

Redes de colaboração: análise da coautoria em comunidades científicas disciplinares

Antonio Raphael Brito Meireles¹, Ingrid dos Santos Vilas Verde¹, Maria Teresinha Tamanini Andrade¹, Hernane Borges de Barros Pereira^{2,3}

¹Instituto de Educação Ciência e Tecnologia da Bahia – Salvador – BA - Brazil

²Programa de Modelagem Computacional – SENAI CIMATEC – Salvador – BA

³Universidade Estadual da Bahia (UNEB) – Salvador, BA – Brazil

{vilas.ingrid, hernanebbpereira}@gmail.com.br,
raphael_phsu@hotmail.com, tamanini@ifba.edu.br

Abstract. *In this article was analyzed the collaboration in disciplinary scientific communities. For that it was used the Network Theory, studied the co-authorship networks of researchers and identified the associated network model and the implications of this model in the communities that belong researchers. The results indicate networks with small world topology and central researchers in networks.*

Resumo. *Neste artigo analisou-se a colaboração em comunidades científicas disciplinares. Para isso foi utilizada a Teoria de Redes (i. e. Teoria de Redes Complexas e Análise de Redes Sociais). Foram estudadas as redes de coautoria formadas por pesquisadores identificando o modelo de rede associado e as implicações desse modelo nas comunidades às quais pertencem os pesquisadores. Os resultados encontrados indicam redes com topologia small world e pesquisadores centrais nas redes.*

1. Introdução

O tema colaboração científica vem sendo discutido e analisado desde o final da década de 1950. Segundo Ziman (1979), a ciência é uma atividade eminentemente social, e, dentro desse contexto, o contato informal entre cientistas, as conversas com os colegas, o trânsito em conferências e seminários e em outras universidades compõem o ambiente onde ela acontece.

Na opinião de Sonnenwald (2006), a colaboração científica pode ser definida como a interação que facilita não só a realização de tarefas, mas, também, o compartilhamento do significado desta tarefa, relacionada a um objetivo maior compartilhado entre dois ou mais cientistas. Segundo a autora, a colaboração ocorre dentro do contexto social da ciência, que inclui elementos como a revisão por pares, sistemas de prêmios, colégios invisíveis, paradigmas científicos, políticas de ciência nacionais e internacionais e, é claro, as normas implícitas ao campo disciplinar e às instituições de pesquisa e/ou universidades.

Na literatura existem vários trabalhos (ANDRADE et al 2014; DIGIAMPIETRI e MARUYAMA, 2014; ROSA et al, 2014; Vanz e Stump, 2010) que estudam o tema redes de colaboração científica e coautoria científica.

O objetivo desta pesquisa é estudar as redes de colaboração, segundo o critério de coautoria, entre pesquisadores de Programas de Pós-Graduação (PPG) enquanto parâmetro para compreender como ocorre a colaboração no contexto acadêmico (em comunidades científicas), de modo a identificar o modelo de rede associado e as implicações desse modelo nas comunidades às que pertencem os pesquisadores.

Este artigo está organizado da seguinte forma: além desta breve introdução na Seção 2 apresentamos a fundamentação teórica; na Seção 3 descrevemos a metodologia; na Seção 4 analisamos os resultados preliminares da pesquisa e na Seção 5 tecemos algumas considerações finais.

2. Fundamentação Teórica

O estudo de redes tem suas origens na teoria dos grafos, um ramo da matemática. Assim, uma rede é um grafo constituído de um conjunto de elementos chamados vértices ou nós, que são ligados por outro conjunto de elementos chamados de arestas que fazem conexões com os vértices (NEWMAN, 2003).

Uma rede social é um conjunto de pessoas ou grupos de pessoas com algum padrão de contato ou interação entre elas (WASSERMAN e FAUST, 1997). Nessa pesquisa, para discutir e caracterizar as relações de colaboração serão utilizadas três medidas de centralidade comumente aplicadas em estudos de Análise de Redes Sociais (ARS): centralidade de grau, centralidade de proximidade e centralidade de intermediação.

