

## **A Rede de Comunicação World Wide Web no domínio \*.pt: Métricas Fundamentais \***

**António Machuco Rosa**, Universidade Lusófona de Humanidades e Tecnologias, Portugal

**Jorge C. Giro**, Universidade Lusófona de Humanidades e Tecnologias, Portugal

### **Abstract**

Nos últimos anos têm sido levados a cabo diversos estudos empíricos que permitem caracterizar as principais métricas topológicas das redes de comunicação. Neste artigo apresentamos os resultados de um estudo da World Wide Web de domínio \*.pt, em particular os que respeitam à função de distribuição, coeficiente de agrupamento e distância nessa rede. Mostra-se que os resultados obtidos estão de acordo com os de outros estudos que se debruçaram sobre a World Wide Web mundial. Finalmente, aponta-se para a o significado que os resultados obtidos adquirem no contexto do estudo dos novos meios de comunicação.

### **Introdução**

Um conjunto de dados estatísticos surgido nos últimos anos foi acompanhado pela elaboração de novos modelos teóricos que provocaram uma verdadeira revolução no estudo e compreensão das redes. Durante bastante tempo, sensivelmente no período que transcorreu entre as décadas sessenta e noventa do século passado, o principal modelo das redes foi o chamado modelo de Erdős (Erdős e Rényi, 1959). Esse modelo baseia-se no conceito mais estilizado de rede: existem nós que se encontram, ou não, ligados entre si. No caso específico do modelo de Erdős, coloca-se inicialmente um número fixo de  $N$  nós e nenhuma ligação. De seguida, conecta-se aleatoriamente com probabilidade  $p$  cada par de nós. Essa probabilidade pode ser calculada como a fracção entre as ligações actualmente existentes e a totalidade das  $n(n-1)/2$  ligações possíveis de existir em qualquer rede. Um exemplo pode ser o da figura 1.

---

\* Este artigo resultou do projecto de investigação *Trends on Portuguese Networks Culture*, projecto financiado pela FCT/POCTI/34436.

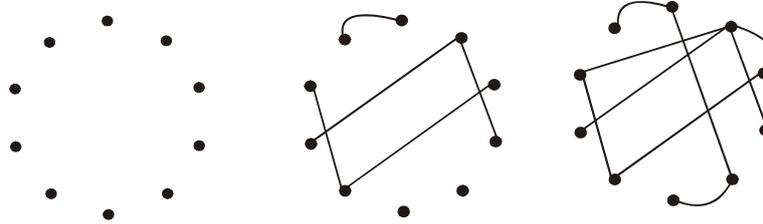


Figura 1. A construção do modelo das redes de aleatórias de Erdős. Parte-se de um conjunto de  $N$  nós fixos que de seguida são conectados aleatoriamente com probabilidade  $p$ . Da esquerda para a direita tem-se  $p=0$ ,  $p \approx 0,1$ ,  $p=0,2$ .

Uma das propriedades fundamentais das redes que o modelo permite investigar é a *função de distribuição* das ligações pelos nós. Como seria de esperar, o modelo prevê uma distribuição binomial das ligações (uma distribuição aproximadamente de tipo gaussiano), que se traduz no facto de a maioria dos nós possuir, em média, aproximadamente o mesmo número de ligações. Quando o número de nós,  $N$ , é grande, a distribuição das ligações é uma distribuição de Poisson:

$$P(k) = e^{-\langle k \rangle} \langle k \rangle^k / k! \quad (1)$$

A figura 2 ilustra essa função de distribuição.

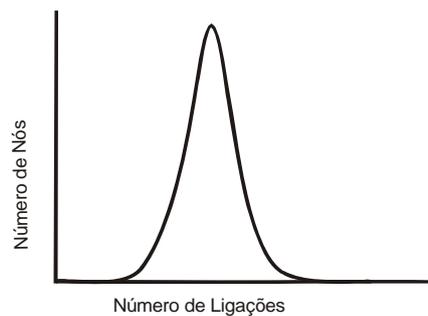


Figura 2. A distribuição de ligações de acordo com o modelo de Erdős: a maioria dos nós possui aproximadamente o mesmo número de ligações.

