

Importância das Colaborações Interdisciplinares nas Redes de Coautoria Científica

Geraldo J. Pessoa Junior¹, Thiago M. R. Dias²,
Thiago H. P. Silva¹, Alberto H. F. Laender¹

¹Universidade Federal de Minas Gerais 31270-901 Belo Horizonte, MG, Brasil

²Centro Federal de Educação Tecnológica 35503-822 Divinópolis, MG, Brasil

{geraldopessoa, thps, laender}@dcc.ufmg.br, thiago@div.cefetmg.br

Abstract. *How important are interdisciplinary collaborations in scientific co-authorship networks? This is a relevant question that has drawn the attention of scholars, since bridging academic relationships contributes to make scientific networks stronger. However, traditional studies have focused on characterizing specific groups rather than on studying a complete and robust scientific community. In this paper, we characterize such relationships by considering the eight major areas of the entire Brazilian scientific coauthorship network. Our results show that the Brazilian coauthorship network grew and became remarkably interdisciplinary.*

Resumo. *Quão importantes são as colaborações interdisciplinares nas redes de coautoria científica? Esta é uma questão relevante que tem chamado a atenção da comunidade acadêmica, já que o estreitamento das relações entre pesquisadores contribui para tornar as redes científicas mais fortes. Entretanto, trabalhos tradicionais têm se concentrado na caracterização de grupos específicos e não no estudo de uma comunidade científica completa e robusta. Neste artigo, caracterizamos essas relações considerando as oito grandes áreas de toda a rede de coautoria científica brasileira. Nossos resultados mostram que a rede de coautoria brasileira cresceu e tornou-se especialmente interdisciplinar.*

1. Introdução

Quão importantes são as colaborações interdisciplinares nas redes de coautoria científica? Esta é uma questão importante e ampla que tem chamado a atenção da comunidade acadêmica, uma vez que o estreitamento das relações entre pesquisadores contribui para tornar as redes de colaboração científica mais fortes. Esforços recentes têm focado nas colaborações externas (por exemplo, cooperação entre grupos de pesquisa, migrações de pesquisadores e influência de áreas de pesquisa) como sendo fatores relevantes na evolução das comunidades científicas [Kato and Ando 2013, Mena-Chalco et al. 2014, Mooney et al. 2013]. No entanto, tais estudos concentram-se principalmente na caracterização de áreas ou grupos de pesquisa específicos (e.g., [Kato and Ando 2013]) em vez de contemplar a análise de uma comunidade completa.

Assim, ao invés de analisarmos cenários particulares de grupos fechados (por exemplo, um determinado departamento acadêmico ou uma área de pesquisa específica), neste artigo avançamos um passo adiante construindo as redes de coautoria global e interdisciplinar de todo o Brasil, usando como fonte de dados a Plataforma Lattes¹. Mantida

¹Lattes Platform: <http://lattes.cnpq.br>

pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a Plataforma Lattes é uma iniciativa de renome internacional [Lane 2010] que fornece um repositório de currículos vitae e grupos de pesquisa integrados em um único sistema. Uma vez que todos os pesquisadores que atuam no Brasil são obrigados a manter seus currículos atualizados nessa plataforma, ela fornece um grande volume de informação sobre as atividades de pesquisa e a produção científica desses indivíduos.

Neste artigo analisamos a importância das colaborações interdisciplinares usando como referência o nível superior do esquema de classificação das áreas do conhecimento proposto pelo CNPq², que considera as seguintes oito grandes áreas: Ciências Agrárias, Ciências Biológicas, Ciências da Saúde, Ciências Exatas e da Terra, Ciências Humanas, Ciências Aplicadas e Sociais, Engenharias, e Linguística, Letras e Artes. Note que este esforço envolve a rede completa de coautoria acadêmica brasileira, compreendendo todos os indivíduos com doutorado (264.731 em junho de 2017) ora atuando em atividades de pesquisa no país. Como veremos, 37,4% de todas as coautorias nessa rede são interdisciplinares, ou seja, envolvem pesquisadores de diferentes grandes áreas, o que enfatiza a importância das colaborações interdisciplinares no cenário atual de ciência e tecnologia.

