



## Proyecto Fin de Carrera

Análisis de redes de interacciones en el  
entorno del movimiento 15M

Ignacio Morer Zapata

Director: Miguel Aguilera Lizarraga

Ponente: Manuel González Bedia

Escuela de Ingeniería y Arquitectura  
Universidad de Zaragoza

Noviembre de 2012



# Análisis de redes de interacciones en el entorno del movimiento 15M

## RESUMEN

El problema de cómo se organizan las redes complejas constituye un tema central en diferentes ámbitos de investigación (biología, ciencias sociales, el estudio de la mente, etc.). Sólo muy recientemente se están desarrollando herramientas para entenderlo gracias a los avances en ciencia de sistemas complejos. Estas herramientas permiten analizar propiedades contraintuitivas y emergentes de sistemas que no se comportan de manera lineal y que son muy difíciles de interpretar por medio de otras metodologías. No basta con identificar y sumar el funcionamiento de las partes del sistema por separado, sino que la interacción entre ellas lleva asociada gran cantidad de información esencial para la comprensión del total.

Mediante la teoría de redes se pueden afrontar estos análisis. En este proyecto se aplicará a un contexto social de actualidad. En concreto, se estudiará el sistema complejo que se forma a partir de las redes de comunicación formadas en torno al movimiento ciudadano 15M. El papel de las nuevas tecnologías (Internet, dispositivos móviles, redes sociales...) en la difusión de información y la movilización política es decisivo a la hora de construir los nuevos movimientos ciudadanos. La población tiene la capacidad de emitir y recibir información de una manera casi inmediata, descentralizada y difusa. El objetivo de este proyecto es entender cómo emergen y qué propiedades tienen los procesos de comunicación característicos del amplio movimiento surgido a partir del 15 de mayo del 2011. Para ello, se emplearán datos extraídos de una de las principales redes sociales utilizadas por los participantes (Twitter).

El trabajo comienza con una fase de introducción a la teoría de redes y a las herramientas para su análisis, completado por una evaluación posterior sobre qué propiedades son de interés en este tipo de redes y qué clase de estudios se pueden realizar en ellas. Después se analizan datos de la red social relativos a las movilizaciones ciudadanas de mayo del 2012 (y otros fenómenos similares) a través de métodos estadísticos y algorítmicos, con el objetivo de caracterizar las redes resultantes. A partir de este análisis se presentan conclusiones acerca de cómo está estructurado el sistema y cómo se comporta de forma dinámica. Los resultados se relacionan con tipos de redes existentes y se trata de identificar diferencias entre procesos de distinta naturaleza.

En particular, la fase de análisis parte del estudio de la topología. Se relacionan los sistemas con dos tipos de redes muy comunes: la red de mundo pequeño y la red libre de escala. Posteriormente, se investiga la división de los sistemas en comunidades y la presencia de mesoescala. Por último, se realiza un análisis de propiedades dinámicas que permite caracterizar los procesos comunicativos en función de las interrelaciones entre sus diferentes escalas temporales.



## Agradecimientos

Este proyecto de fin de carrera ha sido desarrollado desde abril de 2012 hasta noviembre de 2012 en el Grupo de Informática Gráfica Avanzada, dentro del Departamento de Informática e Ingeniería de Sistemas de la Universidad de Zaragoza y el Instituto de Investigación en Ingeniería de Aragón, y en colaboración con la Fundación de los Comunes y el grupo de investigación 15Mdata.

Quiero agradecer el apoyo prestado por toda mi familia. Seguro que tenían tantas ganas como yo de que llegase el momento de entregar este trabajo.

A todos los amigos que han estado cerca desde el principio hasta el final, acompañándome en el proceso y apoyándome en todo lo necesario.

A los compañeros del laboratorio, por los buenos ratos que hemos pasado mientras trabajábamos y por los conocimientos que hemos compartido.

Por último, agradecer a Miguel y Emanuele toda la ayuda prestada, que ha sido imprescindible para llevar a cabo este documento. Además, me han abierto las puertas a todo un mundo de proyectos muy interesantes que espero que podamos compartir en un futuro.



# Índice general

<b>I Memoria</b>	<b>IX</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Movimientos sociales y redes de comunicación . . . . .	1
1.1.1. Tecnopolítica: movimientos sociales y TIC . . . . .	1
1.1.2. Estrategias de organización en movimientos sociales . . . . .	2
1.2. Twitter: la red social de <i>microblogging</i> . . . . .	3
1.3. Motivación y objetivo del proyecto . . . . .	4
1.4. Contenido de la memoria . . . . .	4
1.5. Planificación . . . . .	5
<b>2. Bases teóricas</b>	<b>7</b>
2.1. Sistemas complejos: enfoque tradicional-reduccionista vs. enfoque emergente . . . . .	7
2.2. Teoría de grafos y redes sociales . . . . .	7
2.2.1. Topología. Definiciones y tipos de redes complejas según su estructura . . . . .	8
2.2.2. Análisis de comunidades y mesoescala. . . . .	14
2.2.3. Dinámica . . . . .	15
<b>3. Herramientas</b>	<b>17</b>
3.1. Adquisición de datos . . . . .	17
3.1.1. Recopilando información de la red . . . . .	17
3.1.2. Twitter . . . . .	17
3.1.3. ScraperWiki . . . . .	17
3.2. Tratamiento de datos . . . . .	18
3.3. Análisis y visualización de redes . . . . .	18
3.3.1. <i>Software</i> . . . . .	18
3.4. Algoritmo de detección de comunidades y mesoescala . . . . .	19
3.5. <i>Detrended Fluctuation Analysis</i> (DFA) . . . . .	19
<b>4. Análisis</b>	<b>21</b>
4.1. Selección de datos . . . . .	21
4.2. Análisis de la topología . . . . .	23
4.2.1. Redes de mundo pequeño . . . . .	23
4.2.2. Distribución de grados . . . . .	24
4.3. Análisis de comunidades y mesoescala . . . . .	26
4.4. Análisis dinámico . . . . .	28
<b>5. Conclusiones</b>	<b>33</b>
5.1. Discusión . . . . .	33
5.2. Aplicaciones . . . . .	34
5.3. Trabajo futuro . . . . .	34

# Índice de figuras

1.1. Diagrama de Gantt del proyecto. . . . .	5
2.1. Ilustración conceptual sobre los distintos enfoques. . . . .	8
2.2. Direccionalidad . . . . .	8
2.3. Componente gigante de una red . . . . .	9
2.4. Subgrafos fuertemente conexos . . . . .	9
2.5. Distancia entre dos nodos. . . . .	10
2.6. Coeficiente de <i>clustering</i> . . . . .	11
2.7. Aleatoriedad creciente en función de $p$ . . . . .	12
2.8. $L$ y $C$ en función de $p$ . . . . .	12
2.9. Distribución según ley de potencia, o <i>long tail</i> . . . . .	13
2.10. Redes y resistencia, $r$ . . . . .	15
2.11. Análisis de mesoescala en una red artificial. . . . .	15
4.1. Ejemplo de uso de hashtags . . . . .	22
4.2. $L/L_{random}$ frente a $C/C_{random}$ . . . . .	24
4.3. Distribución del grado de entrada. Red: #es15m. . . . .	25
4.4. Distribución del grado de salida. Red: #es15m. . . . .	25
4.5. Valores absolutos de las pendientes $\gamma_{in}$ y $\gamma_{out}$ . . . . .	26
4.6. Mesoescalas en la actividad neuronal. Fuente: [36]. . . . .	27
4.7. Análisis de mesoescala: #desalojoSol . . . . .	29
4.8. Análisis de mesoescala: #huelgadeclase . . . . .	29
4.9. Análisis de mesoescala: #15mparato . . . . .	29
4.10. Alcance de la correlación frente a $\beta$ . . . . .	31



# Índice de tablas

2.1. Ejemplos de redes reales de mundo pequeño . . . . .	13
4.1. Listado de hashtags relevantes . . . . .	22
4.2. Datos del estudio dinámico. . . . .	31

# Parte I

## Memoria



# Capítulo 1

## Introducción

### 1.1. Movimientos sociales y redes de comunicación

#### 1.1.1. Tecnopolítica: movimientos sociales y TIC

Históricamente, los movimientos sociales han sido la forma mediante la cual la ciudadanía ha mostrado su descontento y ha pretendido defender sus derechos. Por ello, constituyen una herramienta clave para fomentar y consolidar la democracia. Son procesos dinámicos que siempre han estado limitados por las restricciones propias de los procesos de comunicación. De este modo, la participación está estrechamente ligada al acceso a la información y a las formas de interacción individual.

Dado que vivimos en una época en la que la tecnología de la información está experimentando un gran desarrollo es lógico que la forma en que se organizan, expanden y materializan los movimientos ciudadanos evolucione con ella. La transformación de las formas de comunicación ha permitido crear redes más horizontales, distribuidas e interactivas, que están revolucionando el activismo popular. En concreto, el derribo de esas barreras comunicativas ha aumentado notablemente la eficacia en su coordinación y en la difusión del mensaje. Un individuo puede intervenir activamente en estos procesos, gracias a herramientas que proporcionan acceso a información de forma rápida y selectiva y una elevada conectividad.

Denominamos tecnopolítica a la articulación entre el uso estratégico de las tecnologías de comunicación y la organización colectiva [1]. A lo largo de las últimas décadas se han dado múltiples casos de auto-organización traducidos posteriormente en acciones visibles y de gran repercusión, mediados por avances tecnológicos. Entre los ejemplos más destacados se encuentran:

- Protestas anti-globalización en Seattle, noviembre de 1999: Internet y los medios digitales facilitaron la coordinación y cobertura de las protestas. La conexión en red de ONGs locales, ciudadanos y activistas *grass-roots*<sup>1</sup> se tradujo en un entramado global que proporcionó nuevos canales de información, discusión y, en última instancia, de acción [2].
- Reacción a la supuesta manipulación mediática del gobierno de España (13 de marzo de 2004): tras los atentados del 11 de marzo de 2004, se organizó, vía mensaje de texto, una protesta frente a la sede del partido gobernante. Se corroboró que las concentraciones fueron promovidas por actores sociales no vinculados a ningún partido político. Su repercusión se manifestó en importantes incrementos en las comunicaciones por telefonía móvil: durante la tarde del día 13 se registraron aumentos en torno al 30 % en el tráfico de mensajes de texto, respecto a un sábado cualquiera [3]. Paralelamente, la convocatoria saltaba a Internet, donde los portales activistas más frecuentados aumentaban notablemente sus visitas, con amplia participación de colectivos extranjeros.

---

<sup>1</sup> Asociaciones de individuos creadas y sustentadas de forma natural y espontánea.

- Wikileaks y el caso 'Cablegate': esta organización sin ánimo de lucro publica informes y documentos con contenido de interés público, manteniendo el anonimato de las fuentes. El 28 de noviembre de 2010 fueron mostradas en su web unas 250.000 comunicaciones entre el Departamento de Estado de Estados Unidos y sus embajadas, además de ser enviadas a cinco periódicos internacionales. Sufrió una persecución por parte de gobiernos y entidades privadas a la que los usuarios de la red respondieron con acciones de apoyo y protesta.
- Primavera árabe: se produjeron una serie de alzamientos ciudadanos en países del norte de África y Oriente Próximo reclamando libertades y democracia. Internet, desde el comienzo del uso de blogs en Egipto en torno al 2004, fue facilitando los enlaces en red, multiplicados después por las redes sociales. A finales de 2010, la revolución en Túnez supuso el punto de partida de acciones multitudinarias que se contagieron a países vecinos.
- Anonymous: es un seudónimo adoptado por individuos y grupos para realizar acciones de protesta a favor de la libertad de expresión y la independencia de Internet, entre otros temas. Surge de foros de discusión y carece de estructura conocida. Su organización es difusa y se vale de los canales *online* para preparar sus operaciones. Entre ellas, destaca el ataque a sitios *web* de instituciones privadas que actuaron en contra de Wikileaks y su fundador, como PayPal y Mastercard.
- NoLesVotes: es un movimiento catalogado como "ciberactivismo" surgido a raíz de la *Ley Sinde*, aprobada en febrero de 2011. Su objetivo era desanimar el voto a aquellos partidos políticos que apoyaron dicha ley. Tuvo gran repercusión en las redes sociales y sirvió, en cierto modo, de precedente a las muestras de descontento que comenzaron a aparecer los meses sucesivos.
- Democracia Real Ya y 15M: un movimiento auto-organizado se fue formando por miles de personas anónimas en las redes sociales, bajo el nombre de Democracia Real Ya, con clara inspiración en las revueltas árabes y la *revolución islandesa*. Se posicionó en favor de una democracia más participativa, alejada del bipartidismo y del dominio de instituciones financieras y corporaciones, pero sin presentar afinidad política ni sindical. El crecimiento en simpatizantes del movimiento pasó, en gran medida, desapercibido mediáticamente, al tiempo que la campaña se difundía de forma viral. El principal mérito fue construir un estado de ánimo colectivo [1] que se concretó, el 15 de mayo de 2011, en una serie de protestas pacíficas en 70 ciudades y localidades, bajo el nombre de "movimiento 15M".

En este trabajo nos centramos en el movimiento 15M. Su carácter novedoso, su cercanía y la relevancia de las TIC en sus interacciones lo convierten en una buena oportunidad de estudio. Además establece un sentimiento de descontento general que fomenta la aparición de otras muchas movilizaciones de forma simultánea, como acciones en defensa de la educación y sanidad públicas o en contra de los desahucios.

