

O que é um grafo ou rede? De uma forma muito simples não é mais do que um conjunto de nodos (vértices) e ligações (arestas) entre eles, como explica neste artigo José Fernando Mendes. Os exemplos mais comuns de redes incluem a *World Wide Web*, *Internet*, redes sociais, redes biológicas e muitas outras.

JOSÉ FERNANDO F. MENDES

Departamento de Física da Universidade de Aveiro
Campus Universitário de Santiago
3810-193 Aveiro
jfmendes@fis.ua.pt

FÍSICA DE REDES COMPLEXAS

O que é um grafo ou rede? De uma forma muito simples não é mais do que um conjunto de nodos (vértices) e ligações (arestas) entre eles (Fig. 1) [1-5]. Os exemplos mais comuns de redes (grafo é uma designação mais usada na literatura matemática) incluem a *World Wide Web (WWW)*, *Internet*, redes sociais, redes biológicas e muitas outras (Fig. 2). O estudo das redes teve o seu início por volta de 1735 quando Leonard Euler apresentou a solução para o problema das pontes de Königsberg. Esta prova é por muitos hoje considerada o ponto de partida de um ramo da matemática (teoria dos grafos), para o qual pessoas como Erdos Rényi deram contribuições importantíssimas, nas últimas décadas do século XX. Também, durante o século passado, muitos estudos, mais empíricos, foram realizados na área das ciências sociais. Um dos trabalhos percursores nesta área foi o de Milgram [6]. Mais recentemente, com o aparecimento de redes reais com um grande número de nodos e com o aumento da capacidade computacional tornaram-se viáveis estudos das mesmas.

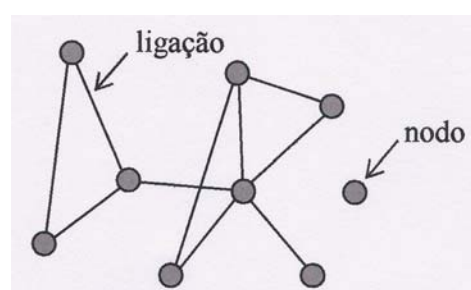


Fig. 1 Representação esquemática de uma rede

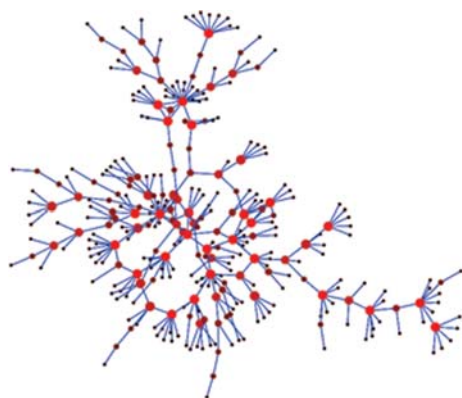
Como pode, por exemplo, a estrutura da rede afectar o tráfego na *Internet*, ou o desempenho de um motor de busca, ou a dinâmica de sistemas sociais ou biológicos? Ou podemos nós, através do conhecimento da rede social correspondente a uma dada sociedade, por exemplo, auxiliar na prevenção da propagação de uma epidemia, ou de um vírus informático na *Internet*? Respostas a estas, e a muitas outras perguntas, têm sido procuradas por uma grande comunidade de cientistas em variadas áreas. No entanto, a compreensão destes sistemas complexos (redes) permanece na sua infância.

simulações numéricas pode ser gerado um número finito de realizações do *ensemble*. Como se sabe da mecânica estatística, os *ensembles* estatísticos são classificados como de equilíbrio ou de não-equilíbrio. No caso de redes, estas são aleatórias de equilíbrio ou de não-equilíbrio, consoante o seu número de nodos é fixo ou cresce.

TIPOS DE REDES

Quando um conjunto de vértices é ligado através de um certo número de ligações, e não levando em consideração outro tipo de aspectos, estamos perante o exemplo mais simples de uma rede (Fig. 1). No entanto, estas podem ser mais complicadas. Pode, por exemplo, haver mais do que um tipo de nodos na rede, ou mais do que um tipo de ligações. Numa rede económica, por exemplo, os nodos podem representar empresas ou bancos, etc., o mesmo se passando com as ligações, que podem ser de diferentes tipos (banco-banco, banco-empresa), ou mesmo terem diferentes intensidades, correspondendo estas, por exemplo, ao volume de negócio entre os nodos. Podem ser dirigidas, ou não, consoante a troca se faz num ou em ambos os sentidos. Por exemplo, a *WWW* é uma rede dirigida, enquanto a *Internet* não o é.