A centralidade de grau é definida pelo número de laços adjacentes que um vértice possui com outros em uma rede. A centralidade de grau foca a importância de um ator nas simples conexões que este estabelece com os atores vizinhos, e é quantificada pelo grau do vértice. Assim, um vértice na rede é mais importante que outro se estabelece um maior número de vínculos com os vértices vizinhos.

Já a centralidade de proximidade é função da maior ou menor distância de um vértice em relação a todos outros em uma rede. A ideia é que um vértice central é aquele que possui maiores condições de interagir rapidamente com todos os outros (SCOTT, 2002; HANNEMAN e RIDDLE, 2005). A centralidade de proximidade de um ator é baseada na proximidade ou na distância. Enquanto a centralidade de grau é medida para os atores adjacentes a um determinado ator, a centralidade de proximidade mostra o quanto um ator está próximo de todos os outros da rede.

A centralidade de intermediação avalia a dependência de vértices não adjacentes de outros que atuam como uma espécie de ponte para a efetivação da interação entre eles (FREEMAN, 1979).

O termo redes complexas refere-se a um grafo que apresenta uma estrutura ou propriedade topológica não trivial, composto por um conjunto de vértices que são interligados por meio de arestas (BARABÁSI, 2003). Essa estrutura não segue um

padrão regular. Os índices comumente usados são o caminho mínimo médio L , o coeficiente de aglomeração C e a curva de distribuição de graus $P(k)$.

Com os índices da teoria das redes complexas é possível caracterizar uma rede como aleatória, livre de escala ou *small world*, que são os modelos mais difundidos.

Uma rede apresenta o comportamento *small world* se seu coeficiente de aglomeração médio for muito maior que o coeficiente de aglomeração de uma rede aleatória e se seu caminho mínimo médio for comparável com o caminho mínimo médio da rede aleatória correspondente. O caminho mínimo é o menor caminho entre dois vértices. O coeficiente de aglomeração, também conhecido como fenômeno de transitividade, ocorre quando um vértice A está conectado a um vértice B, e o vértice B está conectado a um vértice C, aumentando as chances do vértice A também estar conectado ao vértice C. Em outras palavras, a transitividade indica a presença de um número elevado de tríades na rede, ou seja, conjuntos de três vértices conectados uns aos outros.

3. Metodologia

Para realizar a pesquisa proposta, foi selecionado como objeto de pesquisa um Programa de Pós-Graduação (PPG) em Ciência da Informação, na área disciplinar, iniciado em 2001. Foram usados os cadernos de indicadores de Produção Bibliográfica da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), obtidos no formato PDF. Esses cadernos são preparados com base nos dados que os PPGs informam todos os anos e são publicados pela CAPES.

Especificamente foram analisados os pesquisadores e suas produções bibliográficas (artigos em periódicos, trabalhos em anais, livros e capítulos de livro). A partir destes dados foram analisados os pesquisadores e suas produções bibliográficas por meio de redes de coautoria. O *locus* da pesquisa é o PPG selecionado e os sujeitos da pesquisa são os pesquisadores (docentes, discentes e participantes externos) deste programa, considerando o período das avaliações trienais e relatórios da Coleta CAPES. Cabe observar que o período escolhido para análise foi delimitado a partir do início das atividades do Comitê Interdisciplinar da CAPES, neste caso, o ano de 2001 até 2012 (quatro triênios).

Posteriormente foi utilizado o software Ucinet para gerar redes no formato Pajek e por meio destas, aplicar a Teoria das Redes Complexas e a Análise de Redes Sociais (ARS). A partir das redes do PPG foram calculados, com o auxílio dos programas de computador Ucinet e Pajek, os índices e parâmetros das redes sociais e complexas.

Na Seção 4 serão apresentados os resultados das análises encontrados ao longo do período analisado (2001 a 2012).

4. Discussão dos Resultados

Na Figura 1 apresentamos a rede total da Produção Bibliográfica. Os vértices são os pesquisadores e as arestas são as colaborações em coautoria. Os pesquisadores são coautores se publicaram juntos.

