

Xarxes socials

Xarxes socials

**Agustí Canals
Albert Díaz-Guilera
José Luis Molina
Beatriz Pataca**

L'encàrrec i la creació d'aquest material docent han estat coordinats pel professor:
Agustí Canals Parera (2012)

Primera edició: setembre 2012

© Agustí Canals, Albert Díaz-Guilera, José Luis Molina i Beatriz Pataca, del text.

Tots els drets reservats
© d'aquesta edició, FUOC, 2012
Av. Tibidabo, 39-43, 08035 Barcelona

Realització editorial: Editorial UOC
Dipòsit legal: B-21.893-2012

Els textos i imatges publicats en aquesta obra estan subjectes –llevat que s'indiqui el contrari– a una llicència de Reconeixement-NoComercial-SenseObraDerivada (BY-NC-ND) v.3.0 Espanya de Creative Commons. Podeu copiar-los, distribuir-los i transmetre'ls públicament sempre que en citeu l'autor i la font (FUOC. Fundació per a la Universitat Oberta de Catalunya), no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-ncnd/3.0/es/legalcode.ca>

Agustí Canals

Llicenciat i master en Ciències Físiques per la Universitat Autònoma de Barcelona i MBA i doctor en Management Sciences per ESADE (Universitat Ramon Llull). Ha realitzat estudis de postgrau a les universitats nord-americanes de Penn State i Syracuse i ha estat Postdoctoral Visiting Scholar al departament de Management de la Wharton School a la Universitat de Pennsylvania (EEUU) i Visiting Fellow de la Warwick Business School a la Universitat de Warwick (Regne Unit). Durant la seva vida professional ha compaginat l'activitat acadèmica amb la gestió empresarial i la consultoria en els àmbits d'organització, sistemes d'informació i gestió del coneixement. La seva recerca se centra en temes relacionats amb la gestió estratègica de la informació i el coneixement, les xarxes socials i la complexitat a les organitzacions. Actualment és professor agregat de la Universitat Oberta de Catalunya, on va fundar els Estudis de Ciències de la Informació i de la Comunicació, i també professor associat de management estratègic a ESADE. Ha publicat diversos articles especialitzats en revistes internacionals i el llibre "Gestión del Conocimiento" (Ediciones 2000, 2003).

Albert Díaz-Guilera

Llicenciat en Física per la Universitat de Barcelona. Doctor en Ciències per la Universitat Autònoma de Barcelona (1987). Estades postdoctorals en laboratoris Gorlaeus (Leiden, Països Baixos) i "Centre de Physique du Solide" (Sherbrooke, Canadà). La seva recerca se centra actualment en els aspectes generals de la complexitat, sobretot en xarxes complexes. Essent per formació físic estadístic, les seves línies de recerca s'han anat ampliant per abastar els aspectes en diferents camps: biologia, economia, ciències socials, ciències de la computació, la lingüística. Col·laboracions directes amb científics de diferents orígens han estat possible a través d'estades a diferents centres (Matemàtiques de l'Imperial College de Londres, Enginyeria Química i Biològica de la Universitat Northwestern, Ecologia UNAM, Institut de Potsdam de Climatogoy, Potsdam Psicologia, Sociologia de la ETHZ). Autor de prop d'un centenar d'articles en revistes de la física i interdisciplinars. Ha fet més d'un centenar de xerrades en congressos i centres de recerca. Líder del grup de recerca PHYSCOMP2, investigador principal de projectes de governs català i espanyol i de la Unió Europea. Coordinador de la xarxa espanyola "Econosociofísica: Dinàmica i Fenòmens Col·lectius de Sistemes socioeconòmics".

José Luis Molina

Professor titular del Departament d'Antropologia social i cultural de la Universitat Autònoma de Barcelona. És director del grup d'investigació egoLab-GRAFO (HYPERLINK "<http://www.egolab.cat>") www.egolab.cat). Desenvolupa la seva investigació en mètodes i tècniques d'anàlisi de les xarxes personals i les seves aplicacions a l'estudi de fenòmens transnacionals, la migració i els negocis ètnics, entre d'altres.

Beatriz Pataca

Doctoranda del Departament d'Antropologia social i cultural de la Universitat Autònoma de Barcelona. Realitza la seva investigació sobre les xarxes personals de dones immigrades. Actualment és consultora a la UOC i ofereix assessories en organismes públics i privats.

Índex

Capítol I. Introducció a les xarxes	11
(Albert Díaz-Guilera)	
1.1. Introducció a les xarxes	11
1.1.1. La importància de les xarxes o “com Kevin Bacon ha ajudat a curar el càncer”	11
1.1.2. Un món complex	15
1.1.3. Connexions	17
1.2. Xarxes per tot arreu	19
1.2.1. Xarxes tecnològiques	19
1.2.2. Xarxes biològiques	23
1.2.3. Xarxes socials	26
1.2.4. Xarxes d’informació	27
1.3. Què és una xarxa?	32
1.3.1. Nodes i enllaços	32
1.3.2. Tius de xarxes	33
1.3.3. Descripcions.....	35
1.4. Caracterització dels nodes d’una xarxa	36
1.4.1. Grau	37
1.4.2. Camins (<i>paths</i>)	39
1.4.3. Longitud	39
1.4.4. Distància entre nodes	39
1.4.5. <i>Closeness centrality</i>	40
1.4.6. <i>Betweenness centrality</i>	40
1.4.7. <i>Eigenvector centrality</i>	40
1.4.8. <i>Page rank</i>	41
1.4.9. <i>Clustering</i> d’un node	42
1.5. L’estructura de les xarxes	43
1.5.1. Grandària d’una xarxa	43
1.5.2. <i>Clustering</i> d’una xarxa	44
1.5.3. Components	44
1.5.4. <i>Motifs</i>	45
1.5.5. Cliques	46
1.5.6. Comunitats	47

1.6. Universalitat	52
1.6.1. <i>Shortest paths</i> i <i>small worlds</i>	52
1.6.2. Distribució de grau	54
1.6.3. Lleis de potència i <i>Scale-Free Networks</i>	55
1.7. Modelització de les xarxes	58
1.7.1. Erdős-Rényi (ER)	59
1.7.2. Watts-Strogatz (WS)	63
1.7.3. Barabási-Albert (BA)	65
1.7.4. Altres models	67

Capítol II. Anàlisi de xarxes socials 69
(José Luis Molina i Beatriz Pataca)

2.1. Què són les xarxes socials?	70
2.1.2. Xarxes personals	71
2.1.3. Xarxes sociocèntriques o completes	73
2.2. Breu història de l'anàlisi de xarxes socials	74
2.2.1. La sociometria	75
2.2.2. Els investigadors de Harvard i Chicago	75
2.2.3. L'escola de Manchester	78
2.2.4. L'escola de Harvard	80
2.2.5. Camps d'aplicació	80
2.3 Six degrees y capital social	82
2.3.1. Sis graus de separació	82
2.3.2. Capital social	84
2.4. Social networking sites	86

Capítol III. Xarxes, informació i coneixement
(Agustí Canals)

3.1. Introducció	95
3.1.1. Sobre informació i coneixement	95
3.1.2. Xarxes d'informació i xarxes socials	99
3.2. Xarxes d'informació al World Wide Web	100
3.2.1. L'estructura de la World Wide Web	101
3.2.2. Cerca d'informació	107
3.2.3. Altres xarxes d'informació al Web	109

3.3. Difusió del coneixement en xarxes socials	113
3.3.1. Aproximació tradicional a la difusió social	114
3.3.2. Difusió d'informació en xarxa	116
3.3.3. De la difusió d'informació a la difusió de coneixement	118
3.3.4. Influència social	121
3.4. Xarxes de coneixement científic	122
3.4.1. La construcció del coneixement científic: xarxes de citacions	123
3.4.2. La col·laboració científica: xarxes de coautoría	126
3.4.3. L'estructura de la ciència: xarxes de paraules clau	127
3.4.4. Mapes de coneixement: visualització de les xarxes	130
3.5. Les xarxes a l'economia del coneixement	132
3.5.1. El paper estratègic del coneixement	132
3.5.2. Les organitzacions i les xarxes	133
3.5.3. Mercats en xarxa	137
Bibliografia per capítols	141

Capítol I

Introducció a les xarxes

Albert Díaz-Guilera

1.1. Introducció a les xarxes

1.1.1. La importància de les xarxes o “Com Kevin Bacon ha ajudat a curar el càncer”

De tots és conegut aquest fenomen que anomenem de *món petit*. Quantes vegades no ens hem trobat un desconegut o una desconeguda i després de deu minuts de conversa ens en hem adonat que és cosí d'un amic del nostre germà.

Aquest fenomen ha estat portat fins i tot al món de l'espectacle. John Guare al començament de la dècada de 1990 va escriure una obra de teatre titulada *Six degrees of separation* que es va representar durant un temps a Broadway i que més endavant es va adaptar al cinema amb actors de prestigi reconegut com ara Donald Sutherland. Per primera vegada es presentava de manera popular un fet que els sociòlegs coneixien des de feia anys, de fet des de principis del segle XX.

Enllaços recomanats

Entrada a la Viquipèdia sobre l'obra de teatre:

http://en.wikipedia.org/wiki/Six_Degrees_of_Separation_%28play%29

Però va ser un psicòleg social, Milgram, qui, al final de la dècada de 1960, va fer un experiment i va analitzar de manera sistemàtica aquest fenomen. Sense donar-hi un nombre en particular, pensava en la possibilitat que entre qualsevol parell de persones hi hagués un petit nombre d'intermediaris que permetessin fer arribar la informació de l'una a l'altra.

Va buscar una sèrie de persones als Estats Units perquè enviessin una carta personal a un advocat que vivia a prop de Boston, amb la condició que cada carta fos enviada a una persona que coneguessin personalment i, un cop rebuda, aquesta persona l'havia d'enviar a una altra que també conegués personalment, i així successivament. El resultat, sorprenent, és que el valor mitjà del nombre de passos requerits per les cartes (sobre aquelles que realment van arribar) va ser de sis. Així quedava demostrat que tothom als Estats Units podia estar unit a qualsevol altre per un petit nombre de sis passos intermedis. I aquests són els famosos "sis graus de separació". Una altra de les conclusions sorprenents de l'estudi va ser que la majoria de les cartes que van arribar tenien en comú els tres últims intermediaris. Això demostra, a més, com convergeixen trajectòries socials que, en principi, creiem que són absolutament independents. També permet visualitzar l'estructura social com un entramat de connexions i, per extensió, com un graf o xarxa en què els nodes són les persones i els enllaços, en aquest cas, relacions de coneixença personal.

Enllaços recomanats

La versió traduïda al castellà de l'article original es pot trobar en la revista *Araucaria* (2003, núm. 10, pàg. 15):

http://www.institucional.us.es/araucaria/nro10/ideas10_2.pdf

El grau màxim de popularització va arribar precisament amb Internet i la facilitat d'accedir a grans bases de dades, una de les quals anomenada *Internet Movie Database* <<http://www.imdb.com/>>. Actualment, conté més d'un milió i mig de pel·lícules i gairebé dos milions d'actors i actrius. Un informàtic a la Universitat de Virgínia havia estat calculant la cadena mínima de pel·lícules entre qualsevol parell d'actors. I es va adonar de l'estructura rica de la xarxa d'actors, de tal manera que gairebé el 90% dels actors estan connectats, inclòs en Kevin Bacon. Precisament, va ser aquest nom el que va donar lloc a l'anomenat *Oracle of Bacon*. Si per curiositat entrem en aquest web actualitzat cada dues setmanes a partir de les dades de la IMDB, podem trobar el nombre de pel·lícules entre qualsevol parell d'actors, i per a qualsevol parell que triem ens adonarem de com de petit és aquest món. I no sols els més coneguts; podem trobar, a més, a quina distància, en mitjana, es troba qualsevol actor de la resta de la base de dades. I els nombres són inesperadament, o potser ja no tant, petits. El nombre de pel·lícules entre qualsevol parell d'actors es troba per sota de quatre.

Enllaços recomanatsInternet Movie DataBase: <http://www.imdb.com/>**Enllaços recomanats**The Oracle of Bacon: <http://oracleofbacon.org/>

Però altres comunitats també tenen les seves pròpies interpretacions de com de petit és el món professional. Els matemàtics utilitzen el que s'anomena nombre d'Erdős. Aquest matemàtic hongarès, el més prolífic de la història a l'hora d'escriure articles, va col·laborar directament amb més de 500 científics. Aquests tenen l'honor de tenir un *nombre d'Erdős* igual a 1. Els coautors d'aquests tenen el nombre 2, i així successivament. I no cal dir que amb pocs nombres podem descriure completament tota la comunitat de matemàtics del món, llevat, potser, d'algun cas aïllat.

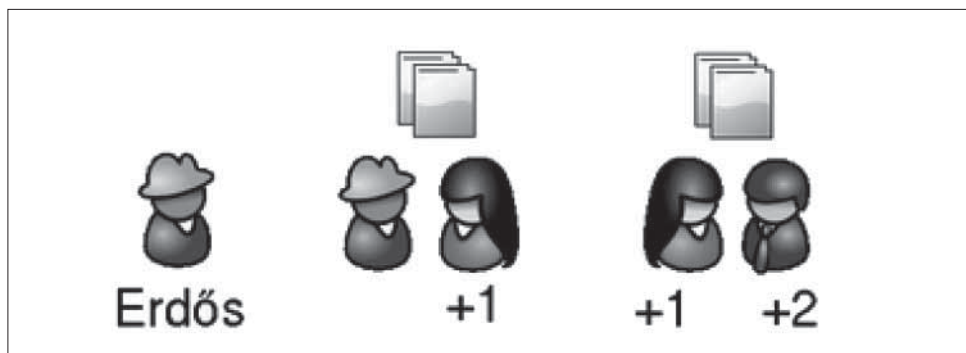
Enllaços recomanatsEl nombre d'Erdős: http://en.wikipedia.org/wiki/Erdős_number

Figura 1.1. Si l'Àlicia col·labora amb el Paul Erdős en un article i amb en Bernat en un altre, però en Bernat no col·labora mai amb l'Erdős, aleshores en Bernat té un nombre d'Erdős 2, ja que és a dues passes d'Erdős.

Aquestes proves confirmen, doncs, la sospita que en la nostra societat tots ens trobem molt més a prop del que en principi sospitem. No solament els actors i els matemàtics, ja que el *New York Times* va proposar l'any 1995, en ple escàndol Lewinsky, a quina distància es trobava qualsevol nord-americà de la famosa becària de la Casa Blanca i, novament, el resultat va ser sorprenentment curt.

Però ara hem de fer una mirada més enllà de les relacions professionals i socials. Hi ha altres sistemes que responen també a aquesta arquitectura? A estructures creades per la natura? A estructures creades per l'ésser humà?

L'any 1998, Duncan Watts i Steven Strogatz publiquen un article a *Nature* que avui en dia tothom considera com el primer de la nova ciència de les xarxes complexes "Collective dynamics of 'small-world' networks". En aquest article analitzen acuradament certes estructures: unes creades per la natura, com ara la xarxa neuronal d'un cuc, anomenat *C. Elegans*; d'altres creades per l'ésser humà, com la xarxa de distribució d'electricitat als Estats Units, i l'altra ja esmentada, creada per les relacions professionals, dels actors i actrius de cinema. I totes comparteixen els fets que hem comentat amb anterioritat, que la distància mitjana entre qualsevol parell de neurones, d'estacions de generació o distribució o qualsevol parell d'actors és molt més petita que l'esperada.

En aquest article no solament es posen de manifest aquestes característiques del que avui en dia anomenem un món petit, sinó que, a més, es proposa un model matemàtic que les explica, i de les quals tractarem amb més detall més endavant.

Enllaços recomanats

L'article pioner de Watts i Strogatz: doi:10.1038/30918

Així, doncs, a partir d'aquestes proves i del mateix model, i d'altres relacions, neix un nou camp de recerca, el de les xarxes complexes. Un món nou que descriu estructures a mig camí entre la regularitat i l'aleatorietat, que trenca la simplicitat topològica que tant agrada als matemàtics, i que troba aquesta simplicitat o bé en la regularitat perfecta o en l'aleatorietat total. I allunyat dels dos extrems simples, enmig, en un terreny inexplorat, ens trobem la complexitat estructural o topològica.

I com a demostració immediata i visual d'aquests fets podeu veure el vídeo anomenat *Connected-How Kevin Bacon cured cancer*. En primer lloc, en aquest documental podem veure com es produiria un experiment semblant al de Mil-

gram en el moment actual. Quina trajectòria seguiria un paquet postal enviat des de qualsevol lloc del món. I seguint amb el documental, com el descobriment inicial de l'entramat de connexions en una xarxa d'actors en què Kevin Bacon tenia un paper important ens ha portat amb el temps a aplicacions en molts camps de recerca, des de les telecomunicacions a l'economia, des de la producció d'energia a la distribució de comerç mundial, i en una d'aquestes aplicacions ens trobem amb la medicina. Com es poden conèixer les interaccions entre proteïnes, els efectes en xarxes gèniques d'expressió han permès començar a dissenyar fàrmacs nous que precisament poden atacar estructures patològiques associades, per exemple, al desenvolupament de cèl·lules cancerígenes, en els seus punts més febles, i ens poden ajudar en un futur no gaire llunyà a lluitar contra el càncer.

Vídeo recomanat

Connected - How Kevin Bacon cured cancer.

Produït per l'ABC, cadena australiana

1.1.2. Un món complex

Aquestes estructures que hem descrit en la secció anterior, que anomenem *estructures complexes* per distingir-les de la simplicitat de la regularitat o del caos organitzatiu més complet, són només una de les mostres que en el nostre entorn abunden les "complexitats". No solament pel tipus d'estructura, sinó també pel tipus de comportament.

Vídeo recomanat

Synchronized fireflies. <http://youtu.be/a-Vy7NZTGos>

Actualment, entenem per *sistemes complexos* els que presenten comportaments emergents en sistemes amb un gran nombre de components o parts. La particularitat dels comportaments col·lectius que apareixen és que no són fàcilment predictibles a partir de les característiques individuals. Cada unitat individual pot presentar una evolució particular, per exemple, una cuca de llum que emet polsos de llum amb una freqüència determinada. Globalment, podem dir que a les ribes dels rius del Sud-est asiàtic hi ha un fenomen en què

al capvespre es pot observar com milers de cuques de llum emeten flaixos de manera sincronitzada. Aquest és un fenomen de comportament emergent, de sistema complex, en què el sistema global podem dir que està sincronitzat, mentre que pels sistemes individuals només podem dir que emeten polsos de llum. Per això diem que el comportament global no és una mera suma dels individuals, una miríada de polsos inconnexos, sinó un conjunt de polsos sincronitzats. Què o qui fa que aquests polsos siguin sincronitzats? La resposta és molt més senzilla del que sembla: són les mateixes cuques de llum que “veuen” els polsos de les seves veïnes i en funció d’això decideixen avançar o endarrerir l’emissió del seu pols. I després d’una estona “comunicant-se” de manera imperceptible entre elles és quan tenen capacitat per a mostrar aquest efecte global. Aquesta mena de comportaments els podem observar a la natura de moltes maneres, com ara el vol de les bandades d’estornells o els bancs de peixos, o fins i tot quan l’audiència d’un concert aplaudeix de manera totalment sincronitzada després de l’espectacle.

Però aquest no és l’únic tipus de comportament emergent que podem trobar a la natura o a la nostra societat. És molt més general del que sembla i apareixen a totes les escales. Les societats són formades per éssers humans, el cervell per neurones, les molècules d’ADN per àtoms, el temps atmosfèric per fluxos d’aire i aigua, les estructures galàctiques a l’Univers...

Vídeo recomanat

TEDtalk, de Steven Strogatz, sobre sincronització (subtitulat en castellà).

http://www.ted.com/talks/steven_strogatz_on_sync.html

Aquests sistemes complexos els estudiem dins la ciència de la complexitat. Aquesta és una disciplina que va més enllà de la divisió tradicional de les disciplines i en pot afectar més d’una en una descripció a diferents nivells, com, per exemple, la psicologia i les ciències socials, quan intentem descriure un sistema social format per persones.

Però, com és fàcil d’imaginar, aquesta descripció qualitativa és lluny de poder ser formalitzada matemàticament, amb models senzills. Les dinàmiques individuals poden ser senzilles d’entendre (una cuca de llum es pot veure com un simple oscil·lador), però sobretot és la interacció entre les parts que normalment es produeix d’una manera no trivial. I, addicionalment, encara ens falta tenir en compte que les interaccions es produeixen mitjançant patrons de connectivitat que no tenen res de trivials, com els descrits en la

secció anterior, o que, lligats amb els exemples que hem vist aquí, anirien de la relació visual que s'estableix entre els animals (cuques, estornells o peixos); l'auditiva, entre l'audiència d'un concert, o familiars o d'amistat, entre els individus d'una societat.

1.1.3. Connexions

I aquest és justament el tipus de complexitat que ens comencem a trobar i del qual hem fet un petit esbós en la introducció. Hem de començar, doncs, descrivint la complexitat estructural o topològica inherent a molts dels sistemes naturals, tecnològics o socials que tenim al voltant. Hem de saber si aquestes connexions són fixes o varien amb el temps, si són totes iguals o si són les unes més fortes que les altres, si malgrat veure-les idèntiques localment resulta que el seu paper en l'entramat global és força diferent, i molts d'altres problemes relacionats amb la complexitat, o no-simplicitat, d'un món ple de connexions.

Veurem alguns exemples més endavant en què la xarxa no existeix en realitat, sinó que és una simple representació d'un sistema complex. Vegem un exemple concret difícil de classificar entre els diferents tipus de xarxa, però que al mateix temps és prou aclaridor d'aquesta interpretació visualitzadora de les xarxes.

Un problema d'escacs

En la figura que ve a continuació ens plantegem el problema d'intercanviar les posicions dels dos cavalls blancs i els dos cavalls negres. Per a resoldre'l no cal ser mestre d'escacs, però el problema no és senzill, i com s'indica en la figura calen quaranta moviments.

I ara us deveu preguntar què tenen a veure les xarxes amb un problema d'escacs. Doncs com hem dit abans, les xarxes a vegades no són més que una manera de representar situacions, relacions, problemes.... En aquest cas particular, el que ajuda molt a resoldre el problema és representar-lo com una xarxa, com es pot veure en la figura següent. Cada node és una casella i cada enllaç entre caselles és un possible moviment del cavall entre aquestes dues. Un cop la representem com una xarxa, visualment podem tenir una resposta evident del problema, perquè podem identificar perfectament els moviments en forma de camí al llarg d'una xarxa i evitar els solapaments en una mateixa posició.

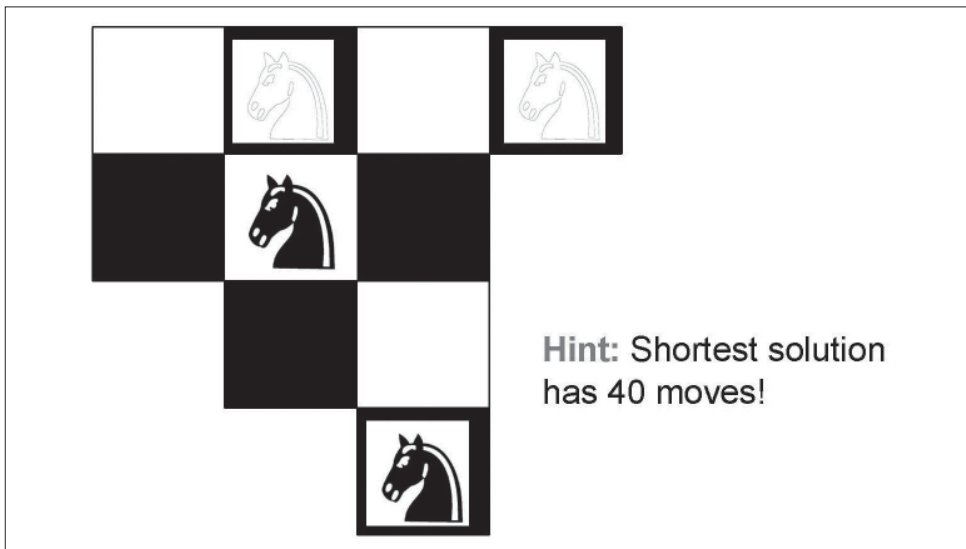


Figura 1.2. Problema d'escacs: intercanvieu les posicions dels dos cavalls blancs i els dos cavalls negres.

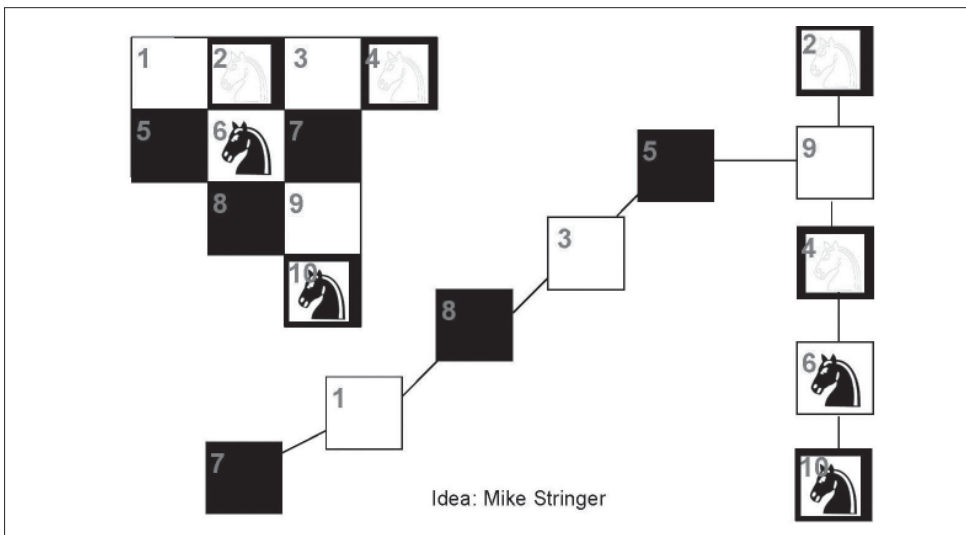


Figura 1.3. Problema d'escacs representat mitjançant una xarxa, en què els nodes són les caselles i els enllaços, els possibles moviments del cavall entre elles.

1.2. Xarxes pertot arreu

Així, doncs, ens trobem al nostre voltant o en l'estructura microscòpica dels organismes multitud d'exemples de sistemes interconnectats de tal manera que podem dir que formen una xarxa. La classificació d'aquests exemples és un punt arbitrari, ja que en podem trobar d'alternatius en la bibliografia existent, i fins i tot podríem dir que hi ha exemples de xarxes que podrien estar classificades en més d'un tipus. Malgrat aquests impediments, establim una classificació temptativa.

1.2.1. Xarxes tecnològiques

En general, aquestes xarxes tecnològiques tenen un suport físic pel qual flueix alguna magnitud. Entre aquests casos tenim: Internet, com una xarxa d'ordinadors i encaminadors per on es transporten bits d'informació; les xarxes de telefonia, per les quals també es transmeten bits, en aquest cas corresponents a veu o dades per a les noves generacions de xarxes digitals; xarxes de distribució elèctrica, per les quals es transporta energia elèctrica, i, finalment, xarxes de transport en què el que es transporta són passatgers.

Internet

La xarxa d'Internet és avui en dia el que podríem dir la mare de totes les xarxes, la que ha permès que en els darrers 25 anys el món hagi canviat de manera tan radical. Com han canviat les nostres comunicacions, les nostres relacions i les nostres aficions. Però hem de distingir entre Internet, la xarxa física de suport, i les xarxes virtuals que han anat proliferant al seu damunt (començant pel mateix World Wide Web, de què tractarem en una de les properes seccions).

Internet no és, en principi, res més que un conjunt d'ordinadors connectats per cables (encara que recentment aquests cables van desaparèixer i donen lloc a multitud de connexions sens fils i en el futur ens portaran al concepte de l'*Internet everywhere*), en què el que s'intercanvien són paquets d'informació. Qualsevol missatge és descompost en paquets que s'envien de manera separada per la xarxa física (poden seguir camins diferents). La manera de trobar la destinació és gràcies a la identificació (l'adreça IP, *Internet protocol address*) de cada ordinador o dispositiu. Justament, el creixement exponencial del nombre de dispositius connectats a Internet ha fet que el nombre d'aquestes possibles

adrees es col·lapsi i sigui necessari adoptar nous protocols de comunicació; per això, ara comencem a parlar d'Internet 2.

Aquesta que hem descrit no és res més que l'estructura que com a usuaris veiem d'Internet. Però aquesta és formada, internament, d'una manera bastant més complexa, perquè, com podem entendre, enviar paquets a destinacions diverses requereix equips intermedis que disposin d'informació, encara que parcial, de l'estructura de la xarxa. Aquests són els encaminadors, que tenen taules d'encaminament dinàmic. A part dels encaminadors, ens trobem els proveïdors de servei, que corresponen a les companyies a escala local; després, les xarxes acadèmiques o de recerca a escala nacional, i, finalment, el que podríem anomenar la columna vertebral de banda molt ampla que intercomunica totes les possibles subxarxes.

Una de les característiques més importants i alhora més sorprenents d'Internet és el caràcter autoorganitzat i espontani de la seva estructura. No hi ha cap autoritat central que dicti la seva manera de procedir i, malgrat tot, s'ha convertit en l'eina més imprescindible avui en dia per a la majoria de nosaltres.

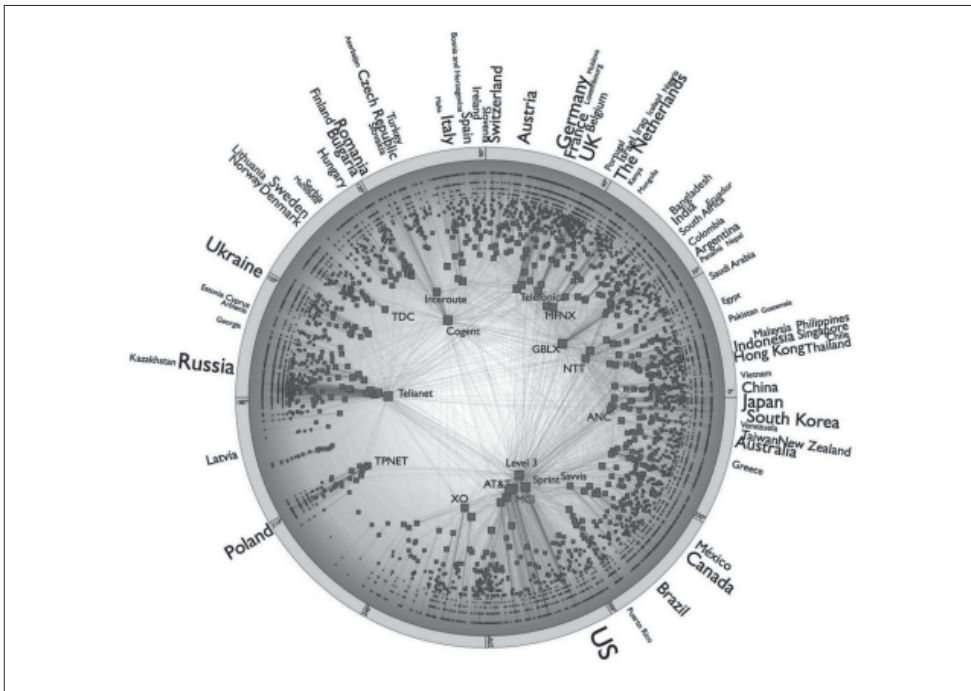


Figura 1.4. Aquesta és una representació d'Internet pel que fa a sistemes autònoms, en què es pot veure l'estructura nacional exterior i l'estructura global més interior.

Xarxa de telefonia

En aquest apartat considerem xarxes de telefonia com les xarxes de cablejat tradicional per les quals es propaguen les nostres converses de veu. Potser ha estat el gran paradigma de les xarxes en el segle XX, poc després de la invenció de la telefonia, fins al començament del segle XXI, quan podríem dir que la telefonia, com a comunicacions de veu, se solapa amb la de dades (Internet). Així, el naixement d'Internet com a gran mitjà de comunicació de masses (a mitjan dècada de 1990) està associat a mòdems de no gaires kilobytes, mentre que en l'actualitat hem passat a encaminadors domèstics amb Internet de banda ampla a alguns megabytes. Però el que ja no podien sospitar les companyies operadores de telefonia fixa és que, amb el temps, moltes de les trucades de veu es fessin utilitzant precisament protocols d'Internet, com ara els *voice-IP*, dels quals potser el més conegut és el programa Skype, adquirit recentment per Microsoft.

Una de les diferències pràctiques entre les xarxes de telefonia i Internet és que mentre que a la segona tot és públic, els protocols, l'emplaçament dels encaminadors, els enllaços físics i lògics, en la primera no, i, en particular al nostre país, hem vist les dificultats dels nous operadors de telefonia per tal de poder accedir a telèfons residencials sense el permís de la companyia que ostentava el monopoli fins no fa gaires anys. Però això és general a gairebé tots els països i així podem dir que no s'han estudiat en absolut les xarxes de telefonia perquè simplement no es disposa de dades.

Xarxa de distribució elèctrica (*power grid*)

En aquest cas, a diferència de l'anterior, ens trobem amb una xarxa de distribució d'energia elèctrica de la qual sí que es disposa de dades. Són un tipus de xarxes fortament condicionades per l'entorn geogràfic i també demogràfic i social.

Ha estat una de les xarxes més estudiades, ja des dels primers treballs de Watts i Strogatz. En aquest cas, es considera la xarxa formada per les línies d'alt voltatge que connecten plantes generadores i estacions d'intercanvi, sense arribar al detall de l'usuari final.

En l'actualitat, hi ha un interès renovat per aquestes xarxes, a causa de les noves fonts de generació d'energia que s'estan implantant de manera general (eòlica, solar...), cosa que dóna lloc al que s'està anomenant l'*smart grid*, ja que tant la generació com el consum (pensem, per exemple, en els cotxes elèctrics) formen xarxes que cada cop arriben a més punts geogràfics. A més, en aquest cas no n'hi prou de conèixer la topologia, sinó que és molt important estudiar

les propietats dinàmiques de la distribució d'energia, que entenem com una dinàmica complexa que té lloc damunt una topologia complexa.



Figura 1.5. Xarxa de distribució d'electricitat a l'Europa occidental

Xarxes de transport

Les xarxes de transport les hem d'entendre a totes les escales, tant a escala metropolitana (la xarxa del metro de Barcelona que mostrem en la figura), com a escala nacional (la xarxa ferroviària o de carreteres), fins a escala mundial (les xarxes de transport aeri o de mercaderies per mar, per exemple).

Òbviament, també és molt important conèixer aquestes xarxes perquè en depèn molt bona part de la nostra economia, de la nostra societat i, fins i tot, de la nostra salut, i si no pensem com s'han propagat els dos casos més importants d'epidèmies a escala mundial, la de la grip aviària i la de la grip A, que ho han fet tan ràpidament perquè les comunicacions entre continents són extremament ràpides gràcies al transport aeri.

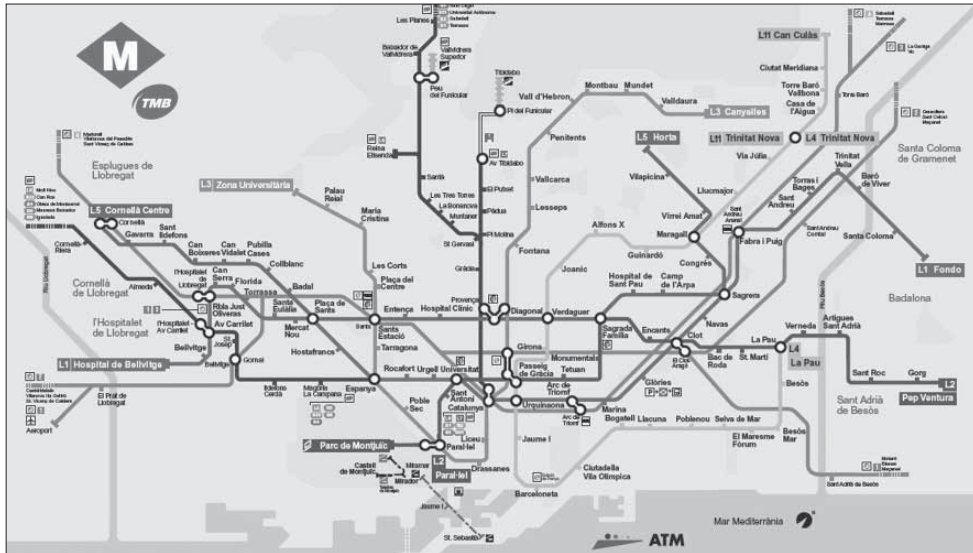


Figura 1.6. Xarxa de metro de Barcelona

1.2.2. Xarxes biològiques

A diferència de les xarxes tecnològiques, les xarxes biològiques majoritàriament són xarxes relacionals, en les quals no hi ha un enllaç físic real entre elements, sinó que el que establim són patrons d'interacció entre elements biològics, com poden ser relacions entre elements que participen en una mateixa reacció química a la cèl·lula o quines espècies es mengen quines altres en una xarxa ecològica. A continuació, considerarem alguns d'aquests exemples.

Xarxes bioquímiques

En general, podem dir que aquestes xarxes representen mecanismes que tenen lloc dins la cèl·lula. Hi distingim:

- Xarxes metabòliques: en aquest cas, la xarxa correspon als components químics que participen en una reacció, i es relaciona cada reacció amb un dels processos que tenen lloc dins la cèl·lula.
- Xarxes d'interacció entre proteïnes: aquest cas correspon a un grau de complexitat superior, ja que les proteïnes no sols interaccionen químicament entre elles, sinó que també ho fan físicament, ja que poden modificar la seva estructura de plegament.

- Xarxes de regulació genètica: finalment, en aquest cas es consideren les interaccions entre diferents parts que formen l'estructura de l'ADN, és a dir, els gens. Identificant aquests gens el que es pretén amb aquestes xarxes és entendre com l'activació o inhibició d'algun (o alguns) d'ells afecta les conseqüents activació o inhibició d'altres.

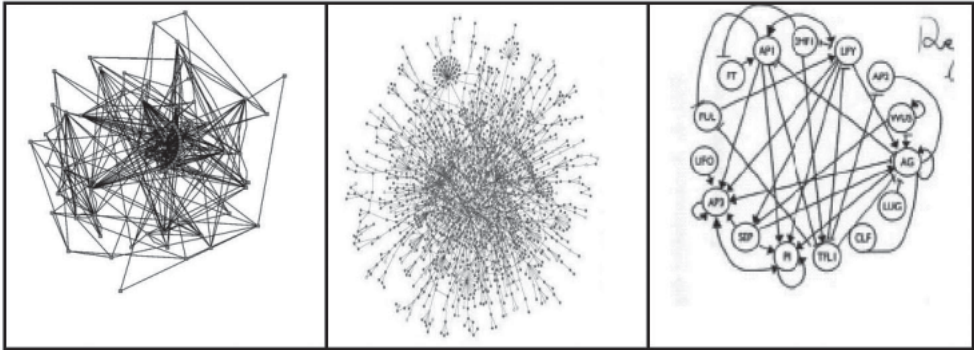


Figura 1.7. Exemples dels tres tipus de xarxes bioquímiques (metabòlica, d'interacció entre proteïnes i de regulació genètica)

Xarxes neuronals

Aquestes són les xarxes que formen les neurones en els sistemes cerebrals dels animals. Que en aquest sistema se l'anomeni *cervell* o no ja és una altra història, perquè normalment es parla de *cervell* en animals vertebrats.

Les neurones són les unitats bàsiques per a processar la informació. Aquesta informació arriba a la part central de la neurona –el soma–, és processada i produeix una sèrie de senyals que es propaguen pels axons per a arribar com a input a altres neurones. Aquesta connectivitat és la que forma la xarxa.

Per a una xarxa senzilla, com la del cuc *C. Elegans* que mostrem en la figura, el paper que tenen les diferents neurones és molt clar, i en tenim moltes dades. En aquesta gràfica es representen amb colors diferents les que són dins els mateixos grups funcionals. Quan passem a animals superiors no és tan fàcil determinar aquests papers i aleshores les xarxes neuronals es passen a considerar d'una altra manera. Per exemple, com a resultat de mesures externes al cervell, es poden mesurar correlacions entre diferents zones corticals i això dona lloc al que anomenaríem *xarxes corticals*, que ja no són xarxes entre neurones individuals sinó que connecten zones funcionals del cervell. Aquestes relacions de connectivitat s'estableixen analitzant les correlacions entre diferents zones del cervell quan l'individu està sotmès a diferents tipus d'estímuls.