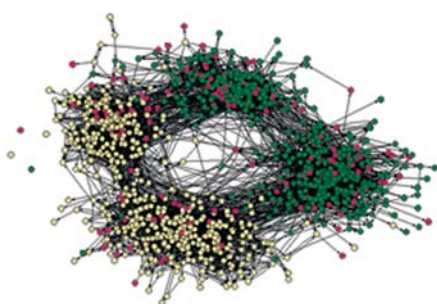
O que são redes aleatórias? Redes aleatórias são *ensembles* estatísticos de redes. Um *ensemble* estatístico é um conjunto de redes particulares, cada uma delas tendo uma probabilidade específica de realização (um peso estatístico). Em estudos empíricos, regra geral, é observado um único membro (uma realização particular) deste *ensemble*. Em



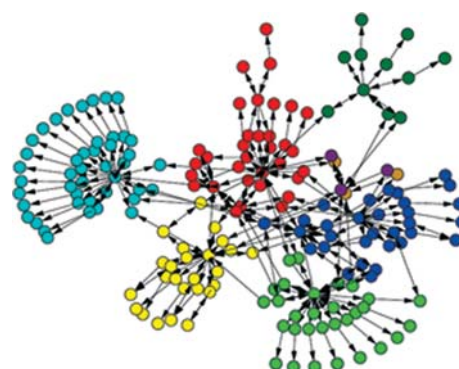
2a)



2b)



2c)



2d)

Fig. 2 a) Rede de contactos sexuais entre indivíduos, por Potterat *et al.* [7]; b) Rede de contágios entre pessoas [8]; c) Rede dos amigos numa escola dos Estados Unidos. As relações de amizade foram obtidas através de questionários, por James Moody [9]; d) Documentos num sítio da *Web* e ligações entre eles. As cores designam diferentes comunidades, por M. Girvan e M. E. J. Newman [10]

Nodo (vértice): característica local de uma rede, pode ser um documento (*Web*), um computador (*Internet*), um actor (filmes), um gene (biologia), etc.

Ligação: a linha entre dois nodos. Podem ser de vários tipos ou intensidades.

Dirigida/não dirigida: consoante aponta num sentido entre dois nodos ou em ambos. Por exemplo, na *Web* uma página pode ter um link para outra e o contrário não se verificar. Na *Internet*, os cabos ópticos transportam informação em ambos os sentidos.

Conectividade ou grau: número de ligações presentes num nodo. Se a rede é dirigida, fala-se de conectividade de entrada e de saída, consoante a ligação aponta para o nodo ou, pelo contrário, sai do nodo e aponta para outro.

Distribuição de conectividades: diz-nos como se distribuem as ligações pelos nodos, dá-nos a probabilidade de um nodo ter k ligações. Numa rede aleatória, a distribuição de conectividades é a fracção média dos nodos com grau k : $P(k) = \langle N(k) \rangle / N$. Aqui $N(k)$ é o número dos nodos de grau k numa rede particular do *ensemble* estatístico. O cálculo da média é sobre todos os elementos do *ensemble* estatístico.

Caminho mais curto: é a menor distância entre dois nodos na rede. Em geral existe mais do que um caminho a ligá-los.

Diâmetro: é o comprimento da maior distância entre quaisquer dois nodos (medida em número de ligações).

Matriz adjacente: esta matriz contém toda a informação sobre uma rede. Uma rede de N nodos tem uma matriz de adjacência de $N \times N$. Cada elemento da matriz de adjacência a_{ij} é igual ao número das ligações que conectam os nodos i e j .

Coefficiente de agregação: a sua definição está relacionada com os ciclos de comprimento três (triângulos de ligações). O coeficiente de agregação local é o número relativo das conexões entre os vizinhos mais próximos de um nodo i , $C_i = 2n_i / (k_i(k_i - 1))$. Aqui k_i é o grau do nodo i e n_i é o número total das conexões entre seus vizinhos mais próximos.

