Àlex Arenas[1], Albert Díaz-Guilera[2], Roger Guimerà[3,4], Marta Sales-Pardo[4]

[1] Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili
[2] Departament de Física Fonamental, Universitat de Barcelona
[3] Institució Catalana de Recerca i Estudis Avançats (ICREA)
[4] Departament d'Enginyeria Química, Universitat Rovira i Virgili

RESUM

Les unitats que formen els sistemes complexos interaccionen entre elles en xarxes que no són ni perfectament regulars ni perfectament aleatòries. L'estructura d'aquestes xarxes determina el comportament dels processos dinàmics que hi tenen lloc i, alhora, és un reflex de l'evolució i les funcions del sistema. Des d'aquesta doble perspectiva, l'estudi de les xarxes complexes és important. En aquest article discutim el resultats clàssics de la teoria de xarxes i algunes de les línies de recerca actualment més actives.

Doi: http://dx.doi.org/10.2436/20.2001.01.13

1 Introducció

La complexitat de sistemes com la cèl·lula, el cervell o la societat és conseqüència d'una sèrie de factors. El primer és la no-linealitat de les interaccions; aquest factor has estat àmpliament reconegut i estudiat des de fa més d'un segle. Un altre factor important, però, és la complexitat de l'*estructura* de les interaccions. Pensem, per exemple, en un sistema social: triem una persona qualsevol i connectem aquesta persona amb els seus amics, i els amics amb els seus amics respectius, i així repetidament. Tot i que és clar que la xarxa resultant d'aquestes relacions d'amistat no és ni totalment aleatòria ni totalment ordenada (figura 1), ja fa vint anys que els investigadors van adonar-se de les importants conseqüències d'aquest fet.

En realitat, les xarxes reals combinen elements d'ordre local (els amics d'una persona determinada són, sovint, amics entre ells) amb elements aleatoris (sovint tenim un amic en un lloc llunyà o que no està relacionat amb la resta de les nostres altres amistats). Les xarxes reals també són molt heterogènies: combinen regions amb una gran densitat de connexions amb regions molt poc denses; i nodes amb moltes connexions amb nodes quasi desconnectats, de manera que les distribucions de connectivitats sovint escalen amb exponents ben definits.

Que les xarxes d'interaccions tinguin aquestes propietats no trivials té implicacions a dos nivells. Per un cantó, l'estructura de la xarxa té un impacte directe sobre els processos dinàmics que hi tenen lloc. Des d'aquesta perspectiva, el fet que les xarxes tinguin una estructura no trivial és un factor que no podem obviar, però que més aviat complica l'anàlisi dels sistemes complexos.

D'altra banda, però, hem de pensar que les xarxes complexes tenen l'estructura que tenen per alguna raó: o perquè hi ha uns mecanismes evolutius ben definits o perquè la xarxa ha de complir una certa funció. Sigui com sigui, si som capaços de caracteritzar l'estructura d'una xarxa estarem per força entenent alguna cosa del comportament del sistema corresponent. Des d'aquesta perspectiva, les xarxes complexes ens ofereixen noves oportunitats per entendre millor els sistemes complexos.

A continuació discutim algunes de les àrees de recerca en què els físics estem fent avançar l'estudi de les xarxes, en aquest doble vessant de *complicació* i d'*oportunitat* per a l'estudi dels sistemes complexos.

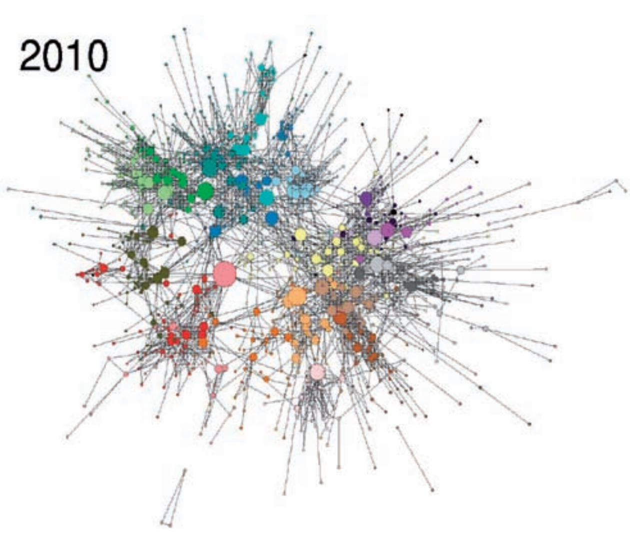

Figura 1: *Xarxa d'intercanvi de comunicacions dins d'una organització amb centenars de persones. Cada node representa una persona de l'organització i cada connexió un intercanvi sostingut i reciprocat de correus electrònics (un mínim d'un missatge intercanviat al mes el 2010). El color de cada node representa el seu departament dins l'organització, i la mida del node representa la seva centralitat a la xarxa.*

