al mapear una red de comunicación entre líderes de una organización, el analista puede definir atributos como títulos (General, Comandante) como límite.
 - De otra manera, si los límites se definen con base en las relaciones, se puede preguntar ¿son relevantes los actores porque están incrustados en un tipo particular de relación? Por ejemplo, los actores dentro de una red financiera son parte de esta por sus vínculos a otros.
 - Otro enfoque puede basarse en eventos o participación de actores en una actividad. Por ello se puede preguntar ¿son los actores pertinentes porque participaron en algún evento, como un ataque terrorista o un crimen? Por ejemplo, los actores que atienden a los mismos eventos sociales pueden ser conectados.

- Además, puede ser una combinación de cualquiera de las categorías anteriores. Por ejemplo, muchas redes incluyen una combinación de relaciones actor a actor (financieras, comunicación), participación en eventos (co-asistencia) y atributos de nodo.

13.1.3 Codificar Relaciones y Atributos

Ya sea que un analista esté trabajando con flujos de datos novedosos o avanzando un proyecto previo, un registro completo de las definiciones de las variables en el conjunto de datos es indispensable para garantizar la confiabilidad y veracidad de dicho proyecto. Es por esto que los analistas a menudo recurren a los libros de códigos como guías de referencia que contienen la ontología de la investigación.

Este manual incluye dos libros de códigos como ejemplo, localizados en el segundo apéndice titulado Libro de Códigos. Las siguientes subsecciones proporcionan ejemplos y preguntas de orientación para como generar códigos.

13.1.3.1 Relaciones

Definidas como tipos de enlaces que conectan a diferentes actores de la red. Al codificar estas variables es importante no sólo definirlas, sino que hay que indicar si son modo-uno o bimodal, dirigidas o no dirigidas, etc. El (límite)[Limites del Análisis] de la red proporciona una guía sobre qué relaciones son cruciales para mapear una red. Sin embargo, el libro de códigos puede ser útil para definir las relaciones, proporcionar ejemplos y denotar explícitamente los tipos de clases de nodos vinculados a través de cada relación. Es relevante considerar las preguntas:

- ¿Qué tipo de relaciones son esenciales para el análisis? Por ejemplo, relaciones actor a actor, participación en eventos y el compartir atributos.
- ¿Cada relación, es de modo uno o de modo dos?

13.1.3.2 Atributos

Características no-relacionales relevantes para cada categoría de nodos que abordara la investigación. Por lo general, los analistas definen antes de recopilar o cuando se lleva a cabo la recopilación de datos, las variables relacionadas con los nodos del proyecto. Como tal, es relevante considerar las preguntas:

- Los atributos en esta red, ¿se deben recolectar al nivel del individuo? ¿Organización? ¿País? ¿Múltiples?
- ¿Qué características no relacionales son clave para describir la red a nivel tipológico, sus subgrupos y los actores de manera individual?

13.1.4 Fuentes de Datos y Métodos de Recopilación

Supongamos que un investigador de redes sociales está estudiando los lazos “financieros” dentro de una determinada organización ilícita. ¿Cómo se podrían recopilar los datos necesarios para este esfuerzo? Es decir, es poco probable

13.1. DESARROLLAR UN PROYECTO INVESTIGATIVO DE REDES SOCIALES223

que los miembros de la red ofrezcan voluntariamente esta información si uno pudiera establecer comunicación con estos. Más aún, si el investigador pudiera entrevistar a los sujetos del estudio con el fin de recopilar vínculos “financieros,” ¿cabría esperar que todos los sujetos interpretaran “financieros” de la misma manera? ¿Cuál sería la mejor manera de interactuar con los sujetos de la investigación (por ejemplo, digitalmente, en persona, a través de la observación)? Como señala Borgatti, “las características de estas preguntas dependerán en última instancia de las características de la población, del tipo de relaciones sociales que se estudien y, sobre todo, de los objetivos de la investigación” (Borgatti, Everett, & Johnson, 2018). Las secciones anteriores han cubierto la importancia de determinar los límites de la investigación, codificar explícitamente las variables e identificar los objetivos del esfuerzo. Aquí, el enfoque cambia nuevamente, esta vez para resaltar fuentes de datos y su recopilación.

13.1.4.1 Cuestionarios

Para el analista de redes, como para otros investigadores trabajando en proyectos de ciencias sociales, el cuestionario suele contener preguntas sobre los lazos del actor, aquí un informante, a otros. Por ejemplo, un informante puede proporcionar datos sobre con quién se comunica o de quién busca información. Wasserman & Faust (1994) proporcionan un esquema de los tipos de preguntas que se pueden utilizar en un cuestionario, el siguiente es un resumen de los puntos hechos por los autores:

- Un listado de personas frente a evocar libremente: Cuando el investigador conoce los miembros de la red antes de desarrollar un cuestionario para capturar las relaciones entre estos mismos, se puede presentar un listado de los demás actores para que los informantes seleccionen. Ejemplos de este enfoque se pueden encontrar en el trabajo de David Krackhardt (p. Ej. Krackhardt1992 o Krackhardt1994). Alternativamente, el investigador puede no presentar una lista y solicitarle al informante que evoque libremente a otros actores con los que comparten un tipo particular de vínculo.
- Elección libre frente a elección fija: El solicitar del informante un número específico de relaciones con otros actores en un cuestionario se conoce como elección fija. Este parámetro establece un número máximo de vínculos que un actor determinado puede tener en la red. Alternativamente, el no especificar un número de vínculos del informante se conoce como elección libre. Ambos enfoques tienen ventajas y desventajas, los investigadores interesados en esta forma de diseño de cuestionarios pueden consultar la literatura.
- Puntuaciones frente a ordenaciones: Una forma de medir la intensidad de los lazos es el requerir que los actores puntee u ordenen a los demás actores de la red. Las ordenaciones requieren que los informantes ordenen sus lazos con otros. El asignar puntuaciones implica que cada actor de un valor a cada relación.

Cabe recalcar que el utilizar cuestionarios es poco común para recopilar datos de las redes oscuras ya que, por definición, los miembros de estas buscan permanecer ocultos. Sin embargo, cuando posible, deben ser considerados ya que brindan datos de primera mano sobre las relaciones dentro de una red.

13.1.4.2 Entrevistas

Las entrevistas se utilizan para recopilar datos relacionales y pueden tener lugar cara a cara o por teléfono. Al igual que los cuestionarios, es posible que las entrevistas formales no sean disponibles para los investigadores de redes oscuras debido a la naturaleza de estas organizaciones. Sin embargo, los investigadores a menudo pueden recopilar datos relacionales y de atributos a través de entrevistas informales o conversaciones con personas en el área de responsabilidad (Cunningham, Everton, & Murphy, 2016). Cabe recalcar que al utilizar entrevistas informales es necesario validar toda información recopilada por este medio.

13.1.4.3 Observación Directa

La observación de los actores es otra manera de recopilar datos valiosos de la red. Los métodos de observación son comúnmente una herramienta en estudios que incorporan redes de afiliación donde el investigador puede requistar la asistencia a un evento o acontecimiento social que se estudia.

13.1.4.4 Recopilación de datos de archivo

Los datos de la red a menudo se recopilan de fuentes de archivo, como registros informáticos o de comunicación, listas de asistencia, informes de situación o de inteligencia, entre otros. Algunas fuentes de archivo son intrínsecamente relacionales, lo que facilita la recopilación e interpretación. Por ejemplo, los registros de asociaciones comerciales se pueden aprovechar para vincular entidades basadas en vínculos “comerciales,” usualmente de modo-uno. Otras fuentes de datos pueden requerir interpretación adicional, como registros de co-asistencia a eventos. En este conjunto de datos modo-dos, los participantes del evento pueden vincularse a un evento y, cuando sea apropiado, entre sí.

Más allá de la abundancia de posibles fuentes de datos, los datos de archivo suelen ser por naturaleza longitudinales, lo que permite el estudio de la dinámica de la red y el cambio social (Borgatti, Everett, & Johnson, 2018). Una vez más, si el enfoque de la investigación son los vínculos financieros, un investigador con acceso a registros financieros puede trazar la progresión de los cambios a lo largo del tiempo.