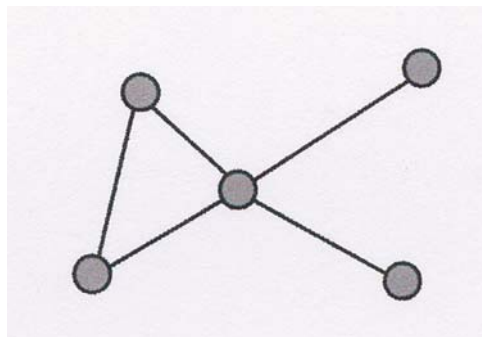


Fig. 3: Ilustração da definição de coeficiente de agregação, C . Esta rede tem um triângulo em seis possíveis, portanto, para o nodo central, $C=1/6$

ALGUMAS REDES REAIS

Nesta secção apresentar-se-ão alguns dos exemplos mais conhecidos de redes reais:

1 - Redes sociais

Uma rede social não é mais que um conjunto de pessoas ou grupos de pessoas, ligados entre eles por relações que podem ser profissionais, familiares ou outras. Os estudos neste campo iniciaram-se nos anos trinta por Moreno (1934) [11]. A introdução de modelos matemáticos foi feita por Rapoport [12], o primeiro autor a realçar a importância da distribuição do número de ligações. Outro estudo importante, nesta área, foi o de Milgram [13] através das suas experiências que levaram ao conceito de "pequeno mundo" (*small world*). Estas experiências estão na base do conceito dos "seis graus de separação".

2 - Redes de informação

O exemplo clássico de uma "rede de informação" é a rede das citações entre artigos científicos. A maioria dos artigos publicados cita outros trabalhos anteriores relacionados ou importantes para este. Estas citações geram uma rede em que os nodos são os artigos científicos e a ligação dirigida do artigo em causa para outro artigo indica que este cita o precedente. A estrutura da rede de citações reflecte assim a estrutura da informação armazenada nos seus nodos: daí a designação "rede de informação". Estas redes têm certas particularidades, que por vezes não surgem noutras. Isso prende-se com o facto de, após a publicação de um artigo que cita outros anteriores, a informação relativa ao número de citações ficar congelada não sendo mais possível alterar esse número. Jaffe e Trajtenberg [14], por exemplo, estudaram a rede das citações entre patentes dos E. U. A., que é semelhante, em alguns aspectos, à das citações entre artigos científicos.

Um outro exemplo muito importante de uma rede de informação é a *World Wide Web*, onde os nodos são os documentos disponíveis e os *links* fazem as ligações entre documentos.

3 - Redes biológicas

(a) - A estrutura de redes neuronais

A estrutura de uma rede neuronal de um organismo minúsculo, o *C. elegans*, contém 282 neurónios que dão forma a uma rede dirigida com uma conectividade média $\langle k \rangle = 14$ [15]. As distribuições de conectividades são exponenciais. O comprimento médio da distância mais curta é 2,65, e o coeficiente de agregação 0,26. Consequentemente, a rede apresenta o efeito de pequeno mundo e o coeficiente de agregação é muito maior do que o valor característico para o caso da rede aleatória clássica, $C = 0,26 \gg 14/282 = 0,05$.

(b) - As redes de reacções metabólicas

Um grande número de sistemas biológicos pode ser utilmente representado como uma rede. Um exemplo valioso de uma rede biológica com uma estrutura topológica extremamente rica é fornecido pela rede das reacções metabólicas [16]. Actualmente, tais redes estão documentadas para diversos organismos. Os seus nodos são substratos - componentes moleculares, e as ligações são reacções metabólicas que conectam os substratos. De acordo com [17], as ligações que apontam para um dado substrato são as reacções em que este participa como um produto.

(c) - Redes de proteínas

Trata-se da rede de interacções físicas entre proteínas (ao contrário das químicas nas redes metabólicas). Este tipo de redes foi estudado por um grande número de cientistas [18].

(d) - Redes genéticas

A rede de regulação genética designa uma outra classe, igualmente importante, de redes biológicas. A expressão de um gene, isto é, a produção através da transcrição e translação de proteínas codificadas pelo gene, pode ser controlada pela presença de outras proteínas. Assim o próprio genoma forma uma rede onde os nodos representam as proteínas e as ligações dirigidas representam a dependência da produção da proteína na presença de outras proteínas (nodos).