### 1.1.2. Estrategias de organización en movimientos sociales

Como punto de partida, nos preguntamos qué tipos de procesos de auto-organización están detrás de las movilizaciones sociales mediadas por herramientas tecnológicas. Los propios activistas han intentado analizarlos y están desarrollando conceptos que les ayuden a categorizar fenómenos auto-organizados en red. En el caso del 15M, se reflexiona sobre el estado colectivo que surge de estos procesos recurriendo al concepto de "clima social" [4]: el ánimo general se enfriá después de períodos activos, pierde visibilidad y frecuencia de acción, pero posee las bases para articularse rápidamente y actuar sin empezar desde cero. Es más, en ocasiones esta relajación se ve como algo necesario para la continuidad y eficacia del movimiento [5]. De aquí podría surgir la coordinación de un primer tipo de movimiento. Aquellas situaciones en las que, de forma abrupta y espontánea, se reúne un elevado número de personas en las redes sociales tratando un tema en concreto. La

información se expande de forma viral y, en última instancia, el resultado es una acción específica con gran poder de expansión y agregación. Este tipo de fenómenos ha sido denominado “enjambre” o “manada”, y presenta un carácter impulsivo y una breve duración como colectivo [6]. Provienen de reacciones a eventos inesperados, como ataques por parte de las fuerzas de seguridad o el desalojo de una acampada.

De forma opuesta, existen acciones programadas cuyo ritmo de actividad es sensiblemente menor. El propio proceso y su objetivo conforman un marco rígido en el que el individuo no puede influir significativamente. Sirva como ejemplo la convocatoria de una huelga general o una gran manifestación.

El periodo de tiempo desde las primeras movilizaciones ha constituido un proceso de aprendizaje para sus participantes, que continuamente experimentan y ponen en práctica nuevas estrategias comunicativas. En uno de los ejemplos más destacados, varios de los colectivos más activos del 15M ponen en marcha la campaña “15MPaRato”, que consigue una gran cantidad de apoyo ciudadano (de participación y económico) para presentar una demanda contra el ex-ministro de economía y ex-presidente de Bankia, Rodrigo Rato, como una de las personas responsables de la crisis económica. Se basaba en la intervención estratégica y calculada de algunos nodos muy bien posicionados en las redes sociales, de modo que actuasen como “catalizadores” capaces de regular el proceso comunicativo. Los activistas habían identificado un problema en el tipo de organización de las movilizaciones y decidieron modular la emergencia del proceso.

## 1.2. Twitter: la red social de *microblogging*

Desde su creación en 2006, Twitter irrumpió como red social en Internet con la novedad de proporcionar al usuario un servicio de *microblogging*. Se puede definir un *microblog* como una herramienta para enviar mensajes cortos a una red de asociados [7]. En este caso, los mensajes (llamados *tweets*) tienen un límite de 140 caracteres y, una vez publicados por un usuario, son recibidos por sus seguidores. Esas relaciones de seguimiento no son automáticamente recíprocas, sino que son direccionales<sup>2</sup>.

Los usuarios han desarrollado formas alternativas de comunicarse mediante la inclusión de caracteres especiales dentro de sus mensajes. Por ejemplo, el carácter '@' especifica un receptor concreto. Escrito antes del identificador del usuario (@identificador) indica que el mensaje pretende ser relevante para ese individuo. Se usa tanto para responder al *tweet* de otro usuario (*reply*) como para hacer mención especial a alguien (*mention*). Además, existe la práctica de reenviar un mensaje escrito por un usuario, conocida como *retweet* (RT). Por tanto, los mensajes pueden clasificarse en tres tipos [8]:

- *Singleton*: mensaje no dirigido y original de un usuario.
- *Reply* o *Mention*: contiene una '@' seguida de un identificador de usuario dentro.
- *Retweet*: aparecen las siglas RT seguidas de '@' y el identificador de usuario del cual proviene el mensaje reenviado.

La evolución de la actividad en Twitter también ha provocado la aparición de etiquetas, llamadas **hashtags**, que tienen la función de reunir mensajes relativos a un tema con la intención de facilitar la búsqueda y discusión sobre el mismo. La forma de “etiquetar” un *tweet* es mediante el carácter '#' seguido de una palabra o palabras sin espacios (por ejemplo, #elecciones20n). Cuando una etiqueta entra dentro de las más empleadas por los usuarios se convierte en lo que en Twitter se conoce como *trending topic*. Los hay a distintos niveles, tanto global, nacional o local, pero sólo los primeros se mantienen en una lista permanentemente dentro de las páginas de inicio de todos los perfiles. Llegar a ser *trending topic* implica haber logrado una enorme difusión y es considerado un éxito por aquellos que lo adoptan.

---

<sup>2</sup> Al contrario que en otras redes sociales, como Facebook, en las que se establece un acuerdo bilateral de relación.

El *microblogging* mantiene una estrecha relación con los movimientos sociales, pues proporciona un medio de comunicación muy rápido (instantáneo desde terminales móviles) y un extraordinario potencial de propagación. Junto con la detección de etiquetas permite seleccionar mensajes relativos a un tema concreto y recopilarlos para construir posteriormente distintos tipos de redes: de intercambio de mensajes, de relación entre *hashtags*... así como redes de seguimiento entre usuarios.

Para este estudio, se considera que el conjunto de mensajes que engloba un *hashtag* corresponde a un proceso comunicativo con identidad propia.

### 1.3. Motivación y objetivo del proyecto

Dada la situación actual de agitación política y la correspondiente respuesta en forma de acciones comunitarias, resulta interesante tratar de describir ese comportamiento colectivo de forma científica. La influencia de las tecnologías de la información y su constante evolución sugieren un cambio de paradigma científico que estudie de manera eficiente estos fenómenos. Los agentes externos a los movimientos generalmente proponen interpretaciones erróneas o sesgadas al intentar analizarlos, al hacerlo dentro de marcos anteriores que no se ajustan a estos nuevos casos. Sirva como ejemplo la escasa repercusión en portadas de prensa escrita de las multitudinarias congregaciones ciudadanas el 15 de mayo de 2011, que no se vió reflejada en la totalidad de medios nacionales hasta tres días después [9]. Por otro lado, la crítica por parte de los colectivos activistas a los tipos de movimientos lleva a preguntarnos si esa distinción se puede corroborar a través de las propiedades de los sistemas.

Nuestro objetivo es comprender cómo funciona un proceso auto-organizado en la red y qué características tiene; cómo un estado de ánimo colectivo, en estado latente, pasa en poco tiempo a transformarse en acciones ciudadanas, carentes de liderazgo institucional. Del mismo modo, encontrar diferencias entre formas de auto-organización supone otra meta para el trabajo.

En particular, nos centramos en las movilizaciones de mayo de 2012. El aniversario del movimiento 15M fue el tema central, pero también se llevaron a cabo muchas otras acciones. Nuestro campo de trabajo lo constituye la red social Twitter, que reúne las condiciones necesarias para ser una muestra representativa del movimiento, tanto por número de participantes como por la accesibilidad a sus datos.

### 1.4. Contenido de la memoria

El documento está dividido en los siguientes capítulos:

1. Introducción: incluye una descripción de la relación de los movimientos sociales y las tecnologías de comunicación. Se comentan brevemente los procesos de organización y se presenta la red social en Internet de la que se extraerán los datos. Se detallan la motivación, el objetivo y la planificación del proyecto
2. Bases teóricas: sistemas complejos, teoría de redes. Se examinan los enfoques existentes que tratan el estudio de los sistemas, así como los conceptos y la tipología de redes relativas al estudio.
3. Herramientas: se presentan las herramientas que nos han permitido desarrollar las distintas fases del proyecto.
4. Análisis de los sistemas-red.
5. Conclusiones: se interpretan los resultados obtenidos, se comentan las posibles aplicaciones del estudio y se mencionan posibles líneas futuras de investigación.

## 1.5. Planificación

La primera fase ha consistido en tareas de documentación sobre teoría de redes y herramientas para su análisis, identificando qué resultaba interesante estudiar y cómo trabajar con la información. Seguidamente, se ha procedido a recopilar datos provenientes de las redes sociales sobre las movilizaciones de mayo de 2012, que luego se han tratado y dispuesto de forma adecuada para su análisis. Por medio de la teoría de grafos y bajo un enfoque, alejado del tradicional, basado en los sistemas complejos y sus propiedades emergentes (ver sección 2.1), se han utilizado métodos algorítmicos y estadísticos para caracterizar las redes resultantes. El diagrama de Gantt correspondiente a la realización del proyecto se muestra en la figura 1.1.

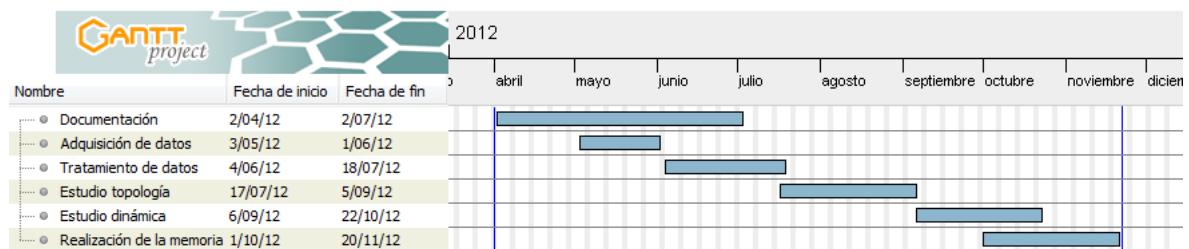


Figura 1.1: Diagrama de Gantt del proyecto.



# Capítulo 2

## Bases teóricas

### 2.1. Sistemas complejos: enfoque tradicional-reduccionista vs. enfoque emergente

El estudio de sistemas complejos ha sido reconocido recientemente como una nueva disciplina científica. Proviene de los avances acontecidos en diversos campos, tales como la física y la antropología. La idea de que toda materia está formada por los mismos bloques (átomos y sus partículas), es uno de los conceptos originales de la ciencia, y su carácter universal complementa las leyes mecánicas. Sin embargo, son también universales los procesos dinámicos que, de alguna manera, hacen evolucionar a estos sistemas. Esto escapa del enfoque científico tradicional, que trata de entenderlos a través de la disección de sus partes como entidades aislables [10]. Bajo esta visión reduccionista, las disciplinas tienden a diverger mientras el conocimiento aumenta a razón del avance específico en cada una de ellas, lo que lleva a su aislamiento progresivo [11]. No obstante, muchas de ellas comparten características y se pueden abordar a través del estudio de los sistemas complejos. De este modo, la aproximación científica a los sistemas particulares se puede hacer dentro de un marco global (figura 2.1).

No se puede modelar un sistema de forma eficiente sin una correcta comprensión de la relación del todo con las partes, así como de las propiedades emergentes de la interacción [12]. Es aquí donde radica la dificultad de su estudio. Por lo tanto, el reto consiste en extraer la información que proporcionan las aristas de unión, tanto desde el punto de vista estático (sus propiedades estructurales, o topología) como dinámico (su evolución en el tiempo), para comprender de qué manera se organizan.

### 2.2. Teoría de grafos y redes sociales

La teoría de grafos es una disciplina de las matemáticas y las ciencias de computación que estudia las propiedades de estructuras formadas por conjuntos de vértices o nodos unidos entre sí por aristas. Aplicada a las redes sociales, permite mapear las relaciones entre individuos, modelando éstos como nodos y sus relaciones como enlaces. De la naturaleza de una red social deriva su clasificación según tres criterios:

- Direccionalidad
  - Dirigida: la relación social no es bidireccional sino que apunta de un vértice a otro, como la red de seguidores en Twitter o el envío de correos electrónicos (ver figura 2.2).
  - No dirigida: existe reciprocidad en el enlace, como los contactos en Facebook o una relación de coautoridad en artículos.

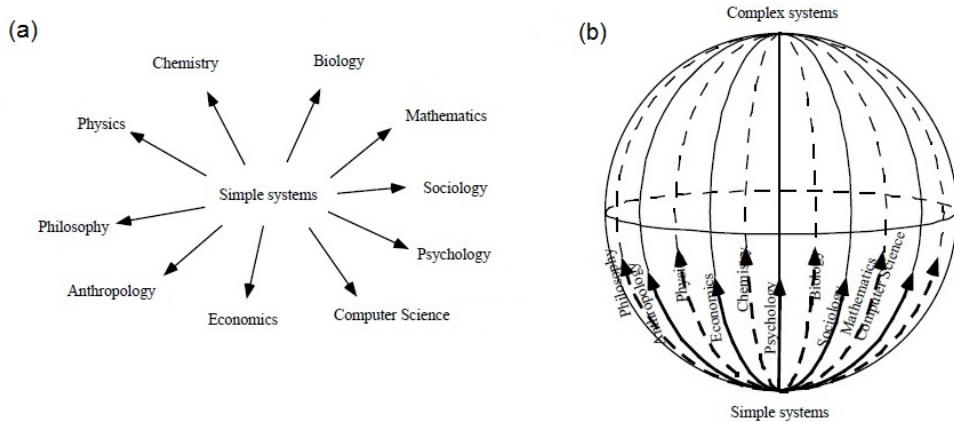


Figura 2.1: Ilustración conceptual: (a) Enfoque tradicional-reducciónista (b) Enfoque emergente. Fuente: [11].

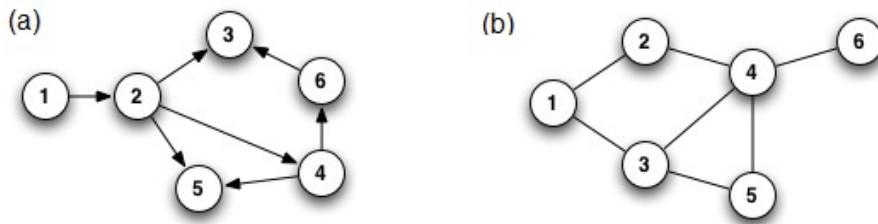


Figura 2.2: Direccionalidad: (a) red dirigida, (b) red no dirigida. Fuente: [13].