Al aprovechar los datos de archivo, los investigadores han encontrado formas muy creativas de descubrir información relacional. Sin embargo, esta fuente no está exenta de limitaciones y posibles fuentes de sesgos. Algunos registros de archivo representan una reconstrucción de los eventos pasados, que pueden contener prejuicios o reflejar las perspectivas de los actores que recopilaron

13.1. DESARROLLAR UN PROYECTO INVESTIGATIVO DE REDES SOCIALES225

dichos datos (Borgatti, Everett, & Johnson, 2018). Además, los datos pueden estar incompletos debido a varias razones; por ejemplo, los lazos comerciales son susceptibles a empresas fachada, los registros de comunicación no toman en cuenta la comunicación cara a cara, entre otros. Como tal, los analistas deben ser conscientes de las posibles deficiencias de datos de archivo y triangular a través de múltiples fuentes verificadas y validadas.

13.1.4.4.1 Registros Propios La utilización de bases de datos propios es un método cada vez más común para recopilar datos de red. Las organizaciones con grandes conjuntos de datos relacionales que involucran datos unimodales o multimodales pueden aprovechar estos datos para estudiar mejor una población. Por ejemplo, los analistas de la policía pueden vincular a las personas con arrestos y generar redes de co-delincuentes. Este enfoque es escalable, ya que muchas organizaciones recopilan grandes volúmenes de registros. Sin embargo, como todos los métodos para recopilar datos presentados en esta sección, este mismo tiene ventajas y desventajas. Por ejemplo, como se señaló anteriormente, los departamentos de policía podrían construir redes sociales utilizando datos de arrestos conjuntos; después de todo, muchos recopilan estos datos para medir el rendimiento de la organización. De tal manera que, al construir un análisis de redes criminales basándose solo en un subconjunto de la población, aquellos que han sido arrestado, los analistas corren el riesgo de construir un círculo vicioso en el que la atención sigue centrada en un subconjunto de delincuentes. Finalmente, al igual que con otras fuentes de datos, la investigación debe impulsar la selección de fuentes de datos, no la disponibilidad de estos.

13.1.5 Conclusiones

Esta sección fue diseñada para guiar a los analistas en el proceso de diseño de una investigación de análisis de redes sociales. Cubrimos cómo construir preguntas de investigación como un primer paso para guiar a los investigadores en el diseño, recopilación y análisis. De manera similar, esta sección abordó el valor de los libros de códigos para construir consistencia en la recopilación y estructuración de datos. Finalmente, esta sección abordó algunas de las fuentes de datos que se utilizan a menudo para recopilar y estructurar los datos de las redes sociales. Sin embargo, esta sección no abordó a fondo temas de confiabilidad, validez y otras consideraciones importantes que los investigadores de redes sociales deben ponderar antes de embarcarse en el análisis. No obstante, el marco proporcionado en esta sección debe verse como un conjunto inicial de principios para educar a los nuevos analistas de redes sociales sobre cómo comenzar a utilizar algunas de las herramientas, técnicas y métodos presentados en este manual.

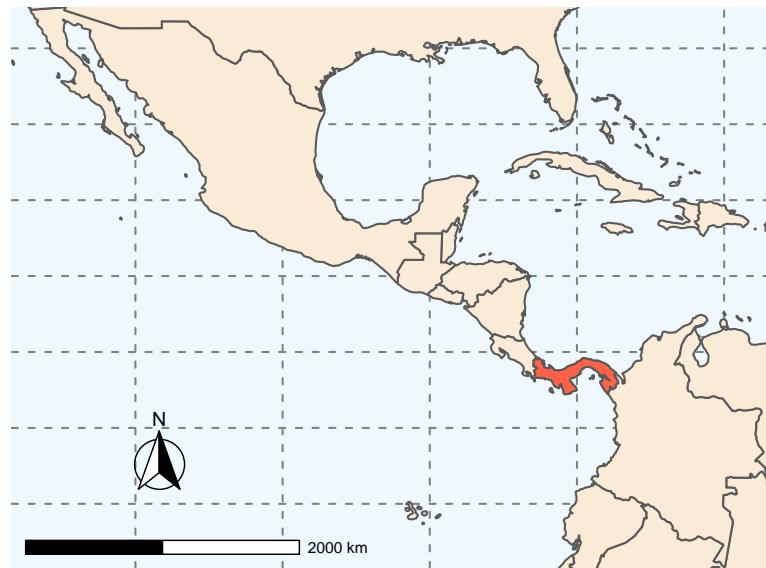
Chapter 14

Ejemplos Practicos

Lo que sigue es una serie de descripciones generales de proyectos de análisis de redes generadas por los estudiantes que han participado en clases realizadas conjuntamente por NPS y ESGE.

Informe de la Red Criminal en la República de Panamá

La red identificada cuyo nombre es “Calor Calor,” la cual mantiene con organizaciones criminales en Colombia y México. Esta organización opera principalmente en Panamá, Centroamérica y Suramérica.



La banda criminal “Calor Calor” se dedica a actividades tales como:

- Robo: Los mismos estuvieron en el robo al Banco Internacional de China por 2.5 millones de dólares en la Provincia de Colón en el año 2005 (“Grupos de panamá,” 2020).
- Homicidios: Tribunales judiciales han confirmado que varios de sus miembros estuvieron involucrados en un homicidio con arma de fuego en el año 2014 en la Provincia de Colón (“Marshall tendrá que pasar 25 años en la celda fría,” 2020).
- Tráfico de Armas: En el año 2015, autoridades de policía han dado con la captura de armas de grueso calibre asociadas con miembros de esta banda criminal (Cañizales, 2015).
- Tráfico de Droga: Se han dedicado al tráfico internacional de drogas desde Colombia para transportarla hasta las costas y puertos de Panamá para luego llevarlas a los carteles de Sinaloa, México (“Mayoría de detenidos en punta coco, ligada a clan úsuga y cartel de sinaloa,” 2019).
- Pandillerismo: Sus actividades de pandillerismo se han extendido a provincias centrales (Elena, 2017).

El (los) objetivo(s) del grupo de investigadores fue definir la estructura de la red y comenzar a cuantificar los lazos ilícitos que unen a miembros de esta organización criminal, Calor Calor, dentro del área de operaciones de la República de Panamá. Este estudio aprovechó los datos de código abierto para los años 2000 a 2020.

14.0.1 Libro de Códigos

14.0.1.1 Relaciones

Todas las relaciones (modo-uno o dos) han sido ponderadas en este análisis, definido como la fuerza o importancia que mantiene un nodo determinado respecto a otros. Los pesos son definidos utilizando la siguiente escala: 3 como alto valor, 2 como valor medio y 1 como valor mínimo.

14.0.1.1.1 Persona-a-persona (modo-uno)

- Tipo de Actividad Criminal: se interpreta como actores conectados por participación conjunta en una actividad que constituye un delito, el cual puede pertenecer a los siguientes tipos o clasificaciones de acuerdo con su naturaleza:
 - Homicidio
 - Drogas
 - Asociación Criminal
 - Robo
 - Blanqueo de Capitales
 - Pandillerismo
 - Tráfico Ilegal de Armas
 - Evasión de Centros Penitenciarios
 - Colaboración desconocida

14.0.1.1.2 Persona-a-Ubicación (modo-dos)

- Provincia: se puede interpretar como la división gubernamental o administrativa en que se divide un espacio territorial de un estado. Esta relación une personas a una ubicación el área de estudio. Para este trabajo se tomaron en cuenta las siguientes provincias de la República de Panamá.
 - Panamá
 - Colón
 - Coclé
 - Panamá Oeste

14.0.1.1.3 Persona-a-Evento (modo-dos)

- Tipo de Actividad Criminal: se interpreta como actores conectados a eventos criminales que constituyen un delito, el cual puede pertenecer a los siguientes tipos o clasificaciones de acuerdo con su naturaleza:
 - Homicidio
 - Droga
 - Asociación Criminal
 - Robo
 - Blanqueo de Capitales
 - Pandillerismo
 - Tráfico Ilegal de Armas
 - Evasión de Centros Penitenciarios
 - Colaboración desconocida