Esse modelo sofreu uma primeira modificação importante quando, em 1998, D. Watts e S. Strogatz (Watts e Strogatz, 1998) propuseram um modelo (chamado dos 'mundos-pequenos') que permite caracterizar de

modo exacto uma outra propriedade fundamental das redes, a *distância*,  $D$ , que informalmente pode ser definida como o número médio de nós intermédios que têm de ser percorridos para ir de um nó a qualquer outro nó (assumindo que esse percurso existe). Watts e Strogatz constataram que a distância nas redes reais tende a ser pequena. Ela é pequena por comparação, por um lado, com a distância existente em redes formais em que o número de ligações é o mesmo em qualquer nó e, por outro, aproxima-se dos valores de distância curta existentes nas redes aleatórias acima definidas. Watts e Strogatz destacaram igualmente uma outra propriedade das redes, o *coeficiente de agrupamento*,  $C$ , que é uma quantidade que mede a existência de triângulos ou ciclos numa rede. Em termos intuitivos, existe um triângulo quando existe uma ligação entre um nó A e um nó B, entre B e um nó C, e existe finalmente o fecho do triângulo com a presença de uma ligação entre A e C. Watts e Strogatz mostraram que o coeficiente de agrupamento tende a ser grande nas redes reais, sobretudo quando comparado com o seu valor nas redes aleatórias. Ocorreu a verdadeira revolução na compreensão das redes com o estudo empírico levado a cabo em 1999 pelo grupo de investigação de A-L. Barabási (Albert *et al*, 1999). O grupo investigou as propriedades de um fragmento da World Wide Web (WWW), a World Wide Web de domínio absoluto nd.edu. e, contrariamente ao que era tradicionalmente assumido com base no modelo matemático de Erdős, verificou-se que a função de distribuição,  $P(k)$ , das  $k$  ligações (*hyperlinks*) entre os  $N$  nós (páginas *web*) estava bem longe de ser do tipo gaussiano. Ela tem antes a forma  $P(k) \sim k^{-\lambda}$ , isto é, uma distribuição sem escala característica, ou em lei de potência, onde a probabilidade de um nó (página) aleatoriamente escolhido receber  $k$  ligações decresce segundo a razão dada pelo expoente de  $k$ . A figura 3 ilustra uma função com uma distribuição sem escala característica.

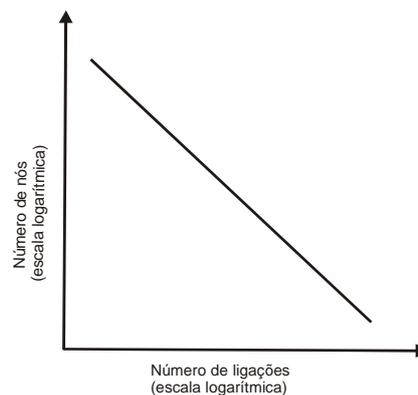


Figura 3. Uma função sem escala característica. O valor do declive da função é dado pelo expoente  $\lambda$ .

Barabási e colaboradores constataram que  $\lambda=2.1$  na rede empírica que investigaram. Verificaram também que a distância é pequena,  $D=11.2$ , e que o coeficiente de agrupamento é alto,  $C=0.11$ . Posteriormente, um estudo mais amplo da World Wide Web baseado nos nós atingíveis a partir de AltaVista obteve também  $\lambda = 2.1$  e  $D = 16.18$  (Broder *et al*, 2000).

Dentro do quadro teórico geral da moderna teoria das redes, foram esses estudos que replicámos para o caso da WWW portuguesa de domínio \*.pt. O nosso estudo teve um objectivo mais amplo, pois procurou-se identificar a distribuição de diversos componentes multimedia na *web* portuguesa, métricas de complexidade, etc. Aqui, demonstramos os resultados respeitantes às principais métricas das redes: função de distribuição, distância e coeficiente de agrupamento. Este estudo tem o interesse de ver se a estrutura da WWW de domínio \*.pt é semelhante à WWW mundial. Além disso, uma estrutura topológica das redes que se afasta dos modelos clássicos tem seguramente consequências ao nível da criação e utilização dos novos *media* em rede. Serão essas consequências que serão sumarizadas na conclusão.