■ Carácter explícito-implícito

- Explícito: los propios individuos declaran la relación. Se da, por ejemplo, entre contactos de correo electrónico o de Facebook.
- Implícito: la relación se deduce del comportamiento. Comprar en e-Bay o interaccionar en foros produce relaciones implícitas.

■ Peso de las aristas

- Capacitada: las aristas llevan asociado un valor que expresa su fortaleza o peso. En la comunicación entre dos personas por correo electrónico, una posible medida del peso correspondería al número de mensajes intercambiados.
- No capacitada: los enlaces sólo representan unión, no poseen ningún valor relacionado (red de seguidores de Twitter).

### 2.2.1. Topología. Definiciones y tipos de redes complejas según su estructura

La teoría de grafos clásica parte de la topología como primer análisis. Es muy útil para caracterizar los sistemas de forma estática y permite relacionarlos con tipos de grafos estudiados. En la naturaleza, las redes que encontramos son, en general, redes de mundo pequeño y redes libres de escala. Estudiamos las propiedades relativas a estas redes para ver si nuestros sistemas poseen estructuras similares.

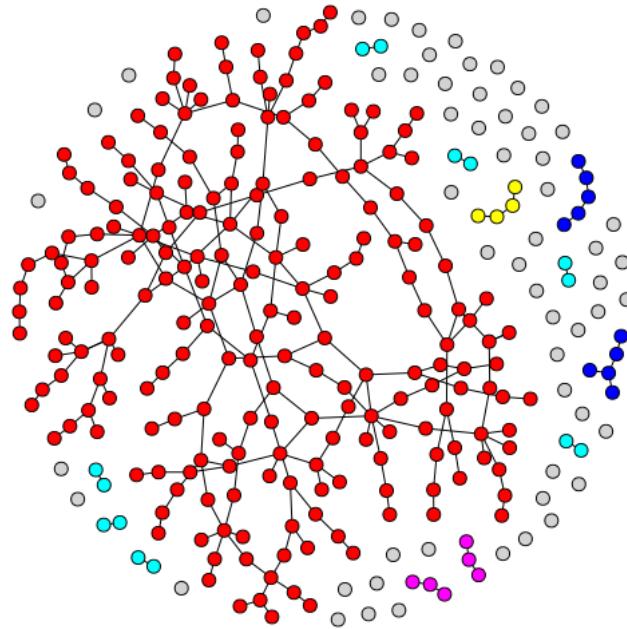


Figura 2.3: Componente gigante de una red. Fuente: [14].

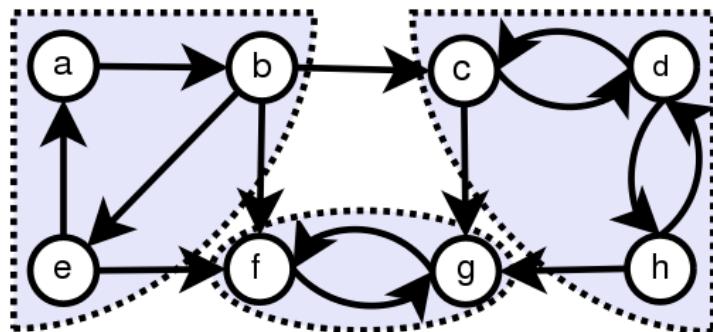


Figura 2.4: Subgrafos fuertemente conexos. Fuente: wikipedia.org.

### Componente gigante

En ocasiones las redes están formadas por más de un componente conexo. Si hay un subgrafo conexo que aglutina la gran mayoría de nodos, es interesante comparar las propiedades de la red global y de este componente, llamado componente gigante (ver Figura 2.3). Las características globales de la red pueden ser muy similares a las que presenta este subgrafo.

Para redes dirigidas, se distinguen dos tipos de componentes conexos:

- Componente fuertemente conexo: si para cada par de vértices  $u$  y  $v$  pertenecientes a un subgrafo existe un camino de  $u$  hacia  $v$  y viceversa (ver figura 2.4).
- Componente débilmente conexo: al sustituir las aristas dirigidas por no dirigidas el subgrafo se mantiene unido, sin restricciones direccionales.

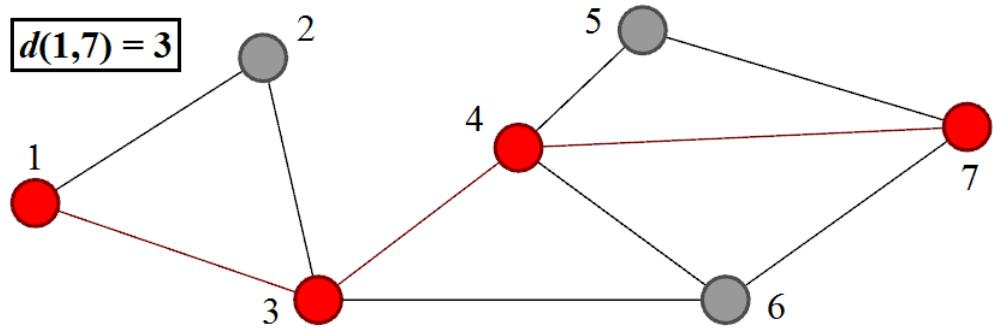


Figura 2.5: Distancia entre dos nodos.

### Distancia promedio

La distancia mide el número de aristas que separan un vértice de otro por el camino más corto posible (ver figura 2.5). Siendo  $G$  un grafo con un grupo  $V$  de  $n$  vértices, y  $d(i, j)$  la distancia entre dos de ellos, la distancia promedio (o longitud de camino media) de la red es:

$$L = \frac{1}{n(n-1)} \sum_{i,j} d(i, j)$$

Expresa el promedio de intermediarios por contacto directo existentes entre todos los pares de nodos del sistema. Si el grafo es dirigido, la distancia promedio tiene dos variantes:

- Distancia promedio dirigida: para trazar un camino es necesario recorrer las aristas en el sentido que indican. Es, por lo tanto, más restrictiva.
- Distancia promedio no dirigida: se calcula igual que en una red no dirigida, convirtiendo todos los enlaces en no direccionales.

### Grado de un nodo

Es un parámetro local de cada nodo, que indica el número de enlaces adyacentes a él.  $k_i$  denota el grado del nodo  $i$ . Cuando se trata de redes dirigidas, el grado está compuesto por la suma del saliente (+) y el entrante (-)

$$k_i = k_i^+ + k_i^-$$

### Coeficiente de *clustering* o de agrupamiento

Mide la densidad de conexiones entre vecinos directos de un usuario (ver figura 2.6). Si se define  $M_i$  como el número de aristas existentes entre vecinos inmediatos del vértice  $i$ , su coeficiente de *clustering* será

$$c_i = \frac{M_i}{k_i(k_i - 1)}$$

$$0 \leq c_i \leq 1$$

Para dar un valor general del agrupamiento local de la red se calcula el promedio

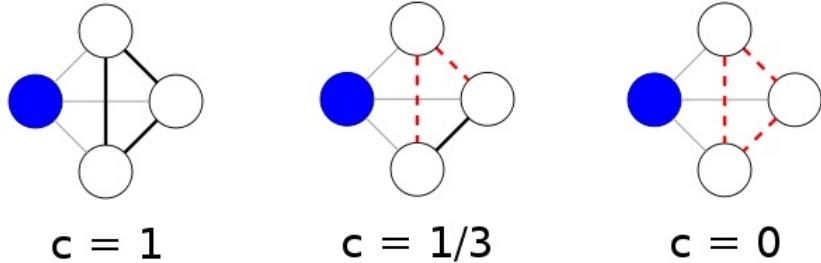


Figura 2.6: Coeficiente de *clustering*. Fuente: wikipedia.org.

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

$$0 \leq C \leq 1$$

### Redes de mundo pequeño

El origen de estos sistemas se encuentra en el experimento del mundo pequeño de Stanley Milgram (1967). Consistió en seleccionar aleatoriamente individuos en el medio oeste americano que debían hacer llegar una tarjeta postal a un extraño en la costa este del país, conociendo únicamente su nombre, ocupación y localización geográfica aproximada. El procedimiento era enviar el mensaje a alguien en su red directa de conocidos que pudiese estar más cerca del objetivo, basándose sólo en esos tres datos. Se procedía así hasta que el mensaje se perdía o llegaba a su destinatario. A pesar de que únicamente 64 de 296 mensajes alcanzaron el objetivo, lo hicieron con una media de entre 5,5 y 6 pasos intermedios.

Para describir este caso científicamente, se parte de los tipos de grafos estudiados tradicionalmente. Con objeto de simplificar su estudio, se distinguía entre redes regulares y aleatorias. Para un determinado tamaño de red  $n$ , y con grado medio  $\bar{k} = k$ , las primeras se distribuyen de manera uniforme y todos sus nodos poseen grado  $k$ . Los grafos aleatorios se construyen según el modelo de Erdős-Renyi<sup>1</sup> [15] que, manteniendo el grafo conexo, alcanzan idéntico número de aristas total. Partiendo del grafo regular, se asigna una probabilidad  $p$  de que una arista conecte aleatoriamente con otro individuo sin duplicar una ya existente [16]. Por consiguiente, se van obteniendo redes con una aleatoriedad creciente con  $p$  (figura 2.7).

El estudio del comportamiento de  $L$  y  $C$  en zonas próximas a los extremos permite extraer empíricamente ciertas conclusiones. Conforme el caso se aproxima a la red regular

$$p \rightarrow 0 \Rightarrow L \sim \frac{n}{2k}, C \sim \frac{3}{4}$$

Cuando se acerca al otro extremo

$$p \rightarrow 1 \Rightarrow L \approx L_{random} \sim \frac{\ln(n)}{\ln(k)}, C \approx C_{random} \sim \frac{k}{n}$$

Por tanto, intuitivamente, cuando  $p \rightarrow 0$  se asocian valores altos de  $L$  (que crece linealmente con  $n$ ) con elevados valores (en torno a  $\frac{3}{4}$ ) de  $C$ . Si  $p \rightarrow 1$  las dos variables parecen alcanzar valores bajos al mismo tiempo ya que  $L$  sólo crece logarítmicamente con  $n$ , y  $C$  cae abruptamente.

<sup>1</sup> El modelo Erdős-Renyi es uno de los más aplicados para crear grafos aleatorios. Este documento se referirá a este modelo cuando se mencione un grafo aleatorio o sus propiedades.

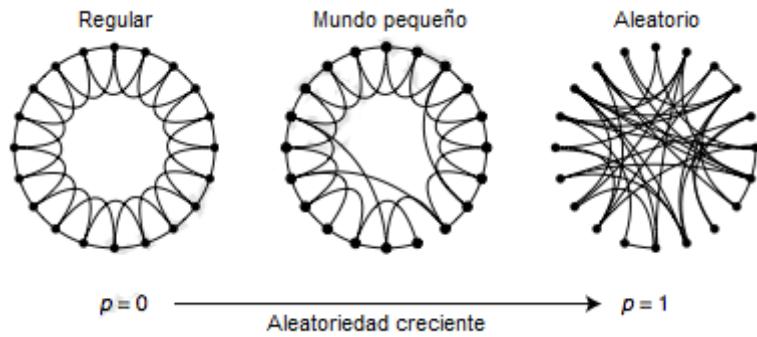


Figura 2.7: Aleatoriedad creciente en función de  $p$ . Fuente: [16].

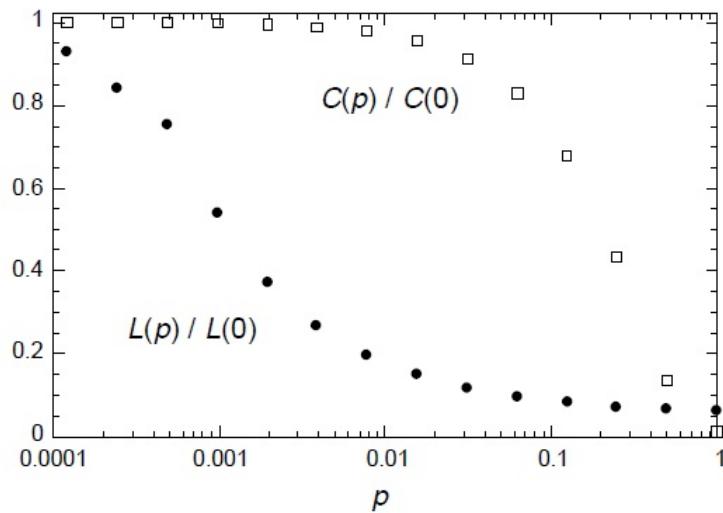


Figura 2.8:  $L$  y  $C$  en función de  $p$ . Fuente: [16].

Sin embargo, la representación de la variación de los coeficientes con  $p$  (normalizados por sus correspondientes valores en grafo regular,  $L(0)$  y  $C(0)$ ) revela un amplio rango en el cual  $L$  presenta valores bajos y  $C$  mantiene valores elevados (ver figura 2.8). Estas redes intermedias se denominan redes de mundo pequeño y presentarán un agrupamiento alto a nivel local, propio de una red regular, y una distancia promedio baja a nivel global, característica de un sistema aleatorio. Comparado con sus valores análogos en el caso aleatorio,  $L$  ha de ser similar o ligeramente superior, y  $C$  ha de ser muy superior.

$$L \gtrsim L_{\text{random}}$$

$$C \gg C_{\text{random}}$$

En la tabla 2.1 se observa la comparación de estos valores en tres sistemas reales.