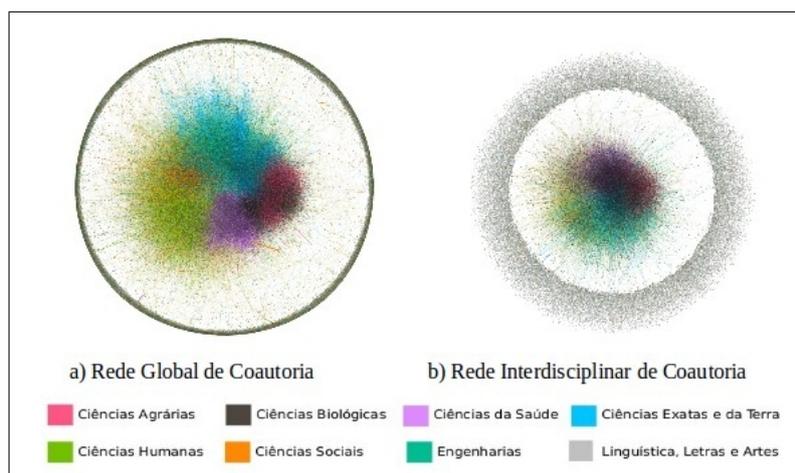


Figura 1. Redes Global e Interdisciplinar de Coautoria no Brasil

Para ilustrar nossa alegação, a Figura 1 revela o volume e a complexidade de tais colaborações nas diferentes comunidades científicas. O grafo à esquerda mostra a rede de coautoria global brasileira incluindo as colaborações dentro de cada grande área, enquanto que o grafo à direita mostra apenas as colaborações interdisciplinares. Como podemos observar, a rede global é mais densa com fortes conexões entre os nós do componente gigante, enquanto que a rede interdisciplinar possui um conjunto maior de componentes isolados, tendo em vista que ao serem removidas as conexões não interdisciplinares ela tornou-se menos densa, aumentando a quantidade de nós isolados que podem ser observados nas suas extremidades. Embora as grandes áreas estejam bem definidas quando se observa mais de perto a rede global, nota-se claramente um componente conectado bem distinto na rede interdisciplinar. Particularmente, existe uma forte sobreposição das grandes áreas, representada por uma mistura maior de cores na proximidade de cada uma

²Este esquema de classificação é organizado nos seguintes quatro níveis (veja <http://bit.ly/1JM2j1k>): *grande área* (e.g., Ciências Exatas e da Terra), *área* (por exemplo, Ciência da Computação), *subárea* (e.g., Teoria da Computação) e *especialidade* (e.g., Linguagens Formais e Autômatos).

delas, sendo que cada cor corresponde a uma grande área.

Conforme mencionado anteriormente, vários estudos tendem a se concentrar apenas no aspecto global, dando a mesma importância para as arestas que conectam partes específicas de uma rede. Em contraste, neste artigo enfatizamos a importância de minerar mais profundamente padrões interdisciplinares para explicar a força dessas ligações internas, isto é, possibilitar uma melhor compreensão da evolução da ciência [Mooney et al. 2013]. Neste contexto, este artigo tem como objetivo apresentar uma análise preliminar da importância das colaborações interdisciplinares na formação de redes de coautoria, tendo como escopo a comunidade científica brasileira.

O restante deste artigo está organizado da seguinte forma. A Seção 2 descreve como foi construído o conjunto de dados usado neste trabalho, a Seção 3 apresenta os resultados análise realizada e a Seção 4 apresenta algumas considerações finais.