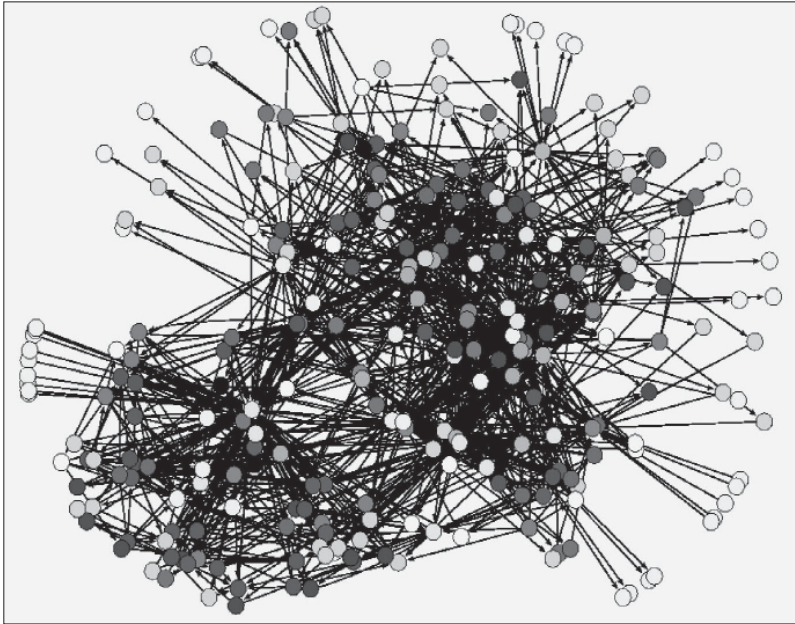


Figura 1.8. Xarxa neuronal del *C. Elegans*

Xarxes ecològiques

En aquest apartat, en l'escala més gran, la que correspon a espècies animals, ens trobem amb els ecosistemes. Un ecosistema és un concepte més general que *xarxa* i inclou multitud d'interaccions. Per tal de restringir-nos a les xarxes complexes que apareixen normalment en el món de l'ecologia ens fixem en les anomenades *xarxes tròfiques*, de les quals tenim un exemple esquematitzat en la figura.

Una xarxa tròfica és una xarxa formada per les interaccions entre espècies, en què unes són l'aliment d'alguna de les altres. Aquesta relació tan senzilla de descriure ha donat lloc a un conjunt de xarxes molt conegudes i que han estat elaborades amb gran paciència pels ecòlegs de camp. El gran avantatge d'aquestes xarxes és que en ecosistemes ben determinats, com, per exemple, un llac, no hi ha gairebé interaccions amb els sistemes exteriors i es poden anar recollint dades amb molta estabilitat temporal.

D'altra banda, aquestes xarxes ens indiquen el grau de sensibilització dels mateixos ecosistemes. I és possible entendre per què la desaparició d'alguna espècie, pel motiu que sigui, pot donar lloc a una superpoblació de les seves preses i a problemes de subsistència dels seus depredadors. Però això seria una conclusió massa senzilla, ja que una de les coses que hem après de les xarxes

complexes és que la interconnectivitat comporta conseqüències que van més enllà dels simples efectes als nostres veïns. I quan diem que una espècie desapareix no tenim ni idea, en principi, dels possibles efectes devastadors que aquest fet pot tenir en tot l'ecosistema.

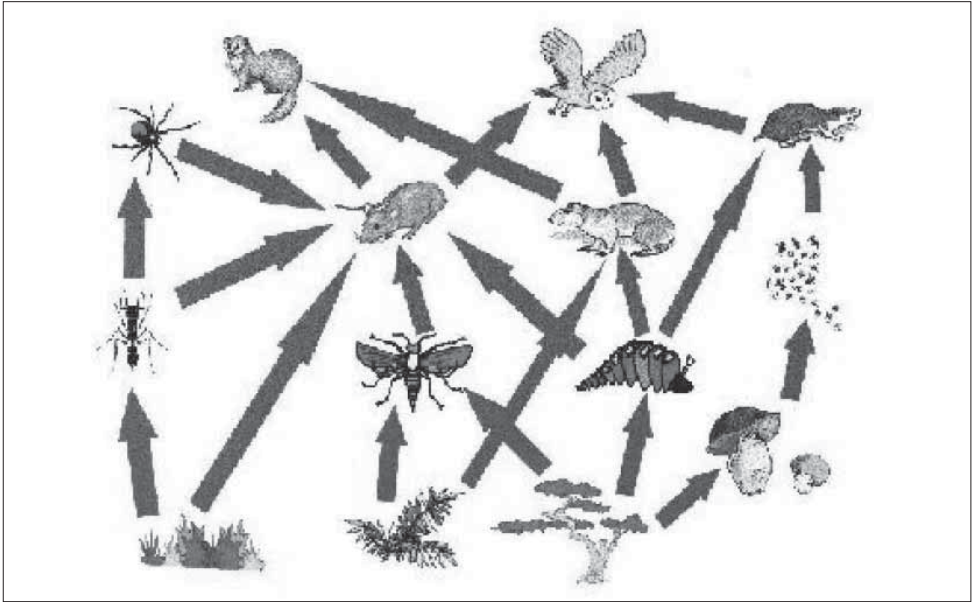


Figura 1.9. Exemple de xarxa tròfica, de qui es menja a qui en un ecosistema.

1.2.3. Xarxes socials

Quan parlem de *xarxes socials*, el primer que ens ve al cap és Facebook, sobretot gràcies a la pel·lícula recent dirigida per David Fincher. Ens creiem, potser, que no som ningú en la nostra societat si no som a Facebook? Però aquesta enorme xarxa amb milers de milions d'usuaris ha desenvolupat certes propietats que fan que de vegades ens la mirem amb força recel. Però ara no discutirem sobre aquesta xarxa tan popular i tan particular, i deixarem per a la segona part d'aquest document la discussió detallada de les xarxes socials.

En un context general, podríem dir que justament les xarxes socials, que han estat tema de recerca en ciències socials durant dècades, mostren la gran evolució que ha sofert aquest camp en els darrers dotze anys. No és difícil imaginar com es feien les anàlisis relacionals fa trenta o quaranta anys, mitjançant qüestionaris. I, òbviament, som conscients de les limitacions que això comportava en termes de temps necessari per a fer-les, de recursos i de biaix, perquè el

mateix qüestionari o qüestionant poden influir en les respostes. Què tenim en l'actualitat? Dades, moltes dades, i és tan fàcil accedir-hi que podem construir la nostra xarxa d'amics a Facebook (vegeu la segona part d'aquest document) o conèixer la xarxa que va donar lloc al fenomen dels indignats a tot l'Estat espanyol a partir del 15-M.

Enllaç recomanat

Anàlisi del moviment 15-M o dels indignats fet al BIFI, a partir de piulades geoposicionades. <http://15m.bifi.es/>

1.2.4. Xarxes d'informació

Per acabar amb aquesta classificació parlem ara de xarxes en les quals el que flueix bàsicament és informació, entesa d'una manera bastant general i, fins i tot, d'una manera abstracta. I recordem que aquesta classificació no deixa de tenir un punt d'arbitrarietat, ja que algunes xarxes podrien ser classificades en més d'un dels apartats.

El World Wide Web

Aquesta teranyina d'amplària mundial és possiblement la xarxa que ens va obrir els ulls a Internet a mitjan dècada de 1990. L'arrel la trobem en el desenvolupament d'una sèrie de protocols, principalment per científics del CERN (laboratori de física d'altres energies) a Ginebra, per a la comunicació interna en el grup d'investigadors. Encara que en aquells temps hi havia idees similars entorn de la comunicació sobre Internet, l'èxit fulgurant del Web (i del protocol que fa servir, l'HTML) es va deure al fet que els creadors van decidir que el programari fos de lliure accés i ús. Aquesta universalització, i d'altres que han succeït posteriorment, és la que ha permès que trenta anys després del naixement d'Internet parlem comunament d'*informació i programari d'accés lliure i obert*.

El World Wide Web té una estructura que s'anomena *dirigida* perquè els enllaços només funcionen en una direcció. Els enllaços (*links*) són els que ens porten d'una pàgina web a una altra. Es tracta, òbviament, d'una xarxa amb milions i milions de documents en què és difícil navegar, però justament aquesta és una de les característiques del Web: administrar la quantitat ingent d'informació i indexar-la perquè sigui útil. I aquí és on entren possiblement les eines amb més potencial de la Xarxa, com són els cercadors. I els grans negocis que s'han anat desenvolupant al seu voltant, com és el cas de Google.

Xarxes de cites, patents i legals

Aquestes són unes altres xarxes en què el que es propaga és la informació. En el cas de les xarxes de cites, són els articles científics els que se citen els uns als altres. Els articles són els nodes i les cites, els enllaços. Com en el cas del WWW, la xarxa també és dirigida. A més, en aquest cas, hi ha una restricció important que és la de la seqüenciació temporal, ja que hi ha clarament un efecte de la data de publicació, i això fa que quan parlem de xarxes aquesta no presenti cicles tancats.

Mirant en perspectiva, de fet, una de les primeres anàlisis de les xarxes de cites (feta per de Solla Price l'any 1965) descriu una distribució del nombre de cites dels articles que té una forma especial, que s'allunya de la gaussiana o normal tradicional, i que veurem en detall més endavant en aquest mateix document. Mentre que en aquest i altres treballs d'aquella època la recollida es feia eminentment a mà, ens trobem avui en dia amb grans repositoris, com ara les versions electròniques de les revistes de més importància, en què la recollida de dades es pot fer de manera automàtica i constitueix, així, una font important.

Aquest tipus de xarxes seran tractades en especial en la tercera part d'aquest document.

Molt semblants en estructura a les xarxes de cites, trobem les xarxes de patents, en què ara els nodes corresponen a patents i els enllaços a com els registres de les patents se citen les unes a les altres. Tenen la característica compartida que són dirigides i que no hi ha cicles tancats.

Finalment, dins d'aquesta categoria podríem incloure les xarxes de cites legals, que serien aquelles en què, per exemple, certes sentències promulgades per un jutge en poden citar d'altres o en què algunes lleis en citen textualment d'altres. En aquests casos també observariem els efectes descrits anteriorment.

Xarxes de correu

Aquesta que descrivim aquí és òbviament un tipus de xarxa que transmet informació entre dos usuaris. Òbviament, considerant tots els servidors de correu d'Internet i tots els usuaris, això forma una xarxa gegantina i indesxifrable. El que es considera, en aquest cas, són xarxes de correu tancades entre usuaris, per exemple, dins una organització.

En podem trobar alguns exemples a la bibliografia, un dels quals és el que en el nostre grup de recerca vam fer sobre la xarxa de correu a la Universitat Rovira i Virgili. Podeu trobar detalls d'aquest treball en l'enllaç següent. En aquest cas, es van recollir tots els missatges intercanviats entre usuaris (anònims) durant els tres primers mesos de l'any 2002. El resultat el podeu veure en la figura que presentem a continuació.

Enllaç recomanat

Alex Arenas; Albert Díaz-Guilera (2009). "Análisis y minería web. Identificación de comunidades analizando el uso del correo electrónico". *El Profesional de la Información* (núm. 18).

<http://www.elprofesionaldelainformacion.com/contenidos/2009/enero/04.html>

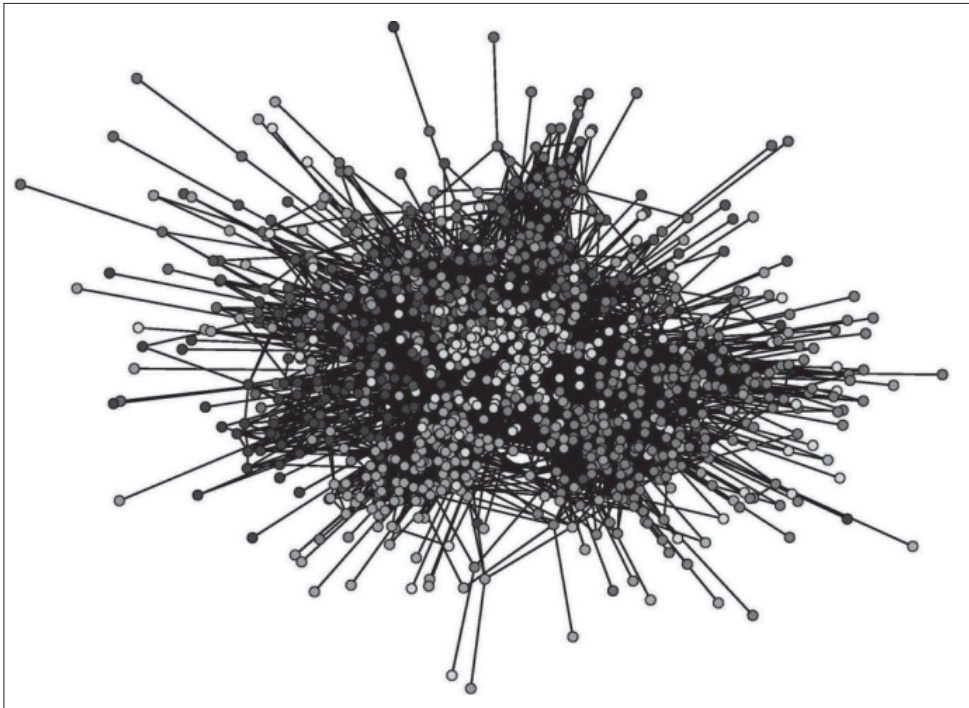


Figura 1.10. Xarxa de correu de la Universitat Rovira i Virgili, recollida durant els tres primers mesos de l'any 2002. Cada node correspon a un usuari (anònim) i hi ha un enllaç si entre dos usuaris s'han intercanviat almenys un missatge durant aquest període. Les tonalitats de grisos corresponen al centre d'adscripció de l'usuari.

Aquesta recollida de dades ens aporta una informació molt valuosa sobre la mateixa organització, ja que permet establir els patrons de comunicació entre els diferents centres, per exemple. En la part del document en què tractarem d'estructures internes explicarem detalladament els resultats en aquesta xarxa de correu.

En la bibliografia trobem analitzats altres exemples de xarxes de correu. Una alternativa ha estat, per exemple, considerar un conjunt d'usuaris i crear les

dades en les seves llibretes d'adreces. Aquests encreuaments també formen una xarxa, però en aquest cas correspondria més a una xarxa social, de contactes més o menys intensos, que no pas a una xarxa d'informació.

Coautories

En aquest apartat parlem d'un tipus de xarxa bastant diferent de les anteriors, perquè novament és una xarxa que tant la podríem classificar aquí com en les xarxes socials, però entenem que té més a veure amb la informació intercanviada per científics que col·laboren en un projecte comú que no pas en una xarxa social.

Ja des del principi es van analitzar xarxes de coautories en bases de dades científiques corresponents a diferents disciplines. Més endavant, descriurem aquests resultats, però el que podem dir a hores d'ara és que, independentment de la disciplina, els patrons de la xarxa són comuns.

Aquí presentem els resultats obtinguts en el nostre grup de recerca corresponents a una sèrie de congressos (FisEs) que van tenir lloc cada any i mig en la comunitat de físics estadístics a Espanya. Considerant cada edició del congrés, es van analitzar els conjunts de treballs presentats i es van establir enllaços entre autors si havien participat en la mateixa contribució. La xarxa que presentem és l'acumulada al llarg de totes les edicions.

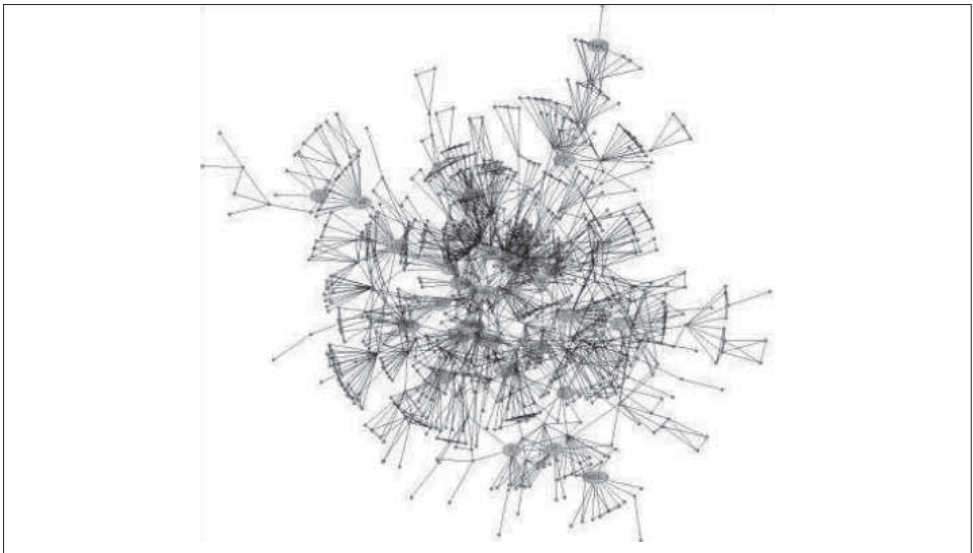


Figura 1.11. Xarxa de coautoría. Xarxa acumulada al llarg de les diferents edicions del congrés de físics estadístics a Espanya (FisEs). Els nodes marcats corresponen als membres del comitè científic.

En aquesta xarxa podem veure que s'articula entorn de certes temàtiques (en aquest cas, la xarxa no és anònima ja que les dades són públiques) i també entorn de certs científics que han tingut un paper important en el desenvolupament de la disciplina a escala estatal.

Consells d'administració

Un dels exemples paradigmàtics que des del principi es va anar considerant en les xarxes, sobretot relacionades amb l'economia, és el de les xarxes que es creen considerant els consells d'administració de grans empreses i corporacions. En aquest cas, fem un primer plantejament prenent com a subjectes els membres dels consells d'administració i els considerem connectats si pertanyen al mateix consell. Sorprenentment, aquests consells tenen una gran força en l'economia nord-americana. I com a xarxa d'informació, aquesta la podem entendre com una mesura del que està passant i passarà en l'economia real.

Sistemes de recomanació

Un altre exemple de xarxa d'informació són les xarxes de recomanació. En aquest cas, la informació és valuosa tant per a l'usuari final com per a l'entitat que la genera. Aquestes xarxes es generen dinàmicament i els mateixos actors no en tenen consciència. Per què cada cop que comprem un llibre a Amazon, com ens diu que els usuaris que han comprat aquest llibre també han comprat aquests altres? Doncs simplement perquè es generen patrons de similitud entre usuaris i entre productes, i és justament aquesta similitud la que estableix les relacions entre els objectes recomanats i aquells a qui se'ls recomanen. Per a tenir una idea de la importància d'aquest tipus de xarxa, fa uns anys una companyia de distribució de DVD que es poden llogar en línia, que ja tenia un algorisme de recomanació per als seus usuaris, va obrir un concurs amb un premi d'1 milió de dòlars per a la persona que trobés el millor algorisme de recomanació de les seves pel·lícules basant-se en les dades acumulades fins aquell moment. I en aquest concurs hi van participar equips de recerca de les millors universitats del món.

1.3. Què és una xarxa?

1.3.1. Nodes i enllaços

Fins ara hem vist alguns exemples del que anomenem *xarxes complexes*, que podem trobar fàcilment al nostre voltant, com ara les xarxes tecnològiques, o que simplement no són res més que una representació adient per a entendre una certa fenomenologia, com ara alguns exemples de xarxes biològiques.

Però, què tenen en comú totes aquestes xarxes? De la més gran a la més petita, hi ha dos elements imprescindibles:

- **Nodes:** són els elements principals de les xarxes. Són les unitats bàsiques de la nostra representació, són els ordinadors d'Internet, les pàgines del WWW o els llibres que comprem a Amazon.com. Depenent de la disciplina, aquests nodes reben diferents noms, però nosaltres mantindrem el de *node* al llarg del document.

- **Enllaços:** són les relacions entre els nodes. Hem vist abans que aquests enllaços poden ser reals, sobre els quals flueix alguna magnitud física, o virtuals, i els interpretem com una relació entre dos nodes, com pot ser un enllaç entre dues pàgines al WWW. Normalment, però, mai no considerem que els nodes tenen autoenllaços ni que entre dos nodes hi ha més d'un enllaç.

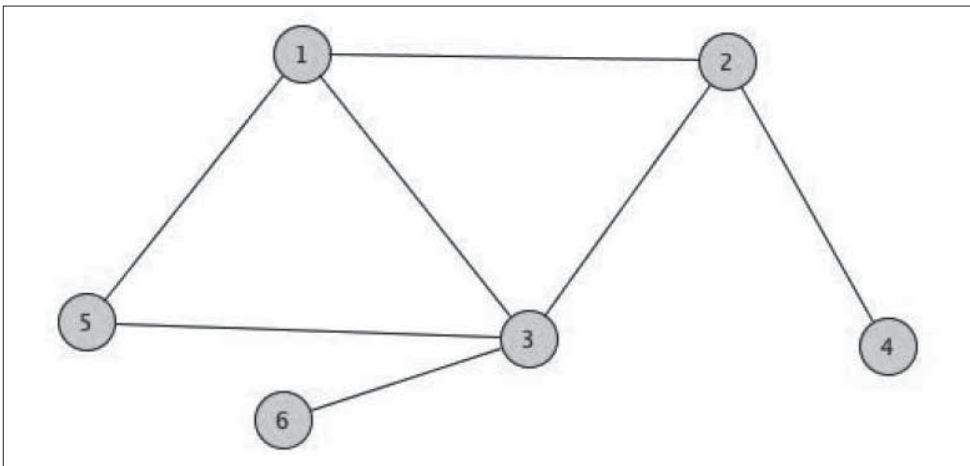


Figura 1.12. Xarxa amb sis nodes i set enllaços. En aquest cas, els enllaços són idèntics i no impliquen direccionalitat.

1.3.2. Tipus de xarxes

L'exemple que hem posat anteriorment correspon al tipus més simple de xarxa i és el que anomenem una *xarxa binària*, que vol dir que els enllaços entre nodes existeixen o no existeixen, no ens importa quantificar els enllaços ni saber-ne la direccionalitat. Això correspondria al cas, per exemple, d'una xarxa de carreteres sense entrar en detalls de la capacitat de la carretera o al fet que un usuari sigui en la llista d'adreces de correu d'un altre sense entrar en detalls de quants missatges s'han intercanviat.

Però, recentment, s'ha posat més èmfasi en les xarxes en què sí importa la direccionalitat i la intensitat dels enllaços. En els exemples anteriors, correspondria a quina seria la capacitat de la carretera entre dos punts o el nombre missatges enviats en una direcció o en l'altra (perquè aquesta magnitud no cal que sigui simètrica). Aquestes són les xarxes anomenades *xarxes pesades*, si tenim en compte el pes o la intensitat, i dirigides, si tenim en compte la direccionalitat.

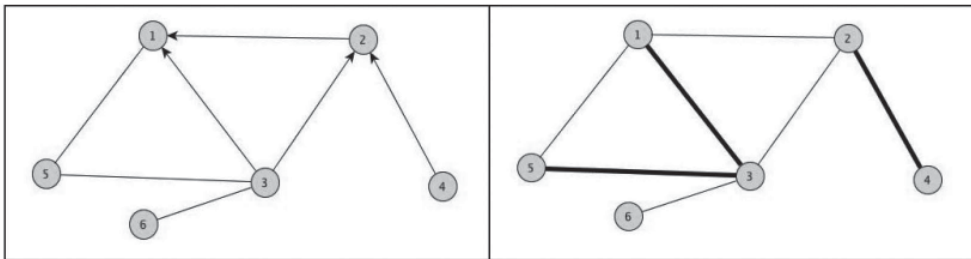


Figura 1.13. Exemples de xarxes no binàries. Esquerra: xarxa dirigida. Fixem-nos que en aquest cas particular, tot i que la xarxa sigui connexa, no es pot arribar al node 4. Dreta: xarxa pesada.

Veurem més endavant, en parlar de la caracterització de les xarxes complexes, que ha de ser diferent per a les xarxes binàries respecte de les xarxes dirigides o pesades. Per exemple, queda clar que si una xarxa és dirigida i pretenem descriure els camins possibles sobre una xarxa pot passar que malgrat formar una component connexa no es pugui arribar a tots els nodes, com passa en la figura amb el node 4. D'altra banda, les xarxes pesades també ens indiquen la diferència entre els tipus d'enllaços entre nodes, per exemple, quan quantifiquem el nombre de missatges de correu intercanviats entre usuaris ens indica amb quina facilitat pot arribar la informació que es propaga per la xarxa.

Xarxes bipartides

Dins dels tipus elementals de xarxes, introduïrem unes que tenen una rellevància especial perquè són formades per dos tipus de nodes i que es poden reconvertir en xarxes ordinàries amb unes característiques especials.

Són les anomenades *xarxes bipartides*, i s'anomenen així perquè són formades per dos tipus de nodes, com podem veure en la figura següent.

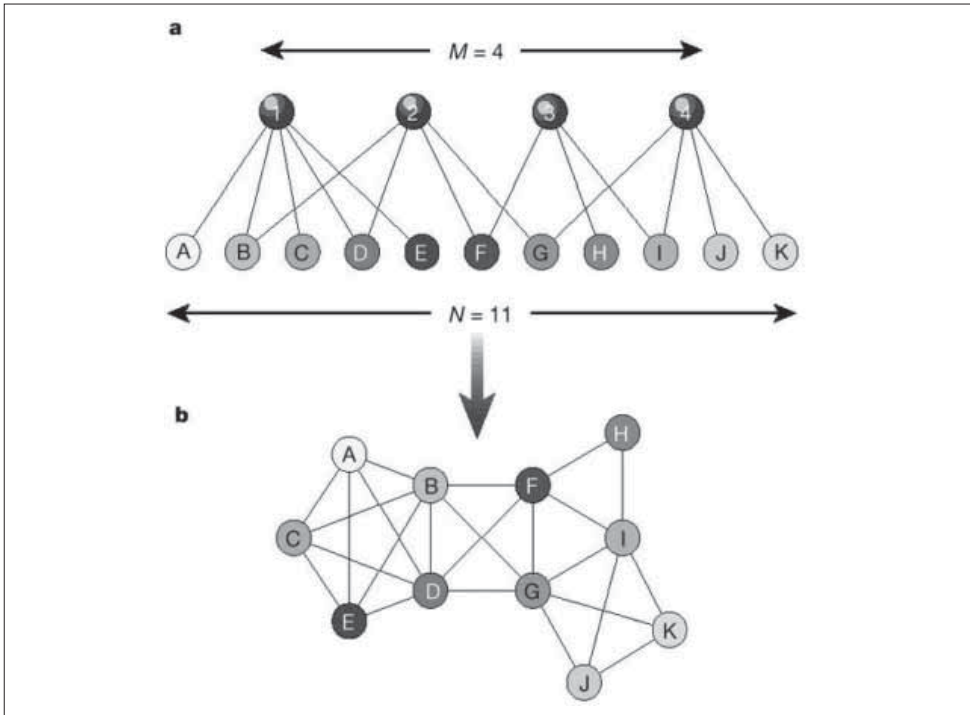


Figura 1.14. a) Exemple de xarxa bipartida formada per dos tipus de nodes. Per exemple, els nodes superiors podrien correspondre a articles i els nodes inferiors, a autors.
b) Xarxa ordinària corresponent a la imatge superior.

Un tipus de nodes (els de colors) corresponen als nodes originals i l'altre, als grups als quals pertanyen. Per a posar-ne un exemple aclaridor, els nodes originals poden ser autors i els grups poden correspondre a articles escrits per aquests autors, aleshores es dibuixa un enllaç entre els autors i l'article en el qual han participat. Això correspon a la part superior de la figura. Aquesta seria la xarxa bipartida amb els dos tipus de node. En canvi, en la part inferior de la figura representem el que s'anomena *projecció de la xarxa bipartida*, és a dir, una xarxa ordinària en què els enllaços estan establerts entre nodes que

pertanyen al mateix grup; en el nostre cas particular, correspon, doncs, a una relació de coautoria entre autors d'articles. En aquest cas, tenim grups de nodes que estan completament connectats, són els que comparteixen com a mínim un article en comú. Aquests grups de nodes completament connectats s'anomenen en la bibliografia cliques.

1.3.3. Descripcions

Fins ara hem descrit les xarxes d'una manera totalment qualitativa. Però si en volem fer un tractament matemàtic o gràfic necessitem fer-ne una representació algebraica. Normalment, el que expressarem és l'existència o no d'enllaços entre nodes, com són aquests enllaços o les llistes que existeixen.

Aquestes representacions serien:

1. Llista d'enllaços: aquesta és la que utilitzen la majoria de paquets de programari, com ara el que s'utilitzarà majoritàriament al llarg d'aquest curs, el GEPHI. Consisteix a fer una llista de tots els enllaços com a parells ordenats de nodes per a les xarxes no pesades. Si la xarxa és no dirigida no cal posar-hi els dos ordres, però si la xarxa és dirigida aleshores sí que cal especificar l'ordre. Per a una xarxa pesada s'escriu una tercera columna amb el pes de l'enllaç.

1	2		2	1		1	2	1
1	3		3	1		1	3	2
1	5		3	2		1	5	1
2	3		5	1		2	3	1
2	4		4	2		2	4	2
3	5		3	5		3	5	2
3	6		3	6		3	6	1
			5	3				
			6	3				
			1	5				

Figura 1.15. Llista d'enllaços de les tres xarxes senzilles representades anteriorment. En el primer cas, cal especificar que la xarxa és *undirected*, ja que qualsevol enllaç es considera en les dues direccions. En el segon, cal dir *directed* perquè en aquest cas importa la direccionalitat. Finalment, en el tercer cas, hi ha una tercera columna que correspon al pes de cada enllaç.

2. Llista de veïns (*edgelist*): aquesta representació s'utilitza només per a les no pesades i fem una llista de veïns (siguin dirigits o no) per a cada node. Aquesta manera de considerar les dades és més eficient per a accedir a les posicions de memòria, sempre que s'utilitzin adreces dinàmiques (apuntadors).

1 : 2, 3, 5	1 : 2, 3, 5
2 : 1, 3, 4	2 : 3, 4
3 : 1, 2, 5, 6	3 : 5, 6
4 : 2	4 :
5 : 1, 3	5 : 1, 3
6 : 3	6 : 3

Figura 1.16. Llista de veïns per a les xarxes no pesades anteriors

3. Matriu d'adjacència: aquesta representació és molt útil en tractaments matemàtics, perquè expressa en forma de matriu les llistes anteriors i val també per a les xarxes pesades. En aquest cas, entenem un element de la matriu $-i, j-$ com l'existència (i el pes) de l'enllaç entre el node i i el node j .

$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 0 & 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 & 2 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$
--	--	--

Figura 1.17. Matrius d'adjacència per a les xarxes anteriors

1.4. Caracterització dels nodes d'una xarxa

En aquest apartat, veurem com caracteritzar el paper dels nodes en la xarxa tant des d'una perspectiva local com global, en particular introduint el concepte de centralitat.

1.4.1. Grau

El grau d'un node és el nombre d'enllaços que el connecten. Normalment, es designa el grau amb la lletra k ; així, k_i correspon al grau del node i . Si tractem amb xarxes no dirigides, parlarem del grau sense cap adjectiu; en canvi, quan parlem de xarxes dirigides hem de distingir entre el grau entrant (*incoming degree*) i el grau sortint (*outgoing degree*).

La mesura del grau d'un node és important perquè ens dona una primera indicació de la seva centralitat; és a dir, com més veïns tingui un node més central serà perquè afecta (o és afectat) més un nombre de nodes al seu voltant. Però aquesta mesura és el que diríem local, perquè ens fixem només en un simple veïnatge del node que considerem. Podem tenir un node, molt connectat, però que el nivell d'influència en la resta de la xarxa del qual sigui molt petit, ja que està "allunyat" de la part més important de la xarxa. Aquests fets són els que ens fan definir altres mesures de centralitat més globals.

En la figura que presentem a continuació tenim una xarxa amb un conjunt de nodes amb diferents nivells de centralitat.

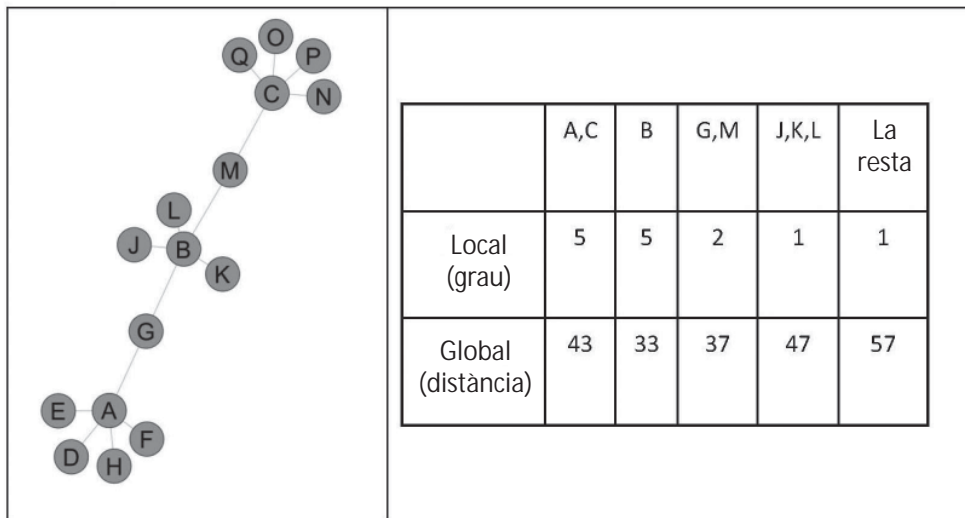


Figura 1.18. Esquerra: xarxa amb setze nodes amb diferent nivell de "centralitat". Dreta: taula amb centralitats local i global dels nodes. La local correspon al grau i la global, a la distància total entre un node i la resta de nodes de la xarxa.

GEPHI

Xarxa generada a partir del fitxer `proves.net` (<http://dl.dropbox.com/u/572806/proves.net>). Hi apliquem posteriorment l'algorisme Yifan Hu proporcional per a la distribució dels nodes en el pla.

En la taula adjunta, trobem el grau de centralitat local que correspon a cada conjunt de nodes (i ens referim a *conjunts* perquè són idèntics a tots els efectes). Podem veure, doncs, que els més centrals amb aquesta mesura seran A, B i C, mentre que la resta són poc centrals a causa del grau baix que tenen. Però queda molt clar en l'estructura de la xarxa que el paper del node B quan ens fixem en el conjunt de la xarxa ha de ser més central que no pas A i C, i el mateix podríem dir de G i M (poc centrals localment) respecte de A i C. En la figura 19 hem inclòs representacions de la mateixa xarxa marcant amb un codi de colors en funció dels diferents tipus de centralitat, en particular del grau.

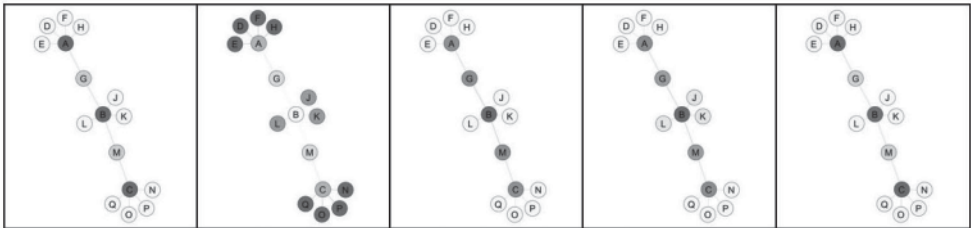


Figura 1.19. Representació de la xarxa amb colors de nodes segons el grau de centralitat: grau, *closeness*, *betweenness*, *eigenvector*, *page rank*. Els colors més clars corresponen a valors numèrics més baixos i els més blaus, a valors més alts.

GEPHI

Executem totes les funcions estadístiques. Anem a representació i triem que representi mitjançant colors cadascuna de les magnituds. Com que podem tenir també un format Taula podem tenir els valors de cada mesura a cada node.

Per a generalitzar aquest concepte de *centralitat* necessitem definir, en primer lloc, el concepte de *camí*.

1.4.2. Camins (*paths*)

En una xarxa binària (sigui dirigida o no, i si ho és tenim en compte el sentit de l'enllaç), definim un camí entre dos nodes com qualsevol seqüència de nodes formada de tal manera que qualsevol parell de nodes consecutius en la seqüència estiguin units per un enllaç de la xarxa. Així, en la xarxa anterior un camí entre M i G seria M, B, G, però també ho seria M, B, L, B, G. I en la xarxa dirigida senzilla introduïda anteriorment entre els nodes 6 i 1 serien 6, 3, 1 o 6, 3, 2, 1, per exemple.

Veiem, doncs, que el concepte de *camí* és força general i que pot incloure nodes que es creuen més d'una vegada o enllaços també que es travessen més d'una vegada. Des d'un punt de vista pràctic, però, només són interessants els camins que anomenem *autoexcloents*, que serien els que no repeteixen nodes ni enllaços.

1.4.3. Longitud

Un cop definits els camins sobre una xarxa el concepte que apareix de manera immediata és el de *longitud d'un camí*, que equival al nombre d'enllaços que travessa un camí. Així, els camins abans esmentats entre 6 i 1 serien de longitud 2 i 3, respectivament.

1.4.4. Distància entre nodes

Però, tal com ens podem imaginar, no tots els camins (infinitos) que es poden definir en una xarxa són igual d'importants. Hi ha un conjunt particular de camins que són els que s'anomenen *camins mínims*, que són els que corresponen a la longitud mínima. Així, podem definir la distància entre nodes com la longitud del camí mínim que els ajunta, i així generalitzar el concepte de *proximitat* i amb aquest el de *centralitat*. En el cas particular que estem considerant podem veure quina és la distància total d'un node a la resta i és el que en la taula hem descrit com a centralitat global, en què podem veure que efectivament el node més central és el B perquè es troba a menys distància (33) de la resta de nodes. Aquesta distància ens permet fer una primera definició de *centralitat global*.

1.4.5. *Closeness centrality*

Aquesta magnitud, com el seu nom indica, es refereix a la centralitat per proximitat, que normalment es dona de manera normalitzada com la distància mitjana del node corresponent a la resta dels nodes. Per al cas del node B, aleshores la seva *closeness centrality* és de 2,2, mentre que la dels nodes més allunyats val 3,8. En el segon panell de la figura 19 presentem novament la xarxa ordenant el color dels nodes segons el valor numèric de la *closeness centrality*.

1.4.6. *Betweenness centrality*

La *closeness* no és l'única mesura de centralitat global disponible. Relacionada també amb els camins mínims ens trobem amb la *betweenness centrality* definida com la centralitat en termes del nombre de camins mínims entre qual-sevol parell de nodes de la xarxa que passen per un node determinat.

A diferència de la *closeness*, per la qual un valor baix indica un alt grau de centralitat global, perquè indica que és a prop de la resta, la *betweenness* elevada indica un alt grau de centralitat. Vegem-ho amb detall. En el fons els camins mínims entre parells de nodes indiquen quin és el flux d'informació en la xarxa quan pensem en algun tipus de senyal que es propaga al llarg d'aquests camins. Aleshores, els nodes que es troben al llarg d'aquests camins són els que reben informació sobre aquest senyal propagat. Per tant, com més camins mínims passen per un node, més informació rebrà de la resta de la xarxa.

En el tercer panell de la figura 19 presentem novament la xarxa ordenant el color dels nodes pel valor numèric de la *betweenness centrality*.

1.4.7. *Eigenvector centrality*

Aquesta és una mesura generalitzada de centralitat associada al grau, ja que considera no solament el nombre de veïns que té un node, sinó quina és la importància relativa dels veïns, de manera autoconsistent. Des d'un punt de vista matemàtic, equival a resoldre un sistema d'equacions algebraïques lineals, de tal manera que la centralitat d'un node és proporcional a la suma de les centralitats dels seus veïns.

Sense entrar en detalls matemàtics (es pot calcular mitjançant GEPHI), la seva importància la podem veure en les xarxes socials, ja que un individu en

una xarxa social pot ser important per la seva mesura, és a dir, per a conèixer molta gent, encara que aquesta gent no sigui tan important en ella mateixa, o per a conèixer poca gent però que sigui força important.

En el quart panell de la figura 19 presentem novament la xarxa ordenant el color dels nodes pel valor numèric de l'*eigenvector centrality*.

1.4.8. Page rank

Si en el decurs dels darrers anys una companyia en el món d'Internet ha creat l'admiració de tothom aquesta és, sens dubte, Google. Encara que avui en dia els seus productes es troben en qualsevol dels vessants de la moderna economia en línia (acabant pel recent Google+), el producte estrella de la companyia ha estat i és el seu cercador. I el seu cercador funciona amb una recepta que es considera tan secreta com la de la Coca-Cola, encara que està basada en una mesura de centralitat semblant a les anteriors. Bé, doncs, aquesta mesura de centralitat és la que s'anomena *Google page rank* o simplement *page rank*.

En el fons, quan fem una cerca a Google el que ens torna el cercador és una llista ordenada d'adreces de pàgines web que contenen la combinació requerida. I aquesta ordenació es fa mitjançant el *page rank* de la pàgina, que ens n'indica la popularitat.