### Distribución de grados

Al construir grafos aleatoriamente es interesante ver cómo evoluciona su conectividad al ir añadiendo aristas, es decir, al ir aumentando el grado medio  $\bar{k}$ . Se emplea una representación gráfica del número de elementos  $N(k)$  que poseen  $k$  conexiones. La probabilidad de que un vértice

	Red	$L_{real}$	$L_{random}$	$C_{real}$	$C_{random}$
(a)	Actores	3.65	2.99	0.79	0.00027
(b)	Red eléctrica	18.7	12.4	0.08	0.005
(c)	<i>C. elegans</i>	2.65	2.25	0.28	0.05

Tabla 2.1: . Corresponden a los siguientes casos: (a) Componente gigante de la red de colaboración de actores de cine de una popular base de datos en Internet. (b) Generadores, transformadores y subestaciones unidas por líneas de alta tensión en EEUU. (c) Representación de la red neuronal del gusano *Caenorhabditis elegans*.

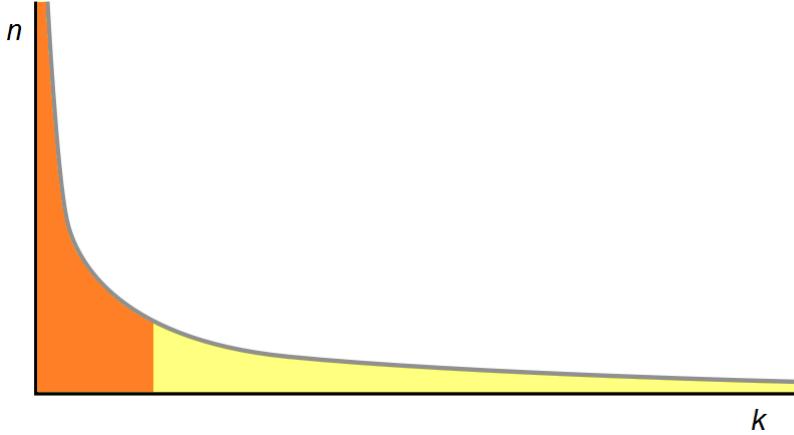


Figura 2.9: Distribución según ley de potencia, o *long tail*

esté dotado de un grado  $k$ ,  $P(k)$ , varía según el modelo analizado. El modelo aleatorio de Erdős-Renyi muestra una distribución de Poisson con el pico en  $\bar{k}$

$$P(k) = e^{-z} \frac{z^k}{k!}$$

### Redes libres de escala

Las redes libres de escala, clasificadas como uno de los tres tipos de redes de mundo pequeño según variables estadísticas [17], están compuestas por una gran mayoría de elementos pobemente enlazados, y unos pocos, llamados *hubs*, con una conectividad muy alta. La distribución de grado responde a una ley de potencias, llamada también “de larga cola” o *long tail* (ver figura 2.9).

$$P(k) \sim ck^{-\gamma}$$

donde  $c$  es una constante de normalización y  $\gamma$  un coeficiente que oscila típicamente entre 2 y 3. Si aplicamos logaritmos a ambos lados de la expresión se obtiene que

$$\log(P(k)) \propto -\gamma \log(k)$$

La distribución queda representada por una recta de pendiente  $-\gamma$ . Existe una tendencia generalizada por parte de los nuevos individuos a enlazarse mayoritariamente con *hubs*. Numerosos estudios sobre redes reales han demostrado comportamientos de este tipo [18][19][20]. Las redes sociales en Internet también suelen presentar una distribución de larga cola [21].

### 2.2.2. Análisis de comunidades y mesoescala.

La división de un sistema en módulos o comunidades, cuyos nodos presentan un alto grado de conexión dentro de su grupo, es importante para descubrir estructuras *a priori* desconocidas que se agrupan por algún tipo de afinidad: geográfica, política... Para determinar la calidad de una partición se usa la función modularidad,  $Q$ . Se define como [22]

$$Q = \sum_i (e_{ii} - a_i^2)$$

donde  $e_{ij}$  es la fracción de aristas total que unen el grupo  $i$  con el  $j$ , y  $a_i = \sum_j e_{ij}$ . Para redes capacitadas [23]

$$Q = \frac{1}{2w} \sum_{ij} \left[ w_{ij} - \frac{w_i w_j}{2w} \right] \delta(C_i, C_j)$$

donde  $w_{ij}$  representa el peso de la arista  $m_{ij}$ ,  $w_i = \sum_j w_{ij}$  suma de pesos de las aristas del vértice  $i$ ,  $C_i$  representa la comunidad  $i$ ,  $w = \frac{1}{2} \sum_{i,j} w_{ij}$  y la función  $\delta(u, v) \begin{cases} 0 & \text{si } u \neq v \\ 1 & \text{si } u = v \end{cases}$ .

Hay diversos métodos de detección de comunidades:

- Algoritmos decrementales, que detectan y eliminan enlaces intercomunitarios [24][25].
- Algoritmos aglomerativos, que fusionan progresivamente nodos y comunidades similares [26].
- Métodos de optimización que maximizan la modularidad [27].

Todos se valen de la modularidad para evaluar sus resultados. Un valor de  $Q$  mayor de 0,3-0,4 es un claro indicador de que los subgrafos obtenidos corresponden a comunidades. Se obtienen divisiones más precisas y rápidas mediante la maximización de la modularidad, por ello se ha utilizado un algoritmo de este grupo, descrito en la sección 3.4.

En ocasiones, existen comunidades dentro de otras más grandes que son imposibles de detectar en la red original. En particular, si una comunidad  $s$  con una suma de pesos internos  $w_{ss}$  cumple

$$w_{ss} < \sqrt{\frac{w}{2}} - 1$$

permanecerá indetectada [28]. Si se introduce un factor  $r$  de aumento de peso a cada nodo, la expresión queda

$$w_{ss} < \frac{1}{2} \left( \sqrt{2w + Nr} - n_s r - 2 \right)$$

Al aumentar  $r$  disminuye el límite de resolución, ya que  $\sqrt{r}$  crece más despacio que  $r$ , por lo que disminuirá el tamaño de las comunidades identificables.

Es interesante incluir aquí el concepto de mesoescala, relacionado con el tamaño de los módulos detectados. Abarca un espectro que va desde la microescala (cada nodo representa su propia comunidad) a la macroescala (toda la red formando un elemento único). El parámetro  $r$ , llamado resistencia o *self-loop*, determinará las divisiones intermedias. Para aumentar el peso de cada nodo, agrega bucles de peso  $r$  (ver figura 2.10). Matemáticamente, se suma a la matriz de adyacencia con pesos de la red,  $W$ , la matriz identidad multiplicada por el factor  $r$  [29]:

$$W_r = W + r\delta_{ij}$$

Cabe mencionar que esta variación topológica no altera ninguna otra propiedad global de la red. Representando el valor de  $r$  frente al número de comunidades obtenidas, se puede obtener información acerca de la estabilidad de las particiones (ver Figura 2.11).

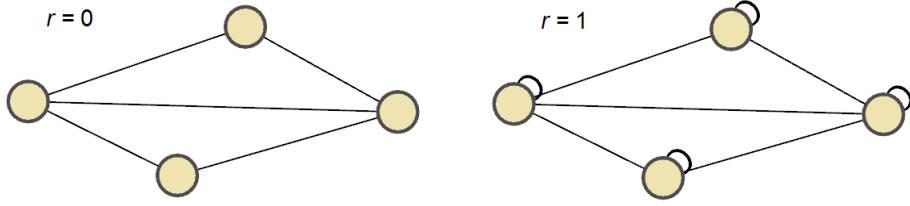


Figura 2.10: Redes y resistencia,  $r$ .

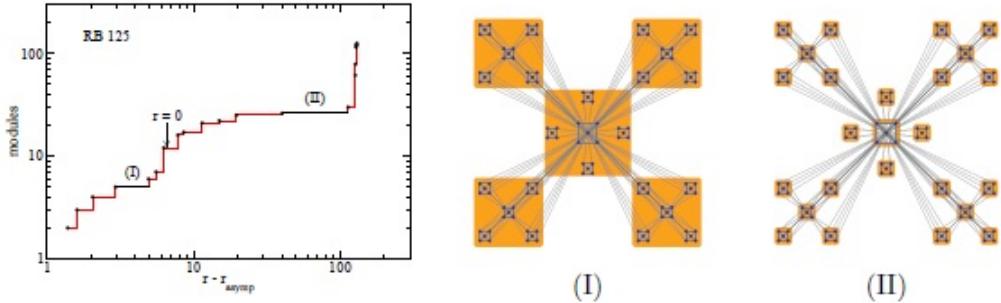


Figura 2.11: Análisis de mesoscala en una red artificial: en la gráfica se observan las dos particiones estables, (I) y (II), y el punto donde  $r = 0$ , correspondiente al sistema sin resistencia. Fuente: [29]

### 2.2.3. Dinámica

La dinámica de los sistemas reales, incluidos los movimientos sociales, no responde a un modelo reduccionista. El estudio de las partes de un sistema por separado lleva a una interpretación errónea por obviar otras propiedades emergentes del mismo que no se aprecian de forma directa, sino que surgen como resultado de la interacción entre los componentes. Por ello, imprimen notables diferencias al sistema en comparación con la suma de dinámicas individuales.

Se han desarrollado herramientas capaces de analizar la constitución dinámica de un sistema. Una de ellas es la caracterización de los procesos dentro de un sistema en términos de las correlaciones temporales entre las diferentes escalas que lo componen, en forma de procesos de tipo  $1/f^\beta$ . El ruido  $1/f^\beta$  hace referencia a la distribución de la densidad espectral, que representa el modo en que se reparte la potencia o energía en una señal:

$$S_{xx}(f) = |X(f)|^2$$

donde  $X(f)$  es la transformada de Fourier de  $x(t)$ . En concreto, decimos que un proceso es de tipo  $1/f^\beta$  cuando su densidad espectral es proporcional a una función de este tipo:

$$S_{xx}(f) \propto \frac{1}{f^\beta}$$

La existencia de ruido  $1/f^\beta$  depende de la interdependencia de las diferentes escalas temporales de un sistema, permitiéndole comportarse como una unidad dinámica coherente, algo que es característico de procesos complejos y auto-organizados [30]. Aplicado a un proceso de movilización, podemos pensar en las escalas más rápidas (conversaciones entre individuos, retransmisión inmediata (*retweet*) de mensajes) o en las escalas más lentas (la gestación de una gran manifestación). Habría otras muchas escalas intermedias, como los ritmos de actividad de usuarios condicionados por sus horarios, ritmos guiados por noticias en los medios de comunicación, o dinámicas propias

de las redes sociales, como los *trending topic* en Twitter. Por ello, para que un proceso de movilización se constituya como una unidad con identidad coherente debe haber un equilibrio en el peso de cada escala. Si dominan las rápidas, el sistema se comportará de manera impulsiva e irreflexiva. Si, por el contrario, dominan las lentas, las identidades creadas suelen ser muy rígidas e incapaces de adaptarse al entorno inmediato, por lo que dificultan la comunicación efectiva.

La caracterización  $1/f^\beta$  mide la influencia de las dinámicas a diferentes escalas de un sistema, para evaluar el peso de cada una de ellas (rápidas, lentas e intermedias). El parámetro  $\beta$  permite clasificar los procesos y relacionarlos con los movimientos sociales:

- Ruido blanco ( $\beta = 0$ ): describe procesos decorrelados, dominados por fluctuaciones aleatorias sin correlación en el tiempo. Dominados por escalas temporales cortas (a alta frecuencia), funcionan espontánea y creativamente pero son incapaces de mantener una identidad estable. Un ejemplo sería la respuesta a un desalojo o una concentración espontánea.
- Ruido marrón ( $\beta = 2$ ): describe procesos fuertemente correlados y predecibles, en los que dominan las escalas temporales largas (bajas frecuencias). Son sistemas inflexibles y, aunque mantienen una identidad coherente, su rigidez les impide adaptarse al entorno. La preparación de una huelga o la gestación de una gran manifestación podrían ser dos ejemplos, pues están estructurados de antemano.
- Ruido rosa ( $\beta = 1$ ): describe procesos cuyas escalas temporales (cortas, medias y largas) tienen una influencia similar. Se alcanza un equilibrio entre los estados desordenados con gran contenido de información (ruido blanco) y estados “con memoria” pero bajo contenido de información (ruido marrón). Esta dinámica, dominada por la interacción de sus componentes, le confiere al proceso estabilidad en su identidad y versatilidad a la hora de responder al entorno.

# Capítulo 3

## Herramientas

### 3.1. Adquisición de datos

#### 3.1.1. Recopilando información de la red

Existen varios métodos para extraer datos de la red.

- *Web scraping*: es una técnica que simula el comportamiento humano en la exploración de Internet. Se basa en recorrer sitios *web* y extraer la información deseada, pasando de la forma desestructurada de las páginas, normalmente en lenguaje HTML, a un formato estructurado que facilite su posterior almacenamiento y análisis.
- *Web crawling*: explora y copia la red de forma automática. Parte de un conjunto inicial de sitios *web*, llamados semillas, y busca URLs dentro de éstas, siguiendo los criterios del usuario. Todo lo explorado se descarga y un motor de búsqueda lo indexa para poder analizarlo después.
- Bases de datos y *Open Data*: existen sitios en Internet donde se pueden consultar, y en ocasiones descargar, información sobre multitud de campos.

#### 3.1.2. Twitter

Twitter proporciona dos APIs para la extracción de información. Por un lado, Search API permite buscar, aplicando los criterios deseados, todos los *tweets* bajo unas limitaciones temporales y cuantitativas. Respecto al marco temporal, devuelve resultados de la búsqueda con un máximo de una semana de antigüedad. Respecto a la cantidad, el máximo de resultados por búsqueda está limitado a 1500. Asimismo, se puede alcanzar un número máximo de llamadas al servidor, establecido en 150 llamadas/hora por cada usuario o IP. En ese caso, una hora después de la última llamada, el contador vuelve a cero y la búsqueda puede proseguir.