2. Metodologia

Para gerar os dados necessários para a nossa análise, utilizamos um conjunto de currículos vitae (versões XML) de 264.731 pesquisadores brasileiros com título de doutor, coletado da Plataforma Lattes em junho de 2017. Posteriormente, os respectivos documentos XML foram transformados em arquivos CSV contendo apenas os campos necessários para identificar as colaborações científicas entre os pesquisadores, reduzindo assim a quantidade de dados a serem tratados. Além disso, também foram desconsiderados todos os currículos cujos pesquisadores não haviam indicado a grande área mais próxima de suas atividades de pesquisa, reduzindo assim o número total de currículos para 245.583.

Para construção e caracterização das redes necessárias ao nosso estudo, aplicamos a estratégia proposta por Dias & Moita [2015] para identificar as colaborações por comparação dos títulos das publicações encontradas nos currículos. Para fins das análises realizadas, foram consideradas apenas as coautorias identificadas a partir de artigos de periódicos, uma vez que este tipo de publicação é o mais comum na maioria das áreas de pesquisa. A distribuição final do número total de pesquisadores por grande área no conjunto de dados considerado é: Ciências Agrárias 24.957, Ciências Biológicas 33.756, Ciências da Saúde 42.392, Ciências Exatas e da Terra 36.752, Ciências Humanas 41.969, Ciências Sociais e Aplicadas 27.137, Engenharias 23.344, Linguística, Letras e Artes 15.276. Como pode ser visto, o número total de pesquisadores em cada grande área é bastante desigual e basicamente reflete a popularidade de suas respectivas áreas de atuação.

3. Análise das Redes

Nesta seção, analisamos as propriedades das redes brasileiras de coautoria global e interdisciplinar construídas a partir de dados da Plataforma Lattes. Além disso, descrevemos como essas redes estão interligadas, mostrando a proporção de colaborações em cada grande área. Por rede global, entende-se aquela cujos vértices correspondem aos 245.583 pesquisadores considerados e as arestas indicam as suas coautorias. Já a rede interdisciplinar possui o mesmo conjunto de vértices, sendo descartadas as arestas que correspondem a coautorias dentro da própria grande área dos pesquisadores. A Tabela 1 mostra os valores de algumas métricas usuais referentes às redes global e interdisciplinar. Primeiro, como se pode observar, 37,4% (758.392 de 2.027.511) das arestas correspondem a colaborações interdisciplinares. Este número é muito expressivo, já que se refere à rede de

uma comunidade científica completa. Além disso, o fato de que uma grande quantidade de arestas do componente gigante³ da rede global ser mantida na rede interdisciplinar é bastante relevante, uma vez que 95,5% dessas arestas pertencem ao componente gigante. Isso reforça o quão próximas e importantes são tais colaborações no contexto acadêmico.

Tabela 1. Comparação entre as Redes Global e Interdisciplinar

Métrica	Global	Interdisciplinar
Número de Vértices	245.583	245.583
Número de Arestas	2.027.511	758.392
Grau Médio	9,9	5,7
Tamanho do Componente Gigante	194.021	145.255
Percentual de Vértices no Componente Gigante	73,3%	54,9%
Arestas no Componente Gigante	2.024.152	754.635
Densidade da Rede	5,79E-5	2,16E-5
Diâmetro da Rede	15	16
Caminho Mínimo Médio	5,4	6,7
Total de Componentes Isolados	65.362	113.509

Como esperado, o número de arestas e o grau médio dos nós diminuíram na rede interdisciplinar, enquanto que o número total de componentes isolados aumentou. Embora os diâmetros e os tamanhos do caminho mínimo médio das duas redes sejam próximos, nota-se que tais valores são ligeiramente maiores na rede interdisciplinar que possui menos arestas, tornando-a mais dispersa. Apesar disso, os tamanhos do caminho mínimo médio de ambas as redes são compatíveis com a propriedade do *mundo pequeno* [Milgram 1967]. Quanto à densidade das redes (a relação entre o número de arestas existentes e o número de arestas possíveis), a rede interdisciplinar tende a ser menos densa, uma vez que somente as colaborações interdisciplinares são consideradas.