Però com es mesura el *page rank* d'una pàgina? Simplement, considerem com si el web fos un sistema en què tenim un navegador (*surfer*) que es mou de manera aleatòria per la xarxa entre nodes que estan connectats (compte amb la direccionalitat dels enllaços o *hiperlinks* del web). Aleshores, la popularitat de la pàgina és simplement proporcional al nombre de vegades que el surfer hi passa en els camins aleatoris. Un dels problemes, però, que aquest caminant es troba si una xarxa és dirigida, com ara el cas del Web, és que no hi ha manera de sortir d'algunes pàgines que no ens porten a cap lloc. La manera de resoldre-ho és considerar que el caminant té una probabilitat p (presa normalment com a 0,15) de recomençar el procés d'un altre node triat de manera absolutament aleatòria. Aquesta magnitud es pot expressar de manera matemàtica com a

$$r_i = \frac{p}{N} + (1 - p) \sum_{j:j \rightarrow i} \frac{r_j}{k_{\text{out},j}}$$

En aquesta expressió veiem que el valor del *page rank* a cada pàgina depèn dels valors als seus veïns. Igual que l'*eigenvector centrality* això s'ha de resoldre de manera iterativa numèricament, però això no és cap problema greu perquè fins i tot per a xarxes tan grans com el WWW el problema computacional convergeix molt ràpidament.

En el cinquè panell de la figura 19 presentem novament la xarxa ordenant el color dels nodes pel valor numèric del *page rank*, utilitzant els càlculs fets amb GEPHI.

1.4.9. Clustering d'un node

En aquest apartat introduïm una mesura que no té res a veure amb la distància ni tampoc amb la centralitat, però que té una gran rellevància de cara a les xarxes socials. És el que s'anomena *clustering*, que per a un node correspon al fet que els veïns d'un node siguin veïns entre ells. Des d'una perspectiva social, per exemple en una xarxa d'amistats, vol dir que és molt probable que els meus amics també siguin amics entre ells. Des d'una perspectiva més matemàtica, el *clustering* té a veure amb el nombre de triangles que es formen en una xarxa. En el cas de la xarxa que hem considerat, el *clustering* de tots els nodes és zero, justament perquè no hi ha triangles. De totes maneres, aquest concepte serà estudiat amb més detall en la segona part d'aquest document.

Encara que sigui per a entendre de manera intuïtiva el que significa el *clustering* considerem una xarxa molt senzilla, formada només per quatre nodes (ABCD) amb enllaços AB, BC, CD, AD. En la figura podem veure clarament que només hi ha un triangle.

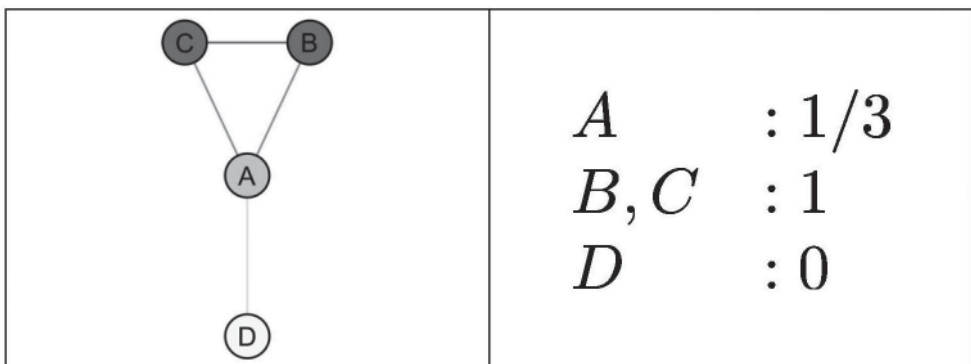


Figura 1.20. Esquerra: xarxa molt senzilla per a veure un exemple de com calcular la clusterització. Dreta: valors de la clusterització per a cada node.

Aleshores, veiem que l'existència d'aquest triangle produeix els valors de la clusterització als nodes indicats a la dreta de la figura. El node B té només dos veïns (A i C) que estan connectats entre ells; per tant, hi ha un triangle en què només n'hi podria haver un, així que la clusterització de B (i també la de C) és igual a 1. En canvi, D no participa de cap triangle però només té un veí; en aquest cas, encara que sigui una indeterminació, s'estableix que la clusterització és 0. Finalment, el node A té tres possibles combinacions de parells de veïns (BC, BD i CD) perquè tots ells són veïns de A. De totes aquestes tres combinacions només en un cas (BC) estan connectats entre ells, cosa que ens dóna una possibilitat sobre tres, així que la clusterització del node A serà d'1/3.

1.5. L'estructura de les xarxes

1.5.1. Grandària d'una xarxa

La mesura que hem estat definint en els paràgrafs anteriors es refereixen a caracterització dels nodes individuals, és a dir, que en el fons el que ens interessa és esbrinar quin és el paper de cada node. En l'apartat que obrim ara i en altres de posteriors el que mirarem és la caracterització global de la xarxa.

En primer lloc, ens fixarem en les distàncies. Si abans hem definit la distància a altres nodes com la caracterització del node, ara utilitzarem la distància entre nodes de manera estadística per a caracteritzar la xarxa. Així, definim:

1. Distància mitjana: aquesta mesura és simplement la distància mitjana entre tots els parells de distàncies entre nodes. Per al cas de la xarxa que estem considerant aquesta és de 3,29.

2. Diàmetre: aquesta mesura ens planteja l'analogia amb un cercle i, com que el diàmetre d'un cercle correspon a la distància màxima entre dos dels seus punts, en el cas de la xarxa també es defineix com la distància màxima entre dos nodes. La xarxa que hem estudiat té un diàmetre de 6.

GEPHI

Executant l'opció "Network diameter".

3. Radi: aquest correspon a la mesura de la distància entre el node més central i el més exterior, que en el nostre cas particular és de 3.

1.5.2. *Clustering* de la xarxa

Així com hem introduït una distància com a resultat mitjà dels valors individuals dels nodes, també es pot introduir un valor mitjà per a l'altra magnitud, que és la clusterització. De fet, hi ha dues possibles interpretacions del que hauria de ser el paràmetre de *clustering* de la xarxa:

1. *Clustering* mitjana: en aquest cas, es defineix la clusterització simplement com la mitjana dels valors en els nodes individuals.

2. *Clustering* global: per a distingir-la de l'anterior, es pot definir també l'efecte global directament, sense fer mitjanes. Així, el que miraríem és el nombre total de triangles en la xarxa respecte al nombre total de triangles que hi podria haver, tenint en compte els enllaços existents.

Val a dir que, en els dos casos, el *clustering* global de la xarxa que estem analitzant és zero perquè no hi ha triangles. En canvi, per a la xarxa dibuixada en la figura 20, aquesta clusterització mitjana és 0,583.

GEPHI

Amb referència a la clusterització, ens calcula els valors individuals i la clusterització de la xarxa està calculada com la mitjana sobre tots els nodes.

1.5.3. Components

En la descripció que hem seguit fins aquest moment ens hem fixat en dues escales d'una xarxa que podríem anomenar *extremes*. Hem mirat quin és el paper de cada node (sia de manera local, com ara mesurant-ne el grau o el *clustering*, sia de manera global, mesurant les diferents caracteritzacions de centralitat) o quines són les característiques de la xarxa en la seva escala total, com ara quina és la seva grandària o quin és el seu coeficient de *clustering* mitjà.

Però en aquests darrers anys hem comprovat que les xarxes complexes tenen estructures a totes les escales que són molt importants, primer perquè ens ajuden en la seva caracterització i segona perquè aquestes estructures són importants per la seva robustesa, la seva flexibilitat o pels efectes sobre la dinàmica que comporten.

Pel seu ordre de més petit a més gran, aquestes estructures que descriurem aquí són: *motifs*, *cliques*, *cores*, i comunitats o mòduls.

1.5.4. Motifs

Quan hem començat a parlar de les xarxes hem dit que són nodes connectats per enllaços. I a partir d'aquesta propietat hem desenvolupat tota una sèrie de mesures; en particular, el grau d'un node que equival al nombre d'enllaços que té. En el fons, podem reinterpretar una xarxa com una col·lecció d'"estrelles" amb un nombre de braços igual que el grau de cada node, en què les estrelles estan unides entre elles.

Aquesta descripció anterior seria en termes de l'estructura local al voltant de cada node, però aquest concepte es pot generalitzar. Si ens fixem en qualsevol parell de nodes, ens adonem que estan connectats (direccionalment si la xarxa és dirigida) o no, és a dir, només hi ha dues possibilitats si no és dirigida (existència o no de l'enllaç) i quatre si és dirigida (existència o no de cada enllaç dirigit). Per tant, aquesta seria l'estructura en el segon nivell a la xarxa.

I aquest concepte es pot generalitzar a qualsevol nombre de nodes, encara que fins ara el que més èxit ha tingut és el dels grups de tres nodes, que presentem en la figura següent.

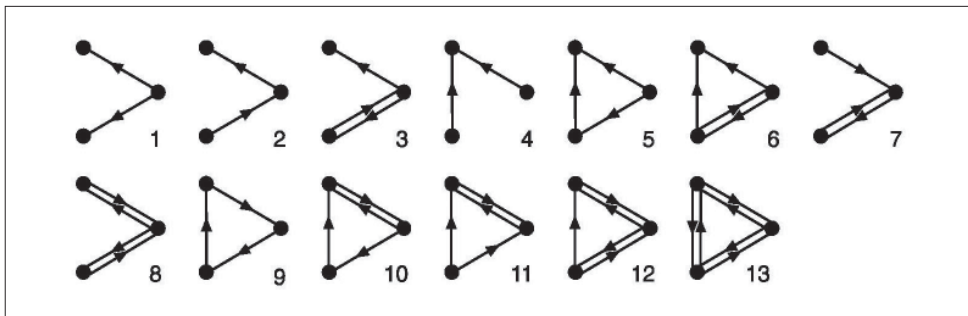


Figura 1.21. Tots els possibles motifs de tres nodes

En aquesta figura podem veure totes les possibles combinacions de grafs de tres nodes, tenint en compte la direccionalitat de l'enllaç. Aleshores, qualsevol conjunt de tres nodes de la xarxa es pot identificar amb un d'aquests *motifs*. Fixem-nos en la xarxa dirigida de l'esquerra de la figura 13. Com que tenim sis nodes, el nombre de triangles que podem formar serà el nombre de combinacions de sis elements agafats de tres en tres, és a dir, vint. D'aquests vint

possibles triangles no tots tenen un mínim de dos enllaços perquè puguem dir que formen un triangle; per exemple, els nodes 1, 2 i 6 no el formen. Dels que sí que en formen un, per exemple, tenim els nodes 1, 2 i 3 que formen un *motif*, segons la figura 21, del tipus 5; els nodes 1, 2 i 4, del tipus 2, i així successivament (tots enumerats a la dreta). Resumint: el tipus de *motif* que observem seria un del tipus 2, 4, 5, 7, 8 i 13, mentre que el tipus 3 ens apareix tres vegades.

De fet, el que importa és aquesta estadística de *motifs*, saber quantes vegades apareix cadascun. I es va veure que en moltes xarxes reals aquests *motifs* apareixen més sovint del que s'esperaria.

Aquesta definició es podria generalitzar a qualsevol nombre, però és fàcil adonar-se que el nombre de *motifs* creix molt de pressa amb el nombre de nodes. Per exemple, en el cas de sis nodes és ja de centenars, cosa que fa que sigui totalment impracticable.

Motifs
Llista de triangles i <i>motif</i>:
1, 2, 3: 5
1, 2, 4: 2
1, 2, 5: 7
1, 3, 5: 12
1, 3, 6: 3
2, 3, 4: 4
2, 3, 5: 3
2, 3, 6: 3
3, 5, 6: 8

1.5.5. Cliques

Aquest és un nou tipus d'estructura a l'escala més petita, que d'alguna manera complementa l'anterior. En aquest cas es parla de *k-clique*, en general, com un subgraf de *k* nodes completament connectat. En particular, els triangles, que analitzem en el context de la clusterització, són *3-cliques*. De fet, des d'un punt de vista pràctic, la clusterització és un nombre baix que indica que el nombre de triangles en una xarxa és relativament baix. I si considerem *cliques* d'ordre superior, la probabilitat de trobar-los encara és més baixa.

1.5.6. Comunitats

A una escala superior, el que s'anomena normalment la *mesoescala*, ens trobem amb estructures que tenen una interpretació fàcil dins de cada disciplina.

Per exemple, en el camp de les ciències de la vida (en què s'anomenen *mòduls* o *grups funcionals*) o en les ciències socials (en què s'anomenen comunitats).

Les estructures anteriors són més importants des d'un punt de vista més matemàtic o formal.

Les comunitats (aquest serà el nom que adoptarem en la resta d'aquest text) tenen un paper important dins la mateixa estructura general de la xarxa perquè en el fons estan associades al mateix procés ordenat de creixement de la xarxa (estudiarem més endavant alguns models de creixement).

En principi, una comunitat es pot definir com un subconjunt de nodes que estan més densament relacionats internament, en comparació de la resta de la xarxa. Però això és una definició molt simple que no pot assegurar la identificació correcta dels grups que formen la xarxa complexa.

El problema de detecció de comunitats és força complicat i ha estat objecte de discussió en diverses disciplines. Es tractaria, bàsicament, d'identificar una partició en què els nodes estiguin agrupats seguint criteris de connectivitat interna. És a dir, atesa una xarxa, com es particiona (proposa una partició) de manera que s'optimitzi una magnitud determinada. En principi, no tenim ni idea de quantes comunitats conté la millor partició i, fins i tot, ens podem trobar organitzacions jerarquitzades o amb solapaments. Encara que en una primera aproximació ens podem oblidar d'aquests efectes.

En els darrers anys (sempre tenint en compte que quan parlem d'una disciplina tan jove, vol dir en els darrers cinc anys), hi ha hagut força intents de fer front al problema computacional d'identificar les comunitats en una xarxa complexa. Aquests mètodes varien molt en termes d'enfocament i aplicació, i això els fa difícil de comparar entre ells. De totes maneres, els lectors interessats en aquestes tècniques (comparar-ne la fiabilitat i el cost computacional) poden trobar referències bibliogràfiques en la bibliografia recent.

La identificació de les comunitats no és simplement un problema qualitatiu, en realitat, la comparació entre els diferents algorismes es fa en termes d'una quantitat que mesura com de bona és una partició determinada. Com que les comunitats de vegades no estan perfectament definides amb una frontera de separació clara entre elles, diferents algorismes poden donar lloc a particions una mica diferents. Així, una mesura que quantifica aquesta exactitud en la partició és molt útil. En l'actualitat, la mesura més comunament acceptada és la que s'anomena *modularitat*. Es basa en la idea intuïtiva que les xarxes aleatòries no mostren estructura de comunitats.

Imaginem-nos que tenim una xarxa arbitrària i una partició arbitrària de la xarxa en un nombre de comunitats N_c . Aleshores, és possible definir una matriu de dimensions $N_c \otimes N_c$ en què els elements e_{ij} representen la fracció del total d'enllaços a partir d'un node i a la partició j i acaba en un node de partició j . Llavors, la suma de qualsevol fila (o columna) de i , $a_i = \sum_j e_{ij}$ correspon a la fracció d'enllaços connectats a i .

Si la xarxa no presenta estructura de comunitat, o si les particions s'assignen sense tenir-ne en compte l'estructura subjacent, es pot estimar el valor esperat de la fracció d'enllaços dins de les particions. És simplement la probabilitat que un enllaç comenci en un node de i , a_i , multiplicat per la fracció d'enllaços que acaben en un node i , a_i . Així, el nombre esperat dels vincles intracomunitat és simplement $a_i a_i$. D'altra banda, sabem que la fracció real dels enllaços exclusivament dins d'una partició és e_{ii} . Així, els podem comparar els dos directament i sumar sobre totes les particions en la xarxa:

$$Q = \sum_{i=1}^c (e_{ii} - a_i^2)$$

Aquesta és la mesura coneguda com la *modularitat*, que per a una partició molt bona, i tenint en compte que la xarxa sigui clarament modular, s'acosta a 1. És a dir, que el valor de la modularitat depèn tant de la xarxa com de la partició que estiguem considerant.

Considerem ara novament la xarxa de la figura 18, que la tenim carregada en el programari GEPHI.

GEPHI

Hem d'executar l'estadística modularitat i ens presenta la millor partició. A l'hora de representar-la, triem *modularity class* i en diferents colors podem veure les comunitats.

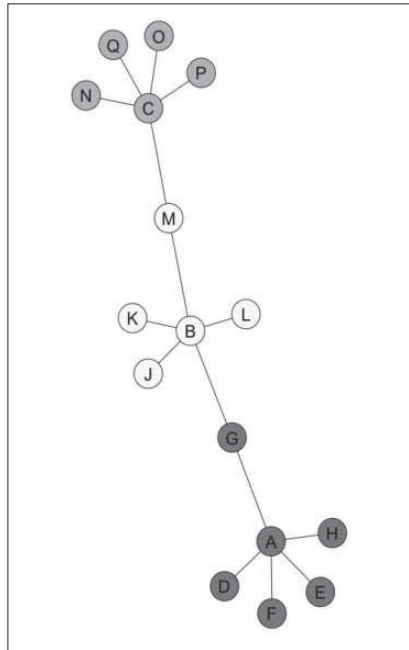


Figura 1.22. Partició en comunitats de la xarxa

Utilitzant els diferents algorismes el que intentem és maximitzar la funció modularitat. En la figura podem veure una partició que té tot el sentit, però podem pensar en particions molt semblants. Ens podem preguntar per què el node G és en la comunitat de baix i no en la intermèdia? O per què el node M és en la intermèdia en lloc de la superior? La resposta a aquestes dues preguntes és molt senzilla, perquè en ambdós casos la modularitat és lleugerament més gran. Però, en canvi, sí que podríem fer els dos canvis alhora, i ens quedaria una distribució equivalent amb exactament el mateix valor de la modularitat, i per tant seria igualment bona.

Parlant de particions, de vegades, no solament estem interessats en la millor, sinó en l'organització jeràrquica de la xarxa en comunitats niades. A continuació, veurem un exemple pràctic d'aquesta mena de particions.

Un dels primers mètodes de detecció de comunitats, proposat per Girvan i Newman, consisteix a dividir la xarxa tallant els enllaços amb els valors més alts de la *betweenness*, justament perquè aquests enllaços són entre zones ben connectades internament. En aquest cas, aquest procediment es pot repetir fins al nivell dels nodes individuals i dona lloc, llavors, a una jerarquia de les comunitats niades. Com es pot veure en la figura que presentem a continuació.

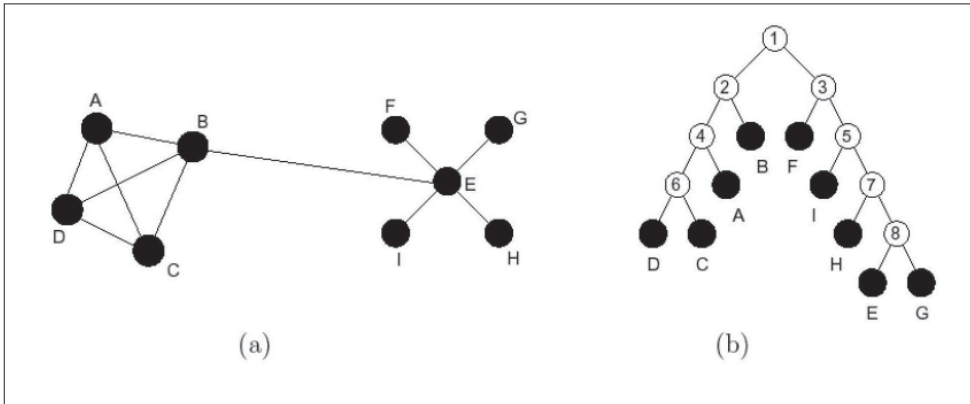


Figura 1.23. Identificació de comunitats. a) Una xarxa amb dues comunitats perfectament identificables. L'enllaç BE és el que divideix les dues comunitats, i és el que té una *betweenness* més gran perquè per anar d'una comunitat a una altra cal passar per aquest enllaç. Per tant, aquest serà el primer enllaç que trencarem. b) Arbre binari que representa la divisió en comunitats. La comunitat global 1 s'ha dividit en dues comunitats (2 i 3). Encara que el procés es fa de manera iterativa, els nodes després de la primera subdivisió hi tenen un paper similar; només el node E té un paper més central en la comunitat 3, i per aquest motiu apareix al final de la ramificació.

L'aplicació d'aquest procediment és molt útil per a la comprensió dels diferents nivells d'organització en una xarxa. En el nostre grup, hem aplicat aquest procediment a la xarxa de correu electrònic de la Universitat Rovira i Virgili.

I el que comprovem és que es manté bona part de l'estructura formal de l'organització però que, al mateix temps, hi ha alguns nodes que no es troben a la comunitat que esperem. Això, sens dubte, és molt valuós com a eina per a la gestió d'una organització.

També com una eina d'identificació de les comunitats de treball i dels líders respectius i, per això, es va aplicar el procediment per a la reunions de física estadística. La xarxa que posem a continuació és el resultat de la partició d'aquest grup de científics, en què podem veure que els nodes verds, identificats com els membres dels comitès científics, es distribueixen per igual entre les diferents branques i apareixen sobretot a les puntes. Això significa que els membres han estat elegits de manera homogènia entre els participants.

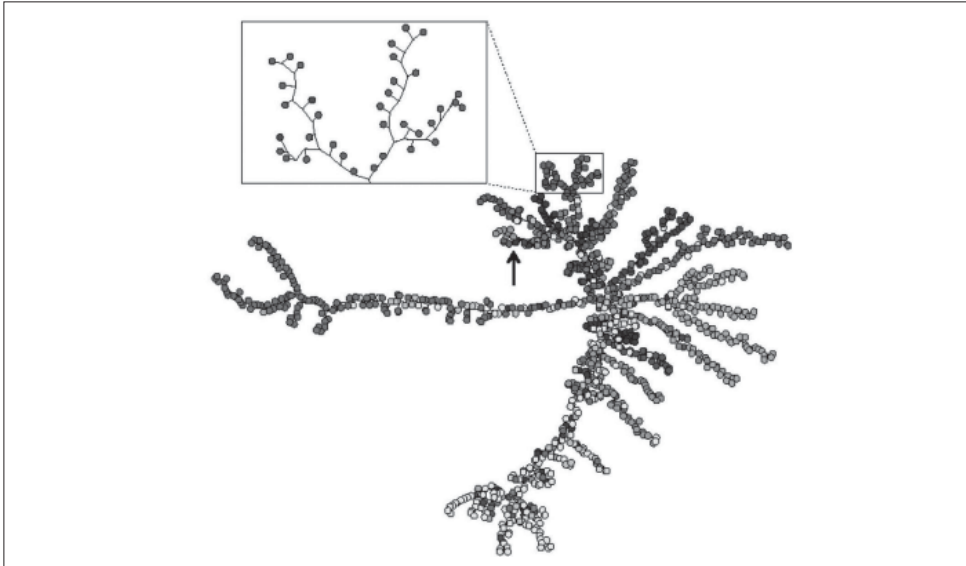


Figura 1.24. Representació mitjançant un arbre binari de comunitats de la xarxa de correu electrònic de la Universitat Rovira i Virgili. Cada branca correspon a una comunitat real i els individus més centrals de cada comunitat són els que apareixen als extrems de les branques.

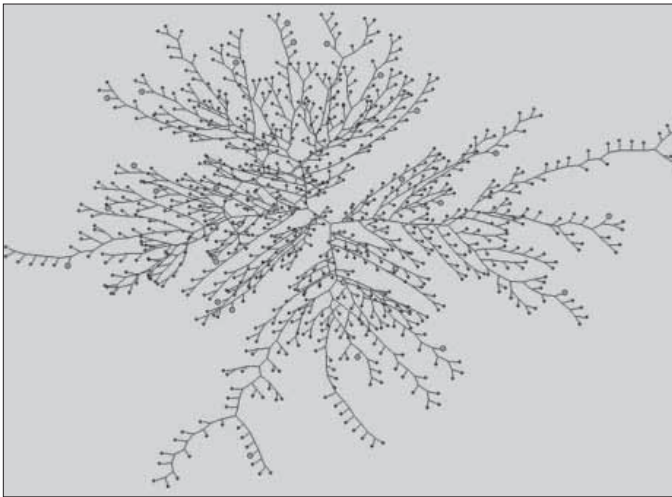


Figura 1.25. Estructura de comunitats de la xarxa de col·laboració de física estadística a Espanya. Les branques petites corresponen als grups de recerca que s'agrupen a les universitats que, alhora, estan estretament agrupades d'acord amb la proximitat geogràfica. Els nodes ressaltats, que corresponen als membres dels comitès científics es presenten principalment a les puntes de les branques, cosa que mostra el seu lideratge en els grups respectius. La distribució homogènia de nodes verds també mostra que han estat escollits de manera uniforme entre els diferents grups que formen la comunitat de física estadística i que aquests membres són els líders dels equips respectius.

1.6. Universalitat

En el camp de les xarxes complexes, a causa de l'interès en la disciplina de la física estadística des del començament, hi ha un concepte que ha acabat tenint un ús força generalitzat. Aquest concepte és el de la *universalitat*. Amb aquesta idea, a la física estadística s'entén que hi ha diferents sistemes físics que, malgrat tenir fenomenologies diferents, es poden caracteritzar mitjançant un conjunt petit de paràmetres equivalents. L'exemple paradigmàtic és el d'un sistema que experimenta una transició de fase, de manera que el que caracteritza aquesta transició és un únic paràmetre. En aquest sentit, quan sistemes físicament diferents són descrits pel mateix valor numèric del paràmetre diem que pertanyen a la mateixa "classe d'universalitat". Veurem en aquest capítol que aquest concepte d'*universalitat* s'estén a les xarxes complexes.

1.6.1. *Shortest paths i small worlds*

En les seccions anteriors hem definit el concepte de longitud en una trajectòria en una xarxa complexa. Entre qualsevol parell de nodes tenim definida una distància que és al llarg del que s'anomenen els *camins mínims* (*shortest paths*, en anglès). Aquesta és una mesura, doncs, característica de cada parell de punts.

També com hem dit abans, a partir d'aquestes mesures es pot definir una distància característica del conjunt de la xarxa, sigui una distància mitjana, un radi o un diàmetre. Si aquesta distància característica de la xarxa és un nombre petit en aquesta xarxa l'anomenarem de *món petit*.

Aquí el concepte *petit* és relatiu, perquè en el fons hem de comparar. Comparem-lo, per exemple, amb el cas d'una xarxa plana en la qual cada node està connectat només als seus quatre propers veïns. En aquestes circumstàncies, no calen fórmules matemàtiques gaire complicades per a veure que si tenim N nodes la distància característica serà de l'ordre de l'arrel quadrada de N . En aquest cas, la frase clau és "serà de l'ordre de...". No pretenem donar-ne un valor exacte, sinó una estimació. Vegem-ho amb un exemple: si tenim 100 nodes en una xarxa 10×10 esperem que la distància mitjana sigui de l'ordre de 10 (es podria trobar entre 8 i 12, però mai ser propera a 50). Estem fent un raonament purament qualitatiu que ens ofereix un ordre de magnitud, encara que no sigui un valor exacte.

Però el fet que en una xarxa de 100 nodes la distància característica sigui de l'ordre de 10 no vol dir que sigui "petita", ja que acabem de demostrar en

aquest cas que correspon a una xarxa reticular. Si aquesta xarxa, doncs, esperem que sigui un món petit, la seva distància característica haurà de ser bastant més petita que la que acabem de descriure.

Pensem en l'altre extrem. La mateixa xarxa de 100 nodes connectats de tal manera que en tenim un de central i tota la resta són perifèrics i formen una gran estrella. Quina és ara la distància característica? Tampoc no cal fer grans matemàtiques per a veure que serà de l'ordre de 2, i ara aquest valor de 2 sí que és molt més petit que 10, i podem dir que es tracta d'una xarxa veritablement "petita".

Però les xarxes reals no són d'una manera ni de l'altra, sempre ens movem entre els dos llindars descrits en els paràgrafs anteriors. Aleshores, si tenim una xarxa de 100 nodes i una distància característica de menys de 5 podem dir que es tracta d'una xarxa "petita".

Com ho generalitzem això a qualsevol tipus de xarxa? Doncs prenent sempre com a referència models reticulars, com l'anterior de dues dimensions, la distància característica és de l'ordre d'una potència (més petita o igual a 1) del nombre de nodes. Per a una xarxa en forma d'estrella i molt centralitzada, la distància mitjana és de l'ordre d'1 (no és funció del nombre de nodes).

D'una manera més rigorosa direm que una xarxa és un "món petit" si la seva distància característica és de l'ordre del logaritme del nombre de nodes. A continuació, presentem un conjunt de xarxes empíriques en què podem veure que efectivament es compleix aquesta llei de dependència en el logaritme del nombre de nodes de la xarxa.

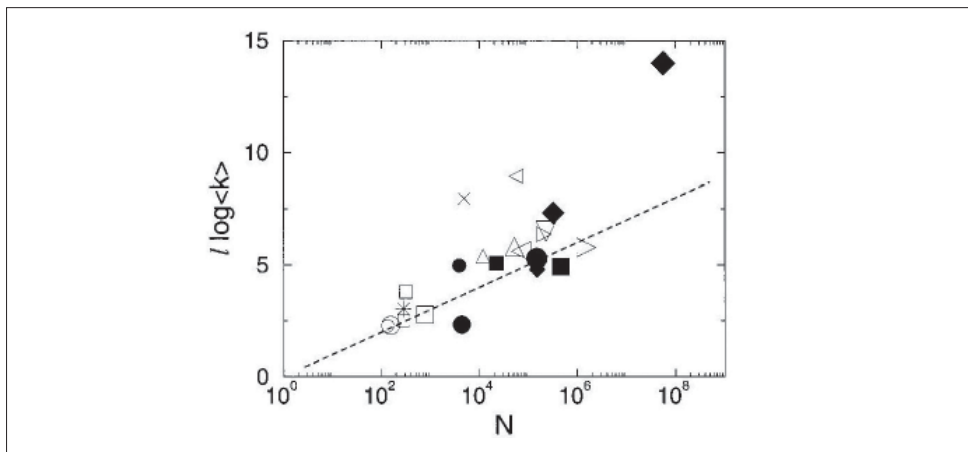


Figura 1.26. Distàncies mitjanes per a un conjunt de xarxes empíriques en funció del nombre de nodes. La línia de punts és $\log(N)$. El factor $\log \langle k \rangle$ és corrector i és el que mostra l'altra dependència, és a dir, com més alt sigui el grau mitjà, més curta és la distància mitjana, tal com lògicament es pot esperar.

Aquesta descripció matemàtica del que és un “món petit” va ser determinant en l’explosió de l’interès de la comunitat dels físics i matemàtics en el camp de les xarxes. Ara no entenem només un món petit com aquell que intuïtivament ens permet reconèixer que som molt a prop en un món connectat, sinó que, atesa l’estructura de la xarxa, el seu patró de connectivitats ens permet dir que veritablement és un món petit. Per tant, ara tornem al nostre exemple inicial de la IMDB amb més d’un milió d’actors i que la distància màxima entre qual-sevol parell d’actors, en termes de pel·lícules, és quatre. Matemàticament, el logaritme decimal de 10^6 és 6 i, per tant, això ja no ens ha de sorprendre.

1.6.2. Distribució de grau

En la secció dedicada a la caracterització a escala individual hem començat definint què és el grau d’un node. I aquest grau, que correspon al nombre de connexions d’un node, ens caracteritza la centralitat del node de manera local.

Si ara, en canvi, volem una descripció estadística de la xarxa i dels graus dels seus nodes, el que farem és un histograma d’aquests graus, és a dir, representar gràficament el nombre de nodes que tenen un grau concret. Si la xarxa és petita, considerarem el grau k com una variable discreta. A continuació, presentem l’histograma de graus de la xarxa de la figura 18.

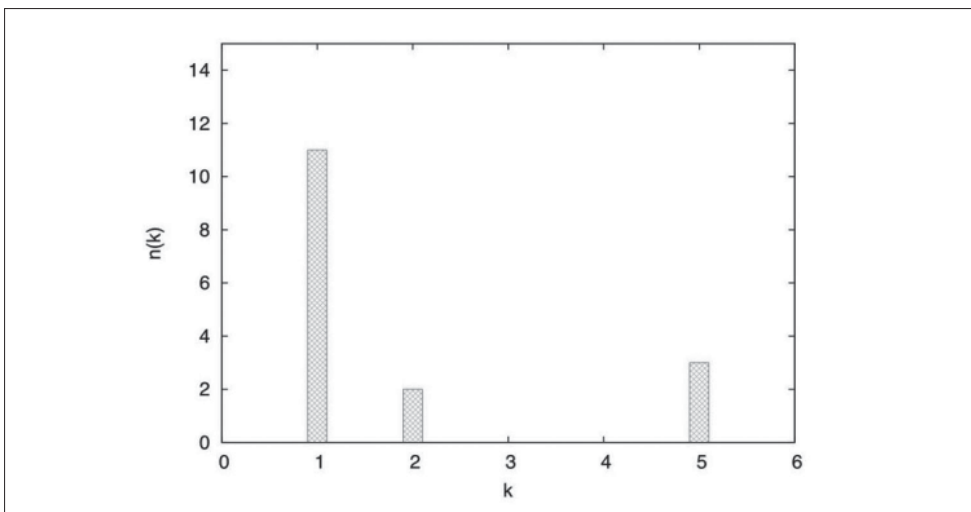


Figura 1.27. Histograma de graus per a la xarxa de la figura 18. Efectivament, tenim onze nodes amb grau 1, dos amb grau 2 i, finalment, tres amb grau 4.

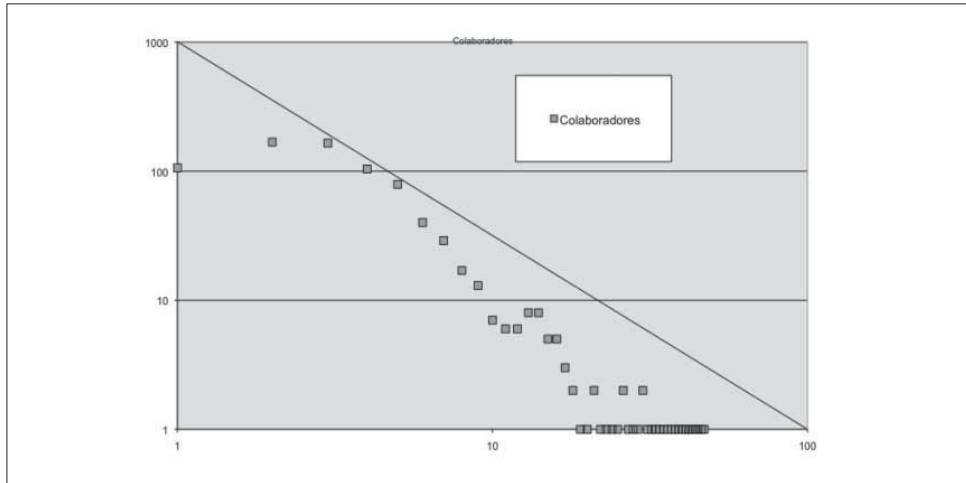


Figura 1.28. Distribució de graus (nombre de col·laboradors) en la xarxa de col·laboració de FisEs. L'escala és log-log, és a dir, representació logarítmica en els dos eixos.

En aquest tipus de figura podem veure, doncs, l'existència de nodes que tenen un grau més elevat que la resta. Ens dóna una idea estadística de la centralitat relativa d'alguns nodes, encara que no ens permeti identificar quins són, però sí conèixer-ne l'existència. En particular, en la segona gràfica podem veure que hi ha un petit nombre de nodes amb moltes col·laboracions i un gran nombre de nodes amb poques col·laboracions.

1.6.3. Lleis de potència i xarxes lliures d'escala

Pocs mesos després del treball pioner de Watts i Strogatz, en què posaven l'èmfasi en la característica de món petit de moltes de les xarxes observades, Albert i Barabási publiquen un article que complementa perfectament l'anterior. Efectivament, les dades sobre les xarxes que es van recollint de manera incessant a partir d'aquell moment són de món petit, però, a més, descobreixen que la distribució de graus no és del tipus que s'esperaria a partir de les "estadístiques" normals.

Què és una *estadística normal*? En els estudis estadístics hi ha una distribució de probabilitat que és tan "normal" que fins i tot té aquest nom. Es tracta d'una distribució que matemàticament té la forma següent:

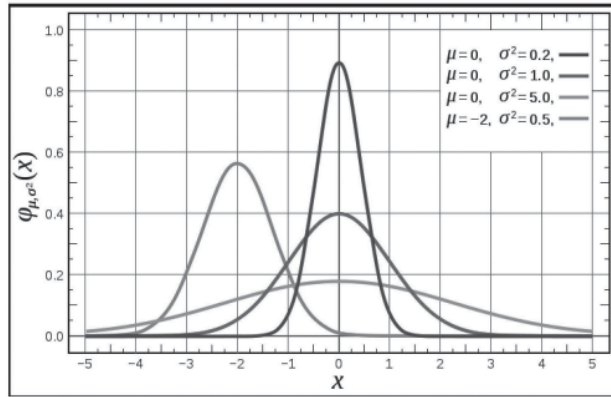


Figura 1.29. Exemples de distribucions “normals”

en què μ és el valor mitjà, al voltant del qual se centra la distribució, i σ és la desviació estàndard, que dóna una idea de l'amplària. I és tan normal que és la que permet explicar molta de la fenomenologia estadística que tenim al nostre voltant, des de la distribució d'alçades en una població fins a la distribució de notes en un examen. La característica principal d'aquesta distribució és que és molt picada al voltant del valor mitjà i decau molt ràpidament quan ens n'allunyem.

En canvi, les distribucions de grau que van trobar Albert i Barabási resulta que no s'ajustaven gens a la distribució normal, com podem veure en les figures que trobareu a continuació i com hem vist també en la distribució de graus en la xarxa de col·laboracions del FisEs.

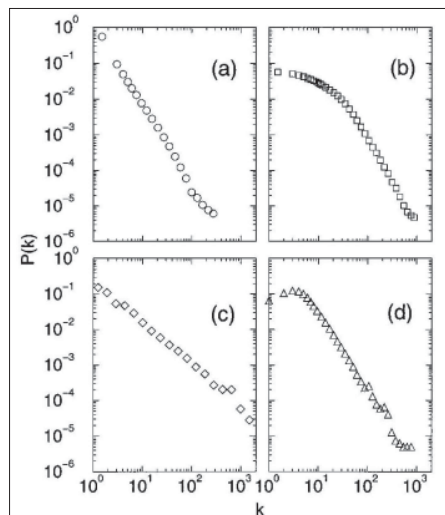


Figura 1.30. Distribucions de grau (normalitzades) per a algunes xarxes reals: a) Internet a nivell d'encaminadors; b) xarxa d'actors; c) xarxa de coautors en la comunitat de física de les altes energies; d) xarxa de coautors en neurociència.

Així, doncs, igual que la xarxa de FisEs, el que podem veure aquí és que en una gràfica log-log les distribucions són properes a una recta, mentre que si fossin una distribució normal decaurien molt més ràpidament, com demostrem en la figura 31.

En aquestes gràfiques hem representat tant una distribució normal com una distribució potencial (de les del tipus observat per Albert i Barabási). Aquí podem veure, sobretot en la gràfica de la dreta, en l'escala log-log, que la distribució normal decau molt més ràpidament que la potencial. De fet, per exemple, amb relació a la distribució d'alçades en una població el que esperem és una distribució picada al voltant d'un valor mitjà, una certa dispersió i un decaïment molt ràpid, perquè, simplement, no esperem en la nostra població la presència d'un individu de més de 2 metres i mig d'alçada.

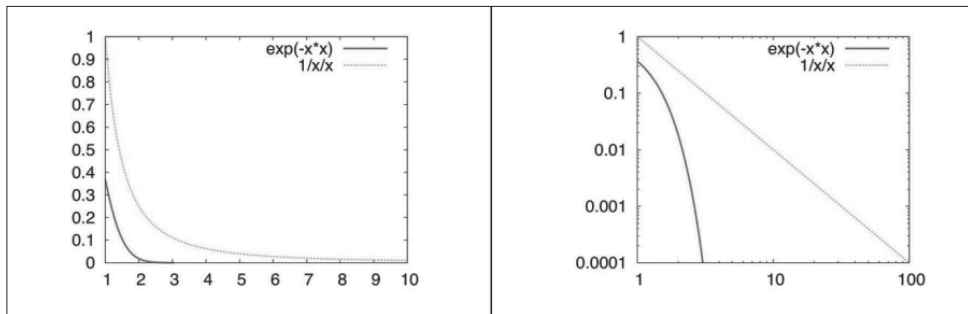


Figura 1.31. Distribucions normal (fosc) i potencial (clar), en escales lin-lin (esquerra) i log-log (dreta)

En canvi, en una llei de distribució potencial aquest decaïment és molt lent i, en termes dels graus, el que podem veure és que en aquestes distribucions hi ha molts nodes amb un grau molt baix, però hi ha un nombre considerable de nodes amb un grau que és d'un ordre de magnitud més gran que el valor mitjà. Què són aquests nodes? Òbviament són nodes amb una centralitat (local) molt elevada i que, per tant, esperem que tinguin un paper molt important en la xarxa; és el que en anglès anomenem *hubs*. I en les distribucions que presentem en la figura 30, en una escala log-log, aquestes s'apropen a una línia recta, i això vol dir que es tracta de distribucions que tenen forma de potència. De fet, si canvia la potència canvia el pendent.