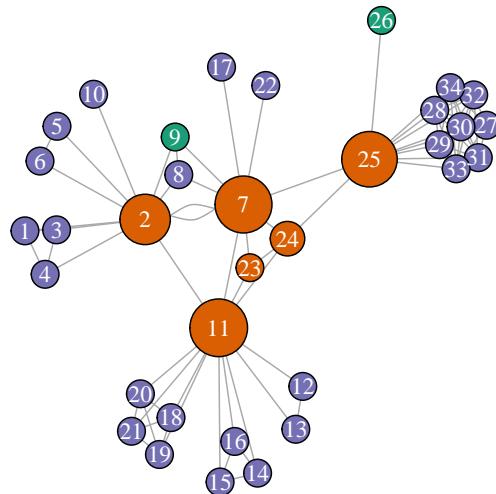
14.0.1.2 Atributos

- Tipo de Actor
 - Personas: Se considera a un ser que pertenece a la especie del tipo humano.
 - Evento: Se refiere a cualquier acto que vaya en contra las normas jurídicas en donde se comete un delito.
- Rol de Persona:
 - Líder: son las personas que dentro de una organización criminal ejercen mando o dan órdenes a sus subordinados.
 - Operador: Se puede interpretar como una persona que obedece órdenes dentro de un grupo delincuencial; es decir, no tiene mando.
 - Colaborador: en este trabajo, se define como una persona que mantiene cierto grado de colaboración, asistencia, apoyo con una organización de forma no precisada.
- Sexo de la Persona:
 - Hombre: Una persona con características masculinas. También se puede entender varón o macho.
 - Mujer: Se refiere a una persona del sexo femenino.

14.0.2 Mapa de red

El siguiente sociograma de red es el resultado de proyectar datos de dos modos y unificar todas las relaciones. Por tanto, el resultado final es una red de actores de un solo modo. Esta banda criminal mantiene 34 personas quienes han estado involucrado en 21 eventos criminales. Esta red mantiene una densidad de 0.148 y un diametro de 8. Lo cual sugiere una red dispersa.

Los nodos se han sido coloreado de acuerdo con la función de cada persona dentro de la red. Los nodos morados son “operadores,” los “líderes” son verdes y los “colaboradores” son rojos. Los datos registrados sugieren que la red Calor Calor está conectada por un conjunto de nodos con alta centralidad de intermediación. Estos nodos parecen formar el tejido conectivo central que une múltiples subconjuntos en la periferia.



Appendix A

Glosario de Términos

Actores (*Actors*): Se refiere a todo tipo de entidades incluidas en la red, tales como: individuos, grupos, etc. Sinónimos incluyen *nodos* or *vértices*.

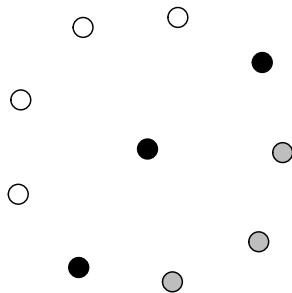


Figure A.1: Actores

Actores Aislados (*Isolates*): Actores que no tienen conexión con otros actores. Ejemplo, actor en rojo:

Centralidad (*Centrality*): Las medidas de centralidad dan una indicación aproximada del poder social de un actor en función de su posición dentro de la red. Un actor central puede ser alguien que tiene numerosos vínculos con otros

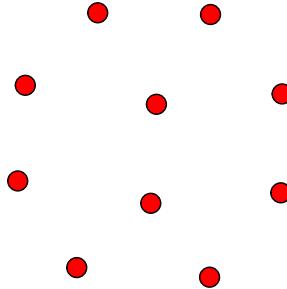


Figure A.2: Aislados

actores (*grado de centralidad*), que está más cerca (en términos de distancia de ruta) a otros actores de la red (*centralidad de cercanía*), que se encuentra en el camino más corto entre dos actores en la red (*centralidad de intermediación*), ó que tiene vínculos con actores altamente centrales (*centralidad de eigenvector*).

Centralidad de cercanía (*Closeness centrality*): La centralidad de la cercanía captura qué tan cerca (en términos de la más corta distancia de ruta) cada actor es a todos los demás actores de una red. Ejemplo, red con nodos dimensionados por centralidad de cercanía:

Centralidad de eigenvector (*Eigenvector centrality*): La centralidad de eigenvector supone que los vínculos con los actores centrales son más importantes que los vínculos con los actores periféricos y, por lo tanto, pondera las conexiones sumadas de cada actor con los demás por sus puntajes de centralidad. Ejemplo, red con nodos dimensionados por centralidad de eigenvector:

Centralidad de intermediación (*Betweenness centrality*): La centralidad de intermediación mide la medida en que cada actor se encuentra en el camino más corto entre todos los demás actores de una red. Ejemplo, red con nodos dimensionados por centralidad de intermediación:

Centralidad de grado (*Degree centrality*): Formalmente, el grado de centralidad de un actor es igual al número de vínculos incidentes ó el número de lazos de un actor. Ejemplo, red con nodos dimensionados por centralidad de grado:

Centralización (*Centralization*): La centralización utiliza la variación en la centralidad de todos los actores dentro de la red para medir el nivel de cen-

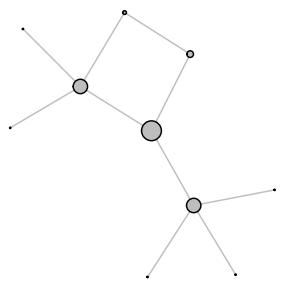


Figure A.3: Nodos dimensionados por cercanía

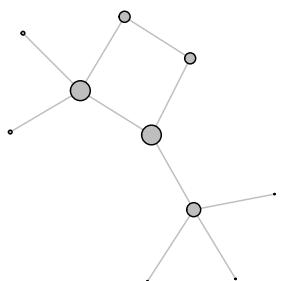


Figure A.4: Nodos dimensionados por eigenvector

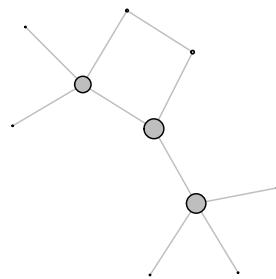


Figure A.5: Nodos dimensionados por intermediación

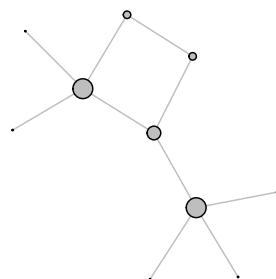


Figure A.6: Nodos dimensionados por grado

tralización. Más variación produce puntuaciones más altas de centralización de red; menos variación produce puntuaciones más bajos. En general, cuanto más grande es un índice de centralización, más probable es que un solo actor sea muy central.

Datos de atributos (*Attribure data*): Características (no-relacionales) de un actor en particular como identidad étnica, raza, años de educación, edad, nivel de salario, etc.

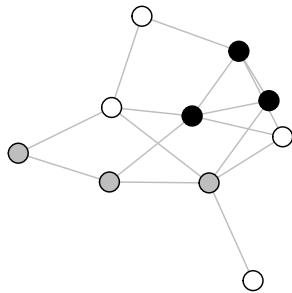


Figure A.7: Nodos coloreados por atributo

Datos de red modo uno (*One-mode network data*): Consiste en un solo conjunto de actores. Por ejemplo:

- Relaciones de persona a persona como amistades, familiares, etc.
 - Relaciones de organización a organización.
 - Relaciones de grupo a grupo.

Datos de red modo dos (*Two-mode network data*): Una red que consta de dos conjuntos de actores. Por ejemplo:

- Relaciones de persona a organización como afiliación.
 - Relaciones de persona a grupo.
 - Relaciones de grupo a nación.

Densidad (*Density*): Conceptualmente, la densidad se refiere al grado de interconexión de una red. Formalmente, es el número de vínculos en una red, expresado como una proporción del número máximo posible de vínculos. Está inversamente relacionado con el tamaño de la red (es decir, cuanto más grande es la red, menor es la densidad) porque el número de líneas posibles aumenta

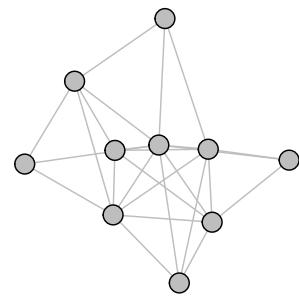


Figure A.8: Red modo uno

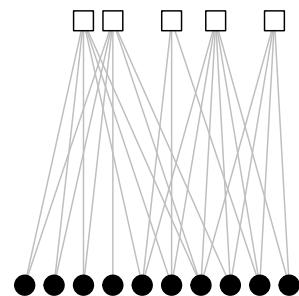


Figure A.9: Red modo dos

exponencialmente a medida que se agregan actores a la red, mientras que el número de lazos que cada actor puede mantener tiende a ser limitado.