#### ***Dados globais respeitantes à WWW de domínio \*.pt.***

Os varrimentos que efectuámos da WWW de domínio \*.pt basearam-se nos domínios registados na FCCN (Fundação para a Computação Científica Nacional), e que descriminamos por ano.

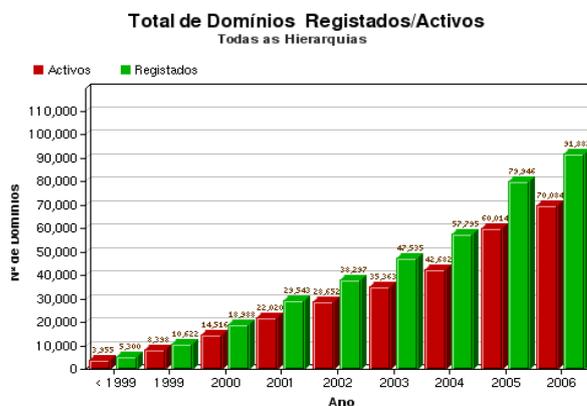


Figura 4. Evolução de domínios registados/ativos no domínio \*.pt.

Fonte: Fonte FCCN: <http://www.fccn.pt>

Levámos a cabo dois varrimentos, um em 2002 e outro mais recente em 2006.

Do ponto de vista metodológico, utilizámos procedimentos *standard* neste tipo de pesquisa. Dado não se poder garantir que a WWW é uma rede conexa (i.e, que existe um percurso entre quaisquer dois nós) e,

por outro lado, dado o domínio \*.pt se encontrar sob a administração centralizada da FCCN, adoptou-se a metodologia de efectuar uma travessia de cada URL registado, transpondo, de seguida, as hiperligações externas de cada sítio para uma matriz de adjacência que permite representar as ligações entre os diversos sítios e calcular as quantidades desejadas. Recorde-se que a matriz de adjacência de uma rede com nós  $N = (1, 2, \dots, n)$  e ligações  $k = (i, j)$  vem definida como:

$$\langle k_i \rangle = \sum_j A_{ij} \quad (2)$$

Portanto, a matriz de adjacência de uma rede é uma matriz  $N * N$  tal que  $A_{ij} = 1$  se  $(i, j)$  é uma ligação actualmente existente de entre a totalidade das  $n(n-1)/2$  ligações possíveis da rede, e  $= 0$  no caso contrário. Por outro lado, o algoritmo utilizada para o varrimento da rede foi naturalmente o algoritmo de busca em profundidade (*Breadth First Search*), o qual progride através da expansão do primeiro nó filho da árvore de busca, aprofundando-se cada vez mais até que o alvo da busca seja encontrado ou até que ele se depare com um nó que não possui filhos (nó folha). Então a busca retrocede (*backtrack*) e começa no próximo nó

Apresentamos agora os dados globais após depuramento das hiperligações e *spidering* completo dos URL's acessíveis de entre os cerca de 70 000 nós activos aquando do varrimento da rede \*.pt levado a cabo em 2006. É importante notar que diversas razões tornam inacessíveis muitos dos URL's aquando de um varrimento, em especial:

1. Sítios inactivos aquando do varrimento.
2. Problemas técnicos com o servidor.
3. Acesso negado (áreas protegidas).
4. URL's que enviam para fora do domínio PT.
5. Tecnologias de construção de páginas que inviabilizam largamente a extracção de hiperligações: *Flash* e páginas geradas dinamicamente através de tecnologias *server-side* associadas normalmente a bases de dados.