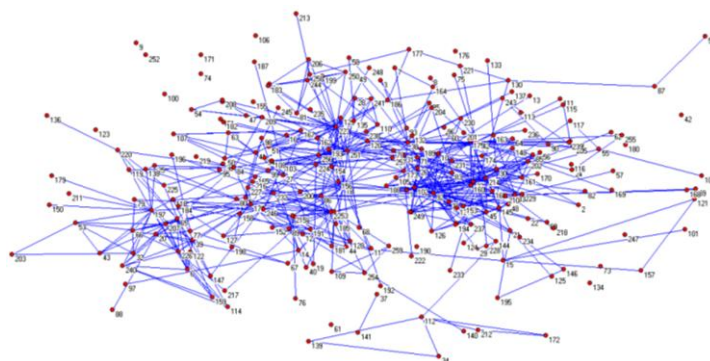


Figura 1. Rede total da Produção Bibliográfica - 2001 a 2012.
Fonte: Elaboração própria a partir de dados da CAPES.

A Tabela 1 apresenta os índices e parâmetros da Teoria de Redes das redes estudadas para a Produção Bibliográfica do PPG. O Cr é o coeficiente de aglomeração da rede aleatória e o Lr o caminho mínimo médio da rede aleatória.

Tabela 1. Índices de redes complexas

ÍNDICES	TRIÊNIO 1	TRIÊNIO 2	TRIÊNIO 3	TRIÊNIO 4	PERÍODO TOTAL
Vértices	39	82	115	136	259
Componentes	13	21	19	10	29
Densidade	0.0526	0.0367	0.0364	0.0269	0.0164
Diâmetro	3	6	8	9	9
Coefficiente de Aglomeração	0.7006	0.7866	0.7620	0.7355	0.7473
Cr	0.0833	0.0178	0.0306	0.0351	0.0181
Caminho Mínimo Médio	1.5783	2.7900	3.0531	4.1095	3.8191
Lr	3.7257	3.6583	3.6015	3.6389	3.9122
Grau Médio	2.0000	2.9756	4.1565	3.6323	4.2471

A Tabela 2 a seguir apresenta os índices obtidos para as centralidades de grau (C_G), de proximidade (C_P) e de intermediação (C_I) dos pesquisadores do PPG. Os

índices abaixo foram calculados para o período total. Os pesquisadores são apresentados com a letra P (pesquisador) seguida do número do seu vértice correspondente.

Tabela 2. Índices de centralidades: Centralidade de Grau, Proximidade e Intermediação. Classificação: D (docente); O (outros: participante externo, examinador externo, coautor, egresso, discente-autor)

Pesquisador	Classificação	C_G	Pesquisador	Classificação	C_P	Pesquisador	Classificação	C_I
P92	D	162	P18	O	1.88	P92	D	17.24
P102	O	116	P92	D	1.88	P253	D	16.27
P214	D	97	P253	D	1.87	P18	O	12.13
P23	O	82	P102	O	1.87	P38	D	10.84
P223	D	75	P173	O	1.87	P178	O	10.56
P253	D	73	P71	O	1.87	P223	D	8.92
P18	O	53	P86	D	1.87	P251	D	7.83
P71	O	51	P251	D	1.87	P142	D	7.03
P10	O	42	P10	O	1.86	P86	D	6.32
P142	D	41	P23	O	1.86	P173	O	5.50

A seguir apresentamos algumas discussões para os resultados das redes dos quatro triênios analisados. Abaixo a rede do triênio 1 (Figura 2).

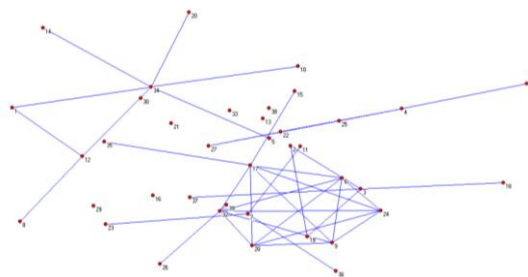


Figura 2. Rede do triênio 1 – 2001 a 2003.

Fonte: Elaboração própria a partir de dados da CAPES.