AS PROPRIEDADES DAS REDES

As redes aleatórias mais simples são o chamado *classical random graph* (Solomonoff e Rapoport, 1950-1952, Erdos e Rényi, 1959-1960 [19], Gilbert 1959). Em termos simples, estas são redes máximas aleatórias sob a restrição de que a conectividade média dos seus nodos, $\langle k \rangle$, está fixa (o número de nodos está também fixo). Por "máximas aleatórias" entenda-se que são redes para as quais a entropia é máxima. Há várias versões de redes aleatórias clássicas:

- O **modelo de Erdos-Rényi**, que consiste no *ensemble*

estatístico de todas as redes possíveis precisamente com N nodos e L ligações, onde cada membro do *ensemble* tem igual probabilidade de acontecer.

- No **modelo de Gilbert**, cada par de nodos (dos N) é conectado com uma probabilidade p . Obtém-se assim um *ensemble* estatístico de todos os gráficos possíveis de N nodos. Os membros deste *ensemble* são pesados com pesos estatísticos. No limite termodinâmico (redes infinitamente grandes), estas duas versões são equivalentes ($\langle k \rangle = p(N-1)$).

A distribuição de conectividades destas redes aleatórias clássicas obedece a uma distribuição binomial, ou uma distribuição de Poisson no limite de N grande: $P(k) \sim \langle k \rangle^k / k!$. A conectividade média $\langle k \rangle$ está fixa. Esta é uma distribuição que decai rapidamente com uma escala natural $k \sim \langle k \rangle$. Uma particularidade desta rede é que todos os seus momentos convergem. O mesmo não se verifica nas redes encontradas na Natureza, com uma dependência em lei de potência.

A maioria das características interessantes das redes reais que atraíram a atenção dos investigadores nos últimos anos prende-se com o facto de estas não serem redes aleatórias no sentido acima definido.

- O **modelo da configuracional** (introduzido por B. Bollobás [6]) é a primeira generalização natural do modelo de redes aleatórias clássicas. De um modo muito simples, o modelo de configuração é uma rede máxima aleatória com uma distribuição dada $P(k)$ da conectividade. Esta rede aleatória complexa de equilíbrio é não correlacionada. A maioria dos resultados para redes complexas são obtidos usando o modelo de configuração. Existe, contudo, uma outra forma, mais tradicional para os físicos estatísticos, de construir *ensembles* de redes. Por vezes, é chamado modelo exponencial. Os membros do *ensemble* estatístico nesta construção são sistemas (conjuntos) de configurações locais dos nodos e ligações. Cada tipo destes conjuntos ("tijolos") tem a sua "energia de excitação". Por excitação térmica podemos obter um conjunto de realizações (redes) do *ensemble*. As energias específicas da excitação determinam os pesos estatísticos destas realizações, isto é, a estrutura da rede aleatória resultante.

REDES DE CAUSA PESADA (FAT-TAILED)

Se uma rede tiver uma distribuição de conectividades que varia de forma suficientemente lenta com o aumento da conectividade, como acontece com a maioria das redes reais importantes, as suas propriedades são bem distintas das caracterizadas por distribuições do tipo Poisson. Em geral, nestes casos os investigadores tentam interpolar as distribuições empíricas da conectividade por dependências específicas do tipo de lei de potência, $P(k) \sim k^{-\gamma}$ (distribuições sem escala). Contudo, um facto importante

observado é que os momentos de ordem mais elevada das distribuições empíricas da conectividade divergem em redes grandes. Esta observação mostra que, com probabilidade apreciável, os nodos de conectividade elevada estão presentes em redes reais, ao contrário das redes aleatórias clássicas.

OS MODELOS DE REDES EM CRESCIMENTO

Nesta secção iremos examinar uma classe de modelos cujo objectivo preliminar é explicar as propriedades de uma dada rede. Nestes modelos, as redes crescem tipicamente pela adição gradual de novos nodos e de ligações feitas de uma forma que de alguma maneira reflecta os processos de crescimento que originam as redes reais. As ligações são adicionadas de forma preferencial.

Começaremos com o modelo introduzido por Price [20] que foi baseado, por sua vez, num trabalho anterior de Simon [21]. Um trabalho semelhante foi apresentado anteriormente por Yule (1925) [22]. O modelo de Barabási e Albert [23], que foi em certa medida o precursor desta área, nos tempos mais recentes, será apresentado de seguida. Outras variantes e generalizações destes modelos foram introduzidas por outros autores [1].