Aquesta va ser una de les grans contribucions d'aquest article d'Albert i Barabási: descobrir que a les xarxes que trobem normalment al nostre voltant tenen una sèrie de nodes molt més connectats del que podríem esperar. Ens ha d'estranyar

això en les nostres xarxes? Per què alguns proveïdors d'Internet tenen moltes més connexions? Per què tantes i tantes pàgines del Web apunten a Google, al *New York Times* o a la Viquipèdia? Per què hi ha individus molt ben connectats en la nostra societat? Doncs no ens ha d'estranyar que la distribució de graus presenti una llei de distribució en forma de llei potencial amb la presència de *hubs*.

El fet que aquesta llei de distribució tingui forma potencial tampoc no ens ha de sorprendre des del punt de vista de la física estadística, ja que hi ha molts fenòmens en la física que estan relacionats amb distribucions potencials, com ara les distribucions de terratrèmols a la Terra o d'allaus en un sistema magnètic. En una escala log-log, les distribucions d'aquests esdeveniments són rectes d'una manera molt aproximada.

Reprement la discussió del començament d'aquesta secció, quan hem parlat del concepte d'*universalitat*, aquí cal dir que, en termes de les distribucions de grau, universalitat vol dir que la recta té el mateix pendent, ja que és l'únic paràmetre que la caracteritza. Sense importar l'origen físic de la xarxa en qüestió, podem dir que el mecanisme que la genera és el mateix (en veurem més detalls en la part de modelització) i que una xarxa com la de coparticipants en una pel·lícula o la d'encaminadors a Internet es generen amb principis microscòpics molt similars. En aquest llenguatge direm que pertanyen a la mateixa classe d'universalitat.

Normalment, les xarxes que presenten una llei de distribució de grau potencial com aquestes se les anomena *lliures d'escala* (*scale-free networks*, en anglès). Aquest nom prové del fet que les distribucions són rectes i no hi observem cap pic en un valor determinat que correspondria a un valor característic. En una distribució normal, tant el valor mitjà com la desviació estàndard tenen dimensions de la mateixa variable i, per tant, són dimensions característiques. En canvi, en una distribució potencial no apareix cap altra dimensió que no sigui la mida del sistema, i per això diem, doncs, que no hi ha escales característiques i que per tant és lliure d'escala.

1.7. Modelització de les xarxes

Per acabar aquesta part del text, presentarem els models que històricament considerem com a claus en el desenvolupament de la disciplina de les xarxes complexes i que expliquen cadascun dels aspectes observats empíricament. Aquesta modelització prové de la necessitat de proposar casos senzills que permetin explicar trets genèrics. No es pretén explicar amb un model matemàtic

senzill la xarxa d'actors ni la d'un ecosistema, sinó per quina raó aquestes xarxes presenten característiques comunes, com ara ser de món petit o el fet que la seva distribució de grau presenti una forma potencial.

1.7.1. Erdős-Rényi (ER)

Aquests dos matemàtics hongaresos van fer nombroses contribucions, a mitjan segle passat, al que en aquell moment s'anomenava teoria de grafs (bé, actualment en entorns més matemàtics es continua anomenant d'aquesta manera, malgrat que aquí utilitzem la variant més interdisciplinària de xarxa complexa). La qüestió que més ens interessa en aquest text és la del model senzill que van proposar, que és el que normalment anomenem *random graph d'Erdős-Rényi*.

En general, podem dir que un *random graph* és una xarxa (o graf) aleatoris, com el seu nom indica. El que distingeix un model particular són les seves caracteritzacions estadístiques que són, òbviament, una conseqüència de la mateixa construcció del graf.

Com es construeix una xarxa aleatòria de N nodes i M enllaços? Tenim els nombres, N i M , però clar ara el que cal és determinar quins són els nodes connectats pels enllaços i de quina manera es fa això determina les característiques estadístiques de la xarxa.

A partir d'aquesta idea el primer que podem pensar és, simplement, a assignar els M enllaços de manera totalment aleatòria, i sense cap altra restricció, als N nodes existents. El resultat seria un graf, que anomenarem $G(N,M)$. Però ens ha de quedar clar que el graf resultant no és res més que una de les possibles realitzacions de graf aleatori de N nodes i M enllaços. En tractar-se, doncs, d'un model hem de tenir en compte que hem de generar un conjunt de grafs amb les mateixes característiques perquè tingui sentit estadístic. En la figura que presentem a continuació hi ha tots els possibles grafs amb quatre nodes i quatre enllaços.

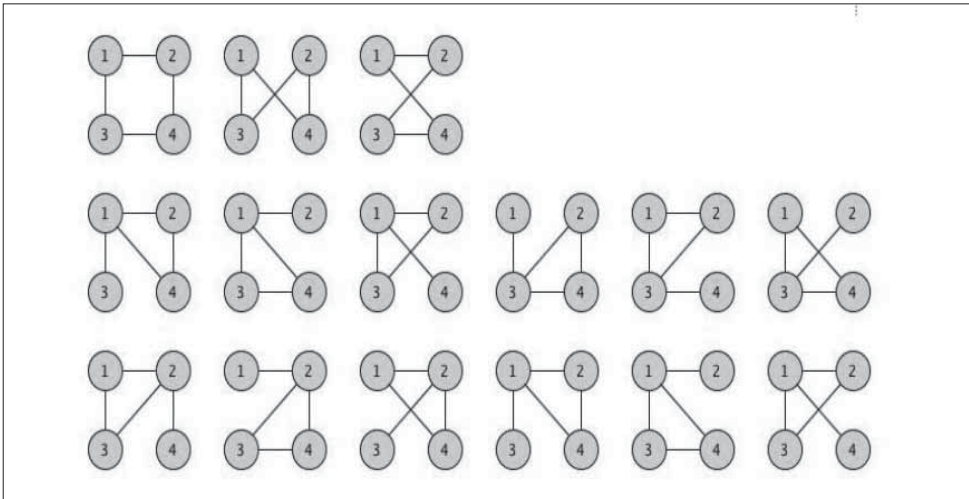


Figura 1.32. Tots els possibles grafs de quatre nodes i quatre enllaços. Els tres de la part superior corresponen als grafs en què tots els nodes tenen exactament un enllaç. I els dotze de la part inferior, als casos en què hi ha un node amb tres enllaços, dos amb dos, i un amb un. Combinatòriament, podem dir que de tots els possibles enllaços entre quatre nodes ($4 \times 3/2 = 6$, perquè el graf no és dirigit) els agafem de quatre en quatre; així, els possibles grafs seran

$$g_{4,4} = \binom{6}{4} = 15$$

Així, doncs, en cas que vulguem analitzar qualsevol característica, per exemple el diàmetre, el que haurem de fer és calcular la mitjana de tots els possibles grafs compatibles amb el nombre de nodes i enllaços. En el cas que hem representat, però, no cal comptar tots els casos perquè, d'independents, només n'hi ha dos. Podríem dir que els tres de la filera superior pertanyen a una classe (A formada per tres elements), ja que són topològicament equivalents, i la resta a una altra classe (B formada per dotze elements). El diàmetre mitjà seria, doncs:

$$\langle D \rangle = \frac{1}{15} (3 * D(A) + 12 * D(B))$$

encara que en aquesta cas la resposta és trivial. Però generalitzar-ho a un nombre arbitrari de nodes i enllaços no és gens trivial.

Aquest model que hem anomenat d'*Erdős-Rényi* admet el que podríem dir una segona representació. Una representació alternativa, que s'expressa com a $G(N,p)$, en què ara el que fixem és també el nombre de nodes, però l'altre paràmetre p és la probabilitat que existeixi cada enllaç i és la mateixa per a tots els enllaços.

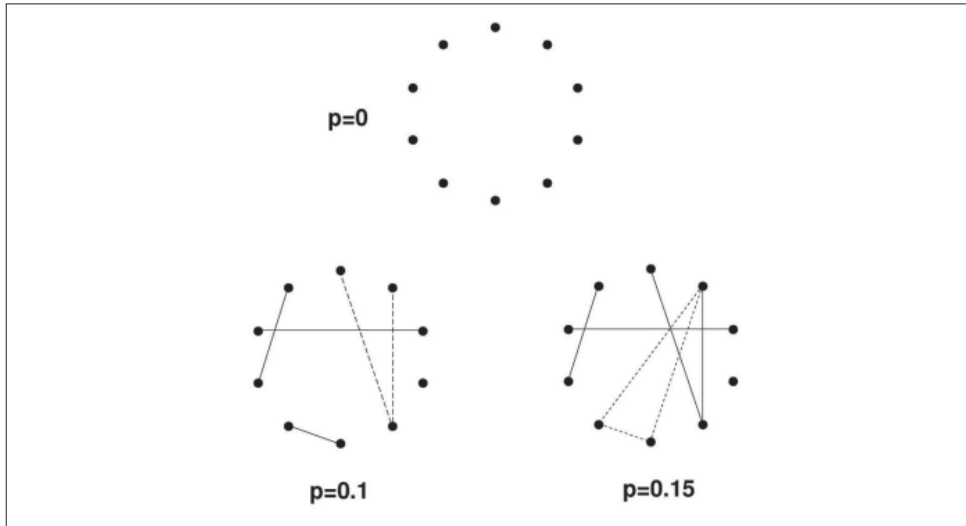


Figura 1.33. Model de graf aleatori d'Erdős-Rényi $G(N, p)$. En una xarxa de deu nodes tenim en principi $10 * 9/2 = 45$ enllaços i cadascun d'aquest existeix, de manera independent, amb una probabilitat p .

GEPHI

Amb aquest programari podem generar grafs aleatoris d'Erdős-Rényi en qualsevol de les dues representacions.

En aquest cas, no fixem el nombre d'enllaços i novament haurem de fer una mitjana sobre els possibles casos. La diferència respecte de l'anterior és que ara no tenim un nombre fixat i tancat d'elements de la mostra estadística, sinó que només tenim una probabilitat que existeixi cadascun dels enllaços. Així, la probabilitat que aparegui un graf del tipus anterior determinat és:

$$P(G(N, M)) = p^M (1 - p)^{\binom{N}{2} - M}$$

Aquesta segona representació dels grafs aleatoris és útil perquè molts dels resultats exactes que es coneixen per a aquest tipus de xarxes s'ha obtingut a partir d'aquesta representació. En particular, atesos N i p és fàcil calcular-ne la distribució de grau, ja que es tracta d'una distribució binomial:

$$P(k_i = k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

i que en el cas d'un nombre de nodes gran, es redueix a la distribució de Poisson:

$$P(k) \approx e^{-pN} \frac{(pN)^k}{k!} = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

que, per a valors grans del grau, decau de manera exponencial i, per tant, molt més ràpidament del que hem observat abans en moltes xarxes complexes.

D'altra banda, també es pot demostrar analíticament que el diàmetre de la xarxa escala com el logaritme del nombre de nodes, i mostrar, així, doncs, la característica de món petit.

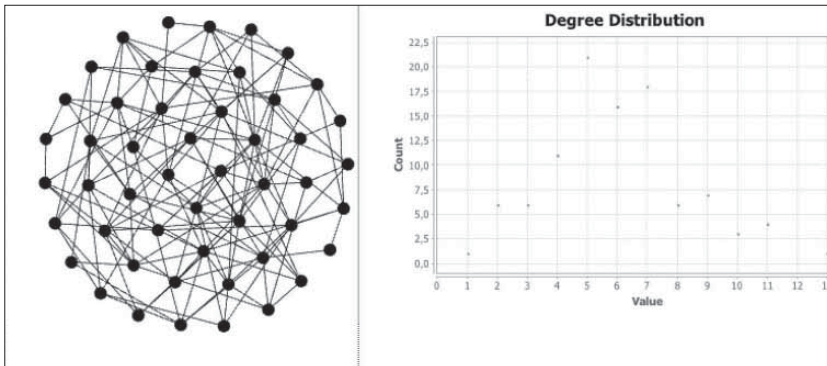


Figura 1.34. Xarxa construïda amb GEPHI seguint el model d'ER amb 100 nodes i 300 enllaços. A la dreta: histograma de la distribució de graus.

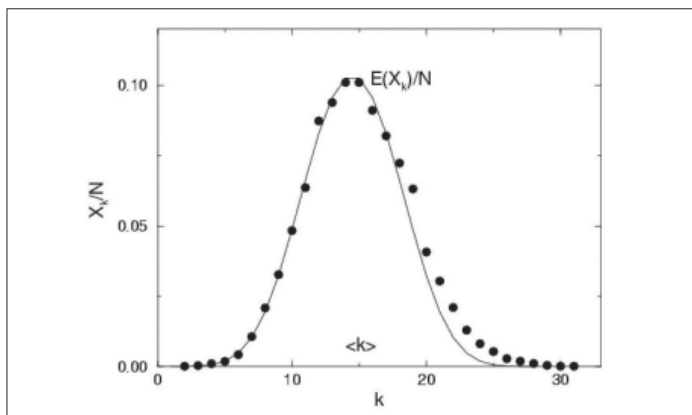


Figura 1.35. Distribució de graus en el model d'Erdős-Rényi. La línia contínua representa la distribució de Poisson.

1.7.2. Watts-Strogatz (WS)

En les seves observacions de les xarxes complexes discutides en apartats anteriors Watts i Strogatz van veure que aquestes xarxes tenien dues característiques comunes: primera, que eren de món petit, és a dir, que les distàncies mitjanes creixen amb el logaritme del nombre de nodes, i segona, que el coeficient de *clustering* era més gran de l'esperat.

Aquestes dues propietats en el fons són relatives. I ens ho hem de mirar en la perspectiva de l'època. Al final del segle XX només es parlava de dos tipus d'estructures topològiques discretes (anomenades xarxes només en entorns de les ciències socials). D'una banda, estructures reticulars, com ara un anell en dimensió 1 o una malla quadrada en dues dimensions. De l'altra, estructures totalment aleatòries, com ara el model de graf aleatori d'ER.

I aquests dos models presenten característiques ben diferents. D'una banda, els models reticulars, com hem vist abans, tenen distàncies mitjanes que creixen amb una potència del nombre de nodes i , en canvi, el seu coeficient de clusterització és força elevat perquè per construcció la densitat local és elevada. D'una altra banda, el model ER és un model de món petit perquè el diàmetre creix com el logaritme del nombre de nodes i la clusterització és molt baixa, de fet exactament igual a p .

Aleshores, Watts i Strogatz es troben amb la paradoxa que les seves observacions de petites distàncies i clusteritzacions altes no es pot explicar amb cap de les dues estructures que acabem de relatar. El model reticular té distàncies grans i *clustering* elevada, mentre que el model d'ER és justament el contrari, distàncies petites i *clustering* baixa.

La seva proposta va ser la d'un model que podia explicar els dos fets empírics simultàniament. D'una banda, era prou aleatori perquè les distàncies fossin petites i, de l'altra, perquè mantingués certa estructura regular per tal de mantenir una *clustering* elevada. Gràficament, el representem en la gràfica següent.

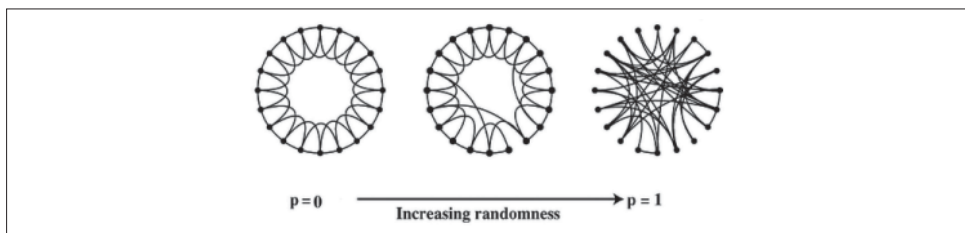


Figura 1.36. Model de Watts i Strogatz. Es parteix d'un anell amb connexions a veïns propers i també als veïns dels veïns. Aleshores, amb probabilitat p es canvia un enllaç de la xarxa reticular original per un enllaç aleatori. Per a valors petits de p el graf és pràcticament regular, mentre que per a p molt propera a 1 el graf és un graf d'ER.

I què guanyem amb aquest nou model de xarxa? La resposta és clara en la gràfica següent.

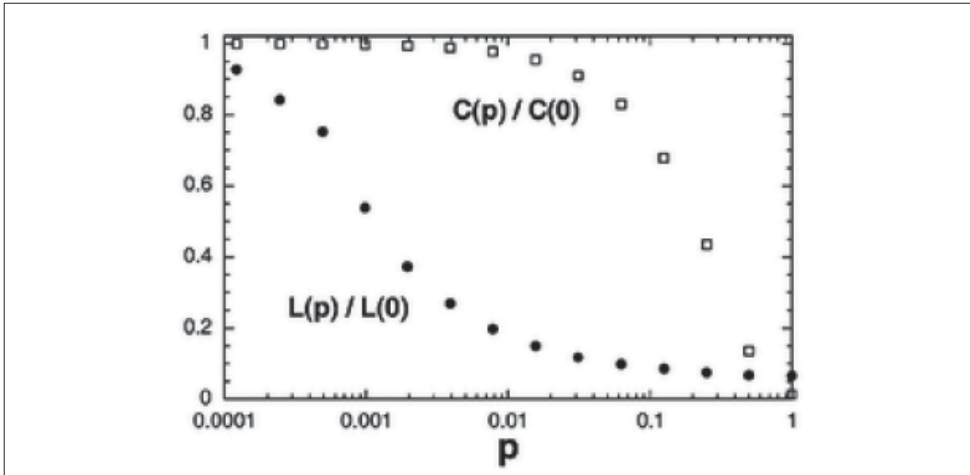


Figura 1.37. Coeficient de *clustering* i distància mitjana entre nodes per al model de WS. Sempre es pren com a referència el valor en $p = 0$. El que hem de destacar és que petits increments de p a partir del 0 fan que decreixi ràpidament la distància, mentre que el *clustering* resta invariant.

Podem veure, doncs, que per a valor de p de l'ordre de 0,01 tenim distàncies petites, mentre que la clusterització continua sent elevada i aquesta és, doncs, la regió en què cauen bona part de les xarxes descrites en la primera part d'aquesta part del text.

Aquest model va ser el punt de partida de l'estudi i modelització de les xarxes complexes. La simplicitat rau en els dos extrems, tant en la topologia regular, reticular, ordenada, com en la totalment desordenada. Aquesta simplicitat que permet càlculs matemàtics relativament senzills. Però el nostre món no és simple. I les xarxes que observem al nostre voltant tampoc. Són complexes i aquesta complexitat la podem entendre a partir d'aquest model com aquest estat intermedi entre el que està ordenat i el desordenat. I els càlculs, malauradament, ja no són tan senzills.

1.7.3. Barabási-Albert (BA)

Hem dit en apartats anteriors que la segona contribució que és considerada com la llavor de l'estudi de les xarxes complexes és el treball de Barabási i Albert que bàsicament descobreixen que xarxes provinents d'origens tan diferents tenen comportaments semblants, en particular ells es van fixar en la distribució de grau, que era potencial.

Bé, aquestes xarxes que analitzen Barabási i Albert són efectivament xarxes de món petit perquè compleixen la propietat que les distàncies són petites i, a més, les clusteritzacions són elevades. Per tant, podríem pensar que es podrien explicar amb el model de WS. Però el model de WS, tot i que explica bé les distàncies i el *clustering*, com hem vist abans, té encara una distribució de graus que és la de Poisson. Per tant, aquest model no serveix per a explicar les distribucions potencials de grau.

La seva proposta rau en el fet que les xarxes apareixen per un procés de creixement. Les xarxes no són productes que existeixen simplement com a agregats, sinó que són el resultat d'una evolució que va afegint nodes a la xarxa, com pot ser qualsevol de les xarxes que hem analitzat prèviament, per exemple, les de coautors, Internet o el WWW. Però la manera d'afegir nous nodes a la xarxa és el que no és aleatori, sinó que compleix una regla, senzilla però que podem dir que té a veure amb la popularitat, entesa de manera molt general. Per exemple, si nosaltres creem una pàgina web, molt probablement afegirem enllaços a altres pàgines perquè són populars, com ara cercadors, el nostre diari favorit o la institució a la qual pertanyem.

Aleshores, Barabási i Albert consideren que la xarxa creix en el sentit que els nodes que s'hi afegeixen tendiran a enllaçar-se amb nodes ja existents a la xarxa però que ho faran amb els nodes que ja tenen un grau elevat (en aquest cas, el grau seria una mesura de la popularitat del node). I aquesta regla és la que anomenen de *preferential attachment* (és a dir, d'afegiment preferencial) i la quantifiquen de manera purament lineal, és a dir, que la probabilitat que un nou node estableixi un enllaç a un node existent és directament proporcional al grau del node ja present a la xarxa.

Seguint, doncs, aquesta regla es poden construir xarxes complexes que siguin a la vegada de món petit, d'alta clusterització i amb distribució de grau potencial, com la que tenim en la figura següent.

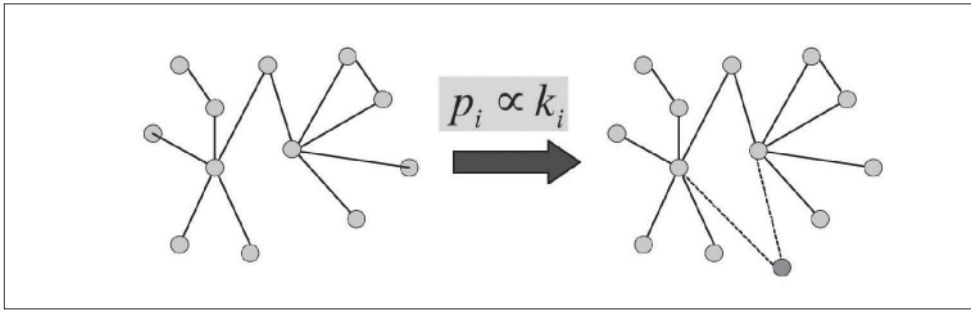


Figura 1.38. Regla del model de Barabási-Albert. El node que arriba es connectarà amb més probabilitat als nodes que ja tenen més connexions. Així, els nous enllaços són els que corresponen a les línies discontinües.

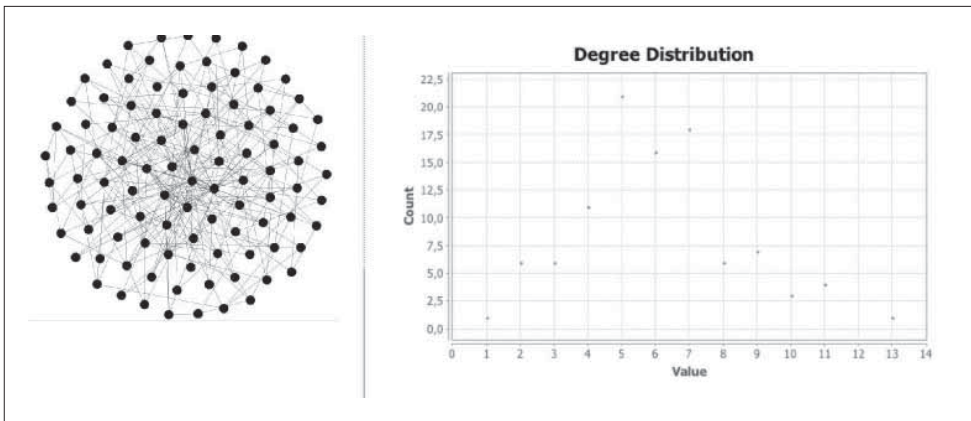


Figura 1.39. Xarxa construïda amb Gephi seguint el model de BA amb 100 nodes i 294 enllaços. A la dreta: histograma de la distribució de graus.

Podem veure en la construcció de la xarxa que, comparada amb l'ER anterior amb el mateix nombre de nodes i pràcticament el mateix nombre d'enllaços, ara aquesta xarxa té una densitat d'enllaços més elevada a la part central. Aquests serien els *hubs*, els nodes amb una connectivitat més elevada, com podem veure en l'histograma de la part dreta de la figura.

Des d'un punt de vista matemàtic, es pot demostrar de manera relativament simple que la distribució de grau associada al model de BA és potencial amb exponent -3 , que la fa de valor molt proper al de les observades empíricament.

1.7.4. Altres models

Com és fàcil d'imaginar, en aquests anys des dels primers models de WS i de BA hi ha hagut multitud de models que han permès explicar bona part de la fenomenologia associada a les xarxes complexes.

Des de models que donen exponents molt més propers als observats, o fins i tot amb un exponent ajustable, que admeten estructures de comunitats, que presenten correlacions entre els graus d'un node i els dels seus veïns.

Aquests models es troben, però, allunyats de l'objectiu d'iniciar els lectors en el món de les xarxes complexes, que és el d'aquest text.

Capítol II

Anàlisi de xarxes socials

José Luis Molina i Beatriz Patraca

Sempre hem viscut en un món socialment connectat. El comerç a llarga distància està documentat des de la Prehistòria, i també els contactes estables entre pobles llunyans, moltes vegades mitjançant socis que feien de ponts entre cultures. Ja a l'antiguitat clàssica trobem nombrosos exemples d'expedicions llunyanes i visitants. Així, per exemple, Xenofont (segle IV-III aC) relata a l'Anàbasi la retirada des de Pèrsia fins a la mar Negra d'un exèrcit grec compost per soldats de múltiples ètnies al llarg de més de 4.000 quilòmetres, o la carta de Plini el Jove a Tàcit per comunicar la mort del seu oncle per l'erupció del Vesubi que va arrasar Pompeia (79 dC), en la qual esmenta que tenia un hoste vingut... d'Hispania.

Disposem de nombrosos documents de l'Edat Mitjana en què s'aprecien mons locals connectats a llarga distància per enllaços matrimonials i comercials, i així successivament fins als nostres dies. La diferència principal resideix que en l'actualitat les interaccions socials queden registrades electrònicament (posicions geogràfiques dels telèfons mòbils, SMS, correus electrònics, plataformes de *social networking*) i és possible visualitzar i gestionar aquestes relacions.

En aquest capítol aprendrem com va sorgir formalment l'anàlisi de les xarxes socials i com s'aplica en diferents camps de recerca social. Com s'ha apuntat al capítol anterior, veurem que l'expressió *Que petit és el món!* és molt més que una frase feta. Finalment, coneixerem com les comunicacions assistides per ordinador han repercutit en la nostra manera de relacionar-nos i com ens permeten visualitzar mapes socials que descriuen les nostres interaccions.

Però abans de tot això, començarem per aprendre què són les xarxes socials i com es constitueixen.

2.1 Què són les xarxes socials?

En l'actualitat, quan usem l'expressió “xarxes socials” o “les xarxes socials” en el llenguatge quotidià es fa referència a llocs digitals per a la creació de relacions socials, popularment coneguts amb el seu nom en anglès: *social networking sites* (SNS). Els SNS es refereixen a plataformes com Facebook i Twitter, entre d'altres. En ser eines ràpides i flexibles de comunicació i de socialització, els mitjans d'informació, alguns líders d'opinió i els partits polítics usen aquests canals de manera intensiva, la qual cosa n'explica també la ràpida popularització.

Vegem per exemple en la figura 1 la representació de 84.397 piulades (*tweets*) enviades durant les eleccions al Parlament de Catalunya del 2010 (Congosto i altres, 2011):

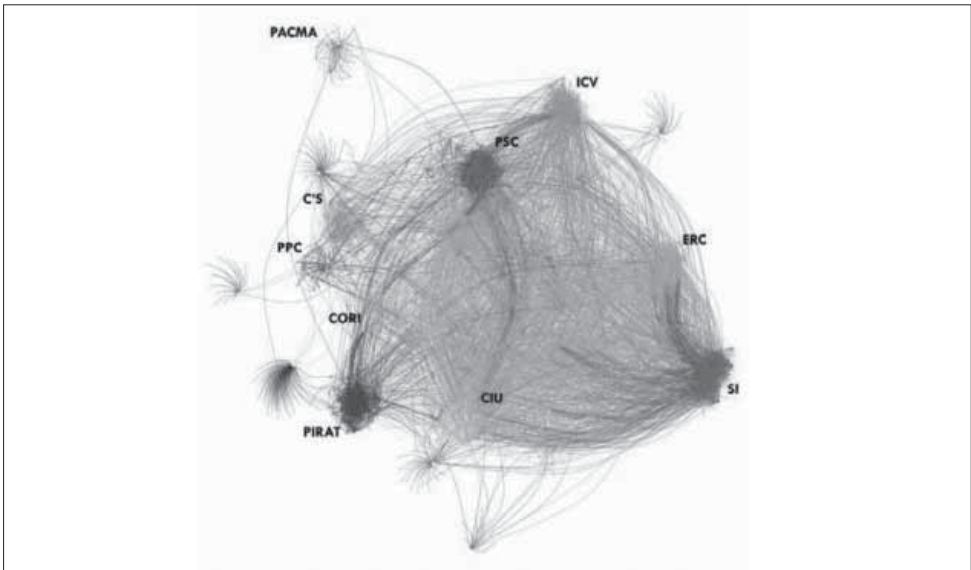


Figura 2.1. Xarxa de piulades (*tweets*) enviades durant les eleccions al Parlament de Catalunya, 2010

S'hi pot veure que hi ha moltes relacions dins de cada partit o formació política, i també nombroses connexions entre grups. A primera vista es pot apreciar que som davant d'una xarxa complexa, amb “comunitats” més o menys properes entre si (per exemple, PSC-ICV, PPC-Ciutadans, o ERC-SI), però amb connexions múltiples que fan del sistema “un món petit”, com veurem més endavant. Sens dubte, quan es fa referència a les xarxes socials se suggereixen realitats semblants a aquesta.

Així, la popularització del terme “xarxes socials” com a *social networking sites* ha reemplaçat l'accepció acadèmica de les xarxes socials com el mesurament i l'anàlisi de les múltiples estructures socials en les quals es desenvolupa la vida social, o “anàlisi de xarxes socials”. L'accepció acadèmica neix als anys trenta del segle XX, quan Jacob Moreno i altres van començar a representar el món social amb punts i fletxes i es va prendre la xarxa com a model de representació de la interacció social. Aquesta perspectiva, que intentava agregar els entramats de relacions a l'anàlisi de les conductes i les característiques dels individus i grups, es va desenvolupar durant el segle XX i va inspirar al començament del segle XXI les primeres aplicacions de *social networking sites*, com Friendster i Myspace, els quals van utilitzar el corpus de coneixement acumulat pels investigadors de xarxes socials fins al moment.

Encara que ambdues accepcions són correctes en l'actualitat, hem de distingir entre una i una altra quan sigui oportú. En tot cas, totes dues es refereixen a relacions entre actors.

Les xarxes socials són conjunts d'actors (persones, organitzacions, famílies, països, etc.) que es vinculen entre si per mitjà de relacions.

Una manera de visualitzar aquestes xarxes és gràcies als grafs, que ens presenten de manera ordenada aquests actors i les seves relacions.

Per a la seva anàlisi, podem dividir les xarxes socials en dos grans grups. D'una banda, hi ha les xarxes personals, o el conjunt dels contactes actius d'una persona en tots els àmbits de la seva vida social. D'una altra, tenim les xarxes completes o xarxes sociocèntriques, que són aquelles que ens presenten el conjunt d'un tipus de relacions entre tots els actors del grup.

2.1.2. Xarxes personals

Penseu en la gent que conegueu pel nom i que us coneguin a vosaltres. Intenteu completar una llista... diguem-ne... d'una dotzena de persones a la taula de la pàgina següent abans de continuar llegint.

Observeu que les primeres que us vénen a la ment són persones importants per a vosaltres o bé persones amb les quals teniu una interacció freqüent o recent. Ara intenteu classificar aquestes persones en àmbits o contextos de relació (família, feina, estudis, oci...). Algunes persones es coneixen, però unes

altres solament es connecten a través de vosaltres. En aquest cas cadascun de vosaltres sou l'ego i les persones que heu nominat són els vostres alteri.

1	7
2	8
3	9
4	10
5	11
6	12

Si representem la xarxa de relacions, ego està connectat a tots els nodes per definició. Per això, se sol representar el graf amb les relacions sense mostrar les relacions d'ego amb els seus *alteri* com podeu veure en la figura 2.

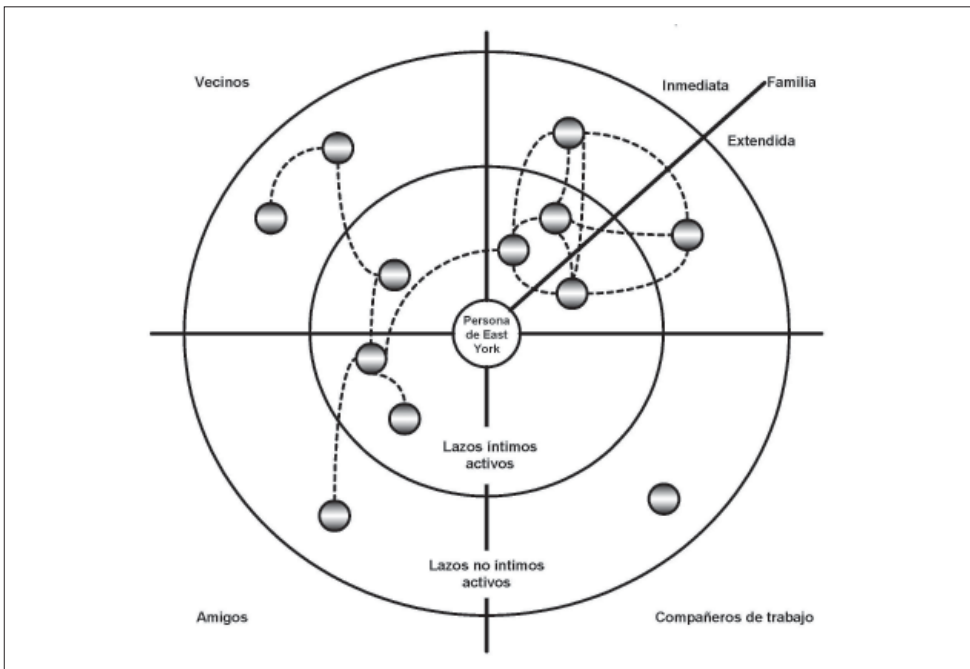


Figura 2.2. Xarxa personal típica d'una persona de Toronto (adaptat de Wellman 1999)

Observeu també que la xarxa personal es divideix entre un nucli de llaços íntims, normalment petit, que es coneixen entre ells i que es mantenen al llarg del temps, i una perifèria de llaços "no íntims" o coneguts, sovint desconec-

tats d'altres alteri d'ego. Aquesta perifèria sol ser molt més gran i canvia ràpidament, encara que en el gràfic solament se'n mostra un esquema ideal o reduït.

Així, les xarxes personals es componen de tots els tipus de relacions significatives per a ego, representant el conjunt dels contextos institucionals en els quals actua.

2.1.3. Xarxes sociocèntriques o completes

La figura 3 ens mostra una xarxa sociocèntrica o “completa” que assenyala els intercanvis verbals professionals entre diferents membres d'un equip de rodatge de cinema. En aquesta xarxa observem com es vinculen els diferents actors entre si (productor, ajudant del director, guionista, actors, maquilladors, etc.) per a la realització d'una escena cinematogràfica.

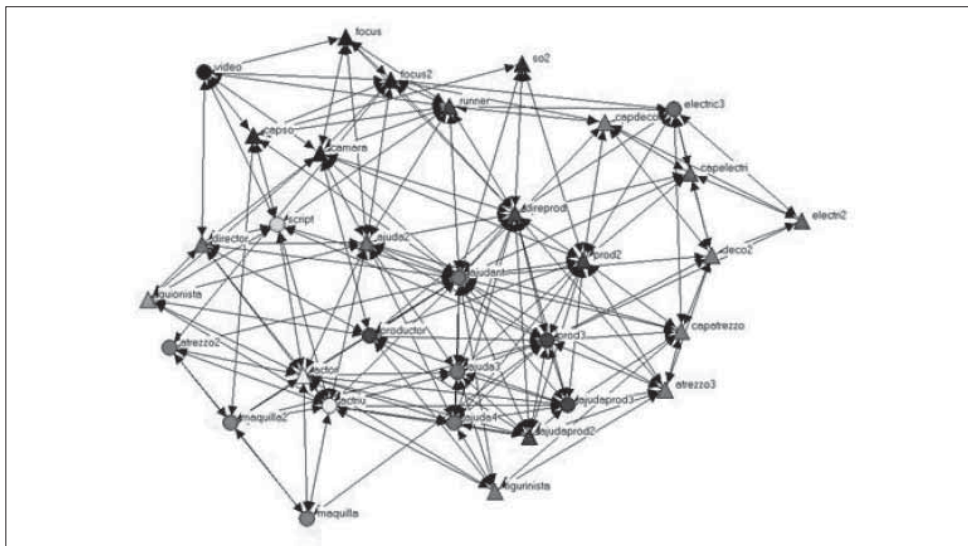


Figura 2.3. Xarxa d'intercanvi verbal durant la gravació d'una escena de cine (Muntanyola i Lozares, 2004)

Aquesta xarxa “completa” es refereix a un tipus de relació, el del treball de cine en aquest cas. Cada ego té altres “xarxes completes” de familiars, de veïns, amics, etc., que no es capturen amb aquesta xarxa.

Ara bé, dins d'aquesta xarxa completa és possible distingir xarxes “egocèntriques”, o el subgraf de cada node amb el seu veïnat o els altres nodes amb els

quals està directament connectat. En la figura 4 es pot veure una xarxa sociocèntrica amb 10 nodes (a dalt) i una de les 10 egonetworks que componen la xarxa.

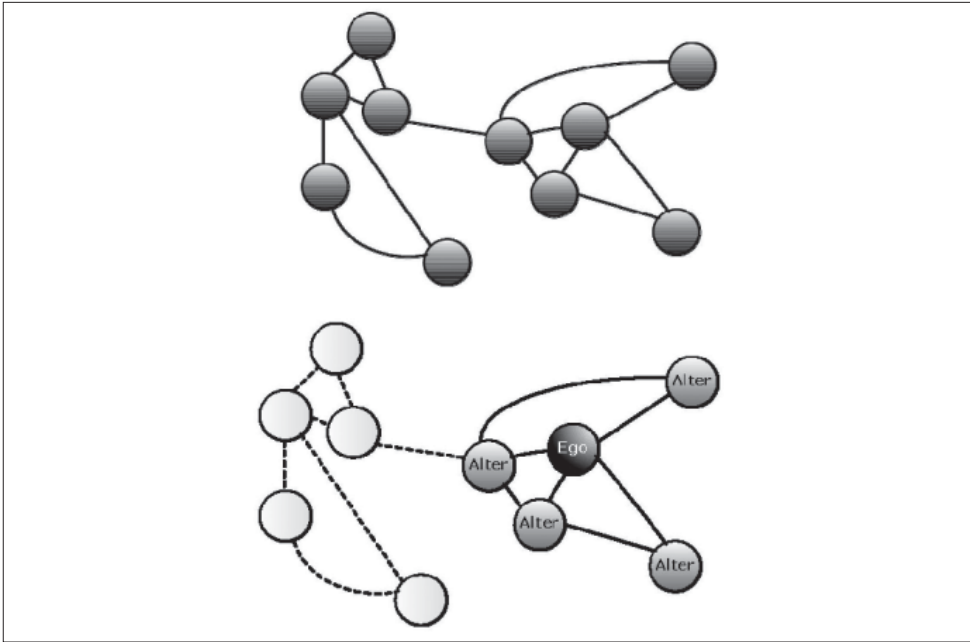


Figura 2.4. Xarxa completa (a dalt) i una de les 10 xarxes egocèntriques que la componen (a baix).

L'estudi de les xarxes egocèntriques per separat és interessant perquè pot veure's en l'exemple que alguns nodes o egos tenen més relacions que uns altres i que algunes d'aquestes relacions poden unir grups o no.

2.2. Breu història de l'anàlisi de xarxes socials

En l'actualitat, l'estudi de les xarxes socials es beneficia dels avenços de la teoria de xarxes descrita en el primer capítol, tant des d'un punt de vista teòric com metodològic i instrumental, i constitueix un camp veritablement transdisciplinari. En aquest apartat revisem les contribucions fetes des de les ciències humanes i socials a aquest camp durant gran part del segle xx fins a l'actualitat.