En segundo lugar, a través del Streaming API se habilita la recogida de *tweets* en tiempo real. Proporciona más información relativa al emisor que el método anterior. No tiene limitaciones de cantidad, llamadas al servidor o tiempo de uso. No obstante, tiene desventajas importantes. Dificulta el seguimiento de información puesto que en la mayoría de los casos no se puede predecir que tipos de datos van a ser relevantes. Se necesita un tiempo para evaluar la importancia de un determinado usuario o *hashtag*. Además, el uso del streaming requiere una conexión permanente tanto en el equipo que recopila como en el origen.

#### 3.1.3. ScraperWiki

Es una herramienta *web* que permite hacer *web scraping* de forma automática y periódica y almacenar datos en archivos descargables. Para el caso que ocupa este documento, la combinación

de este servicio con el Search API de Twitter proporciona la mejor opción. Esto se debe a que había que comprobar la importancia de las etiquetas usadas y ver si tenían buena acogida y mediante Streaming API no se permite acceder a datos pasados.

Realiza un barrido cada hora en el buscador de Twitter recogiendo los mensajes que contengan los datos objetivo (en este caso, los *hashtags* relevantes, ver 4.1). Terminado el proceso, la información se descarga en archivos .csv<sup>1</sup>.

## 3.2. Tratamiento de datos

Para transformar los datos de los *tweets*, adecuándolos a un formato propio para su análisis y visualización, hemos utilizado el lenguaje de programación AWK. Es un lenguaje sencillo y su forma de trabajar con listas indexadas por palabras clave, aunque limitada, es suficiente para nuestro objetivo. Dado que la información por tweet se clasifica en archivos .csv y cada categoría corresponde a una columna, permite de forma sencilla seleccionar los datos de interés y descartar el resto.

El programa realizado se ocupa, para cada listado de *tweets*, de eliminar todo aquel que no contenga una '@' en el campo 'Texto', seguida del identificador del usuario. De este modo se descartan los mensajes que no tienen receptor (los tipo *Singleton*, ver 1.2), ya que no añadirían información a las redes objetivo del estudio. Inspecciona el contenido del mensaje y para cada estructura "@+IDusuario" crea una línea en un nuevo archivo con emisor, receptor y fecha y hora del envío. Es decir, añade un par de nodos unidos por una arista dirigida. Si hay varios mensajes con idénticos participantes el peso de la arista se va incrementando.

## 3.3. Análisis y visualización de redes

### 3.3.1. Software

#### Gephi

Es un programa de software libre [32] que permite importar archivos en varios formatos (entre ellos, el .csv) con datos de un grafo. Sus principales ventajas son la sencillez de manejo, la rapidez en el cálculo de propiedades, la aplicación de filtros para depurar la red y las opciones de visualización. Presenta problemas de computación, en función de la máquina que lo ejecute, al manejar redes de tamaños muy grandes.

Se ha utilizado para crear archivos de grafo a partir de los datos recogidos y tratados. Además, se han calculado propiedades de los sistemas.

#### Radatools

Es un conjunto de programas de distribución libre para analizar redes complejas [33]. Se ha usado principalmente para la detección de comunidades y el análisis de la mesoescala, y se han calculado propiedades que por razones diversas Gephi no podía obtener correctamente.

#### Matlab

Es un popular software matemático que cuenta con su propio lenguaje de programación (lenguaje M). Se ha utilizado para representar los datos y para realizar operaciones matemáticas, como la aproximación por polinomios de una función (sección 4.2.2).

---

<sup>1</sup>Comma-Separated Values: archivo que guarda en filas cada unidad de datos, y distribuye, separados por coma, los distintos campos, como el nombre de usuario emisor, receptor, fecha...

### 3.4. Algoritmo de detección de comunidades y mesoescala

Se pretende diferenciar grupos dentro del sistema global que presenten mayor densidad de conexión dentro de sí mismos, lo que les confiere un carácter comunitario. Este algoritmo busca maximizar la función modularidad

$$Q = \sum_i (e_{ii} - a_i^2)$$

mediante el agrupamiento progresivo de nodos [34].  $e_{ij}$  es la fracción de aristas total que unen el grupo  $i$  con el  $j$ , y  $a_i = \sum_j e_{ij}$ .  $Q$  representa, por la fracción de vértices que se encuentran dentro de subgrupos,  $e_{ii}$ , menos dicha fracción en el supuesto caso de repartir los enlaces aleatoriamente,  $a_i^2$ . La división en comunidades comenzará a ser significativa a partir de  $Q > 0,3$ .

El proceso comienza por aislar todos los individuos, manteniendo los vértices en segundo plano. Se van uniendo todas las posibles parejas de nodos para acabar validando aquella variación que proporcione un mayor  $\Delta Q$ . No se considerarán todas aquellas posibles parejas que en el sistema completo no estén unidas, pues nunca darán como resultado un aumento en  $Q$ .

$$\Delta Q = e_{ij} + e_{ji} - 2a_i a_j = 2(e_{ij} - a_i a_j)$$

Una vez realizada y validada la unión, se actualiza la matriz de elementos  $e_{ij}$ , puesto que ahora dos nodos han pasado a formar parte de un mismo grupo. Cada paso irá formando una nueva partición con un valor máximo de  $Q$  asociado.

El algoritmo funciona de forma equivalente en redes capacitadas, sustituyendo los valores de aristas por los pesos [23].

### 3.5. *Detrended Fluctuation Analysis (DFA)*

Es un método para determinar la *auto-afinidad* estadística de una señal. Es útil para analizar series temporales y localizar correlaciones entre distintas escalas de tiempo. Dada una serie temporal de  $N$  puntos,  $\{z_t, t = 1, \dots, N\}$  se calcula la serie acumulada

$$Z(t) = \sum_{u=1}^t (z_u - \bar{z})$$

donde

$$\bar{z} = \frac{1}{N} \sum_{t=1}^N z_t$$

es la media global. Seguidamente, a lo largo de la serie  $Z(t)$ , se fijan un número  $N_b$  de “ventanas” o “cajas” que contienen todas ellas  $N_\tau$  puntos y cuyo tamaño temporal es  $\tau$ . Dentro de cada caja se elimina la tendencia mostrada por los puntos que la forman mediante la función de la fluctuación sin tendencia (*detrended fluctuation function*), definida como

$$\Psi^k(t) = Z(t) - Z_{fit}^k(t)$$

siendo  $Z_{fit}^k(t)$  un ajuste polinómico de grado 1 realizado por el método de mínimos cuadrados dentro de cada caja, y donde  $k$  indica el número de la caja. Para cada valor de  $\tau$  se calcula la función

$$F(\tau) = \left( \frac{1}{N_b} \sum_{k=1}^{N_b} \frac{1}{N_\tau} \sum_{t=1}^{N_\tau} [\Psi^k(t)]^2 \right)^{1/2}$$

que mide la fluctuación de la raíz cuadrática media. Si la señal presenta un comportamiento de escala, entonces

$$F(\tau) \propto \tau^\alpha$$

donde el exponente de escala o de correlación  $\alpha$  aporta información sobre posibles correlaciones de la señal. El análisis dinámico (sección 4.4) se realiza con la densidad espectral, y la función, dependiente de la frecuencia, responde a

$$S_{xx}(f) = |X(f)|^2 \propto \frac{1}{f^\beta}$$

La relación entre los exponentes  $\alpha$  y  $\beta$  es

$$\beta = 2\alpha - 1$$

donde  $\beta$  determinará el color del ruido.

# Capítulo 4

## Análisis

### 4.1. Selección de datos

Durante el mes de mayo de 2012 se produjeron una serie de movilizaciones relacionadas con el movimiento 15M: el aniversario, fijado el 12 y 15 de mayo, la huelga de educación, campañas contra La Caixa... Se han escogido los *hashtags* representativos de cada movilización. Como ejemplos, para las concentraciones del 12 de mayo de 2012 la etiqueta central fue #12m15m y para la huelga de educación, #22m, entre otras (ver figura 4.1). Cada uno de estos *hashtags* se ha tomado como representativo de un proceso comunicativo, que a su vez ha formado parte de otro proceso más amplio de comunicación social (incluyendo actos en la calle, repercusión en los medios, etc.).

Se clasifican los sistemas correspondientes en tres tipos de movimiento:

- Tipo A: acciones espontáneas o desencadenadas por un factor externo y sin perspectivas concretas más alla del corto plazo. Son procesos que se coordinan rápidamente y que suelen actuar de forma impulsiva. Surgen por diversos motivos, aunque principalmente responden a eventos externos inesperados. Encontramos ejemplos en las reacciones ciudadanas ante el desalojo de una acampada o ante una noticia perjudicial para los intereses de un colectivo. En algún caso, hemos catalogado como tipo A procesos comunicativos que tienen por único objetivo alcanzar una visibilidad en la red notable en unas pocas horas, puesto que comparten características similares en cuanto a organización (rápida y en ocasiones caótica) y duración en el tiempo.
- Tipo B: convocatorias establecidas con antelación y de forma estructurada. Estos procesos suelen ser rígidos, en el sentido en que hay escasas posibilidades de cambio en sus procedimientos u objetivos. Pueden congregar a un elevado número de participantes, pero su actividad suele ser de baja frecuencia y normalmente está condicionada por la planificación propia del movimiento. Se incluye en esta categoría la preparación de una huelga general o la convocatoria de una protesta masiva.
- Tipo A-B: acciones mixtas, bien por su mezcla de planificación e improvisación, bien por tener objetivos tanto a corto como a largo plazo. En ocasiones, se fija una fecha para un evento en concreto que más adelante se complementa con acciones no planificadas. En base a esta clasificación, diríamos que posee elementos de los tipos A y B. Véase como muestra la convocatoria del 12M15M, que derivaba en actos de carácter espontáneo. Asimismo, otros procesos combinan maniobras inmediatas con metas a largo plazo. Es evidente que hay una planificación detrás, pero se realiza de una forma menos visible que los de tipo B. La estrategia de los catalizadores, descrita en la sección 1.1.2, responde a estas características.

El listado completo de hashtags se recoge en la tabla 4.1.

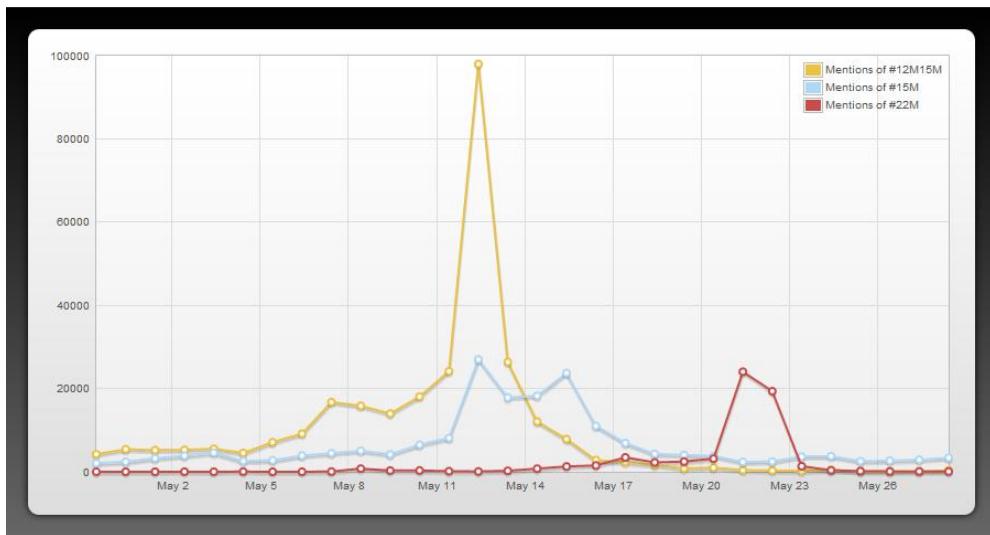


Figura 4.1: : #12m15m, #15m y #22m, en mayo de 2012. Fuente: topsy.com

Hashtag	Descripción	Tipo
#15mSectorRadical	Sectores más ideologizados adoptan esta etiqueta	A
#16m	Acciones del 16 de mayo, surgidas sin planificación previa	A
#alabolsa	Protestas espontáneas en la Bolsa de Madrid el 16 de mayo	A
#asambleafcom	Asamblea en Facultad de Comunicación (Sevilla) sobre recortes	A
#bankiaesnuestra	Respuesta al rescate de Bankia	A
#cierrabankia	Llamada al cierre de cuentas en Bankia	A
#desalojoSol	Desalojo de la Plaza de Sol la madrugada del 13 de mayo	A
#larimaia	Desalojo del centro social La Rimaia, en Barcelona	A
#nurembergfinanciero	Se pide juzgar a los culpables de la crisis financiera	A
#planderescateciudadano	Campaña-respuesta a los rescates bancarios	A
#prima500	Reacción tras sobrepasar la prima de riesgo los 500 puntos	A
#12mglobal	Carácter internacional de las convocatorias del 12 de mayo	B
#22m	Huelga de educación convocada por distintos sindicatos	B
#alaplaza12m	Llamamiento a tomar las plazas	B
#es15m	Aniversario del movimiento, ligado a más protestas	B
#feliz15m	Otra etiqueta del aniversario del movimiento	B
#huelgadeclase	Identifica la huelga de educación del 22 de mayo	B
#yovoy12m	Otro indicador más de las movilizaciones del 12 de mayo	B
#12m15m	Etiqueta central de las convocatorias del 12 de mayo	A-B
#15mpaRato	Nombre de la campaña para juzgar a Rodrigo Rato	A-B
#lacaixaesmordor	Campaña en Barcelona contra La Caixa	A-B
#occupymordor		

Tabla 4.1: Listado de hashtags relevantes

## 4.2. Análisis de la topología

El estudio de la estructura de una red es una primera aproximación a la comprensión de estos sistemas. El análisis más clásico de la teoría de grafos se sustenta en la topología como punto de partida, pues proporciona información básica para su interpretación.