Atendendo a esses factores, obteve-se:

Sites registados/activos	70 084
Sites visitados	14 272
Páginas visitadas:	3 501 975
Total de <i>links</i> válidos localizados	231 853 950
<i>Bytes</i> descarregados:	87774513634 (85 717 298.4K)
Média de páginas por <i>site</i> :	245
Média de <i>links</i> internos por <i>site</i> :	4,56
Média de ligações por <i>site</i>	4,01 (ligações múltiplas para o mesmo <i>site</i> foram colapsadas)

Tabela 1. Dados globais do varrimento da WWW .pt em 2006.

Na tabela, a média de ligações por *site* refere-se às ligações existentes no interior de um URL de domínio absoluto.

### Métricas específicas das redes

Como já se referiu, procurámos obter para o caso específico da WWW de domínio \*.pt as principais métricas das redes. São esses resultados que de seguida demonstramos.

### Função de distribuição

A World Wide Web é uma rede *orientada*, isto é, a existência de uma ligação de um nó A para um nó B não implica a existência de uma ligação de B para A. Convém assim distinguir os *inlinks* (ligações que apontam para um certo *site*) dos *outlinks* (ligações que um certo *site* aponta para outros). Em ambos os casos, trata-se de ver qual é a distribuição das ligações por *site*. Já referimos que estudos anteriores (em particular o que abarcou um fragmento significativo da World Wide Web mundial) encontraram um função de distribuição do tipo  $P(k) \sim k^{-\lambda}$ , com  $\lambda=2.1$ . Atendendo a que essa função é invariante por mudança de escala, seria de esperar que a *web* portuguesa apresentasse uma mesma distribuição com valores semelhantes para o expoente  $\lambda$ . Foi de facto o que se verificou. A WWW de domínio \*.pt possui uma lei em potência com o **valor do expoente  $\lambda$  para *inlinks* = 2.156**. Como se demonstra no gráfico (figura 5), a aproximação a uma lei em potência é bastante boa apesar do pronunciado *cut-off* que o gráfico da função exhibe, decorrente da existência de *sites* com muitos poucos *links*.

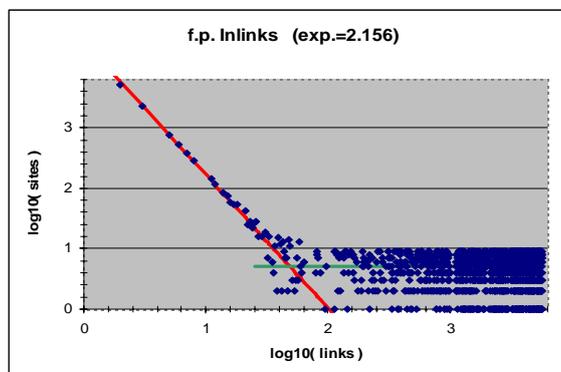


Figura 5. A World Wide Web de domínio \*.pt tem uma distribuição em forma de lei de potência com o valor do expoente  $\lambda$  para *inlinks* = 2.156.

Quando ao valor  $\lambda$  do expoente para *outlinks* ele é igual a 2.203, como de seguida se demonstra.

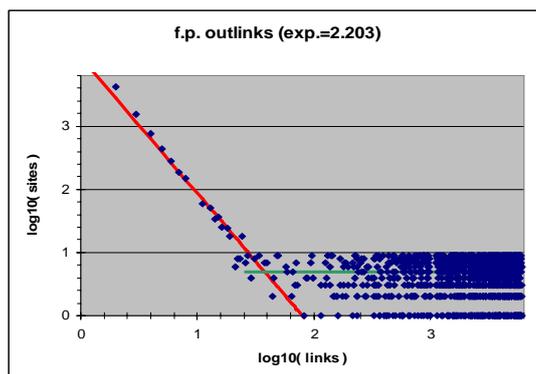


Figura 6. A World Wide Web de domínio \*.pt tem uma distribuição em forma de lei de potência com o valor  $\lambda$  do expoente para *outlinks* igual a 2.203.