2.2.1. La sociometria

Les primeres idees sobre la importància de les estructures de relacions entre individus es troben en el corrent de la psicologia de la Gestalt (Scott, 1991). Kurt Lewin, Jacob Moreno i Fritz Heider, pensadors que van fugir del nazisme i es van establir als Estats Units, van intentar formalitzar l'estudi de les relacions socials en grups petits, iniciant aquesta línia de recerca.

Lewin va proposar la formulació matemàtica del concepte de distància social. Per la seva banda, la proposta de Heider es va basar en la idea que les relacions socials es podien analitzar com a tríades, o triangles de relació entre tres persones, formalitzant unes relacions que tenien influència més enllà de les relacions entre dos individus. Però l'impuls fonamental va venir de la idea de Jakob Moreno de representar les relacions socials com a punts i fletxes, la qual cosa feia possible la formalització de les relacions mitjançant la teoria matemàtica de grafs. Abans de Moreno ja es parlava de teranyines, teixit social i ocasionalment, xarxes de relacions, però aquesta metàfora no s'havia sistematitzat en un diagrama analític. Moreno va fundar el que creia que era una ciència nova, la sociometria, amb la visionària idea que una ciutat com Nova York podria ser estudiada a partir de les seves relacions (Moreno, 1934).

Per a Moreno les configuracions socials eren estructures definides i perceptibles. El mapatge d'aquestes estructures en un sociograma permet a un investigador visualitzar els canals pels quals, per exemple, la informació podria influir en un altre. Moreno va sostenir que la construcció de sociogrames ha permès als investigadors identificar els líders i els individus aïllats per a descobrir factors com l'asimetria i la reciprocitat i les cadenes que existeixen en els mapes de connexió (Scott 1991, pàg. 10).

Jakob Levy Moreno, a més de ser l'iniciador de la sociometria, també va idear el psicodrama per aplicar-la com a psicoteràpia. El psicodrama consisteix en una improvisació teatral en grup, i permet tant el diagnòstic com la resolució del conflicte del pacient. També va ser pioner de la psicoteràpia grupal.

2.2.2. Els investigadors de Harvard i Chicago

De 1927 a 1932, es van dur a terme els estudis de Hawthorne en una fàbrica de la Western Electric Company de Chicago. L'objectiu d'aquests treballs era conèixer la influència en el rendiment i la productivitat dels obrers a causa de factors físics com la il·luminació, les pauses, l'horari dels àpats, així com els factors

socials i psicològics que es recollien en entrevistes i enquestes. W. Lloyd Warner, de la Universitat de Chicago i alumne de Radcliffe-Brown i Elton Mayo, de la Universitat de Harvard, van col·laborar en la realització d'aquests estudis, que van durar diversos anys i que van ser de gran interès per a la teoria de l'organització, la psicologia social, la sociologia i l'antropologia. En el que concerneix a l'anàlisi de xarxes, va ser rellevant la identificació de subgrups en el sistema de relacions.

Els estudis Hawthorne van incloure diversos sociogrames elaborats per l'equip d'investigació. Aquests sociogrames no mostraven les relacions formals que es representen en l'organigrama de gestió, sinó que reflectien les relacions informals existents entre els treballadors. Els estudis Hawthorne van ser la primera gran recerca en la qual es van usar els sociogrames per a descriure les relacions socials en situacions reals, encara que sense citar en aquests diagrames el treball de Moreno. Els subgrups que van identificar els investigadors van ser denominats cliques, petits conjunts de treballadors amb llaços informals en els quals ells mateixos reconeixien els elements que els formaven.

Un clique és un subgraf completament connectat de dos o més nodes. Un node pot pertànyer a diversos cliques.

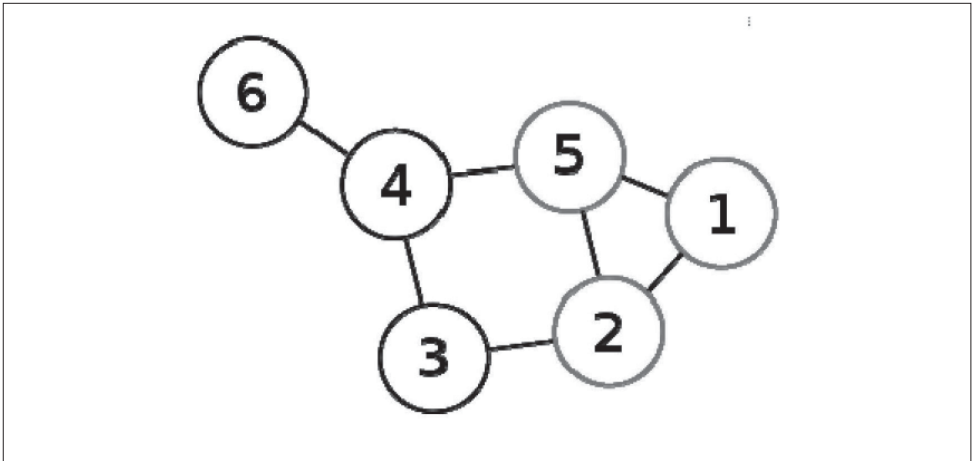


Figura 2.5. En aquesta figura podem observar com 1, 2 i 5 formen un clique dins de la xarxa.

Durant els estudis a la fàbrica de la Western Electric Company, es va descobrir l'anomenat efecte *Hawthorne*, que és la reacció psicològica per la qual els subjectes observats en un experiment tendeixen a canviar les conductes pel simple fet de saber-se observats.

Després de l'experiència de Hawthorne, Warner va tornar a Chicago i va iniciar un estudi sobre comunitats del sud dels Estats Units, continuades per Davis i col·laboradors, aplicant l'anàlisi de xarxes a l'estudi d'una societat complexa (Old City, Davis i altres, 1941). En aquestes recerques es van usar matrius per a la recollida de dades amb l'objectiu de relacionar persones i esdeveniments (i no solament persones entre elles). Aquestes matrius reben el nom de matrius d'afiliació. Així, en la figura següent es pot veure el patró d'assistència a festes i esdeveniments socials (columnes) d'un grup de dones (files). La matriu inferior mostra com reordenant les files i les columnes és possible identificar agrupacions de dones (figura 6).

	1	2	3	4	5	6	7	8
Ann	•		•		•		•	
Beth		•		•		•		•
Chris	•		•		•		•	
Don		•		•		•		•
Ed	•		•		•		•	
Flo		•		•		•		•
Gilll	•		•		•		•	
Hal	•		•		•		•	

	1	3	5	7	2	4	6	8
Ann	•	•	•	•				
Chris	•	•	•	•				
Ed	•	•	•	•				
Gill	•	•	•	•				
Beth					•	•	•	•
Don					•	•	•	•
Flo					•	•	•	•
Hal					•	•	•	•

Figura 2.6. Reordenament d'una matriu d'afiliació (pres de Molina, 2001)

Per la seva banda, George C. Homans va sistematitzar els avenços fets fins al moment en psicologia, sociologia i antropologia en el camp de les relacions en el seu llibre *El grupo humano* (1950). Homans conceptualitza el grup humà com una unitat basada en les relacions cara a cara que s'estructura internament i externament a partir dels patrons de relació. Així, per exemple, va establir que “la interacció freqüent dins d'un grup implica interacció poc freqüent fora d'aquest” o que la “la relació entre dues persones, A i B, està determinada en part per les relacions entre A i una tercera persona C, i entre B i C”. En definitiva, aquests autors van posar de manifest que l'estudi de les relacions era important per a entendre la dinàmica social.

2.2.3. L'Escola de Manchester

A partir de la Segona Guerra Mundial, el concepte de *xarxa social* es desenvolupa en el camp de l'antropologia social britànica per a donar una resposta a les limitacions que l'estructural-funcionalisme presentava enfront dels canvis en les societats tradicionals i l'articulació de les societats complexes. L'escola antropològica de Manchester, fundada per Max Gluckman, va tenir en Elizabeth Bott, John Barnes i J. Clyde Mitchell els seus principals exponents.

El primer a utilitzar el terme *xarxa social* (*social network*) va ser John Barnes (1954) en relació amb el conjunt de relacions que quedaven fora de l'anàlisi institucional en un llogaret de pescadors a Bremnes, Noruega. Així, a l'estudi del territori i del sistema ocupacional, Barnes va afegir el camp social de les relacions personals.

La imatge que tinc és la d'un conjunt de punts, alguns dels quals estan units per línies. Els punts de la imatge són persones que interactuen entre elles. Per descomptat, podem conceptualitzar el conjunt de la vida social com a generadora d'una xarxa d'aquest tipus. Per als nostres propòsits, no obstant això, vull considerar, en sentit ampli, la part de la xarxa total que es deixa enrere quan extraïem les agrupacions i les cadenes d'interacció que pertanyen estrictament al sistema territorial i industrial. (Barnes, 1954)

El 1955, Elizabeth Bott va demostrar que la segregació dels rols conjugals en matrimonis urbans londinencs era una conseqüència del tipus de xarxa social de la parella, de manera que si ambdós cònjuges es relacionaven amb les mateixes persones, els rols eren diferents i, en canvi, si cada cònjuge disposava dels seus propis coneguts amb poques persones en comú, els rols conjugals eren més semblats. Així, l'estructura de les relacions observades a la xarxa

social era capaç d'explicar de millor manera la diferenciació en els rols que les categories socials o els grups institucionals.

Un altre exemple d'aquestes recerques és el que va dur a terme Kapferer (1969) en un grup de quinze treballadors, en què els més joves incrementaven el ritme de treball posant en perill la posició de les persones de més edat incapaces de seguir un ritme tan alt. Abraham, un treballador gran, va acusar el jove Donald de no respectar el ritme establert de treball. Aquest últim va fer una acusació vetllada de bruixeria. Sota aquesta situació s'esperava que els grups tanquessin files entorn d'un treballador o un altre i s'agrupessin d'acord amb l'edat. No va ser així, ja que Abraham va aconseguir la majoria dels suports, fins i tot el de les persones més properes a Donald. Estudiant els continguts de les relacions entre els treballadors en apartats com ara les converses casuals, la conducta jocosa, l'ajuda a la feina, entre d'altres, Kapferer va observar que la conducta a la feina responia al principi de conservar al màxim les inversions fetes en el conjunt de relacions, de manera que la densitat més elevada de xarxes laterals d'Abraham explicava la seva victòria contra Donald.

Una altra contribució interessant va ser la d'A. C. Mayer (1966) amb una recerca sobre una campanya electoral a Dewas, a l'estat hindú de Madhya Pradesh. Hi havia dos partits, el Partit del Congrés i el del Jan Sangh, que estaven empatats. Mentre que Jan Sangh es va concentrar a mobilitzar i fidelitzar persones directament relacionades amb el seu partit, el Partit del Congrés es va dedicar a mobilitzar la gent que estava articulada amb les persones vinculades.

Encara que les aportacions de l'Escola de Manchester van ser molt importants per als plantejaments de l'anàlisi de xarxes, només es van poder analitzar xarxes petites i les línies de recerca no van tenir un desenvolupament com el que era d'esperar. No obstant això, S. F. Nadel va elaborar un corpus teòric (1957) que va exercir gran influència entre els investigadors de l'Escola de Harvard als anys setanta. Encara que el treball de Nadel va ser molt més extens, una mort prematura va impedir el desenvolupament complet de les seves idees.

Siegfried Frederick Nadel va estudiar música a Viena. En aquesta disciplina va arribar a ser considerat com un director d'orquestra prometedora, no obstant això, la seva proximitat amb l'antropòleg Malinowski el va portar a dedicar-se a l'antropologia i va fer monografies de diversos pobles d'Àfrica, com els Nupe. Amb les seves propostes teòriques per a l'estudi etnològic de diverses poblacions africanes, pretenia descriure d'una manera més estructurada les relacions socials.

2.2.4. L'Escola de Harvard

Els investigadors de l'anomenada Escola de Harvard eren fonamentalment sociòlegs amb una forta formació matemàtica, que van reaccionar contra la teoria sociològica dominant en el moment intentant derivar l'estructura social dels patrons d'interaccions socials empíricament observables. Aquests autors van implementar dos avenços molt importants en l'estudi de les xarxes socials. El primer és l'escalat multidimensional, que permet plasmar en un gràfic bidimensional les distàncies socials, i el segon és l'equivalència estructural instrumentalitzada amb la tècnica anomenada *blockmodelling*, que permetia identificar grups sobre la base de l'equivalència de les estructures de relacions (una idea avançada per Nadel).

Les propostes d'aquests autors van poder ser implementades posteriorment amb programes informàtics, i en conseqüència va ser possible conjuminar les aplicacions de la teoria de grafs a les ciències socials realitzades per autors de formació matemàtica com Harary (1963) i així abordar l'estudi empíric de les xarxes socials.

A partir dels anys vuitanta, els treballs de Freeman sobre centralitat (1979), el desenvolupament als noranta d'UCInet (sigles d'*University of California Irvine-net*) per Steve Borgatti i associats i els treballs sobre comunitats als Estats Units per part de Barry Wellman, entre d'altres, van donar l'impuls definitiu a l'associació creada el 1978, l'actual International Network of Social Analysis (INSNA).

El treball que ha desenvolupat la INSNA a través del temps i els treballs que actualment es desenvolupen poden ser visitats al seu lloc web:

<http://www.insna.org/>

Fins i tot és possible seguir Barry Wellman a Twitter:

<http://twitter.com/barrywellman>

Gràcies a la seva pàgina web i al seu blog, podem conèixer les publicacions i les recerques de Steve Borgatti:

<http://www.steveborgatti.com>

2.2.5. Camps d'aplicació

Encara que els camps d'aplicació són múltiples, podem assenyalar alguns dels més importants. El primer són les xarxes d'acció política. En aquest camp

s'emmarquen les contribucions que analitzen les relacions existents entre actors, institucions o organitzacions que es relacionen amb les institucions del poder o amb la lluita pel poder.

Un segon camp d'aplicació és la salut. A partir sobretot de la lluita contra la sida, les contribucions en aquest camp han augmentat a causa que la malaltia es contagia mitjançant relacions personals com ara les relacions sexuals o l'intercanvi de xeringues. També en altres casos d'epidemiologia com el dengue o la malària s'ha aplicat la perspectiva de xarxes socials per a observar-ne l'evolució i la propagació. L'anàlisi d'aquestes xarxes resulta fonamental per a aplicar o millorar les polítiques de prevenció.

La biblioteconomia i la col·laboració científica han experimentat un gran desenvolupament amb l'aplicació d'aquesta perspectiva. L'anàlisi de xarxes socials resulta molt profitós en àrees com la coautoria de textos acadèmics, l'anàlisi de polítiques de recerca i desenvolupament, l'ús de plataformes de xarxes d'Internet per a la difusió del coneixement o la vinculació de documentació electrònica, entre moltes altres possibilitats.

En el número 19 de la revista Redes es poden trobar alguns articles sobre recerques que es duen a terme en l'actualitat sobre col·laboracions científiques, coautoria i citacions.

<http://revista-redes.rediris.es/>

Un altre camp en expansió és el de les xarxes personals. Sota aquesta categoria es reuneixen una sèrie de recerques referents als intercanvis d'ajuda, informació, consell i referències informals. Fenòmens com ara la globalització i la migració en fan unes xarxes deslocalitzades que permeten una anàlisi de la distribució geogràfica de les relacions a través de les fronteres estatals.

Al lloc <http://redes-personales.uab.es> és possible trobar l'Observatori de Xarxes Personals, una font per a la recerca de les xarxes personals a partir de les dades aportades per investigadors de diferents països.

Hi ha molts més camps d'estudi que admeten l'anàlisi de xarxes socials en les seves recerques. Les xarxes socials també s'apliquen en àrees com els ne-

gocis, el màrqueting, el turisme, l'economia formal i els negocis ètnics entre moltes altres àrees.

Un grup d'estudiants de la Universitat de Buenos Aires ha trobat en l'anàlisi de xarxes la metodologia més adequada per a l'anàlisi esportiva. D'aquesta manera, cada setmana analitzen el partit més rellevant de futbol i mitjançant els gràfics obtenen els jugadors més buscats durant el partit, les passades de la pilota i les interaccions entre els jugadors, entre altres temes al·lusius a l'esport. Podeu consultar el seu treball a l'adreça següent:
<http://www.arsfutbol.com.ar/>

2.3. *Six degrees* i capital social

2.3.1. Sis graus de separació

Per introduir-nos en aquest tema, partirem de la citació presa del clàssic estudi de Milgram (1967) sobre el problema del “món petit”:

Fred Jones de Peoria, assegut a la vorera d'un cafè a Tunis, com que necessitava foc per al seu cigarret demana un misto a l'home de la taula del costat. Comencen a conversar. L'estranger és un anglès que resulta que va estar diversos mesos a Detroit estudiant el funcionament d'una fàbrica d'embotellament i reciclatge.

—Sé que és una pregunta ximple —diu Jones—, però alguna vegada et vas trobar per casualitat amb un company que es diu Ben Arkadian? És un antic amic meu, que dirigeix una cadena de supermercats a Detroit...

—Arkadian, Arkadian —murmura l'anglès-. Crec que sí! Un noi petit, molt enèrgic, armant un divertit merder a la fàbrica per una partida d'ampolles defectuoses.

—Fas broma! —Exclama Jones sorprès.

—El món és un mocador, oi?

Tots hem experimentat la mateixa sorpresa en descobrir sovint coneguts comuns amb estranys en llocs no habituals per a nosaltres. L'expressió popular *El món és un mocador* (*small world* en anglès) posa de manifest aquest fet. A la fi dels anys seixanta, Travers i Milgram (1977) van iniciar una sèrie d'experiments en els quals van demostrar que es podia trobar una cadena de coneguts entre dues persones preses a l'atzar de 5,2 passos de mitjana. Com hem vist al capítol anterior, la mesura es repetia en els diferents estudis. Com explicar aquest fet?

Travers i Milgram van fer un experiment que va consistir a demanar a persones voluntàries amb qui s'havia contactat a partir d'un anunci a la premsa de dos estats diferents dels Estats Units que fessin arribar una carta a un destinatari, del qual van proporcionar algunes característiques (professió, ciutat de residència, etc.). La clau consistia a enviar la carta al contacte personal que podria conèixer el destinatari, amb una còpia al centre de recerca. Els resultats van ser que 217 de les 296 persones de partida van enviar el document, i 64 van aconseguir el seu objectiu (29%). La longitud mitjana de les cadenes de coneguts completades va donar el nombre 5,2.

El 1993 la pel·lícula *Un estrany a la família* (*Six degrees of separation*) de Fred Schepisi, amb Will Smith com a protagonista, va popularitzar la mida de 5,2 passos d'un món petit.

L'explicació del fenomen d'un món petit va ser proporcionada per Duncan Watts (1999), entre d'altres. Les xarxes socials no presenten un patró aleatori sinó que, per contra, s'estructuren en comunitats, amb molts llaços interns i *hubs* o persones pont que connecten a la xarxa. Aquestes persones tenen un alt grau d'intermediació. D'aquesta manera, alguns *hubs* amb poques connexions entre comunitats separades permeten que el diàmetre de la xarxa sigui molt petit i fan possible trobar camins de pocs passos que connecten amb qualsevol parell de persones preses a l'atzar.

En la figura següent observem una xarxa de lectors de novel·les històriques. Hi podem veure en diferents colors els diferents cliques que es formen a l'interior de la xarxa. Envoltades per un cercle, podem veure aquelles persones pont que permeten la cohesió de la xarxa pel seu alt grau d'intermediació.



Figura 2.7. Xarxa social amb comunitats que s'enllacen per mitjà de *hubs*.

En la pràctica vivim en aquestes comunitats denses (escola, feina, veïnat), geogràficament properes, però amb un percentatge petit de contactes actius que viuen a molta distància, fins i tot repartits en diferents països. Aquest és el fenomen conegut com a *glocalització* (Hampton i Wellman, 2002). Dins d'aquestes comunitats predomina l'homofília, és a dir, l'existència de característiques comunes a la major part dels membres: gustos semblants, experiències similars, maneres de vestir i de pensar. No obstant això, algunes d'aquestes persones són diferents en el sentit que pertanyen a diverses comunitats al mateix temps. Com cada comunitat té les seves pròpies normes, formals o informals, aquestes persones han de ser competents en diferents conjunts de normes, sovint contradictòries entre elles.

Reflexió:

Aparentment, triem lliurement les nostres amistats i les nostres parelles, però és realment així?

Les persones que tenen un nivell jeràrquic alt en l'estructura social solen exercir aquest rol de connectors. Polítics, directius d'organitzacions de tot tipus, etc. solen tenir molts contactes de característiques molt diverses.

Activitat:

Imagineu que heu de fer arribar un missatge personal a Barak Obama i que solament podeu utilitzar contactes personals. Qui coneixeu que podria conèixer algú que podria fer arribar un missatge a algú perquè el transmetés a Obama?

2.3.2. Capital social

El concepte de *capital social* ha estat conceptualitzat de diferents maneres al llarg del temps, però el podríem definir de la manera següent:

El capital social pot ser definit com el conjunt dels recursos accessibles mitjançant les relacions personals.

D'acord amb Nan Lin (2001), podem diferenciar entre capital humà i capital social. El capital humà inclou els béns materials i simbòlics (com per exemple, reconeixements o diplomes) d'un individu. Els recursos socials són

els recursos als quals aquest té accés mitjançant les seves connexions socials. Nan Lin reconeix que la reciprocitat és un factor fonamental en el capital social, ja que és un principi que atorga cohesió a les xarxes en crear compromisos generats per la solidaritat i l'empatia. Nan Lin conceptualitza el capital social de la manera següent:

El capital social es compon dels recursos que té una xarxa pròpia o d'associacions. L'enfocament de capital social no es refereix als béns que té una persona, però sí als recursos accessibles per mitjà de vincles directes i indirectes. L'accés i l'ús d'aquests recursos és temporal i en préstec, en el sentit que l'actor no els té. Per exemple, la bicicleta d'un amic és un capital social que podem utilitzar per a aconseguir un determinat objectiu, però ha de ser retornada. Una de les conseqüències de la utilització del capital social és l'obligació assumida de reciprocitat o compensació (Nan Lin, 2011, pàg. 21).

El mateix autor assenyala l'existència de dos tipus de networking o d'activitat relacional, el *networking* homòfil i el *networking* heteròfil. El primer es desenvolupa de manera espontània entre les persones que es consideren com a iguals, encara que no es coneguin prèviament. El reconeixement mutu a partir del llenguatge, els gestos, la vestimenta, els gustos o l'experiència fan d'aquest networking un mecanisme d'assegurament dels recursos socials controlats per un grup social. El segon tipus es desenvolupa entre persones de diferent condició, és molt més formalitzat i normalment permet accedir a recursos diferents dels controlats pel grup social original.

Aquest plantejament és una elaboració de la cèlebre proposta de Mark Granovetter, *La força dels llaços febles* (1973), segons la qual aquests llaços febles (llaços entre persones que es coneixen poc i que sovint ocupen posicions socials diferents) poden ser clau en l'obtenció d'una ocupació o, en general, dels recursos que no poden proporcionar els llaços forts. Els llaços forts es basen normalment en la interacció freqüent i la confiança, són reduïts en nombre i estables al llarg del temps. En canvi, els llaços febles són nombrosos, canvien al llarg del temps i representen una relació molt més especialitzada en contextos concrets, com la feina o l'oci. No és estrany que els SNS trobessin en el manteniment fàcil dels llaços febles la base de la seva popularitat. Ens ocupem d'aquest tema en el proper apartat.

Els avenços tecnològics influeixen també en la formació de les xarxes socials i en la variabilitat del capital social. No obstant això, Barry Wellman (1999) i altres autors demostren que les relacions cara a cara i les relacions electròniques, lluny de segregar-se, es reforcen mútuament.

Reflexió:

Imagineu que heu de muntar una festa sorpresa per a algú especial. Quins recursos o accions necessitaríeu dels altres? Com incideix el capital social en la participació comunitària per aconseguir l'objectiu?

2.4. *Social networking sites*

Els *social networking sites* (SNS) han experimentat una gran popularització que propicia un canvi significatiu en la interacció social i en la transformació en les relacions personals.

Els llocs de xarxes socials (SNS) són serveis disponibles a Internet i permeten, a partir de la publicació d'un perfil d'usuari, construir o reflectir relacions articulades entre persones que comparteixen interessos, activitats i recorreguts vitals compartits en una plataforma o pàgina web comuna.

L'anàlisi de xarxes es constitueix com una eina idònia per a estudiar les comunitats sorgides de les interaccions registrades en aquests SNS, els quals proveeixen cada vegada més d'eines d'anàlisi i visualització.

Les xarxes socials complexes han existit sempre però els desenvolupaments tecnològics recents n'han permès l'aparició com una forma dominant d'organització social. Així com les xarxes d'ordinadors connecten màquines, les xarxes socials vinculen gent. Quan les comunicacions assistides per ordinador connecten les persones, les institucions o el coneixement, parlem de xarxes socials assistides per l'ordinador. Sovint, les xarxes socials i les xarxes assistides per ordinador treballen conjuntament. Amb les xarxes d'ordinador la gent es vincula en xarxes socials compartint les situacions que viuen fora de línia per recollir les dels altres que usen aquestes xarxes per a comunicar-se. (Wellman, 2002)

Sens dubte, l'adveniment del Web 2.0 ha contribuït al desenvolupament d'aquestes plataformes. Aquesta arquitectura (O'Reilly, 2005) recupera la idea de concebre Internet com una plataforma enfocada a l'usuari final que genera una dinàmica de participació en la qual el fonamental és la informació. Aquesta informació no es limita al text, com ocorria fa més d'una dècada, sinó que les fotografies, els vídeos i altres llibres gràfics fan aquesta experiència una experiència que crida més l'atenció i més completa.

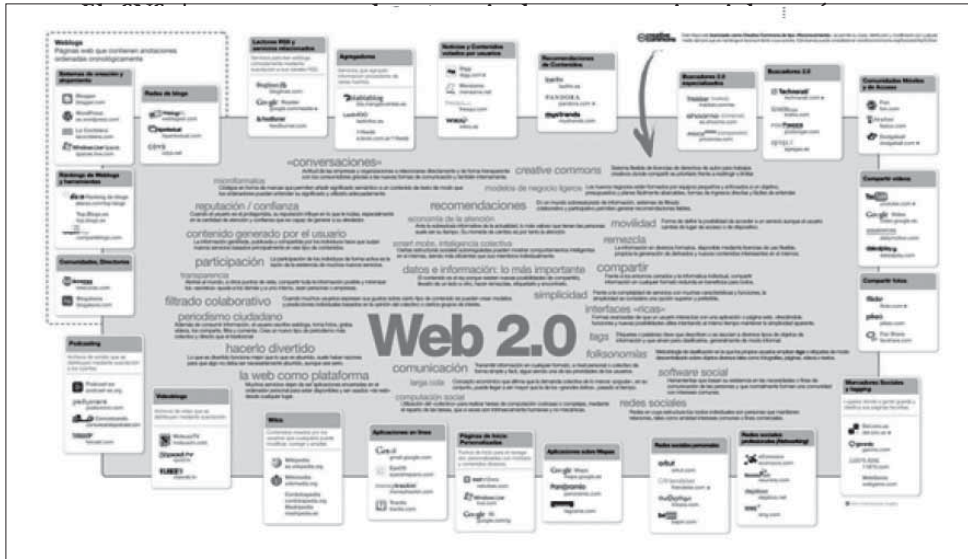


Figura 2.8. Aquest mapa mostra alguns recursos relacionats amb el Web 2.0. En aquest enllaç es pot apreciar millor el gràfic <http://internality.com/web20/>.

Des de fa algun temps conviuen dos termes: el de Web 2.0 i el de Web 3.0. El Web 2.0 s'ocupa de la Internet social. El Web 3.0 es basa en la intel·ligència artificial per a la cerca de dades. Se sol emparentar amb el Web semàntic o fins i tot se solen emprar com a sinònims.

**Per saber-ne més:
Web2.0**

Un concepte fonamental per a entendre els SNS és el que proposa B. Wellman (2002) i que es refereix a l'individualisme interconnectat o individualisme en xarxa (*networked individualism*), en el qual la persona s'ha convertit en la unitat individual de connexió social en substitució de la llar o del lloc de treball i és capaç d'operar les seves pròpies xarxes i determinar-ne el nivell d'interacció en aquestes xarxes. La proliferació de xarxes socials assistides per ordinador fomenten els canvis en el "capital de la xarxa": de quina manera la gent entra en contacte, interactua i obté recursos dels altres.

Per saber més coses sobre l'individualisme connectat i les transformacions de les relacions socials, podeu consultar l'article següent de Barry Wellman: Little boxes, glocalization and networked individualism

Seguint la idea d'Anabel Quan Haase (2004), hi ha tres visions sobre la relació entre la nostra participació a Internet i el capital social. La primera proposa que Internet transforma el capital social en introduir un mitjà alternatiu a l'habitual. La segona visió sosté que disminueix el capital social perquè redueix l'interès per allò local i facilita l'allunyament d'amics i familiars. El tercer enfocament es basa en la complementació dels dos àmbits perquè Internet es barreja ràpidament amb la vida de la gent, facilita la comunicació amb usuaris geogràficament distants, però també promou l'intercanvi amb els més propers i facilita la participació en entorns immediats.

Considerant que les xarxes personals poden tenir vincles dins i fora de l'àmbit d'Internet i dels SNS, l'últim enfocament és el que resulta més útil si volem tenir una visió més àmplia d'una xarxa. Per exemple, una persona pot tenir llaços forts amb repercussions elevades en el seu capital social, encara que amb alguns contactes de la seva xarxa el contacte sigui exclusivament per ordinador. Això té lloc sovint entre algunes xarxes d'immigrants que mantenen les seves xarxes al lloc d'origen gràcies als SNS.

Una de les crítiques més fortes que han rebut els SNS és que, encara que la llista d'amics és gairebé tan gran com vingui de gust (segons els càlculs del mateix Facebook la mitjana de contactes per usuari és de 120 persones), la interacció és superficial i es redueix a un grapat de contactes. Si ho situem al món tangible i en l'àmbit de les relacions no assistides per l'ordinador, ocorre més o menys igual. Conservem i preservem llaços forts i tenim llaços febles gràcies a certs amics o coneguts. Molts usuaris han manipulat les seves pròpies xarxes fent un exercici de reclutament sistemàtic per ser llegits o vistos o per ostentar popularitat. No obstant això, això no vol dir que el seu capital social creixi al mateix ritme que la xarxa que ha creat, ja que la majoria dels seus llaços seran febles o poc interconnectats amb l'àrea més densa de la xarxa. Així, podríem dir que les xarxes socials es resisteixen a mercantilitzar-se completament, perquè, com hem vist, la reciprocitat té un paper en la dinàmica de les relacions.

Reflexió:

Penseu que pertànyer a algun SNS fomenta o facilita la participació en un altre tipus de xarxes que no estiguin connectades per ordinador (voluntariat, equips esportius, etc.)?

Els SNS són molt variats i responen a tot tipus d'interessos. Moltes vegades les seves finalitats són lúdiques, no obstant això també són importants eines de treball que permeten crear col·laboracions en línia en camps tan diversos com les arts i les ciències. Es poden classificar en xarxes horitzontals o generalistes, ja que aglutinen els usuaris sense cap temàtica específica pel mig i les relacions són heterogènies, i en xarxes verticals o específiques, centrades entorn d'un tema en concret, que pot ser lúdic, recreatiu, laboral, acadèmic, esportiu, entre d'altres.

En general el funcionament es basa en llistes de contactes que poden ser simètriques (amics recíprocs) com en el Facebook, o asimètriques (que seguim algú no significa que aquest altre també ens segueixi) com en el Twitter. La interacció es dona a partir de textos, fotografies, enllaços, missatges públics i missatges privats. A continuació es mostra una taula amb alguns SNS destacats:

SNS	Descripció	Registre
Academia.edu	Investigadors i acadèmics	Obert
aNobii	Llibres digitals	Obert
a SmallWorld	<i>Jet set</i> i elits	Només amb invitació
Avatars United	Jocs en línia	Obert
Badoo*	General	Obert a més grans de 18 anys
Blogster	Comunitat de blogs	Obert
Care2	Activisme social	Obert
deviantART	Comunitat artística	Obert
delicious	Marcadors per a compartir enllaços i referències	Obert
Exploroo	Viatges	Obert
Facebook*	General	Obert a més grans de 13 anys
Flixter	Pel·lícules	Obert a més grans de 13 anys
flickr	Fotografies i vídeos	Obert a més grans de 13 anys
Fotolog	Blogs fotogràfics	Obert
Google Buzz	General	Obert
Myspace*	General	Obert mitjançant entorn Google
Tuenti	General	Només amb invitació

Tumblr	Microblogs	Obert
Twitter*	Microblogs	Obert
Windows Live Spaces*	Blog entorn MSN	Obert

Taula 1: SNS amb més de 100 milions d'usuaris. Fonts: Wikipedia i Alexa

D'entre totes els SNS, el que ha tingut el creixement més gran i agrupa un nombre d'usuaris més gran és Facebook. Creat l'any 2005 amb la idea original de compartir contactes entre la comunitat estudiantil de Harvard, s'ha hagut d'adaptar a la demanda. Al començament de 2011 aquesta xarxa tenia més de 600 milions d'usuaris.

David Fincher va dirigir la pel·lícula *The Social Network*, que narra la història de la creació del Facebook. Aquest film es va basar en el llibre de Ben Mezrich *The accidental billionaires* i narra les peripècies de Mark Zuckerberg, creador d'aquest popular recurs. Encara que es reconeix que la història no és del tot certa, ens pot donar una aproximació de la seva creació i de com es va expandir el concepte *facebook*.

Reflexió:

Participeu en alguna d'aquestes xarxes? En una o en més? Varia molt el grau de proximitat que experimenteu amb els diferents contactes?

Com a usuaris individuals, podem tenir més o menys clars els nostres motius per a participar d'un SNS i aquests poden anar des de no perdre el contacte amb amics que viuen lluny, tenir a mà algun company d'estudis per a compartir informació o passar un temps agradable veient què passa amb la vida dels altres. No obstant això, encara que la majoria dels SNS se centren en la interacció d'un individu, és cada vegada més habitual que comerços, partits polítics, organitzacions civils, ajuntaments, centres culturals i esportius, prestadors de serveis, entre d'altres, generin els seus propis perfils i s'integrin a la xarxa amb objectius diversos. Les xarxes socials a Internet no solament agrupen persones, sinó que també vinculen dades, notícies, webs, enllaços i, per tant, les possibilitats d'anàlisi es multipliquen.

Reflexió:**Quins altres usos donaríeu a les SNS?**

La visualització de xarxes socials ens permet obtenir una panoràmica del nostre objecte d'estudi. Mitjançant un patró en què els nodes i els enllaços ens evidencien les seves relacions, podem explorar els actors principals, els grups que el formen, els nodes que queden solts, les parts que componen la xarxa i fins i tot, la qualitat dels vincles que la formen.

En el cas dels SNS, hi ha moltes aplicacions en línia mitjançant les quals podem fer una anàlisi que, encara que una mica limitada i senzilla, ens mostra mapes que ens ajuden a comprendre la formació de les xarxes. L'avantatge d'aquests recursos de visualització és que en la majoria dels casos no es necessiten fitxers de dades o elaborar matrius externes i la seva utilització és bastant senzilla. Alguns d'aquests recursos sí que admeten fitxers externs, la qual cosa permet treballar més sobre les dades. No és d'estranyar que en ser Facebook la plataforma més utilitzada, sigui també la que tingui el nombre de recursos més gran per a visualitzacions en línia.

El mapa següent és una xarxa feta amb Mention map a partir de les dades generades per favstar50, que és una eina que assenyala els missatges més destacats pels usuaris. Mention map és una utilitat dinàmica que mostra una xarxa de les mencions i dels continguts de Twitter. Pot ser útil per a aquells que gestionen continguts amb finalitats comercials, per a conèixer els temes de discussió o per a veure la repercussió dels nostres propis enviaments.



Figura 2.9. Mapa de Twitter fet amb Mention map

En la figura 10 veurem un mapa fet en Facebook sobre la xarxa d'un cineclub. L'aplicació es diu Friend Wheel i està disponible en el propi entorn d'aquesta xarxa social, encara que també es pot usar per a Twitter. És visualment atractiva, es pot variar el nombre de nodes, és estàtica i permet a qualsevol usuari mitjà adonar-se d'una estructura molt bàsica de xarxa.

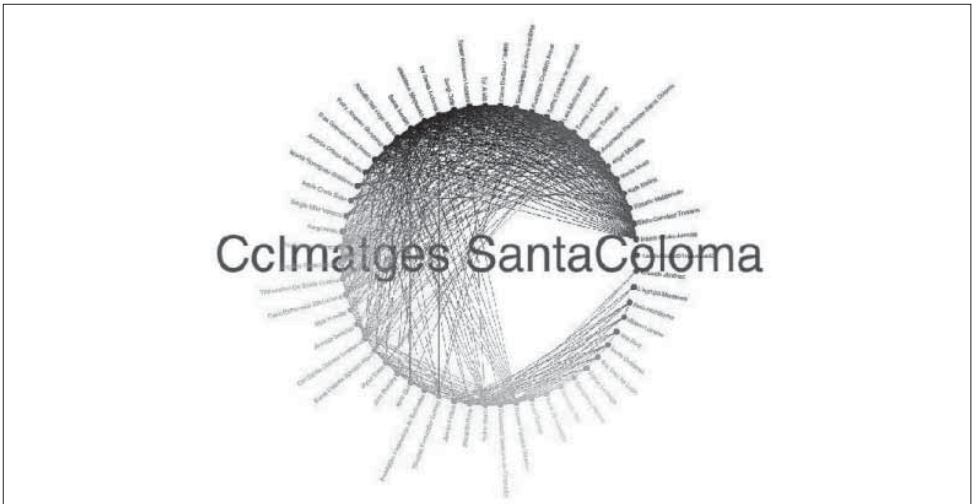


Figura 2.10. Mapa de Facebook fet amb Friend Wheel

Un mapa de recursos web i pàgines sobre un tema específic és el Touchgraph (en la seva versió SEO). N'hi ha una versió de pagament, però amb aquesta n'hi ha prou per a manejar els nodes i els recursos, com el color i la grandària. El gràfic ens permet accedir als clústers i ampliar la informació. En aquest cas, es va fer amb la cerca de Gadafi en plena efervescència del conflicte de Líbia.



Figura 2.11. Mapa de Touchgraph en la versió gratuïta

Enllaços d'interès per a visualitzar xarxes:

Touch Graph

Mention Map

A Facebook, busqueu les aplicacions següents:

Friend Wheel

UpGo!

Touchgraph per a fotos

SocGraph

Activitat:

Trieu algun dels recursos per a fer una xarxa social pròpia. En cas de no pertànyer a cap SNS, trieu algun tema d'interès o algun personatge per a generar una xarxa.

Els recursos que que hem vist fins ara hem vist ens permeten accedir d'una manera molt senzilla i agradable a les nostres pròpies xarxes. No obstant això, moltes d'aquestes eines no ens permeten modificar els resultats o donar més pes a alguns actors i enllaços presents a les nostres xarxes.

Per a la visualització de xarxes hi ha diversos programes com Pajek o UCInet que ens permeten, amb una matriu de dades, gestionar les nostres pròpies xarxes.

Pajek significa 'aranya' en eslovè i és un programa per a Windows que permet l'anàlisi i la visualització de xarxes socials complexes. Va ser desenvolupat per V. Batagelj i A. Mrvar. Està disponible de manera gratuïta per a l'ús no comercial en aquest enllaç: <http://pajek.imfm.si/doku.php?aneu=download>.

Una eina que actualment és molt popular, ja que admet tant matrius de filiació com fitxers automàtics generats per xarxes socials com Facebook, és Gephi. A més de resultar molt útil per a l'anàlisi a través de la filtració d'elements i dels algorismes que ens presenta, resulta visualment molt atractiva.

Activitat:

Seguint el tutorial de Gephi per a fer xarxes de Facebook, que trobareu a l'espai de l'assignatura, proveu a fer la vostra pròpia xarxa. Una vegada que la tingueu llesta intenteu trobar els conceptes següents que hem estudiat: llaços forts, llaços febles, cliques, hubs, etc.

Feu un comentari general sobre les particularitats de la xarxa i el que se'n pot deduir.

Capítol III

Xarxes, informació i coneixement

Agustí Canals

3.1. Introducció

Només cal observar una mica per a veure que les xarxes són ben presents al món que ens envolta, tant a la natura com en entorns més artificials creats per la societat humana. Curiosament, algunes característiques de les xarxes són comunes a tots els àmbits, i per tant resulta útil estudiar-les des d'un punt de vista general i treure'n conclusions que puguin ser aplicades a àmbits molt diversos, des de la biologia cel·lular fins a la sociologia, passant per Internet. Però també hi ha particularitats que distingeixen les xarxes d'algun tipus determinat. Així, per exemple, les xarxes de transport terrestre són diferents de les de transport aeri pel que fa a la distribució de grau. O bé, les xarxes socials solen presentar un nivell molt més elevat de vinculació selectiva (*assortativity*) que les naturals o tecnològiques. És a dir, els nodes d'una xarxa social tenen més tendència a estar connectats amb d'altres que s'hi assemblen o que, si més no, tenen algunes característiques similars.