O triênio 1 é composto por 39 vértices. A partir da Figura 2 é possível observar que a rede apresenta uma conexão isolada entre grupos de vértices. Estes grupos apresentam díades (ligação direta entre dois atores) e tríades (ligação direta ou indireta entre três atores) havendo casos de vários vértices ligados direta ou indiretamente através de um centro transmissor (casos de centralização) onde um nó centralizado recebe conexões dos demais atores. Esse vértice pode servir como intensificador e incentivador de atores distantes ou alterar e interromper a informação entre os vértices da rede.

A densidade, que é a razão entre o número de arestas da rede e o número possível de arestas na mesma, é baixa (0.0526) indicando uma rede de colaboração com

baixo grau de conectividade. O pequeno diâmetro (maior caminho mínimo) 3, revela que há pouca interação na rede, pois é uma rede desconectada.

A rede construída para o triênio 2 (Figura 3) possui um total de 82 vértices. Neste triênio há a presença de um componente maior que comporta a maior parte dos vértices estabelecendo conexões entre si, indicando que houve um avanço na transferência de informações científicas entre os pesquisadores de um triênio a outro. Ainda assim, percebe-se a presença de subgrupos de díades, tríades e atores periféricos que não estabelecem conexões com outros atores. Observam-se também vários grupos de vértices que possuem um centro específico, não estando conectados a um centro em comum e estão interligados entre si.

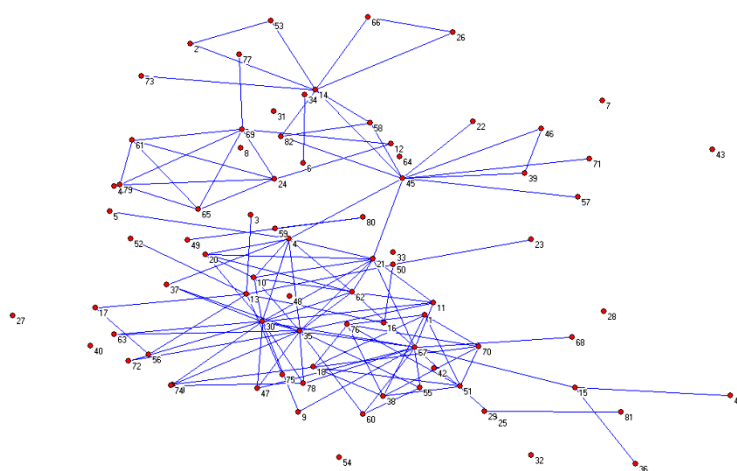


Figura 3. Rede do triênio 2 – 2004 a 2006.
Fonte: Elaboração própria a partir de dados da CAPES.

A densidade baixa (0.0367) indica uma rede esparsa. Houve um pequeno aumento no diâmetro, 6, quando comparado ao diâmetro do triênio 1, indicando que houve um discreto crescimento na interação entre os vértices.

A rede construída para o triênio 3 (Figura 4) possui um total de 115 vértices. É possível perceber que houve um aumento no número de vértices neste período, assim como vem ocorrendo nos triênios anteriores. A rede é conectada, com a presença de apenas um ator periférico que não estabelece conexões na rede. É uma rede descentralizada quando comparada com as redes do triênio 1 e 2 e bem distribuída. A densidade baixa (0.0364) demonstra que a rede continuou se expandindo, a elevação do diâmetro para 8, revela evolução das conexões entre os vértices, conforme vemos em sua estrutura conectada.

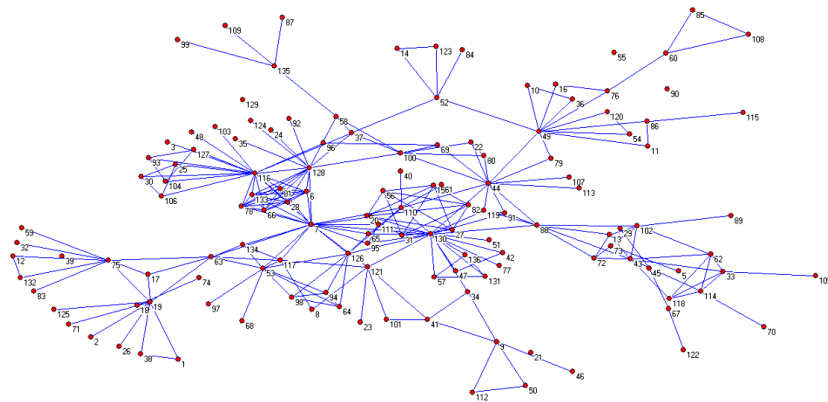


Figura 4. Rede do triênio 3 – 2007 a 2009.
Fonte: Elaboração própria a partir de dados da CAPES.