Se estudian las características descritas en la sección 2.2.1 y se relacionan con los tipos de redes existentes. A partir de este punto, se analizan las características de los componentes gigantes débilmente conectados, pues nos interesa estudiar sólo los grandes sistemas conexos formados, evitando así la aparición de nodos que hayan usado la etiqueta con otros fines<sup>1</sup>. En el anexo ?? se muestra el porcentaje de aristas y nodos que aúnan los componentes gigantes con respecto a las redes completas.

### 4.2.1. Redes de mundo pequeño

La topología de mundo pequeño dota a una red de baja distancia promedio, que favorece la velocidad de transmisión globalmente. Al mismo tiempo, su elevado coeficiente de *clustering* indica que no son redes caóticas, sino que sus vértices tienden a agruparse a nivel local y proporcionan estabilidad al sistema. Por tanto, destaca su **transmisión eficiente de la información y la alta presencia de clusters**.

Para cada *hashtag*, se calculan el coeficiente de *clustering* y la distancia promedio y se comprueba si se aproximan a la topología de mundo pequeño. Las distancias calculadas serán siempre dirigidas (ver sección 2.2.1). Por consiguiente, se necesitan las redes aleatorias equivalentes (con mismo  $\bar{k}$  y  $N$ ) para realizar la comparación. La probabilidad  $p_c$  de que una pareja de vértices esté unida entre sí se define como

$$p_c = \frac{2m}{n(n-1)}$$

Generando grafos aleatorios con los criterios  $p_c$  y  $N$  se obtienen los datos necesarios, recogidos en el Anexo ???. Consideramos las variables  $L/L_{random}$  y  $C/C_{random}$  porque, al dividir por su equivalente aleatorio, se estandariza de alguna forma un análisis que involucra redes de distintos  $\bar{k}$  y  $N$ , cuya influencia puede desequilibrar la comparación entre sistemas.

Las condiciones  $C \gg C_{random}$  y  $L \gtrsim L_{random}$  se satisfacen en todos los casos, aunque algunos sistemas se aproximan más al modelo. En la figura 4.2 se representa gráficamente  $C/C_{random}$  frente a  $L/L_{random}$ . No se aprecian correlaciones entre ambas variables, pero se aprecian diferencias por tipos:

- Tipo A (azul): están agrupados en una zona con valores de  $L/L_{random}$  y  $C/C_{random}$  relativamente bajos. La velocidad de difusión es óptima, pero el escaso agrupamiento local resta consistencia a las redes. Nos lleva a pensar que la rápida coordinación y ejecución restringen el valor de las relaciones  $C/C_{random}$ . Su escasa duración impide que los individuos desarrollen la confianza necesaria para formar “cliques” o grupos afines. Existen dos excepciones, `#desalojosol` y `#planderescateciudadano`. En el caso de `#desalojosol`, dos factores pudieron propiciar un número mayor de agrupaciones locales: la repercusión en los medios, que hizo que la actividad se mantuviera activa durante el día siguiente, y el hecho de que se compartiese el espacio físico. En cuanto a `#planderescateciudadano`, la idea surgió de una campaña llevada a cabo por varios colectivos formados con anterioridad, luego es de suponer que se llegase a un grado alto de *clustering* rápidamente.
- Tipo B (negro): los valores de  $C/C_{random}$  son notablemente mayores que en el tipo A y sus distancias promedio relativas son altas. Su mayor presencia en el tiempo favorece una construcción más estructurada, pero la información, en general, se transporta menos eficazmente. Los dos procesos más representativos del tipo B, `#huelgadeclase` y `#22m`, identificadores de

<sup>1</sup>Por ejemplo, hay usuarios que envían *spam* y que incluyen *trending topics* para darles visibilidad. Estos nodos no estarán conectados al componente gigante.

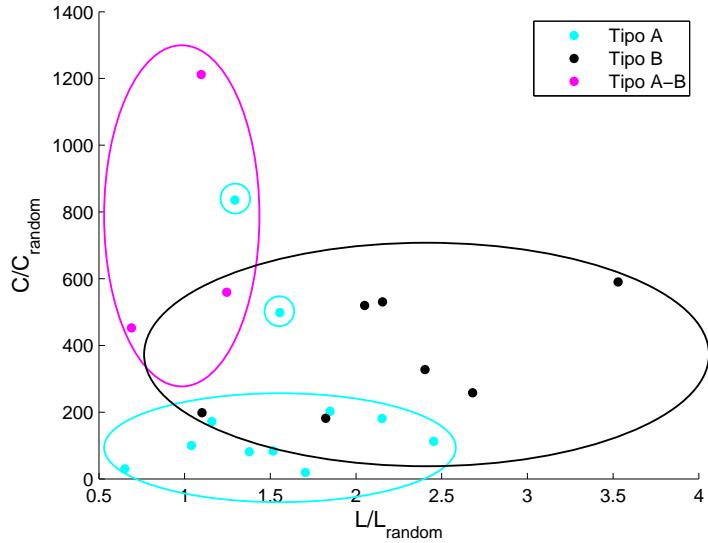


Figura 4.2:  $L/L_{random}$  frente a  $C/C_{random}$

la huelga de educación, y un tercero, #es15m, registran valores entre 2.5 y 3.5 veces mayores que los aleatorios. Aunque no llegan a ser de orden superior, la propagación por estos canales será más costosa que en las otras clases.

- **Tipo A-B (rosa):** llama la atención los valores de distancia promedio, similares a las aleatorias a pesar de ser redes de tamaños grandes (incluida la mayor de todas, #12m15m, con  $N = 32749$  nodos). La abundancia de caminos cortos surgidos de la interacción, desde un punto de vista global, denota una diferencia importante en cuanto a la participación de los individuos. Además, los valores de  $C$  son muy superiores a  $C_{random}$ . Son el tipo de red que mejor representa el concepto de mundo pequeño, permitiendo una propagación rápida y una estructura sólida, que pueden favorecer su permanencia en el tiempo como colectivo.

Se concluye que todos los sistemas poseen topología de mundo pequeño, pero los de tipo A-B poseen las mejores características en cuanto a velocidad de propagación y robustez en su configuración. Los nodos, a través de la interacción, crean canales de forma más variada, reduciendo la longitud de los caminos cortos. Los procesos A no consiguen, en promedio, construir redes interactivas tan sólidas como los otros. Por otra parte, los procesos B requieren más intermediarios para difundir mensajes que lleguen a la totalidad de individuos, que ven su participación más restringida que en los demás sistemas.

#### 4.2.2. Distribución de grados

La distribución según una ley de potencias  $P(k) \sim ck^{-\gamma}$ , propia de las redes libres de escala, indica la tendencia de un vértice nuevo a conectarse en primer lugar a individuos bien comunicados. Esta **conexión preferencial** hace disminuir la distancia promedio, y son una forma de alcanzar la estructura de mundo pequeño que encontramos en la mayoría de ejemplos en el mundo real.

En primer lugar, hay que aclarar lo que significa una arista entrante y saliente en nuestras redes. Una arista saliente de un nodo  $A$  y entrante en  $B$  significa que  $A$  ha publicado un mensaje en Twitter, que, o bien menciona (o responde) a  $B$ , o bien está reenviando algo que  $B$  había escrito previamente. No hay que confundir con la transmisión de información. Un *tweet* de  $A$  puede ser leído por todos sus seguidores directamente, y, de forma indirecta, por cualquiera que lo encuentre mediante una búsqueda, pero esto no implica la creación de enlaces.

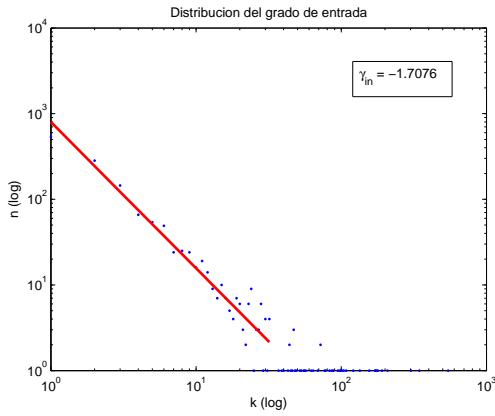


Figura 4.3: Distribución del grado de entrada. Red: #es15m.

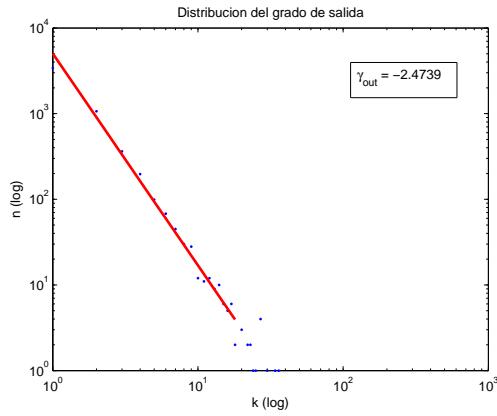


Figura 4.4: Distribución del grado de salida. Red: #es15m.

Un valor alto del grado conlleva un significado para el vértice:

- Grado elevado de salida: significa que ha interaccionado de las formas mencionadas con un alto número de usuarios, lo que le convierte en una “fuente”, que emite mensajes en muchas direcciones.
- Grado elevado de entrada: ha sido muy mencionado o muchos han retransmitido sus mensajes, actuaría como “referencia” en el proceso.
- Ambos grados con valores elevados: la suma de los roles anteriores convertiría a ese nodo en una “autoridad” dentro del proceso de comunicación asociado a ese *hashtag*.

Se carga la distribución de grados entrantes y salientes en Matlab. Se representan en escala logarítmica y se aplica un ajuste polinómico de primer grado. Las rectas se ajustan aproximadamente a los puntos y el valor de  $\gamma$  no difiere mucho de los típicos ( $2 \leq \gamma \leq 3$ ), por lo que se corrobora que son redes libres de escala. Un ejemplo se recoge en las figuras 4.3 y 4.4.

En la totalidad de casos

$$|\gamma_{in}| < |\gamma_{out}|$$

Las rectas de salida son más pronunciadas que las de entrada. El grado saliente presenta en mayor medida esa tendencia de los nodos a vincularse con *hubs*, que hay menor número de ellos y que también hay menos posiciones intermedias. Dado que las aristas que salen son las únicas que conllevan intencionalidad, es lógico que reproduzcan mejor ese comportamiento.

Mediante la representación de la figura 4.5 se analizan las diferencias entre tipos de movimientos:

- Tipo A (azul):  $|\gamma_{in}|$  es más bajo, lo que indica que existen más nodos “referencia” que en los de tipo B pero menos destacados. Su naturaleza dificulta la aparición de vértices jerárquicamente importantes y estables, por eso los *hubs* de entrada aparecen más distribuidos. La menor de las pendientes corresponde con #alabolsa ( $\gamma_{in} = -1,3418$ ). El comportamiento de salida es similar al tipo B. Los valores más bajos corresponden a tres procesos que comparten dos características. Una de ellas es el carácter local, tanto geográficamente (#laRimaia, #asambleaFCom) como ideológicamente (#15mSectorRadical). La otra, un tamaño reducido (tres de las cuatro más pequeñas). Parece lógico que en situaciones así haya una mayor distribución de nodos “fuente”.
- Tipo B (negro): la diferencia se observa en  $|\gamma_{in}|$ , con valores mayores que en el tipo A. Presentan una cantidad menor de vértices “referencia”. Dicho de otro modo, en comparación

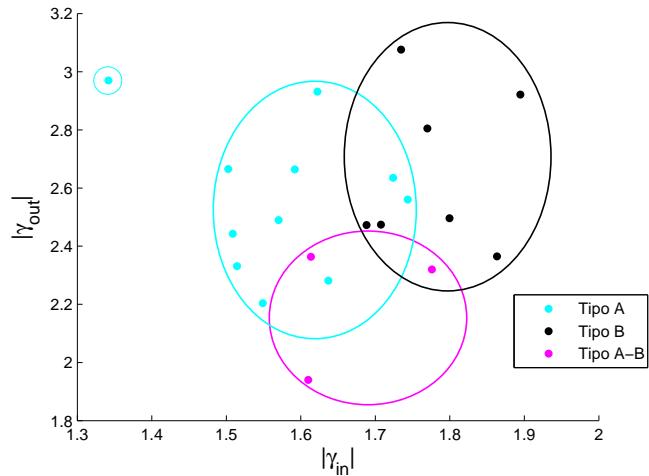


Figura 4.5: Valores absolutos de las pendientes  $\gamma_{in}$  y  $\gamma_{out}$

con procesos A, la recepción de enlaces está más polarizada. Incluso se podría hablar de una jerarquía más acusada. Su naturaleza rígida y estructurada obstaculiza la aparición de otras “referencias” aparte de las que el propio proceso pueda favorecer. El valor promedio de  $|\gamma_{out}|$  es el máximo entre las tres clases, luego los nodos “fuente” interactúan de forma menos variada.

- Tipo A-B (rosa): es significativo que tres redes de tamaños tan superiores (1<sup>a</sup>, 3<sup>a</sup> y 7<sup>a</sup> por orden de número de nodos) presenten valores bajos de  $|\gamma_{out}|$ . Una mayor cantidad de nodos “fuente” intermedios indica que los sistemas acogen mejor la participación y ésta se lleva a cabo de forma más distribuida. No se pueden extraer muchas conclusiones de las posiciones de  $|\gamma_{in}|$  por falta de más ejemplos, pero se intuye que la tendencia del grado entrante se asemeja más al tipo A. En tal caso, denotaría mayor cantidad de *hubs* de entrada, o lo que es lo mismo, una menor propensión a la exclusividad de los nodos “referencia”.

La distribución de ley de potencias se observa siempre, más marcada en el grado saliente por ser intencionado. Los tipos A e intuitivamente A-B, contienen mayor número de “referencias” que el tipo B, de lo que se extrae que la jerarquía está más repartida en las primeras. La existencia de *hubs* de salida, por contra, es similar para procesos A y B y menor que en la categoría A-B, cuyos individuos diversifican en mayor medida su participación.