Té sentit, doncs, un cop sabem com caracteritzar qualsevol xarxa de manera genèrica i coneixem les particularitats de les xarxes socials, aprofundir una mica en l'estudi de xarxes més específiques. Seria impossible voler incloure aquí tots els tipus de xarxes i per això ens cal decidir quines volem tractar més a fons. Una possibilitat seria centrar-nos en les xarxes pròpies d'un àmbit de coneixement específic, però en el marc d'un llibre introductorí ens ha semblat més escaient tractar temes que puguin tenir un interès més general. Per això, hem decidit adoptar una aproximació alternativa, més interdisciplinària, i mirar més en detall aquelles xarxes que tenen alguna relació amb dos dels elements més importants per a entendre la nostra societat: la informació i el coneixement.

Com es relacionen la informació i el coneixement d'una banda i les xarxes de l'altra? Doncs de diverses maneres possibles. Per exemple, l'estructura d'Internet i de les anomenades ja fa uns anys *autopistes de la informació* és clarament en forma de xarxa. Així, les infraestructures tecnològiques que ens permeten difondre grans quantitats d'informació, emmagatzemar-ne i accedir-hi són reticulars. Això té un efecte en les estructures informacionals que hi construïm a sobre –el World Wide Web (WWW) adopta també la forma d'una xarxa– i en els processos que hem de fer servir per a cercar-hi la informació i accedir-hi.

La relació de la informació amb les xarxes, però, no es veu tan sols a les xarxes que anomenem d'*informació*. Qualsevol procés de transmissió d'informació es produeix, generalment, entre agents socials, tant si són persones com grups de persones com és ara les organitzacions. Perquè es produeixi aquesta transmissió d'informació, hi ha d'haver alguna relació entre els agents implicats, o bé ja sigui de confiança, d'interès comú pel tema de la informació transmesa, de simple coneixement mutu o bé simplement de trobar-se al mateix lloc en un moment donat. Totes aquestes relacions, tal com hem vist, ens poden permetre definir xarxes socials. De l'estructura d'aquestes xarxes socials en dependrà molt com es farà la transmissió d'informació.

Un exemple molt interessant de la relació entre el coneixement i les xarxes el trobem a l'estructura del coneixement científic. La ciència és el resultat d'un procés gradual d'acumulació de coneixement, en el qual les aportacions de cada científic es basen en les dels seus predecessors i així successivament. Aquesta relació entre els científics actuals i els que els han anat precedint en el temps es plasma clarament, com veurem més endavant, en l'estructura en forma de xarxa que formen les citacions de la literatura científica. D'altra banda, l'activitat científica es du a terme sovint a partir de col·laboracions entre científics que moltes vegades procedeixen d'institucions diferents i fins i tot de camps diferents. L'estudi de les xarxes que configuren aquestes col·laboracions ens pot dir molt sobre l'estructura de l'activitat de recerca. Fins i tot podem combinar conceptes de l'anàlisi de xarxes amb altres eines de la gestió d'informació per a estudiar l'estructura i l'evolució del coneixement científic i per representar-lo de manera entenedora. L'estudi d'aquestes xarxes permet veure com s'estructura i com evoluciona el coneixement científic al llarg del temps.

Un altre àmbit en el qual la relació entre xarxes i informació i coneixement és clar és el de l'economia i les organitzacions de tot tipus. Tant la creació com la difusió de coneixement a dins de les empreses o institucions o entre elles depèn de l'intercanvi –o no– d'informació entre els seus components. I aquest intercanvi només es pot produir mitjançant els lligams que existeixen entre

ells, és a dir, de les seves xarxes socials, tant si són formals com informals. És bàsic conèixer aquestes xarxes si volem optimitzar la generació de nou coneixement i la seva transmissió i utilització.

En aquest capítol analitzarem aquests i altres casos en els quals, per una banda, conèixer les propietats de les xarxes és fonamental per a entendre els processos relacionats amb la informació i el coneixement, i, per l'altra, les característiques de la informació i el coneixement permeten entendre l'estructura i l'evolució de les xarxes. Començarem, però, per fer un parell d'aclariments conceptuals sobre la visió de la informació, el coneixement i les xarxes socials que adoptarem aquí.

3.1.1 Sobre informació i coneixement*

Sovint els termes *informació* i *coneixement* es confonen i es fan servir de manera indistinta. Algunes vegades, tot i considerar-se coses diferents, s'utilitzen sense unes definicions clares que permetin distingir-los. Això moltes vegades genera confusió. Tot i que en aquest capítol no podem assolir el nivell de detall necessari per a tractar aquest problema a fons, sí que cal que definim breument el que entendrem aquí per *informació* i per *coneixement*.

Com a pas previ necessitarem definir el concepte de *dada*. Podem considerar les dades com a diferències discernibles en els estats del món que ens envolta. Qualsevol cosa que en sabem és a través d'aquestes diferències que percebem. Per exemple, el nivell de claror a l'exterior ens pot indicar si és de dia o de nit; l'alçària de la columna de mercuri del termòmetre ens pot dir la temperatura o el canvi de posició d'un cotxe en el temps ens pot donar idea de la seva velocitat.

D'aquestes dades que percebem, en podem obtenir informació. Constituïxen informació les regularitats significatives que observem a les dades. Els éssers humans, en ser capaços de percebre les dades, les diferències en els estats del món, podem veure-hi regularitats. Aquestes regularitats seran significatives quan ens permetin saber més sobre aquest món que ens envolta. En definitiva, quan ens aportin coneixement.

El coneixement consisteix en un conjunt d'expectatives que tenim sobre com és i com funciona el món que ens envolta i que ens permeten actuar en conseqüència. Inclou, doncs, tant el coneixement que posseïm de manera conscient com aquell que tenim de manera inconscient i que pren sovint la

* La discussió d'aquest apartat està basada en: Boisot y Canals (2004)

forma d'instints o intuïcions que no podem explicar. Tot aquest coneixement és el resultat del que anem aprenent al llarg de la nostra vida. I és precisament per mitjà de la informació que anem modificant la base de coneixement que tenim. Per exemple, la dada sobre l'alçària de la columna de mercuri del termòmetre ens dóna informació sobre la temperatura exterior abans de sortir de casa. Si la temperatura és baixa, hem incorporat a la nostra base de coneixement que "avui fa fred". És a dir, hi hem incorporat l'expectativa que si sortim a fora experimentarem els efectes de la baixa temperatura. Com que, d'altra banda, a la nostra base de coneixement també hi hem incorporat de ben petits que si volem evitar aquests desagradables efectes del fred el que hem de fer és abrigar-nos, abans de sortir ens posarem l'abric.

Així, doncs, podem dir que la informació és una extracció de les dades que, modificant les expectatives d'un agent, té la capacitat d'actuar sobre la seva base de coneixement, tal com es mostra a la figura 3.1

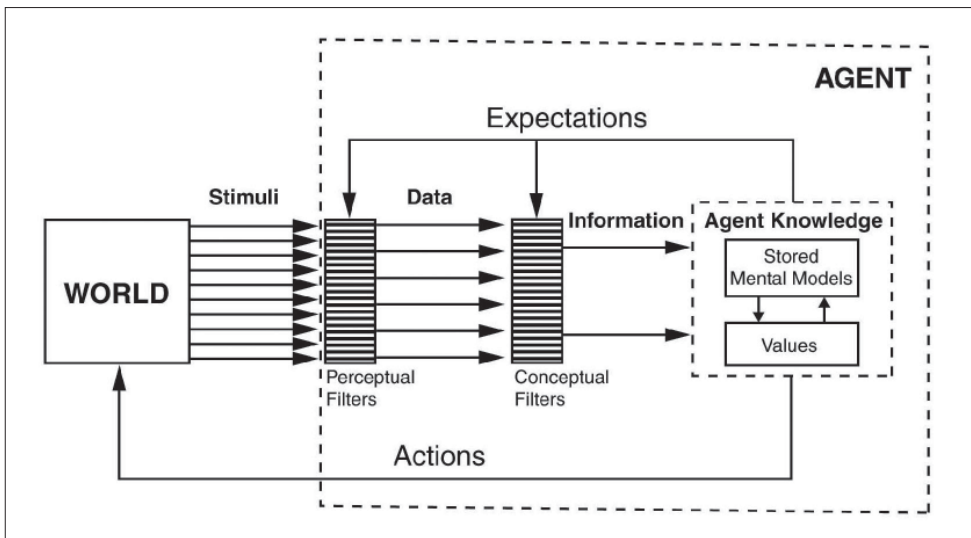


Figura 3.1. Dades, informació i coneixement (Boisot i Canals 2004)

Parlem d'agents en general perquè, segons aquesta manera de definir-lo, el coneixement no és exclusiu dels éssers humans. Qualsevol agent capaç de percebre dades, extreure'n informació i aprofitar-la per a modificar les seves expectatives sobre l'entorn, podem dir que té coneixement. En aquesta categoria podem incloure des dels animals fins a alguns sistemes d'intel·ligència artificial. Fins i tot, per analogia, podem dir que les organitzacions també tenen una base de coneixement que utilitzen per a interpretar el seu entorn i actuar

en conseqüència. En aquest capítol ens centrarem sobretot en les xarxes que tenen a veure amb la informació i el coneixement de persones i organitzacions.

La informació és una extracció de les dades que, modificant les expectatives d'un agent, té la capacitat d'actuar sobre la seva base de coneixement.

Reflexió:

És interessant remarcar que aquesta manera d'entendre el coneixement ens duu a considerar que no tan sols és coneixement el que sabem de manera racional. El fet que quan notem una calor molt intensa a la mà instintivament la retirem del foc es deriva de l'expectativa que ens cremarem i que, per tant, hem de treure-la. És un coneixement que tenim incorporat com a instint gràcies a milions d'anys d'evolució de l'espècie. Els nostres gens, en certa manera, també formen part de la nostra base de coneixement.

3.1.2. Xarxes d'informació i xarxes socials

Les xarxes socials es defineixen com aquelles que tenen com a vèrtexs persones –o com a molt grups de persones– i com a enllaços interaccions socials entre aquestes persones. Segons aquesta definició, algunes de les xarxes de les quals parlarem en aquest capítol es podrien considerar xarxes socials i d'altres no.

Tanmateix, com que bàsicament aquí parlarem de la informació i el coneixement que creen, utilitzen i intercanvien persones i organitzacions, les xarxes que tractarem que es cataloguen com a xarxes d'informació podem considerar que també tenen una certa qualitat de "socials". Encara que en alguns casos no ho siguin pròpiament, responen a interaccions socials entre individus o organitzacions i, per tant, sovint són el reflex de xarxes realment socials subjacents.

3.2. Xarxes d'informació al World Wide Web

Les xarxes d'informació són aquelles que estan formades per nodes que contenen dades. Aquests nodes estan connectats entre ells d'alguna manera i les dades es poden transferir d'uns nodes als altres mitjançant aquestes connexions. Com hem vist anteriorment, la transmissió física de dades té com a conseqüència la possibilitat de transferir informació. Per això d'aquest tipus de xarxes en diem xarxes d'informació.

Les xarxes d'informació solen ser construïdes pels humans, perquè són els éssers humans els que aprofiten la transferència d'informació entre els nodes de la xarxa als quals tenen accés per a adquirir nou coneixement. Tanmateix, cada vegada hi pot haver més agents no humans amb una certa capacitat d'adquirir coneixement que aprofitin les xarxes d'informació sense requerir l'actuació directa dels humans que els han creat. I potser algun dia fins i tot les creïn ells mateixos.

Potser l'exemple més important actualment de xarxa d'informació és el World Wide Web (o WWW o simplement Web), però també ho són les xarxes de correu electrònic, les dels llocs web Facebook o LinkedIn, les de weblogs o les de Twitter.

Reflexió:

Cal remarcar que els SNS el que es coneixen popularment com a xarxes socials (Facebook, LinkedIn, Twitter, etc.) en realitat no serien pròpiament xarxes socials sinó xarxes d'informació, ja que els seus nodes no són persones (o grups de persones), sinó pàgines personals que contenen dades. Ara bé, com que aquests nodes solen representar persones i com que aquestes xarxes sovint reproduïxen els patrons de relació entre persones, s'assimila la xarxa d'informació amb la xarxa social que formen les persones representades a les pàgines corresponents. Per això popularment Facebook o Twitter sovint s'anomenen "xarxes socials".

Com es pot veure amb els exemples anteriors, la distinció que fem entre tipus de xarxes (socials, d'informació, tecnològiques, etc.) no està universalment acceptada i, especialment en l'àmbit popular, podem trobar que una mateixa xarxa es considera de tipus diferents. Això es produeix especialment amb les xarxes d'informació, ja que d'alguna manera se situen a cavall de les socials i les tecnològiques i tenen elements comuns a totes dues.

En aquesta secció aprofundirem en alguns aspectes de les xarxes d'informació més importants i populars avui en dia, el World Wide Web en general i algunes xarxes més específiques que utilitzen el Web com a base. Començarem per veure com s'estructura aquesta base.

3.2.1. L'estructura del World Wide Web

Avui en dia, l'exemple de xarxa més conegut és sens dubte Internet. Tots els que esteu llegint això utilitzeu sovint aquesta xarxa per a treballar, per a estudiar, per a distreure-us, o segurament per a tot plegat.

Tal com hem vist anteriorment, però, Internet és una xarxa tecnològica força complexa, formada per ordinadors i encaminadors (*routers*) que estan units per cables i altres tipus de connexions. Per a la majoria de nosaltres, seria difícil poder interactuar directament amb aquesta xarxa tecnològica. Afortunadament, per sobre d'Internet hi ha una altra xarxa, el World Wide Web, una xarxa d'informació dirigida que ens facilita l'ús de la infraestructura tecnològica subjacent per al que ens interessa: emetre i rebre informació. Als vèrtexs d'aquesta xarxa hi ha el que anomenem pàgines web, que consisteixen en text, imatges o altre tipus d'informació; i les connexions són els hiperenllaços (*hyperlinks*) que ens permeten navegar d'una pàgina a l'altra.

El Web és, de fet, el resultat d'una aplicació informàtica que va desenvolupar entre els anys 1989 i 1991 un grup d'investigadors del CERN (Organització Europea per la Recerca Nuclear), encapçalat per Tim Berners-Lee, per permetre als usuaris compartir informació per Internet.

En aquell moment hi havia diverses idees competint per donar la solució a la necessitat de compartir informació a Internet. El World Wide Web va guanyar la batalla a causa de dos factors. El primer és la facilitat d'ús. Per a posar els documents a disposició pública només s'havien de convertir en pàgines web utilitzant un senzill llenguatge de programació i emmagatzemar-les en un lloc accessible d'un ordinador de la xarxa Internet. I perquè els usuaris hi

* No és estrany que el World Wide Web es desenvolupés en un lloc com el CERN. Ja fa uns anys que la recerca experimental en física d'altres energies es duu a terme al si de projectes internacionals on col·laboren un gran nombre de científics de diverses universitats i centres de recerca. Encara que els experiments estiguin físicament situats a la central del CERN a Ginebra, moltes de les tasques –des del disseny dels instruments fins a l'anàlisi de les dades obtingudes– es fan a partir de la col·laboració d'equips d'investigadors escampats pel món i, per tant, la informació ha d'arribar a tot arreu. És per això que el CERN va ser un dels primers llocs on es va donar la necessitat de difondre una gran quantitat d'informació en format digital a un nombre important de persones repartides per multitud de països i institucions. El WWW va contribuir de manera decisiva a resoldre aquesta necessitat.

tinguessin accés, es va desenvolupar un programa anomenat navegador que permetia connectar fàcilment amb aquestes pàgines web que estaven escampades pels espais públics dels ordinadors connectats a Internet. El segon factor és que les tecnologies en les quals es basava, l'HTML (llenguatge d'etiquetatge d'hipertext o *hypertext markup language*) utilitzat per a generar les pàgines web, i l'HTTP (protocol de transferència d'hipertext o *hypertext transport protocol*), utilitzat per a transmetre les pàgines per Internet, es van posar a l'abast de la comunitat de manera gratuïta.

A la Figura 3.2 es pot veure la xarxa de pàgines del lloc web d'una empresa, que constitueix una molt petita part del Web. Els vèrtexs representen les pàgines, que estan connectades mitjançant hiperenllaços representats per fletxes. Aquests enllaços són majoritàriament unidireccionals, però també n'hi ha alguns de bidireccionals, que representen els casos en els quals cada una de les dues pàgines conté un enllaç que apunta a l'altra, de manera que estan enllaçades mútuament. Els diversos colors dels nodes representen diferents tipus de pàgines.

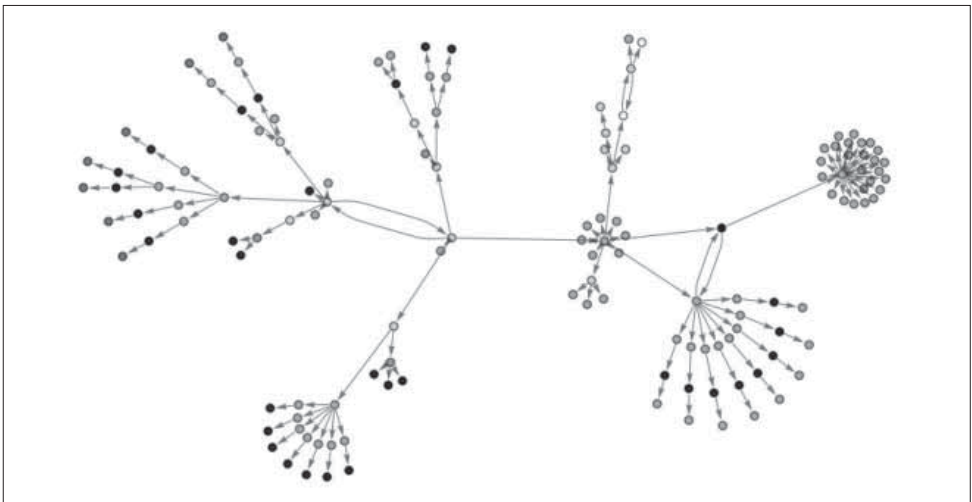


Figura 3.2. La xarxa de les pàgines del lloc web d'una empresa. Els vèrtexs representen les pàgines web, que estan relacionades pels hiperenllaços. Els colors dels nodes representen diversos tipus de pàgines (la pàgina principal, pàgines de text, imatges, formularis, etc.).

Per a veure com es formen aquestes xarxes amb una mica més de detall, utilitzarem el model simplificat de la Figura 3.3, que representa el lloc web d'una institució imaginària: la Universitat Global. Com podem observar, en alguns punts concrets de les pàgines web hi ha hiperenllaços que els connecten amb

altres pàgines mitjançant l'adreça URL incorporada al codi HTML. Aquestes connexions fan que sigui possible representar el lloc web com una xarxa en la qual els vèrtexs són les pàgines web i les connexions entre ells els hiperenllaços.

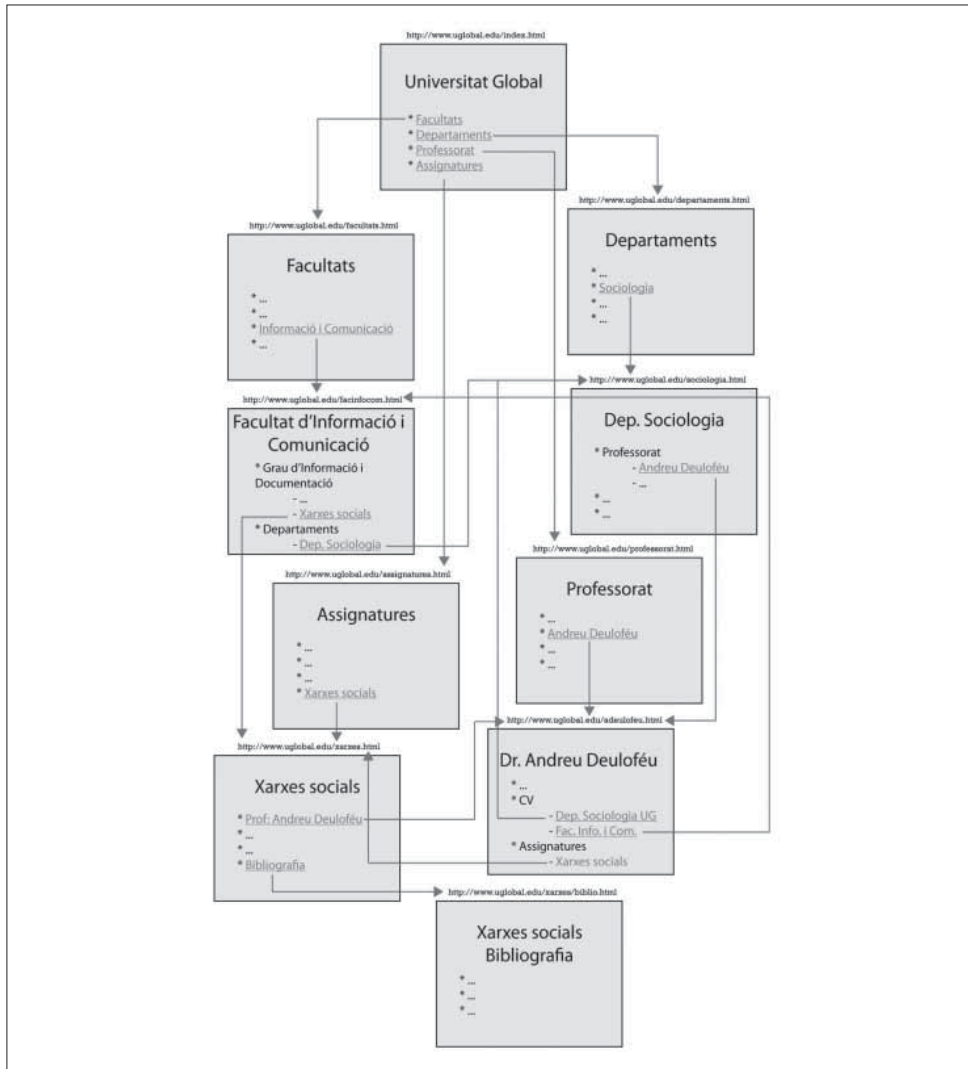


Figura 3.3. Una part del lloc web de la Universitat Global

Repassant els enllaços amb una mica d'atenció, podem veure que la xarxa corresponent al web de la Universitat Global tindria la forma representada a la figura 3.4

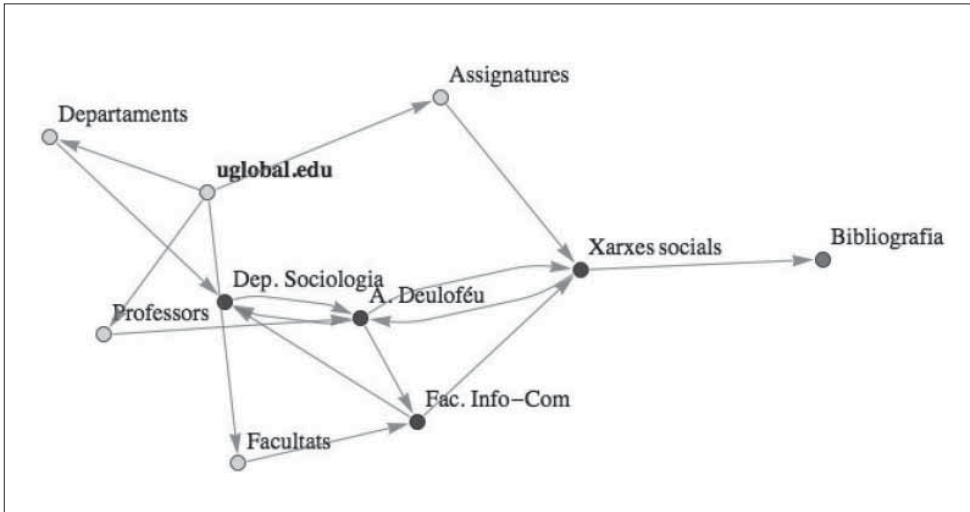


Figura 3.4. Xarxa formada pels enllaços entre les pàgines del lloc web de la Universitat Global

Observem que els enllaços entre les pàgines web són direccionals. És a dir, que el fet que hi hagi un enllaç que des d'una pàgina apunti a una altra no vol dir necessàriament que des de la segona n'hi hagi un que apunti a la primera. De fet, en el nostre exemple això només passa en dos casos: entre la pàgina personal del professor Andreu Deuloféu i la del Departament de Sociologia de la Universitat Global, i entre l'esmentada pàgina personal i la pàgina de l'assignatura de *Xarxes socials*. Aquesta és una característica que es dona sovint a les xarxes d'informació.

Com a conseqüència d'això, l'estructura d'aquesta xarxa, com la de totes les xarxes del Web, és una mica més complexa que la d'altres xarxes en què els enllaços són bidireccionals. Efectivament, encara que tota la xarxa aparentment estigui connectada, no és possible anar de qualsevol punt a qualsevol altre si respectem la direccionalitat dels enllaços. Això sí que passa entre quatre pàgines: la pàgina personal del professor Deuloféu, la de la Facultat d'Informació i Comunicació, la del Departament de Sociologia i la de l'assignatura de *Xarxes socials*. En canvi, si ens fixem en la pàgina d'inici veiem que no és accessible des de cap altra pàgina, però que des de la d'inici es pot anar a qualsevol pàgina. I mirant les pàgines dels departaments o del professorat, podem veure que només són accessibles des d'una pàgina, la principal, i que des de les pàgines dels departaments o del professorat es pot anar a qualsevol pàgina menys a la principal. D'altra banda, la pàgina de la bibliografia de l'assignatura *Xarxes socials* és

accessible des de qualsevol altra pàgina, però des de la pàgina de bibliografia no es pot anar a cap més. Així, seguint aquesta diferenciació, podem dividir la xarxa en com a mínim tres components diferents. El primer és el del conjunt dels nodes que són tots accessibles entre ells. És el que s'anomena *component fortament connectat* i en el nostre cas està format pels vèrtexs de color blau. El segon component consta dels vèrtexs que no són del component fortament connectat, però des dels quals s'hi pot accedir, anomenat *component IN* i representat en el nostre cas pels nodes verds. Finalment, l'anomenat *component OUT* el formen els nodes als quals es pot accedir des del component fortament connectat però que no en formen part, representat pel node vermell en el nostre cas.

Evidentment, el World Wide Web és molt més gran que el nostre petit exemple, però en línies generals segueix aquest tipus d'estructura. La Figura 3.5 és una representació de l'aspecte que tenia la xarxa global del WWW ja el gener del 2005.

Tot i que és molt difícil determinar amb certesa el conjunt de totes les pàgines web existents i els enllaços entre elles, hi ha sistemes que ens permeten aproximar-nos-hi. Els programes informàtics agents o *web crawlers* (també coneguts amb els noms *d'ants*, *web spiders*, *web robots*, *bots* o indexadors) es poden programar perquè vagin rastrejant automàticament el Web i emmagatzemin informació sobre les pàgines que van trobant i els enllaços continguts en aquestes pàgines. Aquesta tecnologia és la que ens permet, per una banda, conèixer l'estructura del Web (o, més ben dit, de la part del Web que l'agent ha pogut rastrejar) i, per l'altra, tenir sistemes eficients de cerca d'informació al Web, tal com veurem en l'apartat següent.

Enllaç:

En el moment d'escriure aquestes línies, algunes estimacions fetes calculen que el Web podria contenir més de 45.000 milions de pàgines web (dades extretes de <http://worldwidewebsize.com>, consultat el 8/6/2011). Cal tenir en compte, però, que aquesta xifra és una aproximació, ja que no hi ha cap sistema que ens permeti determinar amb exactitud el nombre de pàgines existents. A més, el nombre estimat només fa referència a les pàgines anomenades *estàtiques*, les que estan emmagatzemades com a tal als servidors web. No s'hi compten les pàgines dinàmiques, que són aquelles que es van creant en el mateix moment que es necessiten, responen a les peticions puntuals d'un usuari. Exemples de pàgines dinàmiques són els resultats d'una cerca a Google o d'una interrogació a una base de dades en línia. Per la seva naturalesa, el nombre de pàgines dinàmiques possibles és pràcticament infinit.

Podeu veure estimacions de la mida del WWW a:

<http://worldwidewebsize.com>

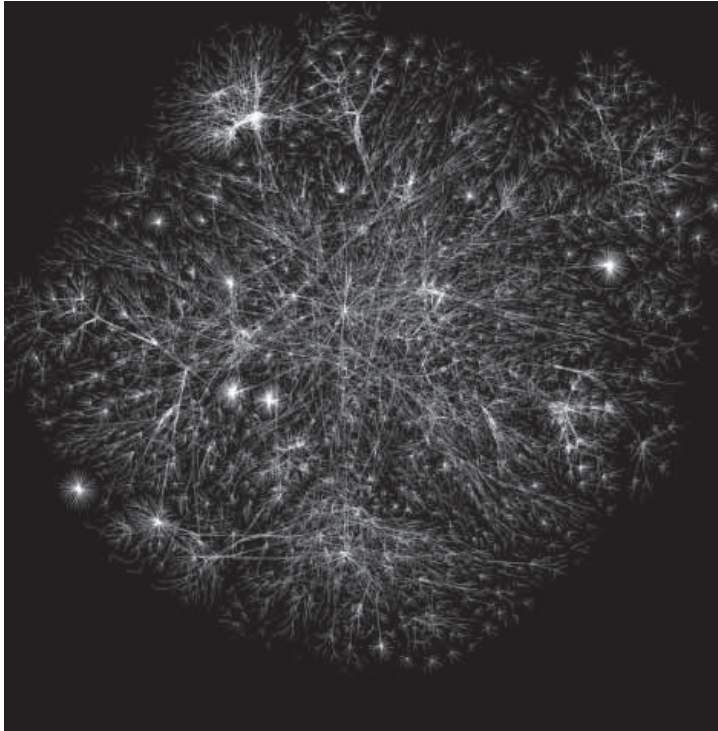


Figura 3.5. El World Wide Web. Representació de la xarxa formada pels llocs web que es relacionen per mitjà d'hiperenllaços. Les dades són del gener de 2005. Els colors representen diferents tipus de llocs web (.com, .net, .org, .cat, .org, etc.).
Font: <http://www.opte.org> llicència Creative Commons.

Tot i que els detalls tècnics poden arribar a ser complexos, el principi general del funcionament d'un agent o *crawler* és força senzill. Utilitzarem l'exemple de la Figura 3.3 per a veure com podria funcionar un agent per obtenir la informació sobre la xarxa. Al començament, l'agent es direccionaria cap a una pàgina web determinada, per exemple la pàgina principal (<http://www.uglobal.edu/index.html>). Un cop a la pàgina, emmagatzemaria la URL corresponent i aniria recorrent tot el text emmagatzemant també els elements que el programador hagi considerat interessants. Per exemple, anar registrant totes les URL que vagi trobant li permetrà tenir una llista de les altres pàgines web a les quals s'enllaça la pàgina principal i, al mateix temps, obtenir noves URL per poder anar a visitar un cop rastrejada la pàgina principal. L'agent també podria anar emmagatzemant tot el text de la pàgina, informació que podria servir per a alimentar un cercador. Després de recórrer tota la pàgina principal, l'agent aniria

a la llista de les URL emmagatzemades, en prendria una qualsevol (per exemple, <http://www.uglobal.edu/facultats.html>) i tornaria a començar el procés. D'aquesta manera, aniria recollint tota la informació sobre pàgines i enllaços que permetria confeccionar el mapa de la xarxa de la figura 3.3. Notem, però, que si en comptes de començar per la pàgina principal hagués començat per la pàgina de les facultats, el nostre agent no hauria pogut arribar mai a la pàgina principal a causa de la direccionalitat dels enllaços, i per tant aquesta pàgina no seria registrada. Aquest és el principal condicionant que fa que no pugem assegurar mai que la informació recollida pels agents sobre el Web sigui completa. Però hi ha altres factors que restringeixen també l'actuació dels agents, com és ara que hi ha servidors que no els deixen accedir a les seves pàgines o que, òbviament, no poden accedir a les pàgines dinàmiques.

3.2.2. Cerca d'informació

Hem vist que el que caracteritza les xarxes d'informació com el Web és precisament que els nodes contenen informació que es pot transferir pels enllaços. Aquesta informació pot ser molt útil a les persones i a les institucions que tenen accés a la xarxa si saben on anar a buscar-la. Però no és fàcil saber quina informació conté la xarxa i a quin node anar-la a cercar. Per això s'han desenvolupat sistemes que faciliten el procés de cerca anomenats motors de cerca o cercadors. En aquest apartat veurem en què es basen els sistemes de cerca d'informació al Web i com aprofiten l'estructura de la xarxa per a optimitzar l'eficàcia i l'eficiència del procés.

El procés de cerca al Web es desenvolupa normalment en dues fases. La primera consisteix a utilitzar agents com els que hem descrit en l'apartat anterior per a rastrejar el Web. Durant aquesta fase, l'agent o agents van recorrent totes les pàgines que troben per enregistrar-ne el contingut i anar-ne creant un índex anotat. Aquest índex pot contenir llistes de les paraules que apareixen a cada pàgina, el lloc i la freqüència amb què hi apareixen i també estimacions de la importància de les pàgines basades en diversos criteris. Llavors aquest índex es deixa emmagatzemat i preparat per a poder-hi accedir

* Podeu trobar un estudi exhaustiu de l'estructura d'Internet i del World Wide Web al llibre de Pastor-Satorras i Vespignani (2004): R. Pastor-Satorras; A. Vespignani (2004). *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge, RU: Cambridge University Press. El tractament es fa des del punt de vista de la física estadística i requereix un cert nivell matemàtic.

en la segona fase del procés de cerca, que és la cerca pròpiament dita. Aquesta fase comença quan un usuari envia una petició de cerca al motor. Llavors el cercador utilitza l'índex creat anteriorment per a localitzar les pàgines que més s'adeqüen a la petició.

Tant en la fase de rastreig com en la de cerca, els motors de cerca poden utilitzar l'estructura de xarxa del Web (que han pogut anar recollint a partir de les URL dels enllaços que apareixen a les pàgines tal com hem vist en l'apartat anterior) per a optimitzar els processos. En la fase de rastreig, per exemple, s'utilitzen algorismes derivats de la teoria de xarxes per a guanyar en eficiència i es fan servir altres estratègies que tenen en compte les característiques de la xarxa, com és ara repartir motors diferents per diverses parts del Web que vagin treballant en paral·lel.

En la fase de cerca els primers cercadors es basaven només en la informació recollida en els índexs. Així, en rebre la petició de cerca amb una paraula, el cercador podia determinar el conjunt de pàgines en què apareixia. Si en la petició hi havia més d'una paraula, podia fer la llista de les pàgines en què apareixien totes dues. Per tal de prioritzar la llista en funció de la rellevància per a l'usuari, es podia basar en el nombre de vegades que apareixien les paraules en cada pàgina. Fins i tot podia emprar criteris una mica més elaborats, com per exemple si les paraules apareixien en els encapçalaments de la pàgina o, en el cas de peticions de diverses paraules, si aquestes apareixien en les pàgines en posicions properes. Tot i així, els motors de cerca que només es basaven en la informació recollida en els índexs i els criteris sobre el text no obtenien resultats gaire satisfactoris, i van ser superats per cercadors de nova generació com Google, més sofisticats, que incorporaven altres criteris com el coneixement de l'estructura de la xarxa.

Els cercadors moderns encara utilitzen els índexs, però només com a primer pas del procés de cerca. Amb la informació recollida en els índexs, determinen un conjunt de pàgines que potencialment podrien ser rellevants per a l'usuari que fa la petició. Aquest conjunt es mira que sigui ben ampli, de manera que contingui la major part de les pàgines que puguin ser d'interès, encara que moltes siguin irrellevants. Així, pot incloure no tan sols les pàgines que contenen les paraules de la petició de cerca, sinó també les pàgines que estan enllaçades amb aquestes. La idea és que aquest conjunt inicial de pàgines es vagi afinant successivament amb l'aplicació de diversos criteris addicionals, alguns dels quals es fonamenten en la topologia de la xarxa.

Aquests criteris poden variar segons el motor de cerca, i les empreses que hi ha al darrere miren d'anar-los millorant contínuament i mantenir-los en secret en la mesura del possible. Se sap que en el cas de Google, un dels criteris més

importants que fa servir és la mesura de centralitat de la xarxa anomenada PageRank (veure capítol 1). Així, un dels factors que determinen el grau de rellevància d'una pàgina depèn del nombre d'enllaços que hi van a parar, però també del nombre d'enllaços que van a parar a les pàgines que hi enllacen. Aquest, però, és només un dels components de la fórmula total que fa servir Google per a determinar el grau de rellevància i que serveix per a ordenar el resultat final de la cerca. Alguns d'aquests altres components poden estar basats en criteris més tradicionals com l'aparició de les paraules de la petició de cerca, la freqüència i la posició amb què apareixen, i d'altres en l'històric de cerques de l'usuari, prioritzant per exemple les pàgines que han estat seleccionades en cerques anteriors.

Com es pot veure, de tots els criteris que utilitzen els cercadors, n'hi ha uns quants que depenen de la petició concreta de cerca que fa l'usuari introduint unes paraules específiques a l'equació de cerca, però n'hi ha d'altres que només depenen de l'estructura d'enllaços de la xarxa. Com que aquesta darrera es coneix a partir de la informació recollida pels agents, els càlculs corresponents al segon tipus de criteris es poden tenir fets per endavant, la qual cosa permet un estalvi de temps considerable en obtenir els resultats finals per part del cercador.

3.2.3. Altres xarxes d'informació al Web

Tal com hem vist als apartats anteriors, el World Wide Web és potser el cas més rellevant del que anomenem *xarxes d'informació*. Tanmateix, no n'és de cap manera l'únic exemple. És més, fins i tot trobem altres xarxes d'informació que utilitzen com a base l'infraestructura d'Internet i del Web però que tenen les seves pròpies característiques. En aquest apartat en veurem algunes.

Les xarxes d'igual a igual o *peer-to-peer* (P2P) han esdevingut força populars en els darrers anys com a mètode per a compartir fitxers entre usuaris d'Internet. En una xarxa d'igual a igual els nodes són els ordinadors dels diversos usuaris, que poden contenir informació en forma de diversos tipus de fitxers, i les connexions entre ells són enllaços virtuals que s'estableixen amb l'objectiu de poder compartir la informació dels fitxers. Aquestes xarxes es fan servir per a compartir per Internet, de vegades de manera il·legal, diversos tipus de continguts com és ara música, pel·lícules o jocs.

La característica diferencial de les xarxes P2P és que la informació en forma de fitxers es transmet entre usuaris finals de la xarxa ("iguals" o *peers*) sense que calgui la intervenció de cap servidor, tal com passa a l'estructura WWW. La Figura 3.6 mostra la diferència entre una estructura de xarxa d'igual a igual i la tradicional client-servidor on la informació és subministrada als usuaris

per un servidor. En una xarxa d'igual a igual cadascun dels nodes té part de la informació, però no n'hi ha cap que la tingui tota. La transmissió d'informació es fa sempre entre dos usuaris finals de la xarxa. El fet que no hi hagi servidors facilita la distribució il·legal d'informació perquè no hi ha servidors que les autoritats puguin obligar a tancar.