O MODELO DE PRICE

Derek de Solla Price apresentou em 1965 [20] aquele que provavelmente terá sido o primeiro exemplo do que hoje é chamado uma rede *scale-free*; estudou a rede das citações entre artigos científicos e encontrou que ambas as distribuições (entrada e saída: número de vezes que um artigo é citado e número das vezes que um artigo cita outros) obedecem a uma lei de potência. De facto, o seu trabalho baseou-se nas ideias de Herbert Simon [21], que mostrou que as leis de potência surgem quando estamos na presença do conceito *rich get richer*, isto é, a quantidade de dinheiro que ganhamos depende da quantidade que já temos. Em sociologia, este efeito é conhecido por efeito Matthew [24]. Também é conhecido por "vantagem cumulativa". Na versão do economista Pareto, seria conhecido pela regra do 80/20. Todos estes mecanismos são equivalentes ao proposto mais recentemente por Barabási e Albert: a **ligação preferencial** [23]. Price foi talvez o primeiro a discutir especificamente a vantagem cumulativa no contexto das redes. A sua ideia era que a taxa com que um artigo começa a receber novas citações deve ser proporcional ao número das que já tem. Isto é fácil de justificar de uma maneira qualitativa. O mesmo argumento pode também ser aplicado a outras redes. A solução exacta do modelo de Price pode ser obtida através de uma equação-mestra.

Consideremos uma rede dirigida com N nodos. Seja p_k a fracção de nodos na rede com conectividade k , de modo que $\sum_k p_k = 1$.

Novos nodos são adicionados continuamente à rede, embora não necessariamente com uma taxa constante. Cada nodo adicionado tem um dado número prévio de ligações começando nele próprio - corresponde ao número dos artigos que cita - e note-se que este valor ficará inalterado durante a evolução da rede.

Por razões de simplicidade seja m (em média) o número dessas citações. A conectividade média da rede é dada por $m = \sum_k k p_k$.

A probabilidade de um artigo, recentemente aparecido, citar um artigo precedente é simplesmente proporcional ao número de ligações que chegam ao artigo existente. No entanto, inicialmente um artigo não tem citações para ele próprio, o que faz surgir um problema. Price resolveu o problema admitindo que o artigo à nascença tem automaticamente uma ligação para si próprio (auto-citação). Portanto, a probabilidade de um artigo receber uma nova citação é simplesmente proporcional a $k+1$. A distribuição para este modelo vem assim dada por:

$$p_k = (1 + 1/m)B(k + 1, 2 + 1/m) \sim k^{-(2+1/m)}$$

onde $B(a, b) = \Gamma(a)\Gamma(b) / \Gamma(a + b)$ é a função Beta de Legendre. Portanto, no limite de N grande, a distribuição de conectividades é do tipo lei de potência (*scale-free*) com um expoente $\gamma = 2 + 1/m$.

MODELO DE BARABÁSI E ALBERT E A IDEIA DE LIGAÇÃO PREFERENCIAL

O modelo que agora apresentaremos consiste numa redescoberta do modelo proposto por Price. Este modelo também é conhecido por modelo com ligação preferencial. A solução exacta deste modelo foi apresentada por Dorogovtsev e Mendes [1] (a solução apresentada por Barabási e Albert [2] era aproximada e apenas válida sob certas condições). A probabilidade de um nodo se ligar a outro de conectividade k é proporcional a k .

Consideremos então uma rede dirigida, e vamos estudar a distribuição do número de ligações de entrada num nodo (*in-degree*), $q(s, t) = k_i(s, t)$. As regras de evolução são as seguintes:

- (1) Por unidade de tempo, um novo nodo é adicionado à rede.
- (2) Simultaneamente, m novas ligações dirigidas provenientes de nodo não especificados são adicionadas.
- (3) As extremidades destas novas ligações são distribuídas de acordo com a seguinte regra: a probabilidade de uma nova ligação apontar para um nodo s é proporcional a $q(s) + A$. O parâmetro $A = ma$ joga o papel de uma atractividade adicional. Desta forma a conectividade total de um nodo s no instante t é dada por $k(s, t) = q(s, t) + m$. Se consi-

derarmos $A = m$, isto é, $a = 1$, então as novas ligações são distribuídas com probabilidade proporcional a $k(s,t)$.