Diámetro (Diameter): El diámetro de una red se refiere a el camino geodésico más largo de una red.

Distancia de ruta (longitud) (Path distance (length)): La distancia entre pares de actores en una red. El camino más corto entre un par de actores se conoce como la distancia geodésica.

Distancia promedio (Average distance): La distancia promedio se refiere a la longitud promedio de todos los caminos más cortos entre todos los actores conectados en una red y puede ser interpretada como un indicador de la velocidad en que la información (y otros recursos) se difunden a través de una red.

Estructura social (Social structure): La estructura social son los patrones duraderos de comportamiento y relaciones dentro de los sistemas sociales (por ejemplo, roles) o las instituciones y normas sociales que se han incorporado a los sistemas sociales de tal manera que dan forma al comportamiento de los actores. Dentro del ARSo, las estructuras sociales se ven en términos de patrones duraderos de lazos entre los actores (es decir, las redes sociales).

Grado promedio (Average degree): El grado promedio es igual al número promedio de vínculos entre todos los actores en una red. A veces se usa una medida alternativa a la densidad de la red porque, a diferencia de la densidad, no es sensible al tamaño de la red.

Tamaño de red (Network size): El tamaño de una red es igual al número de actores en la red.

Tipos de relación (Types of relationships): Valor categorico de vínculo entre un grupo de actores, como: amistad (en rojo), familia (en azul), etc.

Topología (Topography): La topología de la red se refiere a la estructura general de la red. Las medidas de uso común incluyen densidad, fragmentación, tamaño de red y centralización.

Vínculos (Ties): La relación entre dos nodos. Un vínculo puede ser direccionado (llamado *arco*, por ejemplo: $<- | -> | <->$) o no direccionado (llamado *borde*, por ejemplo: —).

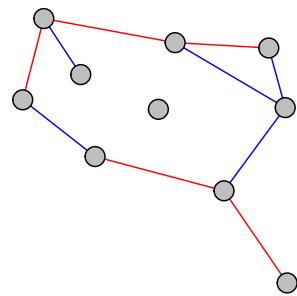


Figure A.10: Relaciones apiladas

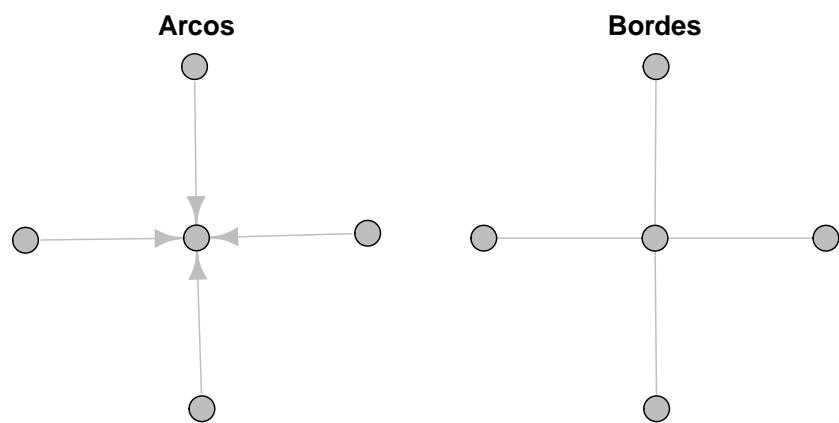


Figure A.11: Red no dirigida (izquierda) y red dirigida (derecha)

Appendix B

Libro de Códigos

B.1 Libro de Códigos de Muestra

El propósito de este documento es establecer un libro de códigos para estudiantes aprendiendo a estandarizar relaciones.

B.1.1 Relaciones

Persona-a-persona (modo-uno)

Las siguientes definiciones son ejemplos de relaciones de modo-uno entre personas. Estas solo deben tomarse como ejemplos, le recomendamos que adapte el contenido para adaptarlo a su investigación.

- **Amistad** (persona-a-persona): definida como dos personas que se declaran explícitamente como amigos, o que se conocen explícitamente como confiados de confianza, en informes o en documentación de inteligencia. Los casos de relaciones familiares autoidentificables en las que dos o más personas se consideran familiares, pero que no están relacionadas por sangre, matrimonio o adopción, también se consideran un ejemplo de este tipo de relación.
- **Asociado criminal** (persona-a-persona): definido como dos individuos que colaboran, ayudan o participan directamente en actividades ilícitas, incluidas acciones criminales fuera de línea y en línea, como compartir sustancias y / o transferir material ilegal.
- **Compañeros de clase** (persona-a-persona): dos personas que asisten, como estudiantes, a la misma institución académica, al mismo tiempo. Una institución académica se define como cualquier institución dedicada a proporcionar educación. Los ejemplos pueden incluir, entre otros, cursos de

capacitación, escuelas primarias, escuelas intermedias, escuelas secundarias y colegios y universidades.

- **Comunicación** (persona-a-persona): definida como la transmisión directa, en persona, de mensajes o información entre individuos o mediante algún tipo de medio, como un teléfono celular o un correo electrónico.
- **Financieras** (persona-a-persona): definido como dos actores, en informes o inteligencia, que se declaran explícitamente como transfiriendo fondos entre sí para cualquier propósito, legal o ilegal. Esta relación, sin embargo, NO incluye la compensación normal entre empleador y empleado.
- **Reclutador** (persona-a-persona): se define como el contacto entre dos personas con el fin de reclutar nuevos miembros para actividades relacionadas con una organización.
- **Relaciones Superior-Subordinado** (persona-a-persona) - Definidas como relaciones entre superiores inmediatos y subordinados en una organización. Esta relación explica los lazos formales / jerárquicos dentro de las organizaciones ilegítimas.
- **Romance** (persona-a-persona): definido como una relación cariñosa y romántica entre dos o más personas. Los ejemplos de esta categoría incluyen las relaciones no matrimoniales, como una en la que dos personas tienen un hijo (o hijos) juntos, que están involucrados en relaciones sexuales entre ellos y aquellos que están saliendo o cortejándose.

Persona-a-Evento (modo-dos)

Las siguientes definiciones son ejemplos de relaciones de modo-dos entre personas y eventos. Estas solo deben tomarse como ejemplos, le recomendamos que adapte el contenido para adaptarlo a su investigación.

- **Evento criminal** (persona-a-evento): definido como la participación de una persona en un evento, reunión, suceso o incidente en el que el propósito, las metas o los resultados del evento relacionados con, o que resultaron en actividades criminales. Esta definición incluye personas que son testigos, perpetradores, co-arrestados y víctimas de incidentes criminales. Los ejemplos pueden incluir, entre otros, crímenes violentos, reuniones formales / informales y eventos de capacitación relacionados con actividades criminales o que resultaron en ellas.
- **Evento no-criminal** (persona-a-evento): definido como la participación de una persona un evento, reunión, evento o incidente en el que el propósito del evento no se relaciona con actividades delictivas. Los ejemplos pueden incluir, entre otros, reuniones sociales, reuniones formales / informales, eventos comunitarios y eventos de capacitación que no están relacionados con actividades delictivas.

Persona-a-Locación or Material (modo-dos)

Las siguientes definiciones son ejemplos de relaciones de modo-dos entre personas y locaciones or materiales. Estas solo deben tomarse como ejemplos, le recomendamos que adapte el contenido para adaptarlo a su investigación.

- **Evidencia** (persona-a-material): definida como la relación física o la asociación directa de un individuo con un elemento de evidencia, como la posesión o el control individual de narcóticos, armas, documentos, etc.
- **Residencia** (persona-a-locación): definida como la propiedad, el alquiler una persona a una casa, un apartamento, un condominio, un piso, una casa de seguridad o cualquier otro tipo de residencia. Esta opción también puede incluir la propiedad, el alquiler y las visitas a residencias en el pasado.