### 4.3. Análisis de comunidades y mesoescala

Los sistemas suelen presentar divisiones en módulos de componentes, aunque a simple vista no se aprecie. En particular, interesa comprobar la presencia de mesoescala en un sistema, pues implicaría que la actividad del mismo se realiza a diferentes escalas. La ventaja de esta característica radica en que **existe actividad significativa en cada escala del sistema complejo**. Si el nivel macroscópico domina la interacción, los componentes por sí solos son incapaces de intervenir de manera significativa en el proceso. Por contra, si domina la microescala, habrá multitud de participantes comportándose de forma caótica y formarán un sistema irracional. La mesoescala enlazaría la actividad a los niveles macro y micro regulando y equilibrando el comportamiento del sistema [35]. La coordinación de la actividad a nivel de mesoscales es crucial para la auto-organización de sistemas complejos. Un ejemplo destacado de este fenómeno es la actividad en el cerebro, en el que asambleas neuronales a diferentes escalas se asocian para producir actividad

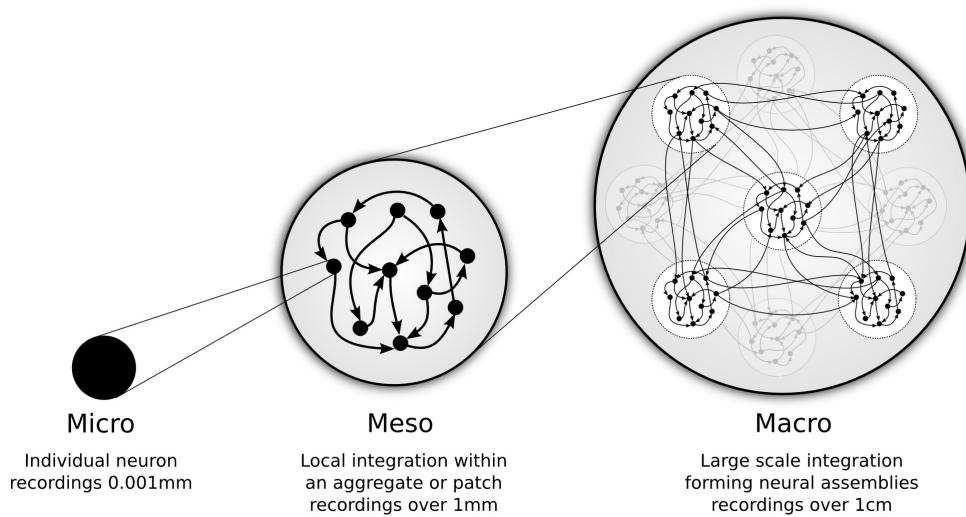


Figura 4.6: Mesoescalas en la actividad neuronal. Fuente: [36].

significativa (ver figura 4.6).

Este estudio es una aproximación preliminar al análisis de mesoescala. Los algoritmos que resuelven el problema con precisión requieren tiempos de computación muy elevados, por lo que se ha decidido utilizar el método descrito en la sección 3.4 para minimizar el tiempo de ejecución a cambio de perder algo de precisión en los cálculos. Hemos escogido una red representativa de cada tipo para estudiar su división en comunidades y la estabilidad de ésta a través de la mesoescala (figuras 4.7, 4.8 y 4.9). Las líneas continuas corresponden a  $Q > 0,3$ , el límite de fiabilidad de las particiones [23]. La línea discontinua vertical marca la posición de  $r = 0$ . En el eje  $x$  se representa la diferencia entre la resistencia y su valor mínimo, para evitar valores negativos que imposibilitarían la transformación logarítmica.

- Tipo A. #desalojoSol ( $N = 3337$ ): presenta escalones al inicio, de longitudes (en  $r$ ) de 0.309 (37 comunidades) y 0.378 (279 comunidades). Tiene cierto grado de mesoescala pero con divisiones muy poco estables. A partir de ahí, no existe una partición razonablemente constante hasta  $r - r_{asymp} \simeq 3$ , que cuenta con 1167 comunidades y que se mantiene durante un gran intervalo. En esta última división hay muchos individuos que se encuentran en comunidades muy pequeñas. En concreto, más de un tercio de la red recae en grupos de menos de 10 componentes.
- Tipo B. #huelgadeclase ( $N = 10018$ ): la primera división fiable comienza con 1891 módulos (comienzo de la línea continua), sin llegar a estabilizarse. La primera vez que parece asentarse corresponde a  $r = 6,61$ , identificando 3678 comunidades. La mitad de los nodos yacen en comunidades de tamaño menor que 10. Demasiados grupos pequeños de elementos están aislados y no pueden influir de manera significativa en el proceso.
- Tipo A-B. #15mPaRato ( $N = 6994$ ): el sistema presenta en mayor media que los anteriores el nivel de mesoescala. Se detecta una primera separación entre 37 y 38 comunidades que se mantiene un intervalo en  $r$  de 1.337. Despues de una acusada subida pronto se vuelve a obtener otra partición que permanece aproximadamente constante en  $1,9 < r - r_{asymp} < 4,5$  con 535 y 609 agrupaciones. Tras unos cortos intervalos en forma de escalera, la función vuelve a asentarse, esta vez ya con 3575 módulos. De esta manera, hay tres divisiones consistentes, lo que implica que hay comunidades de distintos tamaños que interaccionan entre sí en tres escalas diferentes.

El sistema de tipo A demuestra la existencia de particiones en comunidades poco robustas, ya que pequeñas variaciones de  $r$  alteran rápidamente el resultado. El primer periodo estable contiene demasiados grupos y muchos de ellos excesivamente reducidos y aislados, lo que se puede achacar a una organización rápida que no logra plasmarse en subestructuras consistentes. El ejemplo del tipo B se comporta de forma distinta, puesto que las particiones no son fiables hasta cerca de 1000 comunidades. La limitada interacción de los participantes impide que se densifiquen los enlaces dentro de grupos, y que, una vez identificadas de forma correcta, el resultado sea también multitud de participantes solitarios. #15mPaRato sí que presenta separaciones de módulos fuertes que pueden interactuar a varios niveles, algo que dinamiza la existencia del colectivo, al distribuir su actividad entre diferentes escalas del sistema.

#### 4.4. Análisis dinámico

Los sistemas complejos ofrecen muchas opciones de análisis más allá de la topología. Por medio del estudio de variables dinámicas se pueden poner de manifiesto propiedades emergentes de los sistemas que a través del análisis estructural no podríamos detectar.

Se han analizado las correlaciones en frecuencia de los procesos, entendidos como unidades dinámicas compuestas por los nodos participantes. El objetivo es caracterizar la composición y dinámica de cada proceso mediante el análisis de ruido  $1/f^\beta$ .

Mediante la técnica descrita en la sección 3.5, se determina la autocorrelación de procesos tipo  $1/f^\beta$ . Se ha utilizado la actividad de cada *hashtag* como datos para el análisis, medida en *tweets* por segundo. Los datos han sido preprocesados utilizando un filtro Butterworth pasa bajo, con frecuencia de corte de 0.06 Hz, para eliminar ruido a altas frecuencias. Una vez obtenidos los resultados, se clasifican los procesos en ruido blanco, ruido marrón y ruido rosa (ver sección 2.2.3).

La tabla 4.2 recoge los coeficientes  $\beta$  calculados, separados por tipo de ruido, junto con el tiempo de correlación entre escalas, medido en horas. Este último indica el tiempo que el proceso se mantiene como un todo coherente, que no es el mismo que la duración de la actividad del *hashtag*.

La correspondencia entre los tipos de proceso observados bajo nuestro criterio y tipos según el análisis dinámico es:

- Ruido blanco ( $\beta < 1$ ), correspondientes al tipo A: en estos procesos comunicativos dominan las escalas temporales cortas y son volátiles como colectivo. Se puede comprobar en cierto modo un carácter irracional e impulsivo. Su organización es rápida e impide asentar unas bases comunicativas entre los participantes. Al acercarse a  $\beta = 1$  las diferentes escalas temporales parecen incrementar su correlación.
- Ruido rosa ( $\beta \sim 1$ ), correspondientes al tipo A-B: a efectos prácticos, se considerará ruido rosa cuando  $0,8 < \beta < 1,2$ . Mantienen una identidad más duradera a lo largo del tiempo. El equilibrio que se produce entre la interacción a frecuencias cortas, medias y largas provoca que la correlación entre escalas alcance los valores más altos del análisis. Su capacidad de acción a largo plazo es buena, pero además es capaz de adaptarse al entorno fácilmente cuando factores externos influyen en el proceso.
- Ruido marrón ( $\beta > 1$ ), correspondientes al tipo B: presentan tiempos de correlación bajos, que de nuevo disminuyen al alejarse de valores de ruido rosa. Es un buen ejemplo para comprobar que la correlación entre escalas no tiene nada que ver con la duración de la actividad del hashtag. Se desarrollaron en un periodo bastante largo de tiempo y, sin embargo, no demuestran apenas afinidad entre la actividad a distintas frecuencias.

En la gráfica 4.10 se muestra cómo se ajusta la distribución del tiempo de correlación a una función gaussiana, con el pico en torno a los procesos de ruido rosa. La correlación disminuye conforme una de las escalas temporales cortas o largas ejercen mayor influencia.

Hay dos procesos cuya dinámica se opone al tipo que se le había asignado:

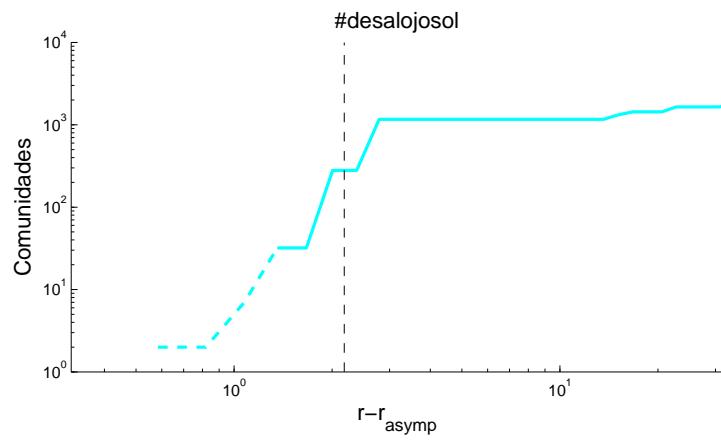


Figura 4.7: Análisis de mesoescala: `#desalojoSol`

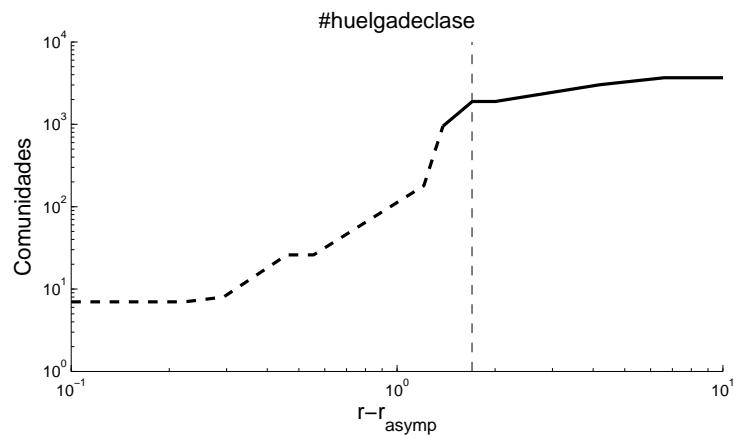


Figura 4.8: Análisis de mesoescala: `#huelgadeclase`

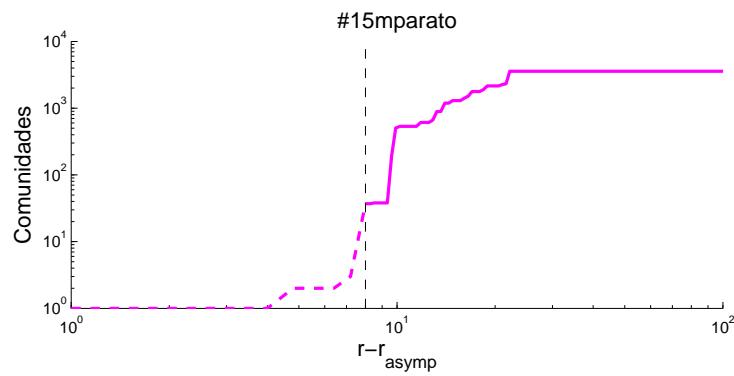


Figura 4.9: Análisis de mesoescala: `#15mparato`

- **#prima500**: a pesar de que la etiqueta se activó cuando la prima de riesgo superó los 500 puntos, los usuarios estaban esperando a que el momento llegara para lanzar sus mensajes, luego en realidad no fue un evento tan sorprendente como creíamos.
- **#cierrabankia**: es una campaña creada por la PAH<sup>2</sup> basada en acciones contra Bankia y, por los datos obtenidos, parece ser que mantuvo un equilibrio entre escalas temporales mayor del que habíamos estimado al clasificarlo como tipo A.

En definitiva, el balance entre la participación a frecuencias bajas, medias y altas sólo se aprecia en procesos de tipo A-B, que coinciden con valores de  $\beta \sim 1$ . Los sistemas se adaptan bien a cambios externos y son capaces de llevar a cabo acciones rápidas, mientras que construyen una identidad colectiva de duración muy superior al resto. Los otros procesos, ya sea por un desarrollo apresurado y escasamente organizado (tipo A) o por una excesiva rigidez que resta variedad e importancia a la participación (tipo B), existen como agrupación de individuos coherente durante un tiempo mucho menor.

---

<sup>2</sup>Plataforma de Afectados por la Hipoteca.