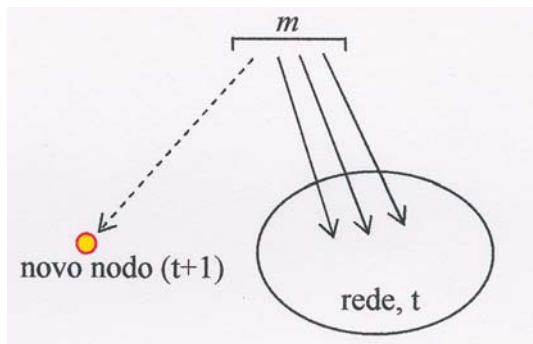


Fig 4: Representação esquemática da distribuição das m novas ligações.

A equação mestra para a evolução da distribuição de conectividade de chegada num nó s no instante t é $p(q, s, t)$ e pode ser entendida do seguinte modo. A probabilidade de uma nova ligação apontar para um nó s é igual a

$$\Pi = \frac{q(s,t) + am}{(1+a)mt}$$

A probabilidade de um nó s receber exactamente l novas ligações das m injectadas é:

$$P_s^{ml} = \binom{m}{l} \left[\frac{q(s,t) + am}{(1+a)mt} \right]^l \left[1 - \frac{q(s,t) + am}{(1+a)mt} \right]^{m-l}$$

Assim, a distribuição de conectividades de um nó particular é dada por,

$$\begin{aligned} p(q, s, t+1) &= \sum_{l=0}^m P_s^{ml} p(q-l, s, t) \\ &= \sum_{l=0}^m \binom{m}{l} \left[\frac{q-l+am}{(1+a)mt} \right]^l \\ &\quad \times \left[1 - \frac{q-l+am}{(1+a)mt} \right]^{m-l} p(q-l, s, t). \end{aligned}$$

Esta equação está sujeita a uma condição de fronteira, $p(q, t, t) = \delta_{q,0}$ (delta de Kronecker). Somando sobre s na equação anterior e considerando t grande, obtém-se a equação integro-diferencial:

$$\begin{aligned} (1+a)t \frac{\partial P}{\partial t}(q, t) + (1+a)P(q, t) \\ + (q+am)P(q, t) - (q-1+am)P(q-1, t) \\ = (1+a)\delta_{q,0} \end{aligned}$$

Levando em consideração a parte estacionária da equação (excluir o termo com a derivada temporal) tem-se uma equação para a distribuição estacionária

$$P(q) = P(q, t \rightarrow \infty),$$

supondo que esta existe. A solução exacta desta equação é dada por [1]:

$$P(q) = (1+a) \frac{\Gamma(1+(m+1)a)}{\Gamma(ma)} \frac{\Gamma(q+ma)}{\Gamma(q+2+(m+1)a)}$$

onde Γ é a função gama. Quando $a = 1$, obtém-se o resultado correspondente ao modelo de Barabási e Albert,

$$P(q) = \frac{2m(m+1)}{(q+m)(q+m+1)(q+m+2)} \sim q^{-3}$$

Uma formulação mais rigorosa deste modelo foi apresentada por Ballobás [6]. Dada a atenção que este modelo despertou na comunidade científica, várias generalizações têm sido propostas. Uma variante, proposta por Dorogovtsev e Mendes [5], consistiu em considerar que a conectividade média da rede cresce com o tempo (crescimento acelerado). Este efeito foi observado, por exemplo, na *World Wide Web* e na *Internet*, o que corresponde ao parâmetro m variar com o tempo. Se admitirmos que m , número de novas ligações por unidade de tempo (por cada novo nó introduzido), cresce com o tamanho da rede, t , na forma t^α , e que a probabilidade de se ligar a um dado nó é proporcional a $q+Bt^\alpha$ (com B constante), então a distribuição de conectividades obedece a uma lei de potência com expoente:

$$\gamma = 2 + \frac{B(1+\alpha)}{(1-B\alpha)}.$$

Muitos outros estudos têm mostrado que na sua maioria, as redes naturais e artificiais apresentam uma estrutura descrita por leis do tipo *fat-tailed* e *scale-free*. Uma das mais importantes e intrigantes propriedades de redes deste tipo está relacionada com o facto destas apresentarem uma grande estabilidade contra ataques intencionais e aleatórios.