Persona-a-Organización (modo-dos)

Las siguientes definiciones son ejemplos de relaciones de modo-dos entre personas y organizaciones. Estas solo deben tomarse como ejemplos, le recomendamos que adapte el contenido para adaptarlo a su investigación.

- **Afiliación con organizaciones ilegítimas** (persona-a-organización): definida como la propiedad, el empleo y / o la membresía de una persona en una organización ilegítima. Una organización ilegítima se define como un grupo u organización cuyos objetivos son de naturaleza criminal (terrorista), o que voluntariamente obtiene recursos materiales y / o no materiales a través de actividades delictivas.
- **Afiliación con organizaciones legítimas** (persona-a-organización): definida como la propiedad, el empleo y / o la membresía de una persona en una organización legítima. Una organización legítima se define como aquella que se abstiene de participar, apoyar o estar involucrada en actividades delictivas. Los ejemplos incluyen, entre otros, las siguientes organizaciones: organizaciones comunitarias, con fines de lucro, organizaciones sin fines de lucro, organizaciones políticas, organizaciones religiosas, e instituciones académicas.

B.1.2 Atributos

Esta sección describe y define los atributos (es decir, las características) de los actores.

- **Afiliación con organizaciones ilegítimas:** definida como la propiedad, el empleo y / o la membresía de una persona en una organización ilegítima. Una organización ilegítima se define como un grupo u organización cuyos objetivos son de naturaleza nefasta (es decir, criminal) o que voluntariamente obtiene recursos materiales y / o no materiales a través de actividades ilegales.

- **Afiliación con organizaciones legítimas:** definida como la propiedad, el empleo y / o la membresía de una persona en una organización legítima. Una organización legítima se define como aquella que se abstiene de participar, apoyar o estar involucrada en actividades delictivas.
- **Dirección:** la (s) ubicación (es) específica (s) en la que reside una persona.
- **Raza / origen étnico:** definido como el grupo racial o étnico con el que el individuo se identifica. Un grupo étnico es un grupo social al que pertenece una persona y que comparte una cultura, religión o idioma común y distintivo.
- **Ideología / creencia:** definida como el sistema de ideas y creencias de un individuo, incluidas las creencias políticas y económicas.
- **Ocupación:** definida como el título funcional actual de empleo de un individuo.
- **Rol:** definido como un rol que un individuo asume en una organización legítima o ilegítima. Los codificadores de datos pueden registrar varios roles dado que la mayoría de los actores tienen más de un rol.
- **Habilidades (tecnológicas):** definidas como habilidades especiales relacionadas con la tecnología, independientemente del dominio o campo que posea (por ejemplo, piratería informática).
- **Habilidades (no tecnológicas):** definidas como habilidades especiales NO relacionadas con la tecnología, independientemente del dominio o campo que posea (por ejemplo, habilidades de las personas).
- **Estado:** definido como la condición física del individuo. El valor predefinido es “Vivo y libre” si falta información o si el estado físico de un actor es “desconocido.” Por ejemplo:
 - Vivo y libre
 - Encarcelado / detenido
 - Muerto

B.2 Libro de Códigos de Noordin Top

El contenido de este libro de códigos es atribuido a la Profesora Nancy Roberts, el Profesor Sean Everton y Dan Cunningham en Laboratorio CORE (Roberts, Everton, & Cunningham, 2014). El libro de códigos corresponde a los datos de las redes de Nordin Top y sus colaboradores en el grupo terrorista Jemaah Islamiya. Estos datos se obtuvieron principalmente de dos informes de International Crisis Group (ICG): “Terrorism in Indonesia: Noordin’s Networks” (International Crisis Group, 2006) e “Indonesia: Noordin Top’s Support Base” (International Crisis Group, 2009). Los dos informes del ICG sobre Noordin contienen datos valiosos de modo-uno y dos sobre una variedad de relaciones y

afiliaciones (amistad, parentesco, reuniones, etc.) junto con datos de atributos importantes (educación, pertenencia al grupo, estado físico, etc.). Los datos fueron inicialmente estructurados y analizados por estudiantes de la Escuela de Postgrado Naval en el curso “Seguimiento y Disrupción de las Redes Oscuras” bajo la dirección del Profesor Sean Everton, Co-Director del Laboratorio CORE y la Profesora Nancy Roberts y luego expandidos y limpiados por la Facultad Asociada para Instrucción, Dan Cunningham.

B.2.1 Relaciones

- **Afiliación Organizacional** (modo-dos 237 x 32): Una organización terrorista/insurgente se define como un sistema administrativo y funcional, cuyo principal objetivo común es la conducta operativa de actividades terroristas/insurgentes, que consiste en miembros voluntariamente afiliados. A los efectos de este ejercicio, las facciones se considerarán separados de su organización principal para evitar que se codifiquen vínculos redundantes. En otras palabras, la mayoría de las afiliaciones de nivel micro están codificadas, mientras que un individuo solo se codifica en la organización principal si no figura como afiliado a una organización secundaria. Las organizaciones afiliadas terroristas/insurgentes, también están codificadas en esta matriz.
- **Afiliación Educativa** (modo-dos 237 x 25): Las relaciones educativas se definen como escuelas donde las personas reciben educación formal, sirven como empleados (maestro, administrador, etc.) y/o participan en instrucción educativa o religiosa adicional en la institución.
- **Compañeros de clase/colegas educativos** (modo-uno 237 x 237): Los compañeros de clase/colegas educativos se definen como personas que reciben educación formal, sirven como empleados (maestro, administrador, etc.) y/o participan en instrucción educativa o religiosa adicional en la misma institución y al mismo tiempo. Es más probable que esta relación refleje relaciones precisas que la matriz de “afiliación educativa” de dos modos, ya que considera el tiempo en el que las personas están presentes en una escuela.

Nota de autor: La codificación en la hoja maestra será diferente de la codificación basada únicamente en el informe del ICG. El primero utilizó información fuera del informe “Noordin’s Networks” del ICG para establecer si las personas podían considerarse compañeros de clase/colegas educativos según la definición.

- **Lazos de Comunicación** (modo-uno 237 x 237): La comunicación interna se define como la transmisión de mensajes entre individuos y/o grupos dentro de la red a través de algún tipo de medio.

- **Medio para comunicación externa** (modo-dos 237 x 5): El medio de comunicación externa se define como los diversos medios utilizados para transmitir información entre personas dentro y fuera de la red. Estos son:
 - Mensajes de computadora, como sitios web (www.anshar.net), cafés de red y correo electrónico.
 - Medios impresos, como la revista KOMPAK
 - Material de apoyo (códigos y contraseñas): para enviar material desde la prisión
 - Desconocido: cualquier tipo de comunicación no definido pero referenciado entre miembros
 - Vídeos: para la fabricación de bombas y la formación para reclutar miembros externos
- **Lazos de parentesco** (modo-uno 237 x 237): El parentesco se define como cualquier vínculo familiar, como hermano, cuñado, sobrino, etc. El parentesco también incluirá los matrimonios actuales y los matrimonios pasados debido a divorcios y/o muertes.
- **Eventos de Entrenamiento** (modo-dos 237 x 15): Participación en cualquier actividad específicamente designada para enseñar el conocimiento, las habilidades y las competencias del terrorismo y la insurgencia. El entrenamiento no incluye la participación en un acto patrocinado por terroristas o actividad muyahidín en lugares como Afganistán, Bosnia, Chechenia o Irak, a menos que la presencia de las personas fuera para participar en un campo de entrenamiento o base específicamente designado en una de estas áreas.

Nota de autor: Las personas que participaron en la “Capacitación religiosa de la embajada de Australia” y/o la “Capacitación para Bali II en el restaurante ‘Selera’” pueden reflejar vínculos redundantes con la “Red de operaciones,” ya que algunas personas participaron en estas capacitaciones específicas de la Embajada de Australia (04 de septiembre) y las operaciones de Bali II (05 de octubre).

- **Lazos de reclutamiento** (modo-uno 237 x 237): Contacto entre dos personas con el fin de reclutar nuevos miembros para actividades relacionadas con el terrorismo o la insurgencia. Las relaciones de reclutamiento solo son relevantes cuando el intento de reclutar miembros ha tenido éxito. No se incluyen los intentos de reclutamiento fallidos.