2 Dinàmica en xarxes

Com hem dit, que les xarxes d'interaccions siguin complexes té un efecte important sobre els processos dinàmics que hi tenen lloc. De fet, un dels treballs que avui considerem fundacionals en la ciència de les xarxes pretenia estudiar, precisament, l'efecte de diferents topologies en la sincronització dels cants dels grills. Aquests insectes mostren un alt grau de sincronització sobre una distància important, talment com si estiguessin connectats per alguna xarxa «invisible». El que en principi volia ser, doncs, un estudi de com grups de grills es coordinen per emetre sons de manera simultània es va convertir en la llavor de la teoria moderna de les xarxes complexes.

Agafem, precisament, el cas de la sincronització per entendre de quina manera l'estructura de les xarxes ha arribat a ocupar un lloc central en l'estudi de la dinàmica dels sistemes complexos. Durant els anys vuitanta s'estudiaven models matemàtics que descrivien com oscil·ladors caracteritzats per una fase eren capaços d'evolucionar col·lectivament a fases i freqüències properes; d'aquesta manera mostraven un alt grau de sincronització (un dels models més utilitzats en la física és l'anomenat de Kuramoto). En aquell moment, es pressuposava un patró d'interacció en què cada oscil·lador estava connectat amb tota la resta; però aquesta estructura no és realista i aviat es va reconèixer que s'havien de considerar altres estructures. Així doncs, en els anys noranta es van considerar estructures de baixa dimensió (1, 2 i 3 dimensions) però encara regulars, és a dir, amb connexions únicament amb els veïns més propers.

És just en tombar el segle que es comencen a considerar estructures amb topologia complexa com les entenem avui en dia: primer afegint algunes connexions aleatòries en una xarxa altrament regular i després introduint heterogeneïtat en la densitat de les connexions i en les connectivitats dels nodes. Avui sabem que hi ha una relació molt estreta entre la dinàmica de sincronització i l'estructura de la xarxa. Per exemple, canviant l'estructura de la xarxa podem fer que un sistema sincronitzi o deixi de sincronitzar, efecte que sembla que és rellevant per entendre l'aparició d'epilèpsia en algunes persones. I un altre exemple: en xarxes amb grups de nodes molt densament connectats internament, la dinàmica de sincronització ve donada per com de complicat és fer sincronitzar un grup amb l'altre; és a dir, l'estructura de grups de la xarxa determina la dinàmica de sincronització dels nodes. De fet, la connexió sincronització-topologia és tan forta que, en última instància, les condicions de sincronitzabilitat d'un sistema es poden expressar en termes de les propietats espectrals de certs operadors que queden definits únicament per l'estructura de la xarxa.

I més enllà de la sincronització? Que n'esperem, en general, d'un procés dinàmic sobre una xarxa complexa? Esperem un comportament qualitativament diferent del que observaríem en una xarxa regular? O del que observaríem en una xarxa perfectament aleatòria? El cert és que qualsevol procés dinàmic es veu quantitativament i qualitativament afectat per l'estructura de la xarxa. Com en el cas de la sincronització, propietats com l'heterogeneïtat en la distribució de connectivitats, l'existència de grups i d'ordre local, o l'existència de correlacions entre les connectivitats de nodes que interaccionen, afecten de manera determinant la dinàmica del sistema.

No és el nostre objectiu discutir aquí exhaustivament tota la llista de fenòmens dinàmics que es veuen afectats per l'estructura de la xarxa, però sí que val la pena discutir-ne un de dramàtic que ha estat estudiat a bastament: el de la propagació d'epidèmies. Mentre que en xarxes regulars i aleatòries hi ha un valor llindar de la infecció per sota del qual una epidèmia queda totalment eradicada (en un procés que és matemàticament una transició de fase contínua), en xarxes més realistes, amb distribucions de connectivitats heterogènies (en què el segon moment de la distribució de grau divergeix), aquest llindar desapareix i mai és possible eradicar completament l'epidèmia. Aquesta és la mena de diferències qualitatives que fan que no puguem ignorar l'estructura de les xarxes quan estudiem el comportament dinàmic de sistemes complexos.

3 Xarxes multicapa

Com hem explicat, un sistema complex sovint es pot representar com una xarxa on els nodes són les unitats del sistema i les connexions les interaccions entre les unitats. Com hem argumentat en la secció anterior (i també ens hi referirem a la propera secció), aquest enfocament ens ajuda a entendre molts aspectes dels sistemes complexos. Ara bé, en molts casos és essencial anar més enllà d'aquesta mena de xarxes i investigar marcs més sofisticats i alhora més realistes. En particular, cada cop és més evident que, en determinats casos, cal tenir en compte que hi ha xarxes amb diversos tipus de connexió, les anomenades *xarxes de xarxes* o *xarxes multicapa*.

