



REDE CI

Quem é Quem na
Ciência da Informação

Introdução à análise de redes

Os conceitos fundamentais da Análise de Redes Sociais (ARS) são: atores, atributos, laço relacional e relação. Para maiores referências ver: WASSERMAN, Stanley; FAUST, Katherine. ***Social Network Analysis: methods and applications***. Cambridge University Press. Structural analysis in social the social sciences series, v. 8, (1994) 1999. 857 p. ISBN 0-521-38707-8. As definições aqui apresentadas foram baseadas nessa publicação.

“Uma *rede social* (do inglês *social network*) consiste de um ou mais conjuntos finitos de atores [e eventos] e todas as relações definidas entre eles”. (WASSERMAN e FAUST)

Um *ator* em ARS é uma unidade discreta que pode de diferentes tipos: uma pessoa, ou um conjunto discreto de pessoas agregados em uma unidade social coletiva, como subgrupos, organizações e outras coletividades. Como se pode ver, o conceito de ator é flexível, permitindo diferentes níveis de agregação, o que permite sua adequação a diferentes problemas de pesquisa. Por exemplo, o ator pode ser tanto um operário quanto toda a empresa . Exemplos de atores são: um indivíduo, um pesquisador, um autor, uma empresa; agências de serviço público de uma cidade; estados-nações do mundo. Na Rede CI o ator é um autor / pesquisador.

Os *atributos* de um ator são suas características individuais. O foco da ARS são os laços relacionais, o que não impede que os atributos sejam analisados, mesmo que com auxílio de outros métodos que estão além da ARS. O conjunto de atributos de uma rede é denominado *composição* da rede social. Na Rede CI, embora ainda não façam parte da base, os atributos seriam: as características individuais (idade, sexo) e profissionais (instituição ao qual está vinculada, formação, etc.).

O *laço relacional* (do inglês *relational tie*) – também denominado simplesmente laço ou *ligação* (do inglês *linkage*) – é responsável por estabelecer a ligação entre pares de atores. Uma *relação* em uma rede (do inglês *relation*) define todo



REDE CI

Quem é Quem na
Ciência da Informação

o conjunto de laços que respeitam o mesmo critério de relacionamento, dado um conjunto de atores.

As relações têm duas propriedades importantes que devem ser consideradas nas pesquisas e que condicionam os métodos de análise de dados disponíveis, que são: direcionamento, podendo ser direcionais, caso no qual têm um ator como transmissor e outro como receptor, ou não-direcionais, caso no qual a relação é recíproca, como no caso da amizade, por exemplo; valoração, podendo ser dicotômicas, o que implica sua presença ou ausência, ou valoradas, com valores discretos ou contínuos.

No caso da Rede CI, os dados de co-autoria são não direcionais e podem ser tratados como dicotômicos ou valorados (número de publicações em co-autoria), dependendo do objetivo do estudo. As redes podem ser representadas na forma de grafos. Os básicos sobre grafos são os conceitos de caminhada, W , (do inglês *walk*), trilha (do inglês *trail*) e caminho (do inglês *path*). Uma *caminhada* é uma seqüência de nós e linhas, que começa e termina em nós, e na qual os nós são incidentes com as linhas que o seguem e o precedem na seqüência. Uma *trilha* é uma caminhada na qual todas as linhas são distintas. Um *caminho* é uma caminhada, no qual todos os nós e todas as linhas são distintos. Uma caminhada fechada (do inglês *closed walk*) é uma caminhada que começa e termina no mesmo nó. Um *ciclo* (do inglês *cycle*) é uma caminhada fechada, de três ou mais nós, no qual todas as linhas são distintas. Um *tour* é uma caminhada fechada na qual todas as linhas do grafo estão incluídas pelo menos uma vez.

A *distância geodésica*, $d(n_i, n_j)$, entre um par de nós é o número de laços que existe no caminho mais curto entre eles, ou caminho direcionado no caso de grafos direcionados, sendo que caso não exista tal caminho a distância pode ser considerada indefinida ou infinita. O *diâmetro* de um grafo é a maior distância geodésica entre todos os pares de nós presentes.