Na literatura já existem alguns estudos acerca da topologia das redes de cada uma das grandes áreas definidas pelo CNPq. Em particular, Mena-Chalco et al. [2014] observaram padrões diferentes em todas elas, comprovando a hipótese por eles levantada que toda a rede brasileira de coautoria científica apresenta uma interdisciplinaridade natural. Por outro lado, Vanz & Stumpf [2012] também mostraram que as principais áreas de pesquisa no Brasil apresentam diferentes padrões de coautoria. Tais evidências mostram a necessidade de uma análise mais profunda dessas colaborações interdisciplinares, particularmente para entender a motivação dos pesquisadores envolvidos [Iglič et al. 2017].

Em particular, investigamos a proporção de colaborações interdisciplinares entre as oito grandes áreas. A Figura 2 apresenta o percentual de colaborações interdisciplinares para cada grande área. Por exemplo, Ciências Agrárias (primeiro gráfico) apresenta um percentual de colaborações com as demais grandes áreas de 58,38% (Ciências Biológicas), 11,62% (Ciências da Saúde) e assim por diante. Observe que as colaborações interdisciplinares mais relevantes envolvem as grandes áreas de Ciências Biológicas (barras pretas), Ciências da Saúde (barras cor-de-rosa), Ciências Exatas e da Terra (barras azuis) e Ciências Agrárias (barras vermelhas), enquanto que as colaborações envolvendo as grandes áreas de Linguística, Letras e Artes (barras cinzas) e Ciências Aplicadas e Sociais (barras laranjas) apresentam percentuais menos expressivos.

O maior percentual de colaborações interdisciplinares ocorre entre pesquisadores das Ciências da Saúde e das Ciências Biológicas (60,53%), seguido pelos percentuais entre os pesquisadores das Ciências Agrárias e Ciências Biológicas (58,38%), Linguística,

³Subgrafo máximo que inclui um caminho que conecta cada par de nós.