Nota de autor: Estos lazos crean una red dirigida y no deben ser simétricos en la matriz. Además, los reclutadores están en las filas mientras que los reclutados están en las columnas.

- **Afiliaciones comerciales** (modo-dos 237 x 11): Definidas como organizaciones con y sin fines de lucro que emplean personas.
- **Afiliaciones de negocios y finanzas** (modo-dos 237 x 12): Definidas como organizaciones con y sin fines de lucro que emplean a personas (incluye la Fundación Durassalam).
- **Tipo de financiación** (modo-dos 237 x 5): Las relaciones financieras se definen como la provisión de fondos (legales e ilegales) para respaldar, mantener y realizar operaciones para la red.
- **Operaciones** (modo-dos 237 x 14): Las relaciones operativas se definen como personas que participan en la preparación, ejecución y/o prestación de apoyo posoperatorio. La preparación debe relacionarse directamente con la operación y puede incluir objetivos reconocimiento, proporcionar refugio seguro para la preparación, contribuir al entrenamiento religioso y/o físico y participar en un robo donde las ganancias financieras un ataque posterior. La prestación de apoyo posterior a la operación, como la ocultación de fugitivos y la eliminación de explosivos, también debe estar directamente relacionada con la operación.
- **Amistad** (modo-uno 237 x 237): Las relaciones de amistad se definen como vínculos estrechos a través del afecto o la estima entre dos personas. Los lazos de amistad no se definen únicamente como reuniones y/o vínculos escolares.

Nota de autor: Las relaciones de amistad pueden ser extremadamente subjetivas si no se establecen explícitamente en el documento. Normalmente, las relaciones implícitas se basan en relaciones estrechas y coherentes a lo largo del tiempo. La relación entre Noordin Top y Azhari Husin, por ejemplo, no se declara explícitamente como una amistad en el documento, pero fueron colaboradores cercanos durante muchos años.

- **Afiliación religiosa** (modo-dos 237 x 8): Las relaciones religiosas se definen como una asociación con una mezquita, iglesia, sinagoga o círculo de estudio religioso. Los círculos de estudio religioso solo se codifican si están separados de otras entidades religiosas (es decir, mezquita). No incluimos las escuelas islámicas aunque asumimos que las escuelas tienen mezquitas. No utilizar las escuelas evita la duplicación de esfuerzos con el equipo que construye los lazos escolares. Además, enumeramos las mezquitas por la ciudad en la que se encuentra. Si había más de uno en una ciudad, agregamos un identificador numérico más el nombre de la ubicación más cercana.
- **Almas gemelas** (modo-uno 237 x 237): Las relaciones de alma gemela se definen como personas que están afiliadas a la misma institución religiosa al mismo tiempo. Es más probable que esta relación indique vínculos

religiosos precisos que la “afiliación religiosa,” ya que considera los períodos de tiempo en los que las personas están afiliadas a instituciones religiosas.

Nota de autor: La codificación en la hoja maestra diferirá considerablemente de la codificación basada únicamente en el informe del ICG. El primero utilizó información fuera del informe del ICG para determinar los plazos en los que las personas estaban afiliadas a una institución religiosa.

- **Lugar logístico** (modo-dos 237 x 35): Las relaciones logísticas se definen como un lugar clave dentro del archipiélago donde ocurrió la actividad logística. La actividad logística se define como proporcionar “casas seguras” para reunirse/esconderse, brindar apoyo material en términos de explosivos, proporcionar armamento o facilitar el transporte de personal o equipo.
- **Función logística** (modo-dos 237 x 4): Las funciones logísticas se definen como el apoyo a operaciones terroristas/insurgentes proporcionando materiales, armas, transporte y refugios.
- **Lazos de mentores: ideológicos, supervisores y tecnológicos** (modo-uno 237 x 237): Las subredes ideológicas, de supervisión y técnicas identifican a las personas en la red de Noordin que poseen la autoridad para dar el “Cómo,” el “Qué” y el “Por qué” al resto de la red. Sentimos que la autoridad religiosa, la habilidad técnica como la fabricación de bombas o la habilidad informática, y el liderazgo comprobado eran habilidades que no se podían transferir fácilmente y valdría la pena centrarse en interrumpir la red.

Nota de autor: Estas redes son dirigidas y no deben ser simétricas. Además, los mentores están en las filas mientras que los mentoreados están en las columnas.

- **Rol** (modo-dos 237 x 16): Definido como el papel que asume un individuo en la red terrorista/insurgencia. Codificado como datos de dos modos (ver también datos de atributos de función principal a continuación) porque algunas personas cumplen más de una función. La presencia de un rol por parte de un individuo se codificó como 1 y la ausencia como 0.
- **Reuniones** (modo-dos 237 x 20): Un evento coordinado y planificado de antemano entre dos o más personas. Las reuniones no incluyen todos los estilos de comunicación. Por el contrario, reunión se refiere a un lugar determinado en una fecha determinada con personas específicas. Las reuniones infieren la necesidad de una decisión, pero los datos no identifican específicamente la decisión o el tema de la reunión.
- **Entidades diversas** (modo-dos 237 x 10): Estas relaciones indican afil-

iación, pero no entran dentro de la definición de otras relaciones. Estas entidades van desde sitios web hasta campamentos militares y organizaciones políticas.

B.2.2 Atributos

- **Nivel de educación:** definido como el título más alto alcanzado, el nivel en el que se enseñó, estudió, participó o asistió. El siguiente es el esquema de códigos:

Código	Description
0	Desconocido
1	Educación Primaria
2	Pesantren
3	Escuela secundaria estatal
4	Algo de universidad
5	Universidad
6	Algo de maestría
7	Maestría
8	Doctorado

- **Contacto con personas fuera de Indonesia:** definido como contacto con personas en diferentes países fuera de Indonesia.

Código	Description
0	Desconocido
1	Afganistán
2	Australia
3	Malasia
4	Pakistán
5	Filipinas
6	Singapur
7	Tailandia
8	Reino Unido
9	Afganistán y Malasia
10	Afganistán y Filipinas
11	Afganistán y Pakistán
12	Afganistán, Malasia y Filipinas
13	Australia y Malasia
14	Filipinas y Malasia
15	Afganistán, Pakistán y Egipto
16	Irak, Afganistán y Pakistán

- **Entrenamiento militar:** definido como el país donde un individuo recibió

entrenamiento militar y alcanzó el estatus de veterano en la lucha en guerras insurgentes/convencionales conocidas.

Código	Description
0	Desconocido
1	Afganistán
2	Australia
3	Indonesia
4	Malasia
5	Filipinas
6	Singapur
7	Afganistán e Indonesia
8	Afganistán y Filipinas
9	Indonesia y Malasia
10	Indonesia y Filipinas
11	Afganistán e Irak

- **Nacionalidad de la persona:** definida como país de nacimiento, ciudadanía o residencia.

Código	Description
1	Afganistán
2	Australia
3	Indonesia
4	Malasia
5	Filipinas
6	Singapur
7	Arabia Saudita
8	Jordania
9	Egipto

- **Estado actual según el artículo del ICG:** definido como la condición física del individuo.

Código	Description
0	Muerto
1	Vivo
2	Encarcelado

- **Rol (Original):** definido como el rol que asume un individuo en la red terrorista/insurgente.

Código	Description
0	Sin información/poca claridad
1	Estratega: Planificador de alto nivel de una red terrorista/insurgente
2	Fabricante de bombas: Individuo que construye bombas
3	Bombardero/combatiente: Individuo que participa en ataques con bombas o que se describe como un combatiente
4	Entrenador/instructor: Individuo que entrena o instruye a nuevos miembros de una red terrorista
5	Terrorista suicida: Individuo que planea o ya ha realizado un ataque suicida
6	Reconocimiento y vigilancia: Participa en la vigilancia y reconocimiento de objetivos
7	Reclutador: Se dedica a identificar y reclutar nuevos miembros (para incluir bombarderos)
8	Mensajero /Intermediario: Facilita las comunicaciones entre miembros
9	Propagandista: Desarrolló campañas de información
10	Facilitador: Asistido en el funcionamiento de la red (especialmente con materiales y finanzas)
11	Líder religioso: Se le proporcionó formación y apoyo religiosos
12	Comandante/Líder táctico: A cargo de las operaciones a nivel local/táctico

- **Rol ampliado (primario y secundario):** Definido como el rol que asume un individuo en la red terrorista/insurgente (incluye roles primarios y secundarios).