Conforme mais abaixo se verá, o valor do expoente de *outlinks* é em geral superior ao de *inlinks*. A diferença numérica não é qualitativamente significativa. Talvez ela decorra de a WWW ser uma rede orientada com uma tendência para, exprimido de forma apenas intuitiva, uma maior divergência da função *inlinks*, traduzida no fenómeno de 'popularidade' de que a função de distribuição dá conta. De facto, numa rede como uma distribuição de ligações em lei de potência, existem poucos *sites* extremamente populares (recebem um número desproporcionadamente grande de ligações), enquanto existe um grande número de *sites* pouco populares. O mecanismo que faz emergir um tal tipo de distribuição é de facto do tipo 'a

popularidade é atractiva' (Dorogovtsev, S., Mendes, 2003), isto é, um *site* (nó) tende a receber mais ligações em função (linear) das ligações que já possui (Barabási *et al*, 1999).

### Distância média

Uma métrica importante é a distância média numa rede. Como já se aludiu, a distância numa rede não é dada pela usual métrica euclidiana. Intuitivamente, é antes o número médio de nós intermédios que é necessário atravessar para ir de um nó a um qualquer outro. Evidentemente que é bastante relevante em termos de acesso e propagação da informação saber se a distância numa rede é curta ou não. Utilizando o conhecido algoritmo de Dijkstra na matriz de adjacência, obtivemos uma distância média entre os *sites* da rede WWW \*.pt de **16.04**. Esse valor compara bem com os obtidos nos outros estudos já referidos (ver tabelas mais abaixo). Ele significa que a *web* portuguesa é um 'mundo-pequeno', isto é, existem relativamente poucos nós intermédios que permitem navegar de um nó a qualquer outro nó.

### Coefficiente de agrupamento

O coeficiente de agrupamento,  $C$ , é a razão entre o número de ligações,  $E_i$ , existentes entre os nós vizinhos de um dado nó, e o número total de ligações possíveis entre esses nós. O valor do coeficiente de agrupamento pode ser obtido pela seguinte fórmula:

$$C_i = \frac{2E_i}{k_i (k_i - 1)} \quad (3)$$

O coeficiente de agrupamento da totalidade da rede é então a média de todos os  $C_i$ .

Obtivemos um valor de  $C$  de **0.12**, que compara bem com o valor obtido para nd.edu (ver tabela mais abaixo). A *web* portuguesa é portanto uma rede como uma alta densidade de ciclos, isto é, é uma rede localmente estruturada. Note-se que o coeficiente de agrupamento é uma quantidade *local* atestando a presença de comunidades de vizinhança (numa rede), enquanto a distância, ao invés, é uma quantidade *global*. Foi este irreduzível afastamento entre local e global que permitiu a Watts e Strogatz compreenderem a propriedade de 'mundo-pequeno', isto é, visto os nós se encontrarem localmente situados, eles não podem perceber aquilo que está para além das condições de acessibilidade da sua

experiência de uma rede: não podem perceber a contracção global da distância levada a cabo por um número reduzido de ligações de longo alcance (cf. Watts, 2003, para a filosofia geral).

Referimos agora de forma sumária os resultados que obtivemos no varrimento levado a cabo em 2002.

### Pesquisa de 2002

A pesquisa da rede que efectuámos em 2006 seguiu-se a uma outra que já tinha sido efectuada em 2002. Apresentamos na seguinte tabela os resultados obtidos nesse ano.

<i>Sites</i> registados	38 297
<i>Sites</i> alcançados	23 902
Total de <i>links</i> válidos localizados:	10 628749
Função de distribuição <i>in</i>	2.1
Função de distribuição <i>out</i>	2.44
Coefficiente de Agrupamento	0.1067

Tabela 2. Resultados de um varrimento da WWW de domínio .pt efectuado em 2002.