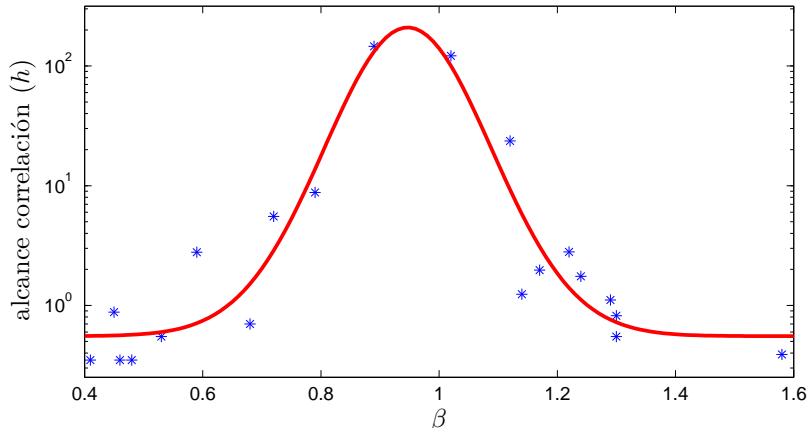


Figura 4.10: Alcance de la correlación frente a  $\beta$

Hashtag	$\beta$	Ruido	Correlación (horas)	Tipo
#alabolsa	0.41	Blanco	0.35	A
#nuremberfinanciero	0.45	Blanco	0.88	A
#desalojoSol	0.46	Blanco	0.35	A
#15mSectorRadical	0.48	Blanco	0.35	A
#laRimaia	0.53	Blanco	0.55	A
#bankiaeasnuestra	0.59	Blanco	2.78	A
#asambleafcom	0.68	Blanco	0.70	A
#planderescateciudadano	0.72	Blanco	5.54	A
#16m	0.79	Blanco	8.78	A
#12m15m	0.89	Rosa	145.78	A-B
#15mparato	1.02	Rosa	121.5	A-B
#lacaixaesmordor	1.12	Rosa	23.64	A-B
#occupymordor				
#alaplaza12m	1.14	Rosa	1.24	B
#Feliz15m	1.17	Rosa	1.97	B
#yovoy12m	1.22	Marrón	2.80	B
#es15m	1.24	Marrón	1.75	B
#cierrabankia	1.29	Marrón	1.11	A
#prima500	1.30	Marrón	0.82	A
#22m	1.30	Marrón	0.55	B
#huelgadeclase	1.58	Marrón	0.39	B

Tabla 4.2: Datos del estudio dinámico.



# Capítulo 5

## Conclusiones

### 5.1. Discusión

Nuestra hipótesis inicial sugería el factor tecnológico como elemento clave en las movilizaciones sociales, en particular las relacionadas con el movimiento 15M. La validez del estudio pasa por comprobar la relación directa que tiene la actividad en la red con lo que sucede en el mundo físico. De este modo, se podrían conectar las propiedades analizadas en la red social con las movilizaciones reales.

La categorización *a priori* de los movimientos bajo nuestro punto de vista coincide en gran medida con sus correspondientes procesos comunicativos en red. A la vista de los resultados obtenidos, deducimos que la auto-organización de los sistemas considerados en el proyecto comparten características estructurales de forma general. En primer lugar, en todos los sistemas se repite la topología de mundo pequeño y libre de escala, lo que les confiere velocidades de propagación altas y un agrupamiento denso a nivel local. La existencia de *hubs*, tanto de entrada como de salida, indica que las redes no son horizontales. Sin embargo, se perciben diferencias entre movimientos que hacen pensar que no todos se traducen finalmente en el mismo tipo de redes. Los procesos están más abiertos a la interacción distribuida cuando no están excesivamente estructurados. Del mismo modo, cuando la participación sí es distribuida, se ha de complementar con cierta organización, algo que es difícil de encontrar en los movimientos rápidamente coordinados y de carácter impulsivo. Los que catalogamos como de tipo A-B representan más fielmente la topología de mundo pequeño, y su distribución de grado de salida y la presencia de mesoescala indican que la interacción se diversifica más y que se desarrolla a varios niveles, actuando como elemento de cohesión en los sistemas.

Por otro lado, el análisis dinámico revela notables diferencias entre procesos. La dinámica de cada red corrobora esa cohesión por medio de la participación activa que sugiere la caracterización estructural. Los procesos de ruido rosa se mantienen como una identidad coherente durante más tiempo, gracias a una actividad distribuida equilibradamente entre escalas temporales cortas, medianas y largas. Al alejarse hacia zonas de ruido blanco o marrón, el tiempo de correlación entre ritmos de actuación desciende drásticamente y, aunque los procesos permanezcan activos, no se constituyen como un colectivo coherente.

Por consiguiente, la auto-organización de movimientos en red no siempre tiene la misma repercusión. La forma en que se desarrolle el proceso influirá en la estabilidad de los individuos como grupo conectado. En concreto, la interacción entre los nodos (tanto en cantidad como en diversidad) y las comunidades es clave. Podemos diferenciar entre comunidades de participación abierta y espontánea, de actividad más impredecible y caótica, y comunidades cerradas y estructuradas en la que la actividad es más previsible y coherente. Las primeras demuestran más creatividad pero son incapaces de mantener su actividad a lo largo del tiempo, mientras que las segundas son más estables pero no consiguen adaptarse a los cambios en su entorno. Los resultados obtenidos sugieren

que la actividad en las escalas temporales y espaciales intermedias son elementos fundamentales para combinar las ventajas de cada uno de estos tipos de organización.

En resumen, concluimos que las dinámicas observadas en el mundo físico y en la red son similares. A través del análisis de un proceso comunicativo en la red, hemos podido caracterizar la actividad de los colectivos asociados a esos procesos e interpretar cómo se organizan y actúan. Por tanto, la tecnología actúa como vehículo para la generación de estados de ánimo comunitarios y potencia su forma de expresarse.

## 5.2. Aplicaciones

El proceso de aprendizaje continuo que se da en torno a la relación entre movimientos sociales y tecnología promueve nuevas formas de análisis. Este trabajo puede servir como medio para evaluar el funcionamiento de redes y las estrategias que se siguen dentro de ellas. Como los sistemas objetivo no son intuitivos, la aplicación de estudios de este tipo puede ser provechosa desde el punto de vista organizativo. La presencia de mesoescala y ruido rosa sugieren que los procesos que funcionan no se controlan de arriba a abajo ni emergen espontáneamente de la congregación de individuos, sino que es necesario un puente entre la actividad a distintos niveles. Los movimientos sociales representan un caso interesante de estudio en este campo, pero otros muchos procesos se valen de las redes sociales para su desarrollo, como puede ser el marketing 2.0, aspectos relacionados con los estudios de mercado, reputación *online* y posicionamiento *web*, etc.

Por otro lado, la aplicación de los procedimientos de este documento en tiempo real puede proporcionar indicadores que monitorizan la evolución de los sistemas, y, en el caso de las movilizaciones sociales, se obtendría información valiosa, como por ejemplo la efectividad del sistema como colectivo.

## 5.3. Trabajo futuro

Se plantean líneas de estudio futuras que pueden seguir lo iniciado en este trabajo. Lo propuesto sería interesante tanto para los movimientos analizados como para futuras movilizaciones.

Una opción sería separar las redes por tipo de actividad (*retweet* y *reply/mention*) para matizar más el significado de los enlaces, así como comparar éstas con las redes de seguimiento. Un planteamiento muy ambicioso incluye el estudio conjunto de distintas capas de actividad (distintas redes en Internet y el espacio físico) que pueda poner de manifiesto relaciones entre ellas. Asimismo, sustituir el componente gigante débilmente conexo por el fuertemente conexo permitiría estudiar los núcleos más centrales de la red.

El análisis de mesoescala se puede mejorar a través de algoritmos que, aunque mucho más costosos en tiempo de computación, proporcionan divisiones más precisas. La parte dinámica ofrece multitud de oportunidades interesantes para el estudio de sistemas complejos. Además de añadir el factor tiempo al análisis estructural [37], la visión de estos movimientos como una mente colectiva habilita conexiones de trabajos de este tipo con los sistemas cognitivos.

Evaluar las dinámicas individuales para relacionarlas con la global sería otro aspecto a considerar.

# Bibliografía

- [1] Toret, J. "Una mirada tecnopolítica sobre los primeros días del 15M", en "Democracia Distribuida", Universidad Nómada (2012).
- [2] GlobaliseThis. "New media for peace", <http://globalisethis.wordpress.com/2009/10/06/1999-new-media-revolution-in-seattle/> (2009).
- [3] Salido, N. "Del 11M al 14M: Jornadas de móvil – ización social", ISBN 84-313-2374-4, págs. 271-284 (2006).
- [4] Fernández-Savater, A. "¿Cómo se organiza un clima?", <http://blogs.publico.es/fueradelugar/1438/%C2%BFcomo-se-organiza-un-clima> (2012).
- [5] Malo, M. y Pérez, D. "Latidos. O del 15M y la revuelta", <http://madrilonia.org/2012/02/latidos-o-del-15m-y-la-revuelta/> (2012).
- [6] Levi S., "Notas para una r-evolución 2. Segunda fase: vicios vs. nuevas virtudes tácticas.", <http://conservas.tk/tacticas/> ,(2012).
- [7] Jansen, B. J., Zhang, M. M., Sobel, K., y Chowdury, A. "Twitter Power: Tweets as Electronic Word of Mouth", Journal of the American Society for Information Science and Technology 60(11), 2169-2188 (2009).
- [8] Kwak, H., Lee, C., Park, H., y Moon, S. "What is Twitter, a social network or a news media?", artículo presentado en el 19th International Conference on World Wide Web (2010).
- [9] Rey, P., "Front page newspaper analysis", <http://numeroteca.org/cat/frontpage-newspaper/> (2011-2012).
- [10] Solé, R., "Redes complejas", Tusquets (2009).
- [11] Bar-Yam, Y. "Dynamics of complex systems", ISBN 0813341213 The Advanced Book Studies in Nonlinearity series, Westview Press (2003).
- [12] Marashi, E. y Davis, J. P. "A Systems Approach for Resolving Complex Issues in a Design Process", Computer-Aided Civil and Infrastructure Engineering, 22 (7), pp. 511-526 (2007).
- [13] Livingston-Gray, S. "Directed Versus Undirected Graphs". <http://think-like-a-git.net/sections/graph-theory/directed-versus-undirected-graphs.html> (2011).
- [14] The iGraph Library, <http://igraph.sourceforge.net/screenshots2.html>.
- [15] Erdős, P. y Rényi, A. "On the evolution of random graphs". Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5: 17–61 (1960).

- [16] Watts, D. J. and Strogatz, S. H. “Collective dynamics of ‘small-world’ networks”. *Nature* 393, 440–442 (1998).
- [17] Amaral, L. A. N., Scala, A. Barthélémy, M. Stanley, H. E. “Classes of small-world networks”. *PNAS*, September 26, 2000.
- [18] Albert, R., H. Jeong and A.-L. Barabasi. “Diameter of the World-Wide Web”. *Nature* 401, 130 (1999).
- [19] Faloutsos, M., Faloutsos P. and Faloutsos, C. *Proc. ACM SIGCOMM, Comput. Commun. Rev.* 29, 251 (1999).
- [20] Jeong, H., Tombor, B., Albert, Z. N., Oltvai, Z. N. y Barabási, A.-L. “The large-scale organization of metabolic networks”, *Nature* 407, 651 (2000).
- [21] Borge-Holthoefer, J., Rivero, A., García, I., Cauhé, E., Ferrer, A., *et al.* “Structural and Dynamical Patterns on Online Social Networks: The Spanish May 15th Movement as a Case Study”. *PLoS ONE* 6(8): e23883. doi:10.1371/journal.pone.0023883 (2011).
- [22] Newman, M. E. J., “Finding and evaluating community structure in networks”, *Phys. Rev. E* 69, 026113 (2004).
- [23] Newman, M. E. J., “Analysis of weighted networks”, *Phys. Rev. E* 70 056131 (2004).
- [24] Girvan, M. y Newman, M. E. J., “Community structure in social and biological networks”, *Proc. Natl. Acad. Sci. USA* 99 7821 (2002).
- [25] Radicchi, F., Castellano, C., Cecconi, F., Loreto, V. y Parisi, D. “Self-contained algorithms to detect communities in networks”. *Proc. Natl. Acad. Sci. USA* 101 2658 (2004).
- [26] Pons, P. y Latapy, M. “Computing communities in large networks using random walks”. *Journal of Graph Algorithms and Applications* (2006).
- [27] Clauset, A., Newman, M. E. J. y Moore, C. “Finding community structure in very large networks”. *Phys. Rev. E* 70, 066111 (2004).
- [28] Fortunato, S. y Barthelemy, M. “Resolution limit in community detection”. 36–41 *PNAS* January 2, 2007 vol. 104 no. 1.
- [29] Arenas, A., Fernandez, A. y Gomez, S. “Analysis of the structure of complex networks at different resolution levels”. *New Journal of Physics* 10 (2008).
- [30] Bak, P., Tang, C. and Wiesenfeld, K. “Self-organized criticality: an explanation of 1/f noise”. *Physical Review Letters* 59 (4): 381–384 (1987).
- [31] <https://scraperwiki.com/>
- [32] <https://gephi.org/>
- [33] <http://deim.urv.cat/~sgomez/radatools.php>
- [34] Newman, M. E. J. “Fast algorithm for detecting community structure in networks” *Physical Review E* 69 (2004).
- [35] Freeman, W. “How Brains Make Up Their Minds”. Columbia University Press (2000).
- [36] Barandiaran, X. (2008) “Mental Life. A naturalized approach to the autonomy of cognitive agents”, PhD Thesis Dissertation, (2008).
- [37] Holme, P. y Saramäki, J. “Temporal Networks”, *Phys. Rep.* 519, 97-125 (2012).