A rede do triênio 4 (Figura 5) contém um total de 136 vértices, o que mostra uma expansão da rede a cada triênio, devido à inserção de novos vértices. Assim como a rede do triênio 3, a rede do triênio 4 possui conexões similares entre os nós, ou seja, não há hierarquia de um determinado vértice, desta forma, a maioria dos vértices de um grupo não possui um centro de referência, sendo bastante distribuída e conectada. A densidade baixa (0.0269) demonstra que a rede continuou se expandindo. O diâmetro se manteve em crescimento, de 8 para 9, revelando que a interação e os processos de colaboração continuaram fluindo de um triênio a outro.

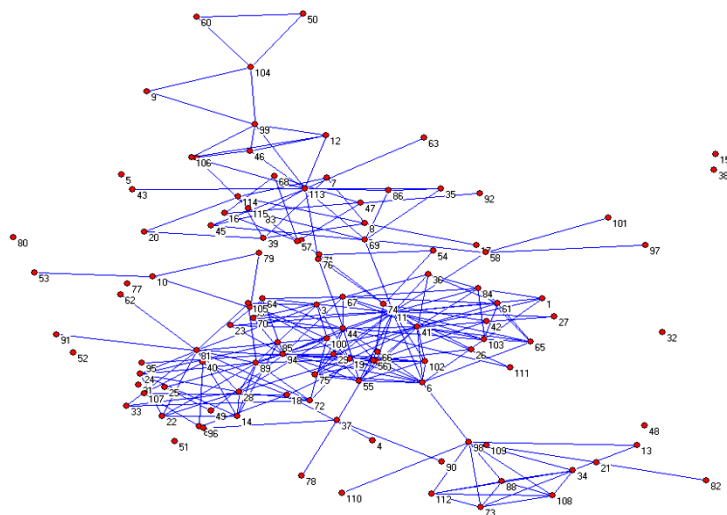


Figura 5. Rede do triênio 4 – 2010 a 2012.
Fonte: Elaboração própria a partir de dados da CAPES.

A rede total (2001 a 2012) da Produção Bibliográfica é composta por 259 vértices. O gráfico desta rede (Figura 1) demonstra que existe uma estrutura de conexão e conectividade que tende a se manter ao longo dos últimos triênios nas redes deste PPG, ou seja, é uma rede conectada. A densidade baixa (0.0164) demonstra que a rede evoluiu bastante, em termos de conexões entre os vértices como consequência do acréscimo de novos componentes ao longo dos anos. O diâmetro (9) indica que de certo modo nas redes de Produção Bibliográfica ocorre interação entre os vértices e os processos de colaboração fluem. Essa rede conectada revela parcerias que se repetem ao longo dos anos, pois um vértice bem conectado tende a fortalecer suas conexões existentes já que ele é o detentor das informações. O fato de ela ser conectada aumenta as chances de interações futuras entre os vértices.

Os dados da Tabela 2 demonstram que dos dez pesquisadores que possuem destaque nos índices de centralidade, 50% são docentes para centralidade de grau, 40% são docentes para centralidade de proximidade e 70% são docentes para centralidade de intermediação. O grupo de vértices dominantes nos triênios analisados é formado em sua maioria por docentes, porém nota-se que existem outros vértices (como por exemplo, participantes externos) que tentam, ao longo dos anos, obterem posições centrais na rede de coautoria do PPG analisado, participando e contribuindo para a evolução da rede e com os processos de difusão do conhecimento. É possível perceber que nas redes de Produção Bibliográfica existe uma interação entre os indivíduos.