### Outros dados

Para efeito de comparação com os nossos resultados, apresentamos de seguida os valores obtidos nos estudos baseado em nd.edu e em AltaVista. No estudo respeitante a nd.edu foram visitadas 269 504 páginas. A função de distribuição resultou ser (para total surpresa, na altura) da forma  $P(k) \sim k^{-\lambda}$ , com  $\lambda = 2.1$  para *inlinks*. A distância revelou ser 11.27, o que levou os autores do estudo a predizer que ela seria de cerca de 19 para a totalidade da WWW. A tabela seguinte resume os principais resultados

Páginas	269 504
Links	1 497 135
Média de <i>links</i> por página	5.55
Distância média	11.27
Função de distribuição <i>in</i>	2.1
Função de distribuição <i>out</i>	2.4
Coefficiente de Agrupamento	0.11

Tabela 3. Valores obtidos para o domínio nd.edu

Em parte motivados pela predição da distância avançada pelo anterior estudo, Broder e colaboradores levaram a cabo uma pesquisa avançada da WWW mundial em 2000 e com base em AltaVista. Um dos seus principais resultados foi mostrar que a WWW mundial não é uma rede completamente conexas. Contudo, assumindo que existe sempre um caminho entre os nós, a distância média revelou ser 16.18. Esse valor compara bem com a predição de Barabási e colaboradores, sobretudo quando se tem em conta que o valor da distância deve ser visto em contraste com o seu valor numa rede aleatória. Assim, por exemplo, a distância seria de 8.32 no caso de a rede nd.edu ser tratada como uma rede aleatória, e de 8.85 em idêntica situação a partir da rede AltaVista.

Adicionalmente, Broder *et al* obtiveram uma função de distribuição sem escala característica, tal como a figura produzida pelos autores demonstra.

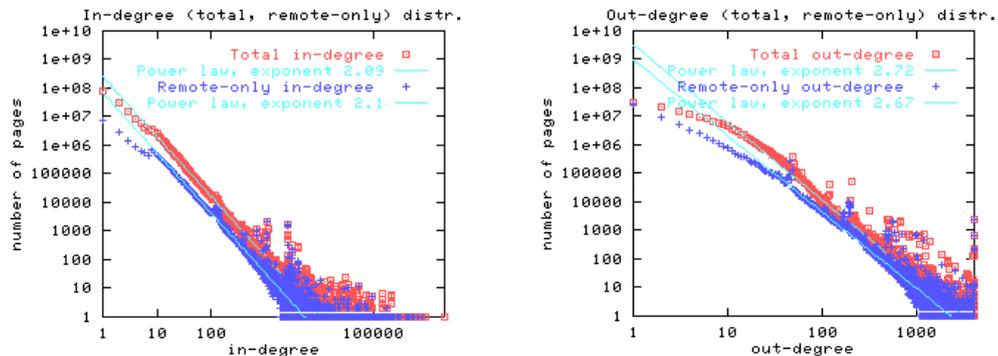


Figura 7. Função e distribuição *in* e *out* obtida a partir de um varrimento de AltaVista. Fonte, Broder *et al*, 2000.

A tabela seguinte resume os principais resultados do estudo de Broder *et al*.

Páginas	203 549 046
Links	2 130 000 000
Média de <i>links</i> por página	10,46
Distância média	16,18
Função de distribuição <i>in</i>	2.1
Função de distribuição <i>out</i>	2.7

Tabela 4. Valores obtidos a partir de AltaVista.

Portanto, os resultados obtidos para nd.edu, para Altavista e para a WWW de domínio \*.pt são qualitativamente similares, mostrando que a WWW enquanto rede tem características robustas

independentes da estrutura fina dos detalhes numéricos. Neste artigo, mostrámos que essas características são a existência de uma função sem escala característica, atestando uma desigualdade profunda na popularidade e condições de acesso a *sítes*, a existência de uma distância pequena, atestando a facilidade na navegação e ao mesmo tempo as constricções, para o bem ou para o mal, da propagação de informação, e ainda, finalmente, a existência de uma estrutura de comunidades medida pelo coeficiente de agrupamento. Na realidade, esses resultados são genéricos, pois prova-se que a maior parte das redes (desde a Internet até redes metabólicas, entre muitas outras) exibem uma distribuição em lei de potência com o expoente  $\lambda$  no intervalo  $2 < \gamma \leq 3$ , tal como exibem distância pequena entre os nós e forte densidade local de ligações. (cf. as sùmulas de Albert e Barabási, 2002, Newman, 2003, Dorogovtsev, S., Mendes, 2003, para um tratamento completo e avançado do tema).