Les xarxes multicapa van ser introduïdes fa dècades en disciplines com ara la sociologia i l'enginyeria. En l'estudi dels sistemes socials, en particular, fa dècades que es va fer evident que calia construir representacions que tinguessin en compte diferents tipus de relacions entre el mateix grup de persones (per exemple, dins d'una organització, les persones amb qui som més amics no tenen per què coincidir amb les persones a qui demanem més sovint consell professional). En aquesta representació, cada tipus de relació defineix una capa de la xarxa multicapa global; i és la xarxa multicapa global, i no cap de les capes per separat, la que determina el comportament del sistema (figura 2).

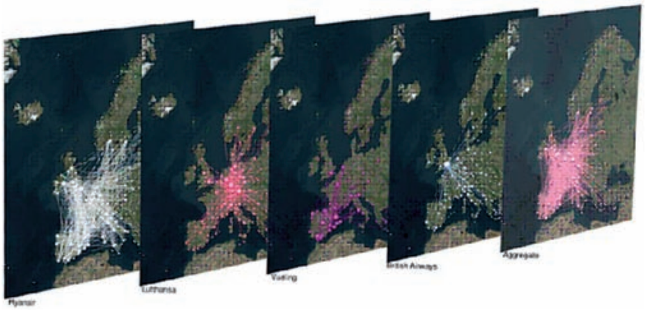

Figura 2: *Xarxa multicapa de transport aeri a Europa. Els nodes de la xarxa representen aeroports europeus i les connexions representen vols directes. Cada capa representa els vols d'una companyia diferent. Imatge gentilesa de Manlio de Domenico, obtinguda fent servir el programa MuxViz (M. de Domenico et al., J. Complex Netw. 3 (2), 159.176 (2015); http://muxviz.net) a partir de les dades a A. Cardillo et al., Sci. Rep. 3, 1344 (2013).*

Si l'ús de xarxes multicapa té una certa història, l'intent de desenvolupar el marc per estudiar-les de manera sistemàtica, estenent les eines existents per xarxes «unicapa», és un fenomen recent. Avui sabem que els mètodes provinents de l'àlgebra lineal ens permeten representar aquests tipus de xarxes multicapa de manera rigorosa. Les xarxes unicapa se solen representar mitjançant l'anomenada *matriu d'adjacències*, on les files i columnes representen els nodes, i els elements de matriu són zero a no ser que la fila i columna corresponents estiguin connectats a la xarxa. D'una manera semblant, les xarxes multicapa es poden representar com a tensors d'adjacència d'ordre superior a dos. I un cop hem representat les xarxes d'aquesta manera, podem aplicar immediatament la maquinària que s'ha desenvolupat per descomposicions tensorials.

Els mètodes més estesos que utilitzen aquest enfocament són generalitzacions de la descomposició en valors singulars (en anglès, *singular value decomposition*): aquestes i altres eines han tingut un gran èxit en moltes aplicacions. Per exemple, els mètodes de descomposició tensorial poden ser utilitzats per extreure comunitats (és a dir, conjunts de nodes que estan connectats densament entre si) o per establir una ordenació dels seus nodes.

Però realment hi ha nova física en les xarxes multicapa? Observem comportaments qualitativament diferents als de les xarxes complexes unicapa? La resposta —cada cop és més clar— és que sí. En trobem un exemple en la teoria de la percolació, que estudia com una xarxa es trenca en trossos a mesura que eliminem algunes de les connexions. Aquesta teoria és important per entendre la robustesa de sistemes complexos i per a processos dinàmics com ara la propagació d'epidèmies. Doncs bé, l'estudi de la percolació en xarxes multicapa ha demostrat que les correlacions negatives entre les connectivitats de les parelles de nodes connectats fan augmentar el llindar de percolació (és a dir, la fracció de connexions que cal eliminar abans no es trenqui la xarxa), mentre que les correlacions positives el fan disminuir. També s'ha demostrat que les xarxes multicapa amb correlacions negatives de connectivitats intracapa són més resistents contra atacs basats que eliminen els nodes més connectats.

I un altre exemple: el de passejos aleatoris en xarxes multicapa amb heterogeneïtat a l'acoblament entre els nodes de la xarxa. En aquesta mena de processos, s'ha demostrat que el temps que necessita un caminador aleatori per visitar la majoria dels nodes depèn de la topologia de les connexions intracapa, de la connexió entre capes, i el tipus del passeig aleatori. De fet, els temps per recórrer la xarxa poden ser més elevats que els necessaris per cobrir cadascuna de les capes per separat (és a dir, el passeig pot ser «infradifusiu») o poden ser més petits que en cadascuna de les capes (és a dir, el camí pot ser «superdifusiu»). Com veiem, doncs, les xarxes multicapa tenen una fenomenologia molt rica que tot just estem començant a explorar.