Esta robustez contra falhas é obviamente importante para as redes biológicas e de comunicações. Isto, em parte, explica por que razão as redes com esta estrutura são comuns na Natureza. A sua enorme estabilidade é uma consequência da sua estrutura. Mas, ao mesmo tempo, apresentam um fenómeno contrastante, a ausência de um limiar epidémico, por exemplo, na propagação de epidemias. Epidemias podem facilmente espalhar-se em redes complexas, sendo este o ponto fraco destas.

REFERÊNCIAS

- [1] Dorogovtsev, S.N. and Mendes, J.F.F., "Evolution of networks", *Adv. Phys.*, **51**, 2002, pp. 1079.
- [2] Albert, R. and Barabási, A.-L., "Statistical mechanics of complex networks", *Rev. Mod. Phys.*, **74**, 2002, pp. 47.
- [3] Newman, M.E.J., "The structure and function of complex networks", *SIAM Review*, **45**, 2003, pp. 167.
- [4] Watts, D.J., *Small Worlds: The Dynamics of Networks between Order and Randomness*, Princeton University Press, Princeton, NJ., 1999.
- [5] Dorogovtsev, S.N. and Mendes, J.F.F., *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, Oxford, 2003.
- [6] Bollobás, B. and Riordan, O.M., "Mathematical results on scale-free random graphs", in *Handbook of Graphs and Networks: From the Genome to the Internet* (S. Bornholdt and H.G. Schuster, eds.), Wiley-VCH, Berlin, 2002.
- [7] Poterat, J. J., *et al.*, "Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs", *Sexually Transmitted Infections*, **78**, 2002, i159-i163.
- [8] Valdis Krebs, <http://www.orgnet.com/contagion.html>
- [9] "High school friendship: James Moody, Race, school integration, and friendship segregation in America", *American Journal of Sociology* **107**, 2001, pp. 679-716.
- [10] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks", *Physical Review E*, **69**, 2004, 026113.
- [11] Moreno, J. L., *Who Shall Survive?*, Beacon House, Beacon, NY, 1934.
- [12] Rapoport, A., "Contribution to the theory of random and biased nets", *Bulletin of Mathematical Biophysics*, **19**, 1957, pp. 257-277.
- [13] Milgram, S., "The small world problem", *Psychology Today*, **2**, 1967, pp. 60-67; Travers, J. and Milgram, S., "An experimental study of the small world problem", *Sociometry*, **32**, 1969, pp. 425-443.
- [14] Jaffe, A. and Trajtenberg, M., *Patents, Citations and Innovations: A Window on the Knowledge Economy*, MIT Press, Cambridge, MA, 2002.
- [15] Watts, D. J. and Strogatz S. H., "Collective dynamics of small-world networks", *Nature*, **393**, 1998, pp. 440.
- [16] Kauffman, S. A., "Metabolic stability and epigenesis in randomly constructed genetic nets", *J. Theor. Biol.*, **22**, 1969, pp. 437.
- [17] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. and A-L. Barabási, "The large-scale organization of metabolic networks", *Nature*, **407**, 2002.
- [18] Jeong, H., Mason, S., Barabási, A.-L., and Oltvai, Z. N., "Lethality and centrality in protein networks", *Nature*, **411**, 2001, pp. 41-42.
- [19] Erdos, P. and Rényi, A., "On random graphs", *Publicationes Mathematicae*, **6**, 1959, pp. 290-297; Erdos, P. and Rényi, A., "On the evolution of random graphs", *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, **5**, 1960, pp. 17-61.
- [20] Price, D. J. de S., "Networks of scientific papers", *Science*, **149**, 1965, pp. 510-515; Price, D. J. de S., "A general theory of bibliometric and other cumulative advantage processes", *J. Amer. Soc. Inform. Sci.*, **27**, 1976, pp. 292-306.
- [21] Simon, H. A., "On a class of skew distribution functions", *Biometrika*, **42**, 1955, pp. 425-440.
- [22] Yule, G., "A mathematical theory of evolution based on the conclusions of Dr. J.C. Willis, F.R.S.", *Philosophical Transactions of the Royal Society of London (Series B)*, 1925, 213:21-87.
- [23] Barabási, A.-L. and Albert, R., "Emergence of scaling in random networks", *Science*, **286**, 1999, pp. 509-512.
- [24] Merton, R. K., "The Matthew effect in science", *Science*, **159**, 1968, pp. 56-63.