### Conclusão

As características da WWW que analisámos apontam para a emergência de meios de comunicação e informação com uma estrutura típica bem distinta dos *media* tradicionais. Essa diferença pode ser abordada a partir de um diagrama que, mesmo se não é explicitamente pensado como uma rede, é geralmente apresentado como a forma genéricas dos meios de comunicação clássicos. O diagrama é o seguinte (cf. Wolf, 1987).

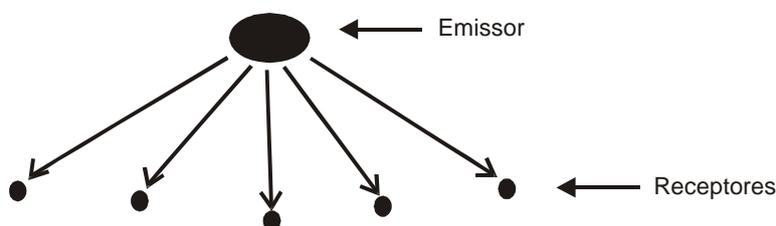


Figura 7. O diagrama dos méis de comunicação de massas clássicos, segundo Wolf, 1987.

O diagrama visa capturar o conceito de *media* estruturado de forma *top-down*, linear e assimétrica, resultando no esquema tradicional emissor  $\rightarrow$  receptores homogéneos. Ora, considerando o diagrama como uma rede, é imediatamente evidente que em tal rede o coeficiente de agrupamento é  $=0$ . Isso aponta, por contraposição, para uma característica fundamental e bem conhecida dos novos *media*: a presença de externalidades traduzidas numa forte densidades de conexões locais horizontais, quantificáveis pelo coeficiente de agrupamento. Por outro lado, estamos evidentemente perante uma distribuição bem distinta da heterogeneidade existente numa rede sem escala característica. Neste último tipo de redes,

existem profundas desigualdades na visibilidade e condições de acesso a sítios, sem no entanto existir uma total orientação para um único ou reduzido número de 'emissores'.

As condições que possibilitam a emergência de novos *media* bem distintos dos meios tradicionais podem ser exactamente analisadas (cf. Machuco Rosa, 2003, 2007). Elas decorrem, ao nível da infra-estrutura física, do facto de o dispositivo tecnológico de 'emissão' e 'recepção' ser o *mesmo*, a saber, o computador universal aberto (susceptível de processar não importa que tipo de formato) e que é economicamente bastante acessível. Ao nível subsequente, o nível lógico dos novos *media*, encontramos plataformas abertas baseadas em protocolos neutrais e disponíveis no domínio público, como os protocolos da Internet e da World Wide Web (TCP/IP e HTTP). Eles permitem o livre acesso e a livre 'emissão' de conteúdos por não implementarem qualquer política discriminatória no seu código-fonte (cf. Lessig, 2001). Decorre dessa arquitectura historicamente contingente o conhecido 'livre acesso' aos novos meios em rede. Portanto, a sua estrutura não é pré-montada no sentido em que, por exemplo, o *design* da rádio, e de seguida o da televisão, o foi pelos poderes políticos de regulação (cf. Starr, 2004). Em sentido preciso, os novos meios crescem por um processo de *auto-organização*. Verdade em geral, essa propriedade foi demonstrada pelo estudo da WWW: enquanto rede, ela é o resultado, permitido pela existência de protocolos e formatos abertos, de inúmeros actos locais de criação de hiperligações. Esses actos são relativamente independentes entre si, e no entanto o seu resultado global (auto-organização) são propriedades estacionárias, precisamente a distância, o coeficiente de agrupamento e, sobretudo, a função de distribuição. A sua emergência atesta uma estrutura bem distinta daquela que é causada pelos mecanismos lineares e assimétricos dos meios de comunicação clássicos. Adicionalmente, a existência de um acesso a páginas constrangido pela profunda desigualdade na visibilidade de cada *site* (de acordo com a distribuição em lei de potência) obriga a repensar as técnicas estatísticas usualmente aplicadas no estudo das audiência nos meios de comunicação clássicos. Elas partem sempre do pressuposto que o sistema escala segundo uma distribuição normal ou gaussiana, a partir de que se obtêm os agregados médios. A constatação que esse tipo de distribuição não se aplica nos novos meios obriga portanto a reequacionar as técnicas de publicidade, de navegação, de disponibilização de conteúdos, etc., que se desenvolvem nesses meios. Seguramente não é por acaso que o algoritmo de busca utilizado por Google (cf. Brin e Page, 1998) induz uma distribuição em forma de lei de potência (cf. Vasquez, 2001).