5. Considerações Finais

Os resultados preliminares obtidos dos índices da Teoria das Redes indicam que as redes de Produção Bibliográfica do PPG analisado, caracterizam-se topologicamente como redes small-world (maior articulação), com coeficiente de aglomeração médio muito maior que o coeficiente de aglomeração de uma rede aleatória (Cr) e o caminho mínimo médio comparável com o caminho mínimo médio da rede aleatória (Lr) correspondente.

Nos primeiros lugares das tabelas de centralidades, estão os pesquisadores considerados relevantes em termos de publicação. Supõe-se que esses pesquisadores com maior interação, trabalham com grupos de pesquisas e têm grande número de colaboradores, conseguindo assim manter seu nível de produção científica de um período para o outro.

De uma maneira global faremos alguns comentários sobre os vértices e as centralidades do PPG analisado para a Produção Bibliográfica, pois constatamos que existem pesquisadores que se destacam em todos os cadernos de indicadores ao longo do período, mas também existem pesquisadores que se destacam em certo período, mas não participaram de outros triênios.

Por exemplo, o vértice 92 se destaca no triênio 1 (posição 1 em centralidade de grau e proximidade e posição 2 em centralidade de intermediação); no triênio 2 (posição 2 em centralidade de grau, 4 em proximidade e posição 7 em centralidade de intermediação); no triênio 3 (posição 1 em centralidade de grau, 2 em proximidade e posição 7 em centralidade de intermediação) e no triênio 4 (posição 5 em centralidade de grau e posição 4 em centralidade de intermediação). Este pesquisador está presente

ao longo do período, ou seja, aparece em todos os triênios entre os dez mais preeminentes.

Já o pesquisador 18 aparece somente no triênio 4 (posição 2 em centralidade de grau, posição 1 em centralidade de proximidade e intermediação) mas não aparece nos outros triênios.

Em síntese, identificamos preliminarmente a estrutura topológica das redes associadas a esta comunidade científica, revelando vértices centrais, que podem estar em contato direto e adjacente com muitos outros vértices, sendo reconhecido pelos outros como um canal de informações.

Referências Bibliográficas

- Andrade, M. T. T., Braga, P. F., Carneiro, T. K. G., Ribeiro, N. M., Moret, M. A., Pereira, H. B. B. (2014), “Contextualized Analysis of Social Networks: Collaboration in Scientific Communities”, *Journal Social Networking*, v. 3, p. 71-79.
- Barabási, A.L. (2003), *Linked: How everything is connected to everything else and what it means for business, science and everyday life*, Plume.
- Digiampietri, L. e Maruyama, W. (2014), “Predição de Novas Coautorias na Rede Social Acadêmica dos Programas Brasileiros de Pós-Graduação em Ciência da Computação”, CSBC, BraSNAM.
- Freeman, L. C. (1979), Centrality in social networks: I conceitual clarification, *Social Networks*, v. 1, p. 215–239.
- Hanneman, R. A. e Riddle, M. (2005), “Introducion to Social Network Methods”, <http://www.faculty.ucr.edu/hanneman>, out.
- Newman, M.E.J. (2003), The structure and function of complex networks, *SIAM Review*. Vol. 45(2): 167–256.
- Rosa, M. G., Fadigas, I. S., Andrade, M. T. T., Pereira, H. B. B. (2014) “Clique Approach for Networks: Applications for Coauthorship Networks Social Networking”, v. 3, p. 80-85.
- Scott, J. (2002), *Social network analysis*, Sage.
- Sonnenwald, D. H. (2006), “Scientific Collaboration”, In: *Annual Review of Information Science and Technology*, Medford, NJ: Information Today, ed. B. Cronin, v. 41, p. 643–681.
- Wasserman, S. e Faust, K. (1997), *Social network analysis: methods and applications*. [reprinted with corrections]. Cambridge: CambridgeUniversity Press.
- Vanz, S. A. S., Stumpf, I. R. C. (2010), “Colaboração científica: revisão teórico conceitual”, *Perspectivas em Ciência da Informação*, v. 15, n. 2, p. 42–55.
- Ziman, J. M. (1979), *Conhecimento público*, Belo Horizonte: Itatiaia.