Código	Description
1	Fabricante de bombas
2	Bombardero
3	Mensajero
4	Facilitador
5	Líder

Código	Description
6	Enlace
7	Líder local
8	Propagandista
9	Reclutador
10	Líder religioso
11	Maestro religioso, mentor, motivador
12	Proveedor de recursos
13	Estratega
14	Bombardero suicida
15	Simpatizante
16	Entrenador
99	No hay datos

- **Función logística:** definida como la provisión de casas seguras, armas, transporte y/o material a la red operativa.

Código	Description
1	Proporcionar una casa segura
2	Suministro de armas
3	Proporcionar transporte
4	Suministro de material
5	Suministro de armas, transporte y material
6	Suministro de armas, material
7	Suministro de transporte, material
8	Proporcionar refugio y transporte
9	Proporcionar refugio, transporte, material
10	Proporcionar refugio, armas, material

- **Habilidades técnicas:** elegimos codificar los atributos técnicos en función de los conjuntos de habilidades explícitos mencionados en el informe del ICG. En contraste con un atributo como “rol,” elegimos solo codificar las habilidades técnicas específicas que se enumeran a continuación y no consideramos cosas como “líder de celda” o “financiero” como un rol técnico. Básicamente, estábamos buscando habilidades relacionadas con la electrónica, fabricación de bombas, diseño de páginas web, etc. A continuación se muestra una lista de los roles y el número correspondiente que recibió cada individuo. En los casos en que alguien tenía múltiples habilidades técnicas, elegimos codificar la que es subjetivamente más relevante para las operaciones terroristas. Por ejemplo, Azhari Husin tiene experiencia en ingeniería mecánica pero también tiene entrenamiento en explosivos. Lo codificaríamos como “fabricación de bombas.” Aquellos sin habilidades técnicas identificadas reciben un 0 en la matriz.

Código	Description
1	Fabricación de bombas
2	Propaganda (producción de videos y CD de video, revistas, diseño de páginas web, etc.)
3	Instrucción de habilidades militares (artes marciales, tiro, etc.)
4	Otro conocimiento avanzado (ingeniería mecánica, agricultura, etc.)
5	Instrucción religiosa

- **Afiliación de grupo principal:** definida como la afiliación de grupo principal de cada miembro de la red, en general.

Código	Description
0	Ninguno (Noordin)
1	Darul Islam (DI)
2	KOMPAK
3	Jemaah Islamiyah (JI)
4	Ring Banten Group (DI)
5	Al-Qaeda

- **Red de Noordin:** un individuo que se considera miembro del grupo de Noordin (Tanzim Qaedat al-Jihad - Organización para la base del Jihad), en lugar de simplemente estar vinculado. Igualmente, si el individuo participó en una operación dirigida por Noordin durante cualquier etapa. Finalmente, si él o ella se declara explícitamente como miembro del círculo íntimo de Noordin y/o está vinculado a Noordin por parentesco o amistad.

Código	Description
0	No es miembro
1	Miembro

- **Original 79:** Definido como las 79 personas enumeradas en el apéndice del Informe del ICG de 2006, “Noordin’s Networks”

Código	Description
0	No incluido en el Apéndice
1	Incluido en el Apéndice

Part V

Referencias

Referencias

- Andrew Papachristos, D. H., & Braga, A. (2013). The corner and the crew: The influence of geography and social networks on gang violence. *American Sociological Review*, 78, 417–447.
- Arquilla, J., & Ronfeldt, D. (2001). *Networks and netwars: The future of terror, crime and militancy*. Santa Monica: RAND.
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgment. In H. S. Guetzkow (Ed.), *Groups, leadership and men*. Pittsburgh, PA: Carnegie Press.
- Azarian, G. R. (2005). *The general sociology of harrison c. White: Chaos and order in networks*. New York: Palgrave Macmillan.
- Bakker, R. M., Raab, J., & Milward, H. B. (2011). A preliminary theory of dark network resilience. *Journal of Policy Analysis and Management*, 31, 33–62.
- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. Retrieved from <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>
- Becker, G. S. (1976). *The economic approach to human behaviour*. Chicago: University of Chicago Press.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008. IOP Publishing. Retrieved from <https://doi.org/10.1088%2F1742-5468%2F2008%2F10%2Fp10008>
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92, 1170–1182.
- Boorman, S. A., & White, H. C. (1976). Social structure for multiple networks II: Role structures. *American Journal of Sociology*, 81(6), 1384–1446.
- Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2018). *Analyzing social networks*. Sage.

- Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25, 163–177.
- Burns, T., & Stalker, G. M. (1961). *The management of innovation*. London: Tavistock.
- Burt, R. (1992). *Structural holes: The social structure of competition*. Harvard University Press.
- Cañizales, A. (2015). 'Calitín' les dio duro a los narcos. *El Siglo de Panamá*. Retrieved from <http://elsiglo.com.pa/cronica-roja/calitin-duro-narcos/23854072>
- Chaudhary, A., Mittal, H., & Arora, A. (2019). Anomaly detection using graph neural networks. Retrieved from <https://ieeexplore.ieee.org/abstract/document/7877466/citations#citations>
- Chavez, M. (1961). *Ordaining women: Culture and conflict in religious organizations*. Cambridge, MA: Harvard University Press.
- Cunningham, D., Everton, S. F., & Murphy, P. (2016). *Understanding dark networks: A strategic framework for the use of social network analysis*. Rowman; Littlefield.
- Davies, A., Gardner, B. B., & Gardner, M. R. (1941). *Deep south: A social-anthropological study of caste and class*. University of Chicago Press.
- Davis, J. (1979). The davis/holland/leinhardt studies: An overview. In P. W. Holland & S. Leinhardt (Eds.), *Perspectives on social network research* (pp. 51–62). New York: Academic Press.
- Decary-Hetu, D. (2014). Information exchange paths in IRD hacking chat rooms. In C. Morselli (Ed.), *Crime and networks* (pp. 218–230). New York: Routledge.
- Devoto, R. (2016). Análisis de redes oscuras - OT sendero luminoso. Retrieved from <http://eie.edu.pe/>
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48, 147–160.
- Doreian, P., & Stockman, F. N. (1997). *EEvolution of social networks*. Amsterdam: Gordon; Breach Publishers.
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge.
- Elena, V. (2017). Cabecillas e integrante de la banda 'calor calor' condenados a prisión. *Panamá America*. Retrieved from <https://www.panamaamerica.com.pa/provincias/cabecillas-e-integrante-de-la-banda-calor-calor-condenados-prision-1084936>

- Emirbayer, M., & Goodwin, J. (1994). Network analysis, culture and the problem of agency. *American Journal of Sociology*, 99(6), 1411–1455.
- Everton, S. F. (2012b). Network topography, key players and terrorist networks. *Connections*, 2, 1–8.
- Everton, S. F. (2012a). *Disrupting dark networks*. Cambridge.
- Freeman, S. C., & Linton C, F. (1979). *The networkers network : A study of the impact of a new communications medium on sociometric structure*. School of Social Sciences Univ. of Calif.
- Garzon, J. C. (2013). La diáspora criminal: La difusión transnacional del crimen organizado y cómo contener su expansión. In J. C. Garzon & E. Olson (Eds.), *La diáspora criminal: La difusión transnacional del crimen organizado y cómo contener su expansión* (pp. 1–26). Washington DC: Woodrow Wilson International Center for Scholars.
- Girvan, M., & Newman, M. (2002). Community structure in social and biological networks. *PNAS*, 99, 7821–7826.
- Gould, R. V. (1991). Multiple networks and mobilization in the paris commune, 1871. *American Sociological Review*, 56, 716–729.
- Gould, R. V. (1993). Trade cohesion, class unity, and urban insurrection: Artisanal activism in the paris commune. *American Journal Of Sociology*, 98(4), 721–754.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 73(6), 1360–1380.
- Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Social Theory*, 1, 201–233.
- Grupos de panamá: Calor calor. (2020). Insight Crime. Retrieved from <https://es.insightcrime.org/noticias-crimen-organizado-panama/calor-calor/>
- Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. University of California.
- Hashimoto, T. B., Nagasaki, M., Kojima, K., & Miyano, S. (2009). BFL: A node and edge betweenness based fast layout algorithm for large scale networks. *BMC Bioinformatics*, 10, 1471–2105.
- Iannaccone, L. R. (1995). Risk, rationality, and religious portfolios. *Economic Inquiry*, 33(2), 285–295.
- International Crisis Group. (2006). *Terrorism in indonesia: Noordin's networks*. Brussels, Belgium: International Crisis Group.
- International Crisis Group. (2009). *Indonesia: Noordin top's support base*. Brussels, Belgium: International Crisis Group.