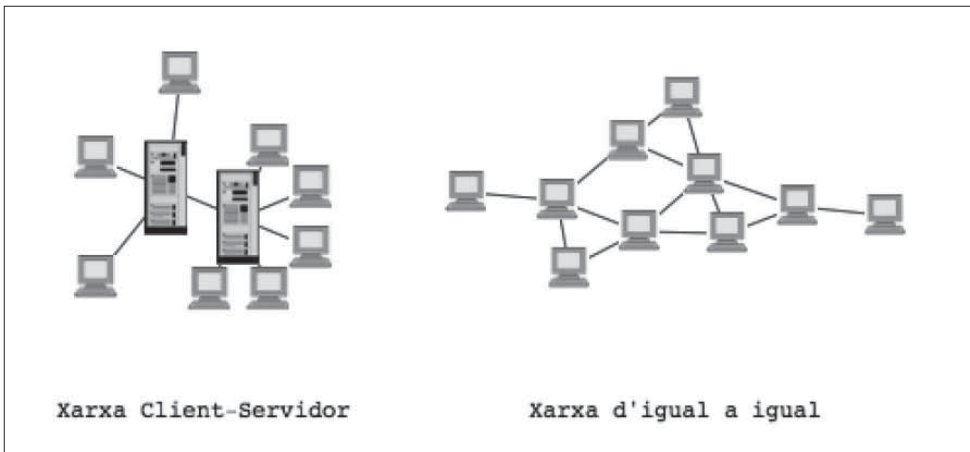


Figura 3.6: Xarxes client-servidor i d'igual a igual

El problema principal d'una xarxa d'aquest tipus és com esbrinar quin node de la xarxa conté la informació que es busca. Al principi això es va solucionar mitjançant una mena de servidors que contenien índexs de què hi havia a cada ordinador membre de la xarxa. Aquesta és la tecnologia que utilitzava la xarxa de compartició de fitxers *Napster* (avui dia reconvertida en un simple web de venda de música en línia). Evidentment, però, aquest sistema era fàcil de controlar per les autoritats simplement intervenint els ordinadors on hi havia els índexs. Per això es van idear sistemes més sofisticats que utilitzaven l'estructura de la xarxa com els que implementen *Gnutella* o *BitTorrent*. El més simple consisteix en el fet que el node des del qual es busca un fitxer "pregun-ti" als nodes amb què està connectat virtualment a la xarxa P2P. Si un d'ells té el contingut que es busca, es transfereix. Si no, aquests nodes veïns de l'inicial, pregunten al seu torn als seus veïns, i així successivament fins que la informació es localitza i el fitxer es pot transferir. Aquest algorisme de cerca, però, presenta alguns problemes tècnics quan es vol utilitzar en xarxes grans i per això s'han ideat estratègies més sofisticades que aprofiten la idea de "hubs" (nodes altament connectats) a les xarxes lliures d'escala presentades al primer capítol

d'aquest llibre. La idea és llançar les cerques primer a la xarxa formada pels *hubs* i d'aquests anar als altres nodes, cosa que permet trobar el fitxer desitjat més eficientment.

Una altra mena de xarxes d'informació que podem trobar al Web són les *xarxes de recomanació*. Molts de nosaltres hem experimentat com en alguns webs de comerç electrònic se'ns recomanen productes que, curiosament, trobem prou interessants, i sovint acabem comprant. Aquests sistemes utilitzen les xarxes de recomanació.

Una xarxa de recomanació és una xarxa bipartida, en la qual un tipus de nodes representa els productes i l'altre tipus les persones. Els enllaços connecten una persona amb un producte que ha adquirit o bé que ha manifestat d'alguna manera que li agrada. La Figura 3.7 mostra una possible xarxa de recomanació d'un web de venda de llibres com és ara Llibres.cat o Amazon.com.

Els anomenats *sistemes de filtratge col·laboratiu* o bé *sistemes de recomanació* són algorismes que, a partir de la informació continguda en aquestes xarxes, poden utilitzar els llocs web comercials per a fer recomanacions als clients. Podem veure la idea general de com funcionen utilitzant l'exemple de la figura 3.7, que representa una part de la xarxa de recomanació d'una botiga en línia de venda de llibres. A la xarxa bipartida hi ha uns nodes que representen clients i uns altres que corresponen a llibres que formen part de l'oferta del lloc web. Cada client està enllaçat amb els llibres que ha comprat. Si ens fixem en els llibres comprats per l'Ariadna, veiem que en comparteix tres amb en Biel. Això pot voler dir que tenen alguns interessos comuns, i que algun dels altres llibres que ha comprat en Biel poden interessar a l'Ariadna, i viceversa. Si a més en Cesc comparteix també tres llibres comprats amb en Biel, i que un dels altres que ha adquirit, *Game of Thrones*, també el té l'Ariadna, encara és més probable que a en Cesc li interessi. Aquest tipus de lògica és el que els algorismes dels sistemes de filtratge col·laboratiu apliquen a la xarxa que construeixen amb les dades de les seves vendes, i d'on surten les recomanacions que rebem.

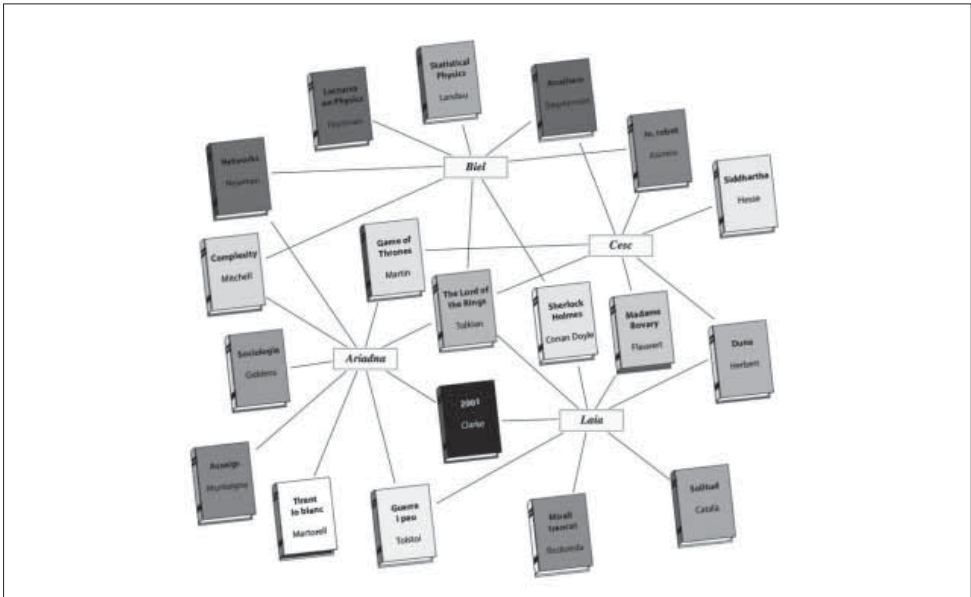


Figura 3.7. Xarxa de recomanació d'una llibreria en línia

Un altre cas interessant de xarxes d'informació a Internet és el de les *xarxes de mitjans socials* com els blogs o bé els microblogs com el Twitter. En aquestes xarxes els nodes són els mitjans socials i les connexions són els enllaços que es fan de l'un a l'altre. Analitzar aquestes xarxes permet veure clarament com són i com evolucionen les posicions dels mitjans socials en funció de les relacions que ells mateixos estableixen. Això és especialment interessant en àmbits específics com el polític.

Enllaç:

<http://politicosphere.net/map/>

A la pàgina web PoliticoSphere.net es pot trobar la representació gràfica de la xarxa de mitjans socials, blogs especialment, sobre temes polítics durant el mes de juny de 2009 als Estats Units. Els colors distingeixen les diferents tendències ideològiques. Per exemple, el vermell correspon als republicans, el blau als demòcrates, el verd als ecologistes i el groc als professionals de la comunicació. El mapa és interactiu i permet identificar els mitjans i destacar-ne les connexions. S'hi pot veure clarament com els mitjans de tendències afins tendeixen a agrupar-se, però que n'hi ha alguns que poden exercir de pont.

Un altre tipus de mitjans socials que té estructura de xarxa d'informació són els wikis. Cada entrada de la Wikipedia, per exemple, té enllaços a altres entrades que es poden rastrejar amb programes agent per a poder conèixer quina estructura de xarxa formen. Darrerament, s'estan fent estudis fins i tot sobre la xarxa bipartida que formen els usuaris que participen en l'elaboració de la Wikipedia i les entrades a les quals contribueixen.

Finalment, no cal dir que actualment són especialment importants al Web les xarxes que s'estableixen entre els espais dels usuaris als llocs web de xarxes socials o *social networking sites* (SNS) com Facebook o LinkedIn, tal com ja s'ha vist anteriorment. Són xarxes d'informació perquè els nodes que s'hi enllacen són, de fet, pàgines web, però també són la representació en línia de les xarxes socials entre les persones.

3.3. Difusió del coneixement en xarxes socials

És evident que a la societat humana, les idees o els comportaments d'uns individus influeixen en els dels altres. Qualsevol innovació, tant si són idees o opinions com la utilització de productes concrets, es transmeten d'uns individus a uns altres i d'aquesta manera s'escampen pel teixit social. Aquest fenomen rep el nom de difusió de la innovació i s'ha estudiat amb profunditat des de mitjan segle passat.

Es considera que una innovació s'ha difós a un individu o a un conjunt d'individus un cop aquests l'han adoptada. És a dir, que no n'hi ha prou que a una persona li arribi informació, per exemple, sobre una nova idea política o un nou producte de neteja perquè considerem que hi hagi hagut difusió. Cal un segon pas consistent que l'individu en qüestió faci seva la idea o passi a utilitzar el producte.

Aquest procés, de fet, té molts paral·lelismes amb el de generació de nou coneixement a partir de la informació que rebem de l'exterior, tal com l'hem presentat a la primera secció d'aquest capítol. Quan ens arriba una informació, utilitzem el nostre coneixement previ per a filtrar-la i decidir si ha de contribuir o no a modificar el nostre pòsit de coneixement i, com a conseqüència, potser també modificar el nostre comportament. Per entendre la difusió del coneixement a les estructures socials, doncs, podem manllevar idees desenvolupades en l'àmbit de la difusió de la innovació.

En aquesta secció veurem com la introducció de l'anàlisi de xarxes socials permet superar el tractament tradicional dels processos de difusió als grups socials en tenir en compte les relacions concretes entre els individus. Aquestes relacions afecten tant la difusió d'informació com el pas posterior de difusió del coneixement.

3.3.1. Aproximació tradicional a la difusió social

Tradicionalment, l'estudi de la difusió social de la innovació s'ha fet prenent com a referència un grup social determinat que es veu exposat a una nova idea o un nou producte i veient com la innovació la van adoptant els membres del grup i quins factors hi poden influir. Sempre, però, considerant el grup com un conjunt d'individus i sense atendre les relacions específiques entre ells. Resumirem aquí algunes de les conclusions més rellevants dels estudis fundacionals d'aquest tema.

Un dels aspectes fonamentals de la teoria de la difusió de la innovació és que no totes les persones reaccionen igual davant d'una innovació, encara que els possibles beneficis que en comportaria l'adopció siguin iguals per a tothom. Així, hi ha individus que de seguida volen provar coses noves, i d'altres que hi són molt refractaris i no decideixen adoptar la innovació fins que ja gairebé tothom ho ha fet, i sovint perquè no tenen cap altre remei. Aquesta actitud està determinada per les característiques psicològiques de cada persona, ja que cadascú té un nivell diferent de tolerància a la incertesa i el risc que representa adoptar la innovació o a la pressió social del grup quan aquest majoritàriament ha decidit adoptar-la o no adoptar-la.

Rogers (2003) distingeix entre cinc grans categories d'individus pel que fa a l'actitud davant de les innovacions: innovadors, primers seguidors, majoria primerenca, majoria tardana i ressagats. Aquesta actitud influeix en el moment de l'adopció, de manera que els individus s'hi van incorporant al llarg del temps en un moment o altre depenent del grup al qual pertanyen. A la Figura 3.8 es pot veure la distribució d'aquests grups en la població seguint una distribució normal (línia vermella) i el percentatge de la difusió que s'assoleix quan cada grup va adoptant la innovació al llarg del temps (línia blava), suposant que tothom l'acabi adoptant i no hi hagi ningú que es tiri enrere i descarti la innovació després de provar-la.

* Vegeu el llibre *Diffusion of Innovations* d'Everett Rogers (2003), un tractat molt complet de la teoria tradicional de la difusió de la innovació.
E. M. Rogers (2003). *Diffusion of innovations*. Nova York: Free Press.

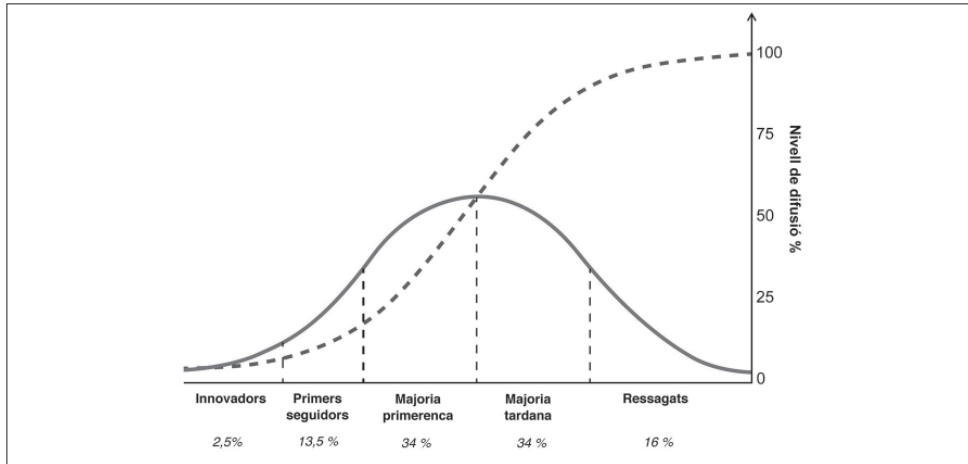


Figura 3.8: La difusió de les innovacions segons Rogers (2003). La línia contínua representa la distribució dels individus en categories pel que fa a l'adopció i la línia discontinua el percentatge de la població total que ha adoptat en el temps a mesura que les diferents categories s'hi van incorporant (en el supòsit que ningú no desfà l'adopció i tota la població l'acaba adoptant).

El ritme d'adopció és la velocitat relativa amb la qual els membres del grup social adopten la innovació. Se sol mesurar amb el temps que un cert percentatge del grup social triga a haver adoptat la innovació i depèn de la naturalesa de la innovació, de les característiques dels individus i del mecanisme d'adopció. Es diu que el procés d'adopció ha assolit la massa crítica quan ha arribat a un percentatge de difusió determinat a partir del qual la difusió continuarà fins a arribar a tota la població.

Un dels mecanismes proposats per explicar com es produeix la difusió es basa en el concepte de llindar d'adopció. La idea és que cada individu decideix adoptar la innovació quan el nivell de difusió en tot el grup arriba a un punt determinat. Aquest punt és diferent per a cada persona i depèn de les seves característiques psicològiques i de la utilitat (sobre la base del balanç cost-benefici) que té per a ella adoptar la innovació. Així, els innovadors serien els que tenen el llindar més baix, i els ressagats els que per a adoptar la innovació necessiten que ja estigui difosa a gairebé tota la població. El llindar de cada individu no és fix, sinó que es pot anar modificant amb el temps a causa,

* Models semblant a aquest que presentem i d'altres una mica més sofisticats que també es poden adaptar a la difusió de les innovacions i el coneixement, s'han estudiat en profunditat en l'àmbit de l'epidemiologia per a investigar els processos de contagi a les epidèmies. Els epidemiòlegs han estat també pioners en l'ús de l'anàlisi de xarxes socials per a entendre com s'han escampat malalties com la sida i la SARS.

per exemple, de l'accés a informació addicional sobre la innovació o bé per la influència dels mitjans de comunicació o de líders d'opinió.

Analitzant aquests mecanismes, és clar que l'entorn social té un paper molt important en la difusió de la innovació i també de tot tipus de coneixement. Això fa pensar que té sentit incorporar a la teoria la consideració de l'estructura social concreta que envolta cada individu, és a dir, de la xarxa social. Als apartats següents veurem com la xarxa social influeix tant en la difusió d'informació com en la decisió d'adopció de les innovacions o la incorporació de nou coneixement.

3.3.2. Difusió d'informació en xarxa

Hem vist que als models tradicionals s'assumia que cada individu prenia decisions en funció d'estats generals del sistema, com per exemple el percentatge d'adopció total del grup, la qual cosa només seria possible si aquest individu tingués informació sobre l'estat de tots els altres. Això en realitat no és així. Cadascun de nosaltres només interacciona amb un grup més o menys limitat de persones, i la informació que tenim ens arriba a través d'aquests contactes, tant sobre ells mateixos com sobre d'altres amb els quals ells tenen contacte directe o indirecte. És evident, doncs, que qualsevol model de difusió d'informació o de coneixement que vulgui ser una mica realista haurà de tenir en compte l'estructura de les xarxes socials on cada individu està immers.

Un dels models més senzills que podem utilitzar per a descriure la difusió d'informació en xarxes és l'anomenat *model SI* en la versió en xarxa. Aquest model també prové de l'epidemiologia, ja que de fet la transmissió d'informació entre persones té una dinàmica força similar a la transmissió de malalties contagioses. Considerem un grup social determinat i la xarxa social que formen on dos individus estan connectats si interaccionen amb una determinada freqüència. Suposem que quan un individu té una determinada informació (tant si és certa com un simple rumor), al cap del dia (o de la unitat de temps que vulguem considerar) la transmet amb una probabilitat donada a cadascun dels seus veïns a la xarxa. Un cop un dels veïns ha rebut la informació, aquest segueix el mateix procediment amb els seus, i així successivament. Els nodes de la xarxa poden estar en dos estats diferents: sense tenir la informació o tenint-la. Un cop tenen la informació, encara que se'ls transmeti altra vegada ja no els afectarà i continuaran en el segon estat. Amb el temps, la informació es va estenent per tota la xarxa si està connectada, o bé per tot el component al qual hi havia el primer node que obté la informació. Evidentment, aquest model és aplicable tant a les xarxes socials tradicionals com a les associades als SNS.

Enllaç:

Simulació parametrizable del model SI desenvolupat per Lada Adamic (Universitat de Michigan):

<http://www-personal.umich.edu/~ladamic/netlearn/NetLogo412/BA-Diffusion.html>

Al mateix web es poden trobar altres simulacions i demostracions molt interessants sobre temes relacionats amb les xarxes.

Un cop hem vist la dinàmica d'aquest model senzill, podem aprofundir en alguns aspectes interessants del procés. Per exemple, com que per a difondre's a tota la població una informació ha d'anar del node inicial a tots els altres per camins formats per les connexions de la xarxa, podem deduir que aconseguir-ho serà molt més fàcil en "mons petits" o xarxes en les quals la distància mitjana entre els nodes sigui petita.

D'altra banda, sabem que obtenir o difondre una informació ràpidament pot donar avantatges individuals considerables en activitats econòmiques o en altres aspectes de la vida. Podem preguntar-nos, doncs, si hi ha individus privilegiats a causa de la posició que tenen a la xarxa social. Sembla, per exemple, que la centralitat ha de ser important. Per a veure-ho, farem servir la figura 3.9, on reproduïm els gràfics utilitzats al capítol 1 per a veure els valors dels diversos tipus de centralitat per a una xarxa d'exemple.

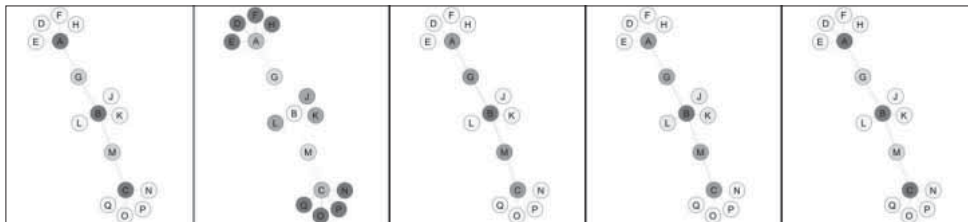


Figura 3.9: Representació de la xarxa amb colors de nodes segons el grau de centralitat que tenen: grau, *closeness*, *betweenness*, *eigenvector* i *Page Rank*. Els colors més clars corresponen a valors numèrics més baixos i els més foscos a valors més alts.

Els nodes amb grau més alt, per exemple, en tenir un nombre més gran d'enllaços, estan en condicions de rebre –o emetre, si escau– informació per més vies, i per tant, rebre'n més o rebre-la més aviat. Una situació similar, o fins i tot més acusada, és la dels nodes amb valors *eigenvector centrality* o *PageRank centrality* més alts, ja que en aquests casos no tan sols el node en qüestió té

més enllaços, sinó que els seus veïns també. Els individus amb *closeness centrality* alta són més “a la vora de tota la xarxa”. Per tant, podem pensar que una informació que surti de qualsevol punt de la xarxa hi arribarà relativament de pressa. Valors alts de *betweenness centrality* fan més probable que qualsevol informació que viatgi per la xarxa hagi de passar pels nodes que els tenen, i per tant els individus que ocupen aquests nodes probablement obtindran molta informació, i també tindran la possibilitat d’impedir que els altres la rebin si decideixen no transmetre-la. Però això darrer ja va més enllà del senzill model que hem descrit aquí, i comença a tenir relació amb el coneixement que la informació pot transmetre.

3.3.3. De la difusió d’informació a la difusió de coneixement

A l’apartat anterior hem vist un senzill model de com es pot difondre la informació per mitjà d’una xarxa social. Ara bé, si el que volem és entendre com es transmet el coneixement o com es difonen les idees, els comportaments o les normes socials, cal que els models incorporin una mica més de complexitat. Perquè es difongui coneixement no n’hi ha prou que es transmeti la informació. Cal que el receptor modifiqui la base de coneixement que té (entès en sentit ampli, com l’hem definit a la primera secció d’aquest capítol) sobre la base de la informació rebuda, la qual cosa farà que canviï, poc o molt, la manera d’interpretar la realitat que l’envolta i, com a conseqüència, el comportament. A continuació veurem alguns dels aspectes importants que hem de tenir en compte si volem entendre aquests processos.

Hem vist anteriorment que a la teoria tradicional de difusió d’innovacions no n’hi havia prou que arribés informació sobre una innovació a un individu perquè aquest decidís adoptar-la. Que això passés depenia de la utilitat que hi percebia el possible adoptant i de la psicologia de l’individu, però també del grau de penetració de la innovació en el grup. Tot això s’incorporava a l’explicació del mecanisme de difusió basada en un lllindar d’adopció dels individus. Aquesta idea de lllindar es pot traslladar als models de difusió en xarxa d’una manera molt més realista.

A l’hora de decidir, per exemple, donar per bona una informació que rebem, incorporar-la al nostre coneixement i modificar el nostre comportament en conseqüència, normalment no tenim informació sobre si ho han fet o no els diferents membres del nostre grup social. Però sí que solem conèixer el que han fet els individus amb els quals tenim més contacte, els nostres veïns a la xarxa social. Per això té sentit pensar en un mecanisme en el qual

el llindar que fa que passem a donar per bona una informació o adoptar un comportament sigui respecte dels nostres contactes directes. La Figura 3.10 representa un exemple d'una xarxa social en la qual s'està difonent una informació que pot canviar el coneixement dels seus integrants. Quan això passa, els nodes corresponents passen de color vermell a color verd. Ens focalitzem en l'individu representat pel node central i representem només la seva part de la xarxa social (ell, els seus veïns i els enllaços als veïns dels seus veïns). Aquest individu suposem que té un llindar del 50%. És a dir, incorporarà la informació quan ho hagin fet la meitat o més dels seus veïns a la xarxa. A mesura que passa el temps i que els nodes van canviant de color, arriba un moment en el qual tres dels cinc veïns del node focal ja són de color verd, i això fa que al període següent ell també canviï de color.

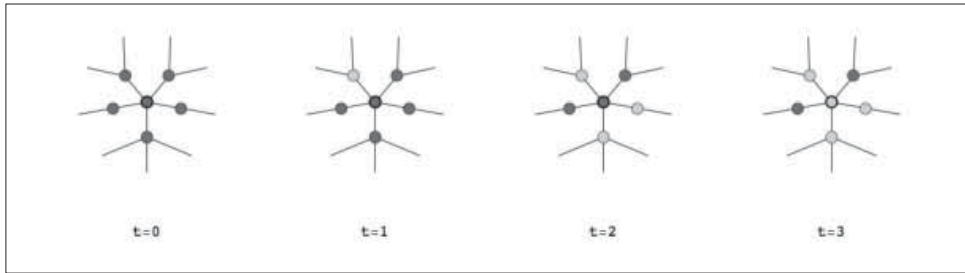


Figura 3.10: Difusió de coneixement segons un llindar. El node central incorpora la informació a la seva base de coneixement després que 3 dels seus 5 veïns ho han fet, ja que té un llindar del 50%.

Una altra consideració important que cal fer quan estem parlant de difusió de coneixement és la complexitat del procés. Tenint en compte el que hem vist a la primera secció d'aquest capítol sobre la relació entre dades, informació i coneixement, el coneixement com a tal no es pot transferir directament. El que passa és que un agent que posseeix un determinat coneixement i el vol transmetre, el codifica d'alguna manera i genera una informació que, mitjançant un substrat físic de dades, podrà transmetre a un altre agent. Aquest segon agent després haurà de descodificar aquesta informació rebuda per tal d'incorporar coneixement a la seva base. És a dir, que un individu a partir del seu coneixement genera informació perquè un altre la utilitzi per a generar-se nou coneixement, tot amb l'objectiu que el nou coneixement generat pel segon individu serveixi per al mateix que servia el del primer. Per això, simplificant, diem que s'ha transferit el coneixement, encara que probablement les estructures mentals que emmagatzemen el coneixement en qüestió en un i altre individu poden ser molt diferents.

En el nostre model de difusió d'informació en xarxa, que es fes efectiva la transmissió era simplement una qüestió aleatòria, com si es tractés d'encomanar-se un refredat en entrar en contacte amb el virus. Quan modelitzem la difusió de coneixement, hem de tenir en compte que el node de la xarxa que vol emetre el coneixement ha de ser capaç de codificar la informació corresponent d'una manera que el receptor la pugui descodificar. Això, evidentment, dependrà de la base de coneixement total prèvia de tots dos. Per exemple, la informació continguda en un article escrit per un especialista en física quàntica només podrà ser descodificada per un altre físic amb prou coneixements per a entendre'l; o bé, per a poder entendre la major part de les obres d'art del renaixement italià cal tenir un cert coneixement de la religió catòlica. A les xarxes dels mitjans socials això també s'experimenta sovint. La majoria de nosaltres, per exemple, hem rebut alguna piulada de Twitter que no entenem de res encara que sapiguem llegir-ne les paraules. Ens falta el coneixement del context que comparteixen els "iniciats" en el tema en qüestió.

Alguns models de difusió del coneixement en xarxa incorporen això suposant que cada individu de la xarxa té una certa base de coneixement, que no cal que s'especifiqui. El que sí que cal especificar és una "distància" entre el coneixement d'aquests individus. Aquesta distància seria molt petita entre dos economistes nascuts a Nova York i educats a Harvard, i seria molt gran entre un d'aquests economistes i un lama tibetà. Llavors la probabilitat de transferir coneixement entre dos nodes serà inversament proporcional a la distància entre els seus coneixements. Si aquesta distància entre un node i algun dels seus veïns és molt gran, pot ser que tot i superar el llindar pel que fa al nombre de veïns que tinguin un determinat coneixement, aquest no se li acabi transferint perquè malgrat que la informació des del veí distant li arriba, no és capaç de descodificar-la i detectar que efectivament ha superat el llindar.

En la societat moderna el contacte directe amb la nostra xarxa social no és l'única manera de transmetre coneixement. Els mitjans de comunicació i la difusió mitjançant publicacions escrites i Internet té un paper molt important. Podem incorporar també aquests efectes al nostre model. Podem recuperar, per exemple, el concepte de *massa crítica* i considerar que, quan un cert coneixement està difós a la població més enllà d'un determinat nivell, la massa crítica, es difon a tots els individus que encara no el tenien mitjançant, per exemple, els mitjans de comunicació. Tot això, independentment de la transmissió directa mitjançant els enllaços de la xarxa social. D'altra banda, podem considerar que hi ha un segon "efecte llindar" que pren com a referència no els veïns directes a la xarxa, sinó els líders d'opinió del grup social. Així, si un determinat percentatge dels líders d'opinió tenen un coneixement, podem su-

posar que es difon, per exemple, mitjançant publicacions o per Internet. Qui són els líders d'opinió? Doncs una bona aproximació al nostre model podria ser considerar que són els nodes de la xarxa amb grau més elevat. Recordem que en la discussió sobre hubs i autoritats del capítol 1, els nodes amb centralitat molt elevada tenen un paper molt important a les xarxes lliures d'escala, com solen ser les xarxes socials.

3.3.4. Influència social

Globalment, la difusió de coneixement (és a dir, d'idees, valors, etc.) té efectes importants en el comportament dels individus que formen la societat. Anomenem *influència social* l'efecte pel qual els pensaments, els sentiments o les accions d'un individu es veuen afectats per altres individus. Aquest fenomen té l'arrel en les necessitats psicològiques dels humans de complir les expectatives dels altres. Això inclou la necessitat de tenir raó (influència social informativa) i la d'agradar als altres (influència social normativa).

Evidentment, la influència social sobre un individu la poden exercir amb més força, d'entre tots els seus *alteri*, aquells que per a ell són més importants. Això inclou tots aquells que tenen un cert prestigi social, però també, i especialment, aquells amb els quals interacciona més sovint, és a dir, els que té més a prop dins de la xarxa social. És lògic, doncs, que l'estructura de la xarxa social tingui un paper en com es produeix la influència social. Aquí ens centrarem en aquest aspecte.

Un dels tipus de coneixement (com sempre, en el sentit ampli apuntat anteriorment) que sembla més propens a difondre's per les xarxes socials són les opinions. Els exemples d'algunes opinions que de cop i volta prenen una inusitada popularitat a la xarxa de Facebook com a "m'agrada" o "no m'agrada", o a la de Twitter com a *trending topics* semblen donar suport a aquesta idea. De fet, s'han elaborat teories de difusió d'opinió a les xarxes socials amb models semblants als que hem descrit a la secció anterior que prediuen resultats força ajustats al que passa a la realitat. Per exemple, es veu que diverses opinions poden coexistir en el temps a la xarxa social i que normalment les opinions es localitzen a zones concretes de la xarxa. D'altra banda, les opinions poden canviar amb una certa facilitat i es confirma la influència en la formació d'opinió dels individus amb alt grau de connectivitat, que sembla que actuen com a líders d'opinió.

Nicholas Christakis i James Fowler, investigadors de la Universitat de Harvard i de la Universitat de Califòrnia San Diego respectivament, van dur a

terme una sèrie d'estudis empírics amb diversos grups de persones sobre els quals tenien informació relativa a les xarxes socials que formaven i a una sèrie d'altres característiques individuals per a mirar d'estudiar la influència social per mitjà de les xarxes. Van arribar a la conclusió que les xarxes socials "transmeten" inclús estats i comportaments com l'obesitat, l'hàbit de fumar, el consum d'alcohol i fins i tot la felicitat. I no tan sols a una distància d'un enllaç, sinó a distàncies de fins i tot tres enllaços. És a dir, que una persona té una probabilitat més alta de, per exemple, ser obesa no solament si els seus veïns a la xarxa social són obesos, sinó també si la majoria dels veïns dels seus veïns ho són, encara que els veïns directes no ho siguin.

Aquests resultats s'han presentat algunes vegades amb enunciats del tipus "l'obesitat s'encomana per mitjà de la xarxa social", la qual cosa sembla com a mínim poc rigorosa. Però sí que és veritat que les dades suggereixen clarament que per la xarxa social es transmet alguna cosa que provoca aquest efecte. Sembla difícil que sigui un "virus de l'alcoholisme" o un "elixir de la felicitat". Però probablement sí que els comportaments, les emocions, els valors, etc. que observem al nostre entorn ens afecten en la nostra base de coneixement i, per tant, modifiquen també el nostre comportament.

En aquesta secció hem vist com l'estructura de la xarxa social té una incidència en els processos de difusió del coneixement i d'influència social. Notem, però, que aquests processos poden també, al seu torn, modificar l'estructura de la xarxa social. Si mitjançant els enllaços de la xarxa ens arriben noves idees polítiques o religioses, per exemple, i nosaltres les fem nostres, és molt possible que comencem a relacionar-nos més sovint amb gent que comparteixi les mateixes idees i, per tant, la nostra xarxa social canviï. De fet, podem dir que el coneixement i la xarxa social actuen simultàniament l'un sobre l'altra i això els fa canviar alhora. Diem que coevolucionen.

3.4. Xarxes de coneixement científic

La ciència ha estat durant molts segles el paradigma de la creació i la difusió de coneixement. Actualment, totes les organitzacions es preocupen per la generació i la transmissió del coneixement i per això la gestió del coneixement

* El llibre *Connected* de Christakis i Fowler (Christakis and Fowler 2009) és una bona obra de divulgació sobre com ens afecten les xarxes socials escrita per científics especialistes en el tema. N. Christakis; J. Fowler (2009). *Connected: The Amazing Power of Social Networks and How They Shape Our Lives*. Londres: Harper Press.

ha esdevingut molt popular. La raó de ser dels científics és precisament això, la generació de coneixement a partir de la investigació i la difusió posterior a la comunitat científica. No és estrany, doncs, que hagi desenvolupat un dels primers sistemes de gestió del coneixement: la publicació científica.

Perquè siguin admesos per la comunitat científica, els resultats del treball dels investigadors s'han de publicar en articles en alguna de les revistes científiques que hi ha. A banda d'assegurar que el coneixement generat es transmeti a tota la comunitat, el sistema incorpora un mecanisme d'assegurament de la qualitat: la revisió dels articles abans de ser publicats per altres científics de l'àmbit, el *peer review*. Encara que de vegades s'arriben a publicar resultats erronis o fins i tot frauds, la veritat és que aquest sistema aconsegueix minimitzar aquests casos.

Però hi ha més aspectes interessants en el sistema de publicacions científiques. Generalment, els autors que signen un article realment han contribuït d'una o altra manera a la investigació, i és poc freqüent que algú que hagi fet una contribució important no figuri com a autor. Els autors, a més, signen amb la seva filiació, amb el nom de la institució o institucions a les quals pertanyen. Sovint, els articles inclouen també un conjunt de paraules clau que defineixen aproximadament el tema de l'article.

El que s'afirma en un article científic ha de tenir una base sòlida. Si és de collita pròpia de l'autor, ha de ser justificat adequadament, i si és un coneixement desenvolupat per un altre científic, se n'ha de citar la font, normalment un altre article científic. D'aquesta manera se n'assegura la qualitat i es reconeix la contribució al treball propi dels científics que l'han precedit.

Totes aquestes característiques de la literatura científica fan que els especialistes en ciències tinguin a l'abast una gran quantitat de dades que permeten, aplicant les tècniques de l'anàlisi de xarxes, saber més sobre com funciona aquesta impressionant maquinària de generació i difusió de coneixement que és la ciència.

3.4.1. La construcció del coneixement científic: xarxes de citacions

Utilitzant una coneguda metàfora del filòsof Bernard de Chartres, el famós físic i matemàtic Isaac Newton va comparar els científics a "nans enfilats a les espatlles de gegants". Era la seva manera de dir que qualsevol avenç que faci un científic es recolza en el treball de tots aquells que l'han precedit. Aquesta naturalesa acumulativa de la ciència es reflecteix perfectament en el sistema de citacions de les publicacions científiques.

La major part dels articles científics tenen un apartat de bibliografia al final on citen una sèrie d'articles anteriors. El fet que un article aparegui a la bibliografia significa normalment que el contingut és rellevant per a l'article que el cita, o bé perquè exposa resultats anteriors en els quals es fonamenta, perquè n'ha extret informació d'alguna mena, perquè ha influït en el treball, o fins i tot perquè hi és criticat.

A partir de les citacions és possible construir una xarxa, en la qual els nodes són els articles i els enllaços, dirigits, van d'un article que cita a l'article citat. A la Figura 3.11 es pot veure un exemple de *xarxa de citacions* en l'àmbit de les ciències de la salut.

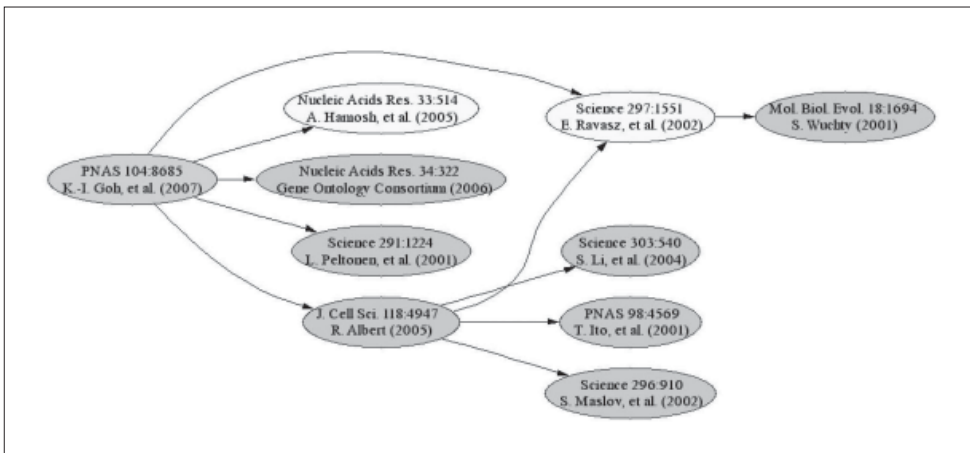


Figura 3.11: Xarxa de citacions de l'àmbit de les ciències de la salut.
Font: Goh i altres (2007).

El primer a estudiar les xarxes de citacions va ser el físic, historiador de la ciència i científic de la informació Derek John de Solla Price durant la dècada dels anys seixanta del segle passat. Aquest àmbit, el de la ciència de la informació (i, més específicament, la bibliometria o la cienciometria), és probablement, doncs, el primer en el qual es va utilitzar l'anàlisi de l'estructura de les xarxes (Price 1965). Els primers estudis de Price van utilitzar les bases de dades de citacions que van constituir l'embrió del que és avui el *Science Citation Index* (SCI). Evidentment, els índexs de citacions com SCI (i els seus complements SSCI, dedicat a les ciències socials, i HCI, dedicat a les humanitats), Scopus, CiteSeer i fins i tot Google Scholar no es desenvolupen en principi per a conèixer l'estructura de la xarxa de citacions. L'objectiu és ajudar els investigadors a saber qui ha citat un article determinat per poder trobar literatura rellevant sobre

el tema, i darrerament també per donar un element d'avaluació de l'impacte de la producció científica d'investigadors i institucions. El supòsit és que el nombre de citacions que rep un article és una mesura de la rellevància que té. Tot i així, les dades que contenen són molt útils per a reconstruir la xarxa de citacions i esbrinar-ne les propietats.

Enllaç:**Thomson Reuters – Science Citation Index:**

http://thomsonreuters.com/products_services/science/science_products/a-z/science_citation_index/

Tal com es pot veure a la Figura 3.11, les xarxes de citacions són, igual que les xarxes del Web, dirigides. Una citació representa un enllaç que surt de l'article citant i va a parar a l'article citat. Al contrari que en el cas del Web, però, les xarxes de citacions són acícliques. Això vol dir que no hi ha camins tancats d'enllaços dirigits que surtin d'un node i hi tornin. La raó per la qual a les xarxes de citacions això no pot passar és que els articles s'elaboren en un moment donat en el temps. I en aquell moment, evidentment, les citacions que es facin només poden ser, en principi, d'articles ja publicats. Per tant, tots els enllaços van cap enrere en el temps i no poden tornar mai endavant i tancar un camí.

Una altra característica rellevant de les xarxes de citacions és que la distribució de grau que tenen sol seguir una llei de potència. És a dir, que són xarxes lliures d'escala. Efectivament, hi ha pocs articles que tinguin moltes citacions i molts que en tenen molt poques o cap. Aquesta situació acaba produint el fenomen dels "científics estrella", que acaparen gran part del prestigi en la seva disciplina, ja sia per la qualitat del seu treball o bé per altres circumstàncies.

Una altra possibilitat de l'estudi de les xarxes de citacions és fer un canvi del nivell d'agregació. Si, per exemple, en comptes de fer l'anàlisi per articles els agrupem per disciplines o subdisciplines, construïm una nova xarxa on aquestes siguin els nodes i ens fixem en els fluxos de citacions entre elles, podem detectar-ne fàcilment el grau d'interdisciplinarietat, que vindrà donat pel grau de citacions que hi hagi a disciplines diferents.

Les xarxes de citacions que hem vist són les més simples que podem construir a partir de la informació que ens donen les citacions dels articles, però n'hi ha d'altres com les *xarxes de cocitació*. Es considera que dos articles són cocitats si se citen tots dos en un tercer article, evidentment posterior. Les xar-

xes de cocitació, doncs, es construeixen connectant entre ells els articles que han estat cocitats. Els enllaços en aquest cas són no dirigits, perquè la relació de cocitació és simètrica. La idea en construir aquest tipus de xarxes és poder detectar els articles que s'ocupen de temes relacionats, ja que se suposa que seran més cocitats i que apareixeran propers a la xarxa. A les xarxes de cocitació és possible definir la força de la relació entre dos articles en funció del nombre d'articles posteriors que els cocitin.

Una altra variant de les xarxes basades en les citacions són les *xarxes d'encavalcament bibliogràfic*. En aquest cas, es considera que dos articles estan relacionats quan citen els mateixos articles. Aquí també hi ha la possibilitat de definir la força de la relació en funció del nombre d'articles que citen tots dos. Els parells d'articles que presenten encavalcament bibliogràfic fort comparteixen una gran part dels seus fonaments, i per tant són dins del mateix tema de recerca. Per això aquestes xarxes s'utilitzen de vegades per a detectar l'evolució de les disciplines científiques i la possible aparició de nous camps.