Algumas medidas obtidas da ARS

Um *grafo não direcionado* $G(N,L)$ (do inglês *undirected graph*) consiste de dois conjuntos de informações: um conjunto de *nós* (do inglês, *node*), representado por $N = \{n_1, n_2, \dots, n_g\}$ e um ou mais conjunto[s] de *linhas ou arestas* (do inglês *edge*), representadas por $L = \{l_1, l_2, \dots, l_L\}$ entre pares de nós. O número total de nós existente em um grafo é representado por **g** e o número total de linhas por **L**. A linha (não direcionada) responsável por ligar os nós n_i e n_j é representada por $l_k = (n_i, n_j)$, sendo que $l_k = l_q = (n_i, n_j) = (n_j, n_i)$.

Em um grafo (não direcionado), o grau (do inglês *nodal degree*), denotado por $d(n_i)$, é o número de linhas incidentes em um nó, ou ainda, de forma equivalente, o número de nós adjacentes a ele. O grau de um nó pode variar de **0**, caso no qual o nó é isolado, até **g - 1**, caso no qual o nó está em contato com todos os demais nós do grafo. No caso de um grafo direcionado, considera-se o grau de entrada (do inglês *indegree*), $d_i(n_i)$, e o grau de saída (*outdegree*), $d_o(n_i)$, de acordo com a direção dos arcos que chegam ou parte do nó. No caso da Rede CI, as 3 medidas são idênticas, uma vez que trata-se de grafos não direcionados.

Pode definir algumas medidas de importância de um nó em uma rede, como grau centralidade (*degree centrality*); grau de proximidade (*closeness centrality*) e grau de intermediação (*Betweenness centrality*). Assim, para cada participante são apresentadas algumas medidas de importância na ARS, como grau centralidade (*degree centrality*); grau de proximidade (*closeness centrality*) e grau de intermediação (*Betweenness centrality*). Na ARS, essas características estruturais mostram a posição de cada nó (pesquisador)

O *grau de centralidade para um ator* é dada por $C_D(n_i) = d(n_i)$; ou seja, é simplesmente o grau do nó. Tal medida pode ser normalizada, a fim de ter um



REDE CI

Quem é Quem na
Ciência da Informação

valor entre 0 e 1 e permitir a comparação entre atores de redes diferentes, dividindo-se o grau do nó pelo grau máximo que um nó pode ter, ou seja o número de nós no grafo menos 1 (o próprio nó), chegando-se a $C'_D(n_i) = \frac{d(n_i)}{g-1}$, sendo $0 \leq C'_D(n_i) \leq 1$.

O *grau de proximidade* de um ator mede o quanto o nó que representa o ator está próximo de todos os demais nós da rede. Para calcular a centralidade de proximidade se soma a distância geodésica do nó em relação a todos os demais nós do grafo e depois inverte-se, uma vez que quanto maior a distância menor a proximidade, chegando-se à seguinte fórmula $C_C(n_i) = [\sum_{j=1}^g d(n_i, n_j)]^{-1}$. A

normalização do índice, para o mesmo ficar na faixa de valores entre 0 e 1, o que pode ser feito multiplicando-se $C_C(n_i)$ por $g-1$, obtendo-se

$$C'_C(n_i) = \frac{(g-1)}{[\sum_{j=1}^g d(n_i, n_j)]} = (g-1)C_C(n_i).$$

O *grau de intermediação*, por seu turno, analisa o quanto um nó está no caminho geodésico entre outros nós. Seja g_{jk} o número de caminhos geodésicos (mais curtos e de mesmo tamanho) que ligam os nós j e k , e $g_{jk}(n_i)$ o número de tais caminhos, no total de g_{jk} , que passa pelo nó n_i . O índice de centralidade de

intermediação $C_B(n_i) = \sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}}$, que mede, para um nó n_i , a soma de

probabilidades de o mesmo estar no caminho geodésico entre todos os demais nós do grafo. Para normalizar tal índice, deve-se dividi-lo pelo seu máximo possível, que é o número de pares de nós no grafo que não incluem n_i , ou seja $(g-1)(g-2)/2$, chegando-se ao seguinte índice de centralidade de

$$C'_B(n_i) = \frac{C_B(n_i)}{(g-1)(g-2)/2}.$$



A ARS utiliza softwares para visualização de grafos a fim de se obter imagens das redes sociais. Geralmente no caso dos softwares disponíveis para ARS, os dados são fornecidos através de matrizes mas podem ser exibidos através grafos. A visualização gráfica por si só pode oferecer novas informações e *insights* para pesquisadores.