- Jacomy, M. (2011, June). ForceAtlas2, the new version of our home-brew layout. *Gephi*. Retrieved from <https://gephi.wordpress.com>
- Kleinberg, J. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46, 604–632.
- Knoke, D. (1990). *Political networks: The structural perspective*. Cambridge University Press.
- Knoke, D., & Yang, S. (2007). *Social network analysis*. Sage Publications, Inc.
- Knuth, D. E. (1993). *The stanford GraphBase: A platform for combinatorial computing*. Addison-Wesley.
- Koschade, S. (2006). A social network analysis of jemaah islamiyah: The applications to counterterrorism and intelligence. *Studies in Conflict and Terrorism*, 24(3), 559–575.
- Krackhardt, D. (1992). The strength of strong ties: The importance of philos in organizations. In N. Nohria & R. G. Eccles (Eds.), *Networks and organizations: Structure, form and action* (pp. 261–239). Boston: Harvard University Press.
- Krackhardt, D. (1994). Graph theoretical dimensions of informal organizations. In K. M. Carley & M. J. Prietula (Eds.), *Computational organization theory* (pp. 89–111). Hillsdale, NJ: L. Erlbaum Associates.
- Krebs, V. E. (2002). Mapping networks of terrorist cells. *Connections*, 24, 43–52.
- Levitt, S. D., & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. Harper Collins.
- Lin, N. (2001). *Social capital*. Cambridge University Press.
- Lofland, J., & Stark, R. (1965). Ecoming a world-saver: A theory of conversion to a deviant perspective. *American Sociological Review*, 30, 862–875.
- Marshall tendrá que pasar 25 años en la celda fría. (2020). *El Siglo de Panamá*. Retrieved from <https://www.pressreader.com/panama/el-siglo/20200801/281479278750886>
- Mayhew, B. (1980). Structuralism versus individualism: Part i, shadowboxing in the dark. *Social Forces*, 59(2), 355–374.
- Mayhew, B. (1981). Structuralism versus individualism: Part II, ideological and other obfuscations. *Social Forces*, 59(3), 627–648.
- Mayoría de detenidos en punta coco, ligada a clan úsuga y cartel de sinaloa. (2019). *La Prensa de Panamá*. Retrieved from https://www.prensa.com/impresa/panorama/Mayoria-Punta-Coco-Usuga-Sinaloa_0_5234476558.html
- McAdam, D. (1986). Recruitment to high risk activism: The case of freedom summer. *American Journal of Sociology*, 92, 64–90.

- McAdam, D. (1988b). Micromobilization contexts and recruitment to activism. *International Social Movement Research, 1*, 125–154.
- McAdam, D. (1988a). *Freedom summer*. Oxford University Press.
- McCormick, G., & Owen, G. (2000). Security and coordination in a clandestine organization. *Mathematical and Computer Modelling, 31*, 175–192.
- McCulloh, I., & Carley, K. M. (2011). Detecting change in longitudinal social networks. *Journal of Social Structure*.
- Meghanathan, N. (2016). Assortativity analysis of real-world network graphs based on centrality metrics. *Computer and Information Science, 9*, 7–25.
- Meyer, J. W., Boli, J., Thomas, G. M., & Ramirez, F. O. (1997). World society and the nation-state. *American Journal of Sociology, 103*(1), 144–181.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, 83*(2), 340–363.
- Milgram, S. (1974). *Obedience to authority*. Princeton, NJ: Princeton University Press.
- Monge, P. R., & Contractor, N. S. (2003). *Theories of communication networks*. Oxford University Press.
- Morselli, C. (2009). *Inside criminal networks*. Springer.
- Morselli, C., & Petit, K. (2007). Law-enforcement disruption of a drug importation network. *Global Crime, 8*(2), 109–130.
- Newman, M., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physics Review, 69*, 1–16.
- Nohria, N. (1992). Is the network perspective a useful way of studying organizations? In N. Nohria & R. G. Eccles (Eds.), *Networks and organizations: Structure, form, and action* (pp. 1–22). Boston, MA: Harvard Business School Press.
- Nooy, W. de. (2011). Networks of action and events over time: A multilevel discrete-time event history model for longitudinal network data. *Social Networks, 33*(1), 31–40.
- Nooy, W. de, Mrvar, A., & Batagelj, V. (2005). *Exploratory social network analysis with pajek*. Cambridge, U.K.: Cambridge University Press.
- Ognyanova, K. (2016). Network visualization with gephi. *Sunbelt 2016*.
- Pescolindo, B. A., & Georgianna, S. (1989). Durkheim, suicide, and religion: Toward a network theory of suicide. *American Sociological Review, 54*(1), 33–48.

- Podolny, J. M., & Page, K. L. (1998). Network forms of organization. *Annual Review of Sociology*, 57–76.
- Powell, W. W. (1990). Neither market nor hierarchy: Network forms of organization. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior: An annual series of analytical essays and critical reviews* (pp. 295–336). Greenwich, CT: JAI Press, Inc.
- Powell, W. W., & Smith-Doerr, L. (1994). Networks and economic life. In N. J. Smelser & R. Swedberg (Eds.), *The handbook of economic sociology* (pp. 268–402). Princeton, NJ: Princeton University Press.
- Raab, J., & Milward, H. B. (2003). Dark networks as problems. *Journal of Public Administration Research and Theory*, 13, 413–39.
- Roberts, N. (2014). Introduction to social network analysis. Retrieved from <http://nps.edu/web/core>
- Roberts, N., Everton, S. F., & Cunningham, D. (2014). The noordin top network.
- Ronfeldt, D., & Arquilla, J. (2001). What next for networks and netwars? In D. Ronfeldt & J. Arquilla (Eds.), *Networks and netwars* (pp. 311–361). Santa Monica, CA: RAND.
- Scott, J., & Carrington, P. J. (2011). *The SAGE handbook of social network analysis*. SAGE Publications Ltd.
- Simmel, G. (1950). *The sociology of georg simmel*. (K. H. Wolf, Ed.). The Free Press.
- Smith, C. S. (2010). *What is a person? Rethinking humanity, social life, and the moral good from the person up*. Chicago; London: The University of Chicago Press.
- Snijders, T. A. B., Bunt, G. G. van de, & Steglich, C. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), 44–60.
- Snow, D. A., Zurcher, L. A., & Ekland-olson, S. (1980). Social networks and social movements: A microstructural approach to differential recruitment. *American Sociological Review*, 45, 787–801.
- Stark, R., & Bainbridge, W. S. (1980). Networks of faith: Interpersonal bonds and recruitment to cults and sects. *American Journal of Sociology*, 85(6), 13767395.
- Turner, J. H. (2006). *Sociology*. Pearson Prentice Hall.
- Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, 61, 674–698.

- Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, 11(2), 447–504.
- Valente, T. W. (2010). *Social networks and health: Models, methods, and applications*. Oxford.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- White, H. C. (1992). *Identity and control: A structural theory of social action*. Princeton, NJ: Princeton University Press.
- White, H. C. (2008). *Identity and control: How social formations emerge*. Princeton, NJ: Princeton University Press.
- White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social structure from multiple networks i: Blockmodels of roles and positions. *American Journal of Sociology*, 81, 730–780.
- Zimbardo, P. G. (1972). Pathology of imprisonment. *Society*, 9, 4–6.