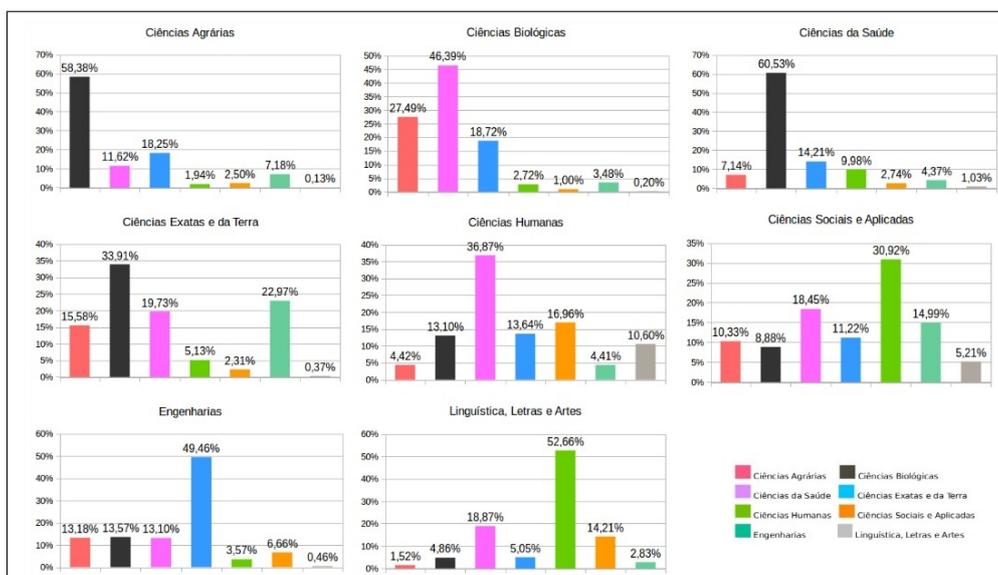


Figura 2. Percentual de Colaborações Interdisciplinares por Grande Área

Letras e Artes e Ciências Humanas (52,66%), Engenharias e Ciências Exatas e da Terra (49,46%), Ciências Biológicas e Ciências Agrárias (46,39%), e Ciências Humanas e Ciências Biológicas (36,87%). Nota-se que tais valores são bastante consistentes com a realidade dos pesquisadores envolvidos nessas colaborações, uma vez que essas grandes áreas possuem uma base teórica bastante próxima (por exemplo, há uma forte interação entre as grandes áreas de Engenharias e Ciências Exatas e da Terra). Vale ainda ressaltar que os pesquisadores da grande área de Ciências Biológicas são os que mais se envolvem em colaborações interdisciplinares. De fato, a grande área de Ciências Biológicas não só provê uma forte base teórica para a grande área de Ciências da Saúde, mas também serve como fundamentação para novas disciplinas como a Bioinformática.

Em relação à diversidade das colaborações, as Ciências Sociais e Aplicadas e as Ciências Humanas podem ser destacadas como as mais democráticas entre todas as grandes áreas, já que seus pesquisadores tendem a colaborar mais amplamente com colegas de outras áreas. Em contraste, os pesquisadores das Ciências Agrárias, Ciências da Saúde e Linguística, Letras e Artes tendem a concentrar as suas colaborações com colegas de uma grande área específica, no caso Ciências Biológicas (as duas primeiras) e Ciências Humanas, como mostra a Figura 2.

4. Considerações Finais

Neste artigo, procuramos salientar a importância das colaborações interdisciplinares nas redes de coautoria científica. Para isso analisamos as oito grandes áreas do conhecimento da rede de coautoria científica brasileira. De modo geral, nossos resultados mostram que essa rede tem crescido, tornando-se mais interdisciplinar. Particularmente, mostramos que 37,4% de todas as colaborações são interdisciplinares. Além disso, nossos resultados mostram o quão integradas são as grandes áreas, enfatizando a força das colaborações entre elas. Considerando o potencial interdisciplinar das grandes áreas, Ciências Biológicas, a terceira mais populosa, destaca-se como a mais democrática. Porém, cabe ressaltar que o tamanho das grandes áreas não é fator decisivo para promover a interdisciplinaridade.

Como trabalhos futuros, pretendemos analisar como ocorre a formação daqueles docentes que transferem conhecimento para outras áreas através de novas colaborações [Mooney et al. 2013, Silva et al. 2016]. Uma outra perspectiva consiste em explicitar a força das colaborações [Brandão et al. 2016] para, por exemplo, destacar os principais atores (pesquisadores ou grupos de pesquisadores) que tornam uma rede mais integrada, como também para identificar aquelas colaborações que em função de sua força dão origem a novas comunidades acadêmicas [Leão et al. 2017]. Além disso, evidências obtidas nos levam a especular que há um compartilhamento do conhecimento interdisciplinar implícito entre as regiões do país. Assim, para melhor explicitar essas colaborações, seria importante identificar aquelas que demonstram a relevância das ligações externas para compartilhamento do conhecimento [Furtado et al. 2015].

Agradecimentos

Trabalho apoiado pelo projeto MASWeb (Proc. FAPEMIG/PRONEX APQ-01400-14) e por auxílios individuais de pesquisa concedidos pelo CNPq e CAPES. O primeiro autor agradece à UFV pela licença concedida para aperfeiçoamento profissional.

Referências

- Brandão, M. A., Diniz, M. A., and Moro, M. M. (2016). Using Topological Properties to Measure the Strength of Co-authorship Ties. In *Proc. of the V Brazilian Workshop on Soc. Netw. Anal. and Mining*, pages 199–210, Porto Alegre, Brazil.
- Dias, T. M. R. and Moita, G. F. (2015). A method for the identification of collaboration in large scientific databases. *Em questão*, 21(2):140–161.
- Furtado, C. A., Davis, Jr., C. A., Gonçalves, M. A., and de Almeida, J. M. (2015). A Spatiotemporal Analysis of Brazilian Science from the Perspective of Researchers' Career Trajectories. *PLOS ONE*, 10(