As conclusões referidas aplicam-se sobretudo à WWW. Esta é uma rede cliente/servidor. Elas ainda se tornariam mais claras se se procedesse a um estudo das redes de *peer-to-peer* (Emule, BitTorrent, etc) e das chamadas redes de *social networking* (myspace.com, etc). Algumas métricas já existem para esse tipo de redes (cf. Ripenau *et al*, 2002). Atendendo à recente explosão das redes de *social networking*, um trabalho futuro deveria incidir sobre essas redes, com possível enfoque nas redes ligadas sobretudo ao

universo de língua portuguesa. De forma ainda mais avançada, sugere-se a continuação do estudo das redes de informação e da comunicação levando a cabo uma análise mais fina que distinga entre diversos tipos de nós e entre diversos tipos de ligações.

### **Bibliografia**

- Albert, R., Jeong, H. & Barabasi, A.-L. (1999). Diameter of the world-wide web. *Nature*, 401, 130–131.
- Albert, R. & Barabasi, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74, 47–97.
- Barabasi, A.-L., Albert, R. & Jeong, H. (1999). Mean-field theory for scale-free random networks. *Physica A* 272, 173–187.
- Brin, S., & Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Proceedings of the 7th International World Wide Web Conference*, obtido a 8/27/2007. Available at: <http://infolab.stanford.edu/~backrub/google.html>
- Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., & Wiener, J. (2000). Graph structure in the web. *Computer Networks*, 33, 309–320.
- Dorogovtsev, S. & Mendes, J. F. (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford: Oxford University Press.
- Erdős, P. & Renyi, A. (1959). On random graphs. *Publicationes Mathematicae*, 6, 290–297.
- Lessig, L (2001). *The Future of Ideas - The fate of the commons in a connected world*. New York: Random House.
- Machuco Rosa, A. (2003). *Internet - Uma História*. 2ª edição, revista, corrigida e ampliada, Lisboa: Edições Universitárias Lusófonas.

Machuco Rosa, A. (2007). Teoria dos meios clássicos de comunicação de massas *versus* novos meios. *Actas do V Congresso da SOPCOM*, Braga, 6-8 Setembro 2007.

Newman, M.E.J. (2003). The structure and function of complex networks. *SIAM Review*, 45, 167-256.

Ripenau, M. *et al* (2002). Mapping the gnutella network: Properties of large-scale peer-to-peer systems and implications for system design. *IEEE Internet Computing Journal*, 6, 50-57.

Starr, P. (2004). *The Creation of the Media: Political Origins of Modern Communications*. New York: Basic Books.

Vázquez, A. (2001). Growing networks with local rules: Preferential attachment, clustering hierarchy and degree correlations. arXiv:cond-mat /0211528.

Watts, D. & Strogatz, S. (1998). Collective dynamics of small-world networks. *Nature*, 393, 440-442.

Watts, D. J. (2003). *Six Degrees: The Science of a Connected Age*. New York: Norton.

Wolf, M. (1987). *Teorias da Comunicação*. Lisboa: Presença.