Ja fora de l'àmbit estricte de la ciència, però encara en el de l'aplicació pràctica, trobem les *xarxes de citacions de patents*. L'estructura d'aquestes xarxes reflecteix l'organització i l'evolució de la tecnologia de manera semblant a com les de citacions ens permeten conèixer l'organització i l'evolució de la ciència. Per exemple, poden ajudar a detectar les dependències entre diverses tecnologies aparentment no relacionades.

3.4.2. La col·laboració científica: xarxes de coautoria

Tot i que les xarxes que hem vist a l'apartat anterior deriven d'activitats amb un component social important com la ciència i la tecnologia, normalment es cataloguen com a xarxes d'informació i no socials. Ara veurem un altre tipus de xarxes rellevants a l'àmbit de la ciència que sí que se solen considerar plenament xarxes socials: les *xarxes de coautoria*.

En una xarxa de coautoria els nodes són els autors d'articles científics i les connexions s'estableixen entre els autors que han publicat algun treball conjunt. Aquest tipus de xarxes són molt útils per a detectar les dinàmiques de col·laboració acadèmica entre els científics individuals i també per a detectar les comunitats o clústers d'investigadors que s'estableixen. Si a més es recull informació sobre la filiació de cada autor, també es pot estudiar la col·laboració entre els diversos grups de recerca i també entre les institucions dels quals els investigadors formen part.

A la figura 3.12 es pot veure un mapa de col·laboració en el camp de la ciència de les xarxes basat en la xarxa de coautories. S'hi poden veure alguns dels clústers d'investigadors més importants en aquest àmbit, alguns dels quals catalans.

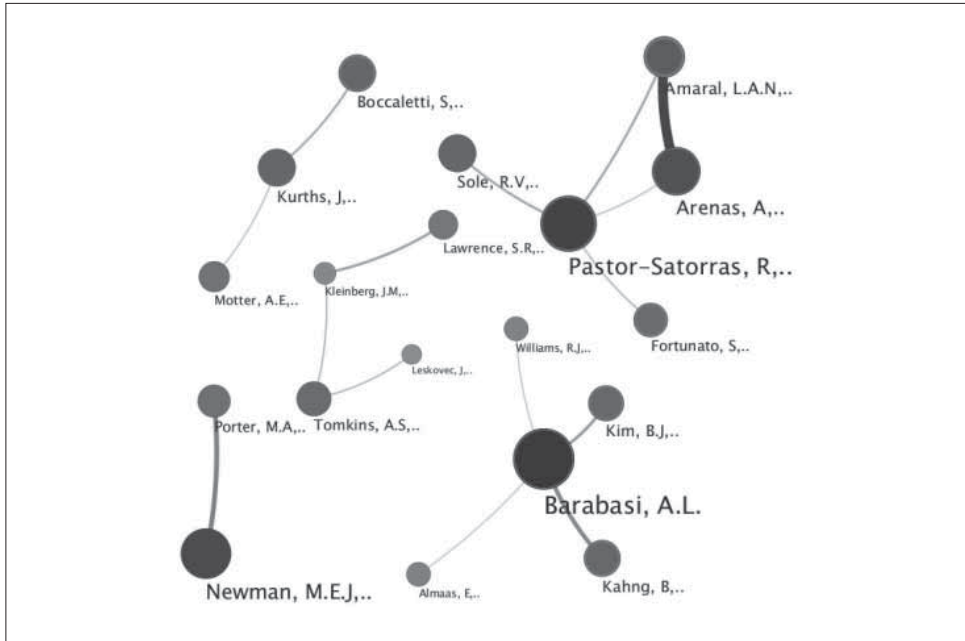


Figura 3.12: Mapa de col·laboració en el camp de la ciència de les xarxes.
 Font: Massimo Franceschet (Universitat dels Estudis d'Udine), <http://users.dimi.uniud.it/~massimo.franceschet/networks/nexus/bibliometrics.html>

3.4.3. L'estructura de la ciència: xarxes de paraules clau

Fins ara hem vist la utilitat per a l'estudi de l'activitat científica de les xarxes basades en informació continguda en els documents científics com les citacions i els noms i les filiacions dels autors. En aquest apartat veurem que també podem construir xarxes d'interès a partir del propi contingut dels articles: les *xarxes de paraules clau*.

Aquestes xarxes es construeixen a partir dels índexs de paraules clau, que són similars als índexs que elaboren els agents rastrejadors del WWW amb les paraules que troben a les pàgines web i que hem mencionat a la segona secció d'aquest capítol. Els índex de paraules clau d'un conjunt de documents (en el cas que ens ocupa, articles científics, però la tecnologia es pot aplicar

perfectament a altres àmbits) consisteixen simplement en una llista de termes rellevants amb les aparicions d'aquests en els documents i són similars als índexs de contingut que podem trobar al final de molts llibres i que relacionen les paraules clau amb les pàgines del llibre on es troben. Així, localitzant una paraula clau a l'índex podem saber en quins documents apareix.

La selecció de les paraules clau que han d'aparèixer a l'índex es pot fer de maneres diverses. Un possibilitat és aprofitar les paraules clau que en moltes revistes científiques l'editor demana als autors que incloguin al principi de l'article. Aquestes paraules clau indiquen força bé el tema principal de l'article, però de vegades no poden reflectir tot el contingut. Per això es pot recórrer a d'altres tècniques per a elaborar llistes de paraules clau més exhaustives. D'una manera semblant a com actuen els programes agents dels cercadors del Web, podem utilitzar programes informàtics que vagin recorrent el text de tots els articles i vagin emmagatzemant les paraules que trobin i en quin document s'ha trobat. Segons el nivell d'exhaustivitat que vulguem aconseguir, podem aplicar aquest procés només als resums dels articles, o bé al text complet. El problema d'aquest procediment és que el nombre de paraules que apareixen finalment a l'índex és molt elevat i algunes paraules tenen poc de clau. Això ja està bé si l'índex està pensat per a la recuperació d'informació, però no és gaire útil per a la construcció de xarxes que permetin analitzar l'entramat de la ciència. Per això s'utilitzen tècniques específiques que permeten filtrar les paraules que no aporten informació i obtenir un índex de paraules realment rellevants.

A partir de l'índex es pot construir la xarxa de paraules clau, que és una xarxa bipartida on els vèrtexs d'un tipus són articles i els de l'altre tipus paraules clau. Els enllaços són entre cada paraula clau de l'índex i els articles en els quals apareix. Cal no confondre aquestes xarxes amb les que es construeixen a partir d'hiperenllaços entre pàgines web que hem presentat anteriorment.

Una de les aplicacions de les xarxes de paraules clau és que permeten localitzar els articles que tracten sobre el mateix tema. A partir d'un article concret i coneixent la xarxa que relaciona paraules clau i articles, podem trobar tots els altres que comparteixin un elevat nombre de paraules clau i que, per tant, tindran una temàtica similar. Això té aplicacions interessants en els sistemes de cerca, ja que un cop s'han recuperat els documents que s'ajusten a les paraules de la petició de cerca, s'hi poden afegir aquells que no s'hi ajusten però que comparteixen moltes paraules clau amb els primers, i que per això poden resultar igualment rellevants per a l'usuari. Aquesta aplicació és, en certa manera, similar a la dels sistemes de filtratge col·laboratiu que hem presentat a la segona secció del capítol.

Una altra aplicació és veure quina relació hi ha entre els diversos temes de recerca dins d'una disciplina científica. A partir de la xarxa de paraules clau podem determinar quins termes surten conjuntament en una gran part dels articles, de manera similar al que fèiem a les xarxes de cocitació. Això ens permet construir noves xarxes on els nodes són les paraules clau i les connexions entre elles existeixen quan es troben simultàniament en una quantitat important d'articles, les *xarxes de coocurrència*. A la Figura 3.13 podem veure'n un exemple elaborat pel grup de recerca KIMO de la UOC en el marc d'un projecte de recerca actual sobre mapes de coneixement. La xarxa s'ha construït a partir de la xarxa de paraules clau extreta dels articles més rellevants de l'àmbit de la ciència de la informació entre els anys 2006 i 2010. A la figura es representen només les paraules clau que apareixien amb connexions més fortes.

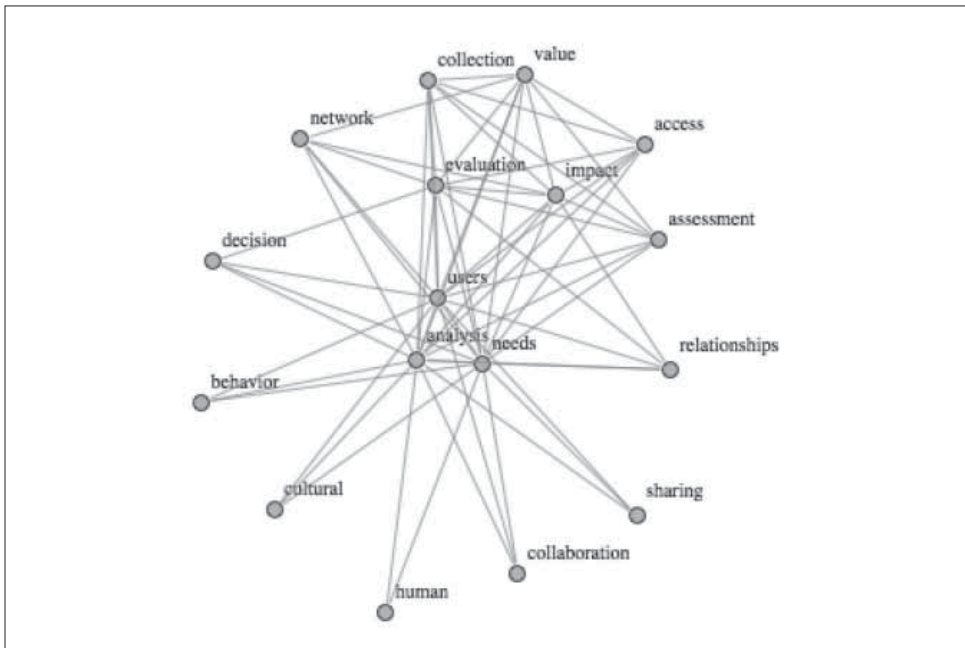


Figura 3.13: Extracte de la xarxa de coocurrència de paraules clau als articles de les principals revistes de l'àmbit de la ciència de la informació entre els anys 2006 i 2010.
Font: Grup de recerca KIMO (UOC)

3.4.4. Mapes de coneixement: visualització de les xarxes

En aquesta secció hem vist que molts aspectes de la generació i la difusió de coneixement científic tenen relació amb xarxes d'informació i xarxes socials i que, per tant, els podem estudiar aplicant les tècniques de l'anàlisi de xarxes. Això, combinat amb la disponibilitat de programes informàtics sofisticats i potència de càlcul a preu assequible, ha donat lloc a un gran desenvolupament en l'àmbit de la cienciometria i, per tant, cada vegada es disposa de més dades interessants sobre l'organització i les dinàmiques de l'activitat científica.

En els darrers anys, la combinació de les tècniques esmentades anteriorment i algunes altres amb el desenvolupament també impressionant de l'àrea de la visualització d'informació ha donat la possibilitat de fer representacions gràfiques molt útils de l'activitat científica impensables no fa gaire temps. És el que es coneix com a *mapeig del coneixement científic* o *visualització d'àmbits de coneixement* (Börner, Chen et al. 2003).

A les figures adjuntes es poden veure un parell d'exemples del potencial que ofereix aquest nou àmbit de coneixement. La Figura 3.14 mostra la xarxa de cocitació construïda a partir d'una base de dades d'articles relacionats amb els àmbits de la bibliometria, la cienciometria i la visualització d'informació. A més de la xarxa, es mostra també el nombre de citacions de cada autor mitjançant una barra vertical. La barra està ombrejada de manera que els tons més foscos representen citacions més antigues i els més clars citacions més recents.

A la figura 3.15 es pot veure un altre tipus de representació feta a partir de la xarxa de paraules clau de la mateixa base de dades d'articles. Es tracta d'utilitzar la tècnica dels mapes autoorganitzats (un tipus de xarxes neuronals) per a generar una representació de les disciplines i subdisciplines científiques i les seves relacions mitjançant una metàfora que compara la ciència amb un paisatge geogràfic.

* Trobareu una bona revisió de l'estat de l'art d'aquest tema a l'article de Börner i altres (2003), aparegut a l'*Annual Review of Information Science and Technology*. També publicat per la mateixa Katy Börner, professora de la Indiana University, trobareu una excel·lent mostra de mapes de coneixement al llibre *Atlas of Science* (Börner 2010).

K. Börner (2010). *Atlas of Science: Visualizing What We Know*. Cambridge, MA: The MIT Press.

K. Börner; C. Chen i altres (2003). "Visualizing Knowledge Domains." *Annual Review of Information, Science and Technology* (vol. 37, pàg. 179-255).

Un recull en línia de mapes de la ciència: <http://scimaps.org/>

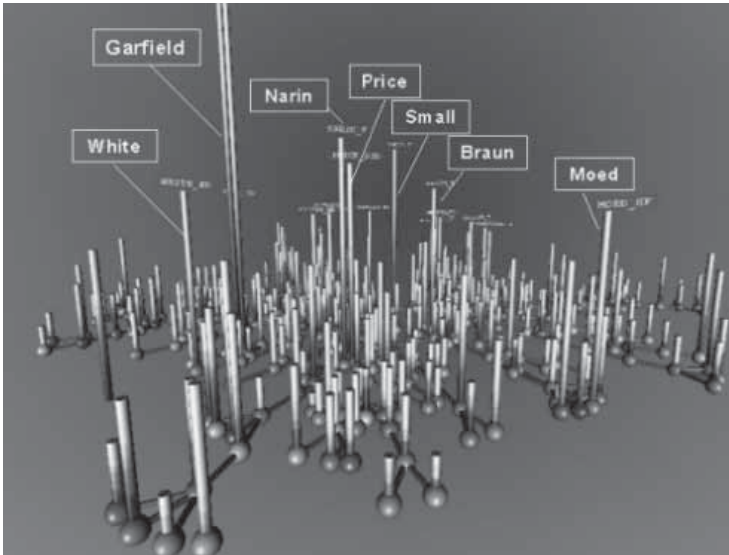


Figura 3.15. Mapa autoorganitzat (self-organizing map) basat en l'anàlisi de la xarxa de paraules clau que utilitza la metàfora del paisatge per a representar les disciplines científiques. Font: Börner, Chen i altres (2003)

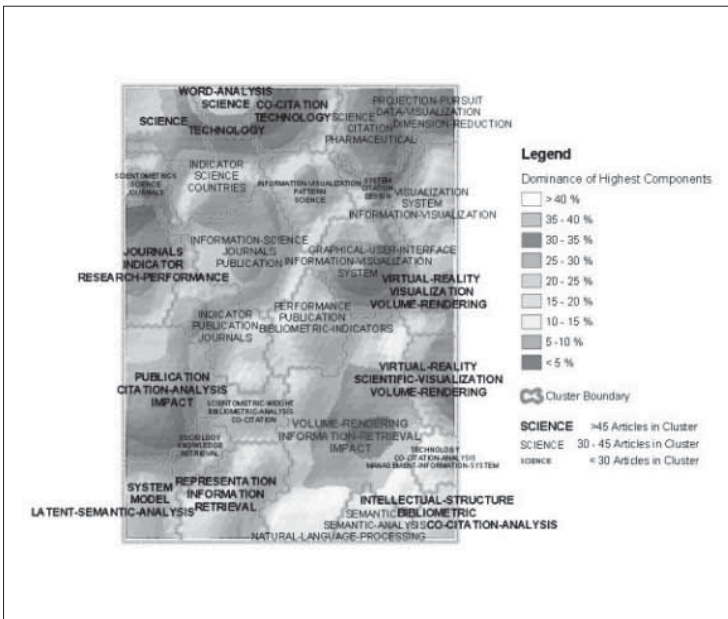


Figura 3.14: Mapa d'una xarxa de cocitació combinada amb una indicació del nombre de citacions de l'autor mitjançant una barra vertical. Font: Börner, Chen i altres (2003)

3.5. Les xarxes a l'economia del coneixement

Si alguna cosa caracteritza l'economia actual és el paper fonamental que hi tenen la informació i el coneixement. Donat l'important paper que hem vist que les xarxes tenen en diversos aspectes relacionats tant amb una com amb l'altre, no és estrany que valgui la pena dedicar un espai a revisar diverses idees sobre les xarxes a l'economia del coneixement.

De fet, les xarxes socials sempre han estat importants a l'esfera econòmica. Les transaccions que s'hi fan, tant al mercat com dins d'organitzacions més o menys jeràrquiques, sempre requereixen que en paral·lel hi hagi relacions socials. Així, l'activitat econòmica sempre s'ha beneficiat de les xarxes socials i alhora ha ajudat a constituir-les.

Ara bé, en les darreres dècades la confluència de tres tendències ha fet que, a més d'incrementar el paper de la informació i el coneixement, el rol de les xarxes socials i també d'informació encara fos més rellevant. Primer, les tecnologies de la informació i la comunicació (TIC) han facilitat fins a extrems impensables fa uns anys la nostra capacitat de processar i transferir informació. Segon, el procés de globalització ha contribuït decisivament a fer caure barres de tot tipus, de manera que el món està molt més connectat. I tercer, cada vegada més la informació i el coneixement són els recursos més decisius per a generar avantatges competitius a les empreses. Tot això fa que sigui difícil entendre alguns aspectes de l'economia actual sense tenir algunes idees sobre quin paper hi tenen les xarxes.

En aquesta secció en veurem unes quantes, tant des del punt de vista de les organitzacions com des del dels mercats on interaccionen. Començarem, però, aprofundint una mica sobre el paper estratègic del coneixement.

3.5.1. El paper estratègic del coneixement

Durant els anys vuitanta l'èmfasi en l'àmbit de l'estratègia empresarial es va posar a mirar cap a l'exterior. El cèlebre model de les "cinc forces" de Michael Porter (1980), per exemple, centrava la formulació de l'estratègia competitiva en l'estudi de cinc elements externs del sector: els proveïdors, els clients, els competidors, els possibles nous entrants i els possibles substituïts. El coneixement sobre aquests elements requeria sobretot una gran quantitat d'informació sobre l'entorn. Aquest interès estratègic i el desenvolupament de les tècniques de vigilància tecnològica i intel·ligència competitiva va contri-

buir a fomentar l'aprofitament de les xarxes que permetien que la informació arribés a l'empresa.*

Més endavant, en la dècada dels noranta, el pensament estratègic va incorporar amb força la idea de l'empresa com un conjunt de recursos que es combinaven per aconseguir els seus objectius (Barney 1991). Això fa que també es miri cap a dins de l'organització, on resideixen els recursos que han de ser la font dels avantatges competitius. De mica en mica es va prenent consciència que entre tots els recursos el coneixement és probablement el més important (Kogut and Zander 1992; Grant 1996) i de la importància de les xarxes per a generar i incorporar el coneixement necessari per a l'empresa. La popularització de les idees de la gestió del coneixement va contribuir a aquesta tendència.

Així, doncs, des del punt de vista teòric macro, queda establerta la importància de la informació i el coneixement i de les xarxes que permeten obtenir-los. Als apartats següents veurem com aquesta importància es trasllada a aspectes més micro.

3.5.2. Les organitzacions i les xarxes

Les xarxes són per tot arreu, i les organitzacions, tant privades com públiques, no en són l'excepció. Els membres de les organitzacions formen xarxes socials entre ells i amb persones de fora de l'organització i les organitzacions com a tals formen xarxes socials amb altres organitzacions. Ens podem preguntar quines són les estructures de xarxa més adequades per a les organitzacions i per als individus que les formen.

Donada la importància del coneixement a les organitzacions, podem mirar de trobar alguna resposta a aquesta pregunta per la via de la incidència de les xarxes socials en la generació i la transmissió d'informació i coneixement.

Suposem que tenim dos tipus ideals ben diferents d'estructura de la xarxa social d'una organització representats pels esquemes de la figura 3.16. A la primera, que anomenarem estel per la forma que té, hi ha un node central pel qual ha de passar per força tota la informació que circula per la xarxa. La segona, en canvi, que té forma de diamant, està altament cohesionada: tothom està connectat amb tothom. Evidentment, són casos extrems que poques vegades es donen al món real, però ens seran útils per a discutir uns quants conceptes.

* Per una introducció al camp de la intel·ligència competitiva, vegeu el llibre d'Ortoll i Garcia-Alsina (2008): E. Ortoll; M. Garcia-Alsina (2008). *La intel·ligència competitiva*. Barcelona: UOC.

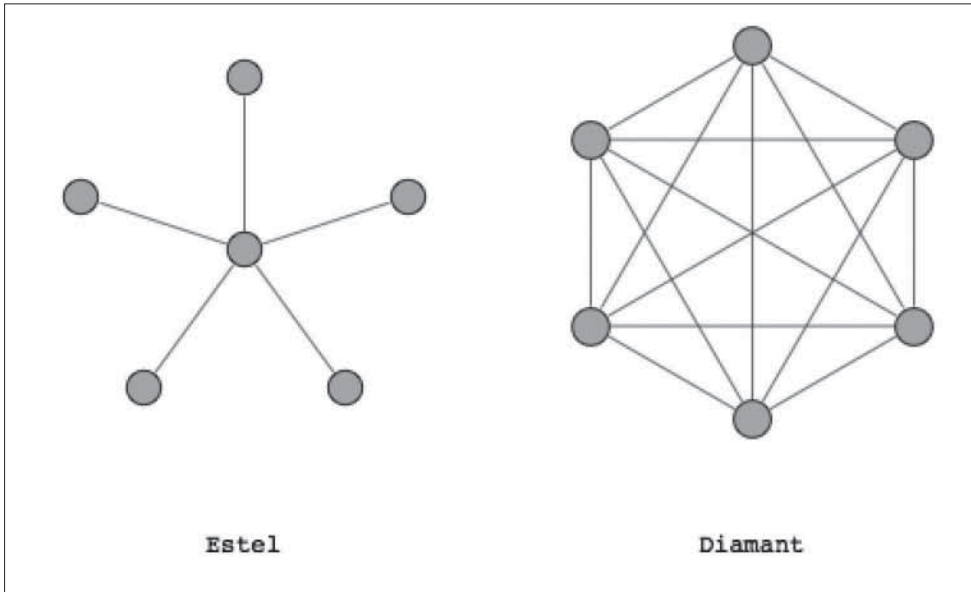


Figura 3.16: Dos models ideals d'estructura de la xarxa social a les organitzacions

Serà millor per a l'organització una estructura de la xarxa social dels seus membres tipus estel o tipus diamant? I per als seus membres individualment? Recordant el que hem vist sobre la difusió d'informació en xarxes, sembla clar que aquesta es difondrà molt més fàcilment a l'estructura tipus diamant. D'entrada, cada node té moltes més vies d'entrada i sortida d'informació i tothom té connexió directa amb tothom, la qual cosa n'assegura una difusió ràpida. Però a més, a l'estructura tipus estel tota la informació ha de passar per un únic node. Segons la capacitat del node central per a processar-la, pot convertir-se en un coll d'ampolla i dificultar molt la fluïdesa del procés. Així, doncs, si a l'empresa o institució li interessa fomentar la difusió d'informació, sempre és millor una estructura tipus diamant, però si necessita controlar els fluxos d'informació, millor que tingui una estructura de xarxa social tipus estel, sempre que l'individu central sigui l'adequat. Cal recordar que l'estructura de la xarxa social no té per què coincidir (i sovint és així) amb l'estructura formal reflectida a l'organigrama.

Des del punt de vista dels individus, la cosa pot ser diferent. Mentre que a l'estructura tipus diamant tothom està en les mateixes condicions, a l'estructura tipus estel l'individu que ocupa la posició central sembla clar que té una situació privilegiada, la qual cosa es pot confirmar calculant qualsevol dels índexs

de centralitat per als diferents nodes. Tota la informació que es mou a la xarxa passa per ell, i ell la coneix abans que ningú més (excepte l'emissor). Si això no li provoca problemes de sobrecàrrega, té un avantatge clar, ja que pot utilitzar la informació, i fins i tot té el poder de decidir quina informació transmet i quina no segons els seus interessos. Novament, si els seus interessos coincideixen amb els de l'organització, no és un problema, però això no sempre és així.

Situacions com la que es produeix a l'estructura tipus estel amb el node central reben el nom de forats estructurals (*structural holes*) segons la definició del sociòleg de la Universitat de Chicago Ronald Burt (1992). La idea és que un element de la xarxa que connecta amb dos nodes o grups de nodes que no estan connectats per cap altra via aprofita (o "omple") un forat estructural, cosa que li confereix el control sobre la comunicació entre les dues parts.

Si anem una mica més enllà de la difusió d'informació i ens fixem en els processos de creació i transmissió del coneixement, les nostres conclusions sobre els dos models alternatius de xarxa social no seran gaire diferents, perquè com hem vist és necessari que es transmeti informació per a transferir coneixement. Ara bé, la generació de coneixement a les organitzacions és un procés complex que sovint requereix l'acció conjunta de grups d'individus que puguin compartir fàcilment informació i també transmetre's entre ells coneixement tàcit. Per això, en organitzacions que basen la seva activitat en la generació de nou coneixement com les empreses innovadores, les estructures del tipus diamant poden ser, no tan sols millors, sinó imprescindibles. Fins i tot des del punt de vista individual de l'individu central d'una estructura inicialment del tipus estel, pot ser beneficiós a la llarga renunciar a la posició privilegiada que té per anar cap a una estructura més tipus diamant que afavoreixi la creació de coneixement, de la qual es pot beneficiar tothom.

A la literatura sobre aprenentatge organitzacional es fa sovint la distinció entre l'exploració de noves possibilitats i l'explotació del que ja es coneix (March 1991). Totes dues són necessàries però és molt difícil aconseguir que les estructures d'una organització s'adaptin bé a totes dues. Sovint el que va bé per a l'exploració no va bé per a l'explotació. Quan es tracta d'explotar el conegut, una estructura social tipus estel pot tenir tot el sentit, perquè el que interessa és optimitzar l'eficiència dels processos establerts. I això es pot fer bé en una estructura jeràrquica i burocratitzada. En canvi, per a l'exploració, que requereix la generació de nou coneixement, les estructures tipus diamant semblen les més adequades. Les organitzacions que millor balancegen les dues activitats ho fan sobre la base de la convivència de dos tipus d'estructures. Cada part de l'organització adopta la que va millor a la seva activitat d'explotació

o d'exploració. És probable, doncs, que a la xarxa social de moltes organitzacions hi hagi parts que tendeixen més a estructura estel i d'altres a estructura diamant.

Un exemple molt estudiat de grups que necessiten una alta capacitat d'aprenentatge i de generar coneixement són les anomenades comunitats de pràctica (Wenger 1998). A les comunitats de pràctica que funcionen se sol establir una xarxa social molt densa (tipus diamant) amb lligams forts en la qual la transmissió de coneixement, especialment coneixement tàcit, és més fàcil.

Tot i els avantatges que hem vist per a la generació de coneixement, però, una estructura cohesiva també pot tenir algun inconvenient. Quan un grup està molt cohesionat, la densitat de comunicació amb els membres del grup fa que es pugui reduir molt la connexió amb l'exterior. Això pot dificultar que entrin idees noves, que són essencials per a la creació de nou coneixement. Per això són importants els enllaços febles que s'han presentat al capítol 2.

Diversos estudis sobre els fluxos de coneixement en organitzacions innovadores, com és ara les que se situen a clústers geogràfics especialitzats en noves tecnologies (Saxenian 1994), confirmen que aquest necessari ingredient de coneixement nou provinent de l'exterior sol venir d'enllaços febles a les xarxes socials personals dels treballadors. Tanmateix, aquest tipus d'aportacions també el pot fomentar l'organització de manera més o menys formalitzada. Aquesta és la intenció que hi ha al darrere de l'establiment d'algunes aliances empresarials i de la presència d'alts càrrecs d'algunes empreses en els consells d'administració d'altres i que dona lloc al que es coneix com a *conselleries creuades*. Tots dos fenòmens es poden estudiar mitjançant l'anàlisi de les xarxes socials que produeixen. En el primer cas són xarxes on els nodes són organitzacions que estan connectades si tenen algun tipus d'aliança, i en l'altre xarxes bipartides on hi ha uns nodes corresponents a les empreses i els altres als directius. Un directiu estarà connectat a les empreses en les quals participa com a membre del consell d'administració.

En els últims anys les TIC han facilitat que empreses i institucions puguin establir formes de coordinació amb altres organitzacions que ja no segueixen els patrons tradicionals d'acords contractuals entre estructures burocràtiques, sinó que tenen un caràcter molt més de sistema social informal en el qual la confiança, la presa de decisions distribuïda i la negociació constant són elements fonamentals. Això ha donat lloc a sistemes relativament estables d'actors autònoms però interdependents que no són ni organitzacions jeràrquiques ni mercats i s'anomenen *organitzacions xarxa*. Aquest tipus d'organitzacions han demostrat que poden ser més eficients que les tradicionals en l'adquisició de coneixement i en la resolució de problemes gràcies a la seva actuació descen-

tralitzada, i també més eficaces en la resolució de problemes perquè poden abordar-los i trobar-hi solucions de manera col·lectiva.

3.5.3. Mercats en xarxa

Les xarxes no tan sols tenen un paper important dins de les empreses. A l'economia actual, sobretot després de la popularització d'Internet, les xarxes a la banda del mercat són cada vegada més rellevants i s'han de tenir molt en compte en el disseny de les estratègies i en el dia a dia. Tractar això en profunditat aniria molt més enllà dels nostres objectius, però sí que en mencionarem breument alguns aspectes interessants.

All final del segle passat, l'anàlisi d'alguns mercats peculiars va dur els economistes a definir un fenomen que, tot i no ser nou, es va veure considerablement intensificat per la presència d'Internet. Es va anomenar efectes de xarxa (*network effects*) o *externalitats de xarxa* (Shapiro and Varian 1999). Es produeixen efectes de xarxa quan el benefici que obté un agent econòmic que consumeix un bé o servei depèn del nombre d'altres agents que el consumeixen. Un exemple tradicional d'efectes de xarxa és el dels teclats que fem servir actualment majoritàriament als ordinadors, que s'anomenen *tipus QWERTY* per la disposició de les tecles de la part superior esquerra. La configuració actual dels teclats es va dissenyar al segle XIX per a les màquines d'escriure i amb l'objectiu d'evitar que les palanques s'enganxessin i produïssin embussaments interns si els mecanògrafs anaven massa de pressa. La configuració QWERTY exigeix que els dits s'hagin de moure per tot el teclat per a utilitzar les tecles més freqüents i, per tant, alenteix l'escriptura però aconseguia evitar els embussos de tecles. Evidentment, avui dia amb els ordinadors aquest teclat ineficient no té cap sentit perquè no hi ha palanques que es puguin enganxar. S'han proposat alternatives molt més eficients a la disposició de les tecles com el teclat Dvorak, però no han aconseguit desplaçar el QWERTY, precisament per efectes de xarxa. Un cop el QWERTY s'ha establert com a estàndard, els processos de producció dels teclats s'hi han adaptat i, encara més important, els usuaris hem invertit temps i esforç a aprendre a teclejar ràpid. Canviar l'estàndard comportaria costos importants de redisseny de sistemes productius per part dels fabricants i d'aprendre una nova configuració pels usuaris, després de "desaprendre" l'anterior. Mentre la majoria dels fabricants continuïn oferint el teclat QWERTY i la majoria dels usuaris estiguin acostumats a treballar amb la seva disposició de tecles, no hi ha cap al·licient per a canviar malgrat l'increment d'eficiència global que a la llarga s'obtidria amb l'adopció d'un nou estàndard. Això són els efectes de xarxa.

Es poden observar els efectes de xarxa en alguns altres casos com l'adopció la dècada dels vuitanta del format VHS com a estàndard de vídeo enfront del Betamax, que era tècnicament millor. Però és amb el gran impuls de les TIC i especialment amb l'arribada d'Internet quan els efectes de xarxa es veuen potenciats. La dominació gairebé total del mercat dels sistemes operatius per a PC a la dècada dels noranta per part de Microsoft amb el Windows, que era manifestament inferior a algunes alternatives, és deguda bàsicament als efectes de xarxa produïts per la incompatibilitat entre sistemes. Un cop la majoria dels usuaris utilitzava el sistema operatiu de Microsoft (bàsicament, a causa que IBM el va adoptar per als seus PC), era millor per a qualsevol usuari fer el mateix, ja que tant el maquinari com el programari eren més barats a causa de les economies d'escala, i a més podia compartir els seus fitxers amb més gent.

A Internet, a causa de la possibilitat que dona als usuaris de connectar-se entre ells a escala global, els efectes de xarxa encara són més potents, ja que els beneficis de triar l'actor dominant per l'usuari es fan més grans. El lloc web de subhastes en línia eBay gairebé no té competència. Evidentment, qualsevol usuari que vulgui posar a la venda un objecte sap que és a eBay on hi haurà més possibles compradors i amb la competència podrà obtenir un preu més alt. D'altra banda, els compradors saben que a eBay trobaran una gamma d'objectes molt més àmplia que enlloc més. De la mateixa manera, ara mateix és molt difícil que triomfi cap alternativa a la Wikipedia com a enciclopèdia col·laborativa en línia, i el domini dels llocs de xarxes socials més populars, Facebook en l'esfera personal i LinkedIn en l'esfera professional, és cada vegada més aclaparador.

Algunes empreses van més enllà fins i tot, i aprofiten la posició dominant que els efectes de xarxa els donen en un mercat per intentar entrar i a la llarga dominar-ne un altre. Seria el cas, per exemple, de Google, que recolzant-se en la posició preponderant que té en l'àmbit dels cercadors ha anat estenent la seva l'activitat a la publicitat en línia i a tota una oferta de serveis en núvol com Gmail, GoogleDocs, GoogleCalendar, etc. Totes aquestes situacions són pròpies del que s'anomena mercats en xarxa (*networked markets*), una àrea de recerca de gran activitat actualment.

Més enllà de beneficiar-se dels efectes de xarxa, les empreses poden aprofitar les xarxes d'altres maneres. Un exemple n'és el màrqueting viral, que consisteix en la utilització de les xarxes socials per aconseguir objectius de màrqueting com és ara notorietat de marca o vendes de productes per mitjà d'un procés viral d'autoreplicació de manera similar a com es transmeten els virus biològics o informàtics. A les xarxes socials tradicionals és el que se sol conèixer com de boca a orella, però és a Internet on l'efecte del màrqueting viral s'ha demostrat més poderós. L'objecte que es transmet pot ser molt variat, des de missatges SMS fins a imatges, programari, jocs o videoclips.

Finalment, cal mencionar que les empreses també utilitzen les xarxes socials per a obtenir informació rellevant sobre els mercats i així ajustar millor els seus productes i serveis a les necessitats dels consumidors. Hi ha especialistes en anàlisi de l'opinió publicada que són capaços de recollir i interpretar l'estat d'opinió sobre un producte o servei o sobre la imatge d'una empresa mitjançant l'anàlisi de les contribucions personals als llocs web participatius. Avui dia moltes empreses creen les seves pàgines als llocs de xarxes socials com Facebook per poder tenir més fàcilment accessible aquest tipus d'informació. I les xarxes socials no tan sols els serveixen per a monitorar els seus mercats. Cada vegada més, les empreses utilitzen les xarxes socials per a rastrejar el mercat de treball a la recerca de possibles candidats o per a la valoració de les candidatures existents. És una mesura intel·ligent, doncs, tenir cura amb el contingut que es penja a la pàgina personal dels llocs web de xarxes socials.

* El llibre de Shapiro i Varian (1999) *Information Rules* conté una bona introducció a l'economia de la xarxa.

C. Shapiro; H. R. Varian (1999). *Information rules: a strategic guide to the network economy*. Boston, MA: Harvard Business School Press.

* Per a una introducció al màrqueting viral, vegeu el llibre de Sílvia Sivera (2008): S. Sivera (2008). *El màrqueting viral*. Barcelona: EDIUOC.

Bibliografia per capítols

Capítol II

Barnes, John (1954). "Class and committees in a Norwegian Island Parish". *Human Relations* (vol. 7, núm. 1, pàg. 39-58).

Congosto, M. L.; Fernández, M.; Moro, E. (2011). "Twitter y Política: Información, Opinión y ¿Predicción?". *Cuadernos de Comunicación Evoca* (núm. 4). Comunicación política 2.0.

<http://www.evocaimagen.com/cuadernos/cuadernos4.pdf> [consulta: 23-04-2011].

Davis, A.; Gardner, B. B.; Gardner, M. R. (1941). *Deep South. A Social Anthropological Study of Caste and Class*. Chicago: University of Chicago Press.

Kapferer, B. (1969). "Norms and the Manipulation of relationships in a Work Context". A: M. J. Clyde (ed.). *Social Networks in Urban Situations: Analyses of Personal Relationships in Central Towns*. Manchester: Manchester University Press.

Granovetter, M. (1973). "The Strength of Weak Ties". *American Journal of Sociology* (núm. 78, pàg. 1360-1380).

Harary, F.; Norman, R. Z. (1963). *Graph Theory as a Mathematical Model in Social Science*. Ann. Arbor: Univ. of Michigan, Institute for Social Research.

Lin, Nan (2001). *Social Capital: A Theory of Social Structure and Action*. Cambridge: Cambridge University Press.

Muntanya, D.; Lozares, C. (2004, gener-febren). "El poder del ejemplo: un análisis reticular del rodaje de una escena cinematográfica". *Redes. Revista para el Análisis de Redes Sociales* (vol. 10, núm. 5).

Molina, J. L. (2001). *El análisis de redes sociales. Una introducción*. Barcelona: Edicions Bellaterra.

Moreno, J. L. (1972 [1934]). *Fundamentos de sociometría* [*Who shall survive? Foundations of Sociometry*]. Buenos Aires: Paidós.

Nadel, S. F. (1966 [1957]). *Teoría de la estructura social* [*Theory of social Structure*]. Madrid: Guadarrama.

O'Really, T. (2005). "What is Web 2.0, Design patterns and Business Models for the Next Generation of Software". Document Word en línia: <http://www.elisanet.fi/aari-set/Multimedia/Web2.0/What%20Is%20Web%202.doc>.

Scott, J. (1991). *Social Network Analysis. A Handbook*. Londres: Sage Publications.

Wellman, B. (1999). "The Network Community: An Introduction". B. Wellman (ed.). *Networks in the Global Village*. Boulder, Colorado: Westview Press.

Wellman, B. (2002). "Little boxes, Glocalization and Networked Individualism?". [en línia]

http://www.chass.utoronto.ca/%7Ewellman/netlab/PUBLICATIONS/_frames.html

Capítol III

- Barney, J.** (1991). "Firm Resources and Sustained Competitive Advantage." *Journal of Management* 17(1): 99-120.
- Boisot, M. H. i A. Canals** (2004). "Data, information and knowledge: have we got it right?" *Journal of Evolutionary Economics* 14: 43-67.
- Börner, K.** (2010). *Atlas of Science: Visualizing What We Know*. Cambridge, MA., The MIT Press.
- Börner, K., C. Chen, et al.** (2003). "Visualizing Knowledge Domains." *Annual Review of Information, Science and Technology* 37: 179-255.
- Burt, R. S.** (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA, Harvard University Press.
- Christakis, N. i J. Fowler** (2009). *Connected: The Amazing Power of Social Networks and How They Shape Our Lives*. London, Harper Press.
- Goh, K.-I., M. E. Cusik, et al.** (2007). "The human disease network." *Proceedings of the National Academy of Sciences* 104(21): 8685-8690.
- Grant, R. M.** (1996). "Toward a Knowledge-Based Theory of the Firm." *Strategic Management Journal* 17(Winter Special Issue: Knowledge and the Firm): 109-122.
- Kogut, B. i U. Zander** (1992). "Knowledge of the firm, combinative capabilities, and the replication of technology." *Organization Science* 3(3): 383-397.
- March, J. G.** (1991). "Exploration and Exploitation in Organizational Learning." *Organization Science* 2(1): 71-87.
- Ortoll, E. i M. Garcia** (2008). *La intel·ligència competitiva*. Barcelona, UOC.
- Pastor-Satorras, R. i A. Vespignani** (2004). *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge, U.K., Cambridge University Press.
- Porter, M. E.** (1980). *Competitive Strategy: Techniques for Analysing Industries and Competitors*. New York, NY., The Free Press.
- Price, D. J. d. S.** (1965). "Networks of Scientific Papers." *Science* 149(3683): 510-515.
- Rogers, E. M.** (2003). *Diffusion of innovations*. New York, Free Press.
- Saxenian, A.** (1994). *Silicon Valley versus Route 128*.
- Shapiro, C. i H. R. Varian** (1999). *Information rules: a strategic guide to the network economy*. Boston, MA, Harvard Business School Press.
- Sivera, S.** (2008). *El màrqueting viral*. Barcelona, EdiUOC.
- Wenger, E.** (1998). *Communities of practice: learning, meaning and identity*. Cambridge, UK, Cambridge University Press.