(Baseado em MATHEUS, Renato Fabiano; SILVA, Antonio Braz de Oliveira e. **Fundamentação teórica para a análise de redes com ênfase na Análise de Redes Sociais**: ARS, Belo Horizonte, 2005. Não publicado.)



Análise de redes de co-autoria

A estrutura da base de dados da Rede CI, já na fase atual, permite o estudo da colaboração entre os pesquisadores na forma de redes de co-autoria. No entanto, recomenda-se a leitura de alguns artigos sobre o assunto, uma vez que "colaboração" e "co-autoria" são termos com conteúdos distintos, isto é, nem toda colaboração resulta em uma publicação em co-autoria e nem toda publicação em co-autoria é fruto de uma colaboração. Sobre esse assunto, ver:

KATZ, J. Sylvan; MARTIN, Ben R. What is Research Collaboration? Research Policy, v. 26, p. 1-18, 1997. Disponível em <<http://www.rvm.gatech.edu/bozeman/rp/read/41407.pdf>>.

Outros artigos já utilizados nos trabalhos do Netic sobre redes sociais foram:

Sobre redes de co-autoria:

NEWMAN, M. E. J. From the Cover: The structure of scientific collaboration networks. Proceedings of The National Academy of Sciences of USA (PNAS), [on line], v. 98, p. 404-409, jan. 2001. Disponível em: <<http://www.pnas.org/cgi/reprint/98/2/404>>.

KRETSCHMER, Hildrun. Author productivity and geodesic distance in bibliographic co-authorship networks, and visibility on the Web. Scientometrics, [on line], v. 60, n. 3, p. 409-420, jan. 2004. Disponível em: <<http://www.springerlink.com/app/home/contribution.asp?wasp=463ed3eb2dfb4aa89ed6796b917f7991&referrer=parent&backto=issue,13,22;journal,12,53;linkingpublicationresults,1:101080,1>>.

MAHLCK, Paula; PERSSON, Olle. Socio-bibliometric mapping of intra-departmental networks. **Scientometrics**, [on line], v. 49, n. 1, p. 81-91, 2000. Disponível em: <<http://www.springerlink.com/app/home/contribution.asp?wasp=463ed3e>



REDE CI

Quem é Quem na
Ciência da Informação

[b2dfb4aa89ed6796b917f7991&referrer=parent&backto=issue,5,9;journal,47,53;linkingpublicationresults,1:101080,1](http://www.redesocial.org.br/revista/index.php?option=com_content&view=article&id=125&Itemid=125&lang=pt-br&ref=parent&backto=issue,5,9;journal,47,53;linkingpublicationresults,1:101080,1)>.

MOLINA, José Luis; MUÑOZ, Juan Manuel; DOMENECH, Miquel. Redes de publicaciones científicas: un análisis de la estructura de coautorías. **Revista Hispana para el análisis de redes sociais**, v. 1, n. 3, jan. 2002. Disponível em <<http://revista-redes.rediris.es/>>.

YOSHIKANE, Fuyuki, KAGEURA, Kyo. Comparative analysis of coauthorship networks of different domains: The growth and change of networks. *Scientometrics*, [on line], v. 60, n. 3, p. 435-446, jan. 2004. Disponível em:

<<http://www.springerlink.com/app/home/contribution.asp?wasp=463ed3eb2dfb4aa89ed6796b917f7991&referrer=parent&backto=issue,15,22;journal,12,53;linkingpublicationresults,1:101080,1>>.

Sobre análise de redes na Ciência da Informação

OTTE, Evellen; ROUSSEAU, Ronald. Social network analysis: a powerful strategy, also for information sciences. *Journal of Information Science*, Thousand Oaks, v. 28, n. 6, p. 441-453, 2002. (somente o *Abstract* em <<http://jis.sagepub.com/cgi/content/abstract/28/6/441>>

MARTELETO, Regina Maria. Análise de redes sociais: aplicação nos estudos de transferência da informação. *Ciência da Informação*, Brasília, v. 30, n. 1, p. 71-81, 2001. Disponível em <<http://www.ibict.br/cienciadainformacao/include/getdoc.php?id=553&article=261&mode=pdf>>

MARTELETO, Regina Maria; SILVA, Antonio Braz de O. Redes e Capital Social: o enfoque da informação para o desenvolvimento local. *Ciência da Informação*, Brasília, v. 33, n. 3, p. 41-49, 2004. Disponível em <<http://www.ibict.br/cienciadainformacao/include/getdoc.php?id=1264&article=563&mode=pdf>>