4 Inferència en xarxes

Si bé, com hem explicat, l'estudi i la caracterització de xarxes complexes han constituït una àrea de recerca activa d'ençà dels últims anys del segle XX, no ha estat fins fa poc que els investigadors ens hem començat a plantejar les xarxes no com a eines *descriptives*, sinó com a eines *predictives*. Es tracta, en definitiva, d'analitzar la xarxa d'interaccions d'un sistema i fer-la servir per inferir, per exemple, com evolucionarà el sistema en el futur.

Els primers treballs d'inferència en xarxes es van fer per resoldre el problema d'identificar errors en una xarxa; és a dir, identificar tant connexions espúries que apareixen en una observació donada, però que no existeixen en realitat, com connexions que existeixen a la realitat però que no apareixen a l'observació. Aquest és un problema ubic en les xarxes biològiques, en què dades experimentals lligades a tecnologies d'alt rendiment tenen taxes d'error sovint per damunt del 20 %.

A primera vista, el problema pot semblar irresoluble: com podem, a partir d'una sola observació d'una xarxa, inferir què és el que està malament en aquesta observació? Però la teoria de la probabilitat ens permet respondre aquesta pregunta de manera rigorosa. El nostre objectiu és obtenir probabilitat $P(X = x/A^O)$ que un cert observable X de la xarxa prengui el valor x donada la nostra observació de les connexions entre els nodes de la xarxa, A^O. En particular, podem calcular la probabilitat que una connexió donada existeixi: en aquest cas l'observable és, senzillament, la connexió en qüestió. Suposem que tenim un conjunt de models $\{M\}$ que són potencials candidats per descriure l'estructura de la nostra xarxa. Aleshores podem escriure

$$P\left(X = x \mid A^O\right) = \sum_{\{M\}} P\left(X = x \mid M\right) P\left(M \mid A^O\right) \qquad (1)$$

on la suma és sobre tot l'espai de possibles models. És a dir, la probabilitat de $X = x$ és la suma sobre tots els models possibles de la probabilitat que el model sigui el model correcte, donada la xarxa observada, per la probabilitat de $X = x$ segons el model.

Aquesta equació és molt semblant a la que trobem en física estadística quan volem calcular la probabilitat que un observable prengui un cert valor: el model faria el paper de la «configuració» (per exemple, l'orientació dels espins en un sistema magnètic), i el logaritme de la probabilitat del model (amb un signe negatiu) faria el paper de l'energia. D'aquesta manera, l'equació (1) es pot interpretar com la mitjana de $P(X = x/M)$ sobre models i es pot avaluar numèricament utilitzant estratègies clàssiques de mostreig de configuracions (com l'algoritme de Metropolis) desenvolupats dins l'àmbit de la mecànica estadística. De fet, la similitud formal entre el problema d'inferència en xarxes i la física estadística és fruit de les connexions profundes que hi ha entre la física estadística i la teoria de la probabilitat entesa com a extensió de la lògica.

Així doncs, resolt el problema de detectar errors, podem tornar a la qüestió de si és possible fer servir les xarxes com a eines predictives. Per fixar idees, considerem el problema següent. Tenim un equip de persones treballant en un projecte durant un any. Al cap d'uns mesos, preguntant els membres de l'equip ens assabentem que hi ha parelles de persones dins l'equip que no tenen una bona relació entre elles: tenen un conflicte d'alguna mena. La pregunta és si podem fer servir la teoria de xarxes per predir, al cap d'uns mesos més, quines d'aquestes situacions conflictives s'hauran resolt soles i quins conflictes nous hauran aparegut. I la resposta és que sí: si representem l'equip com una xarxa i les relacions conflictives com a connexions, saber quins conflictes es resoldran és el mateix que saber quines connexions desapareixeran de la xarxa original. I les relacions que desapareixeran seran les que, per començar, tenien menys probabilitat d'existir: els «errors» de la xarxa original.

Aquest exemple que sembla tret d'un llibre d'Asimov és, de fet, un exemple real: efectivament, la inferència en xarxes s'ha fet servir amb èxit per predir conflictes en equips de treball. I potser el més sorprenent és que, com que la teoria és tan general, es pot fer servir per a problemes tan diversos com predir les decisions d'un tribunal, si a algú li agradarà una pel·lícula o no abans que la vegi o, fins i tot interaccions nocives entre medicaments. Avui en dia, les xarxes són poderoses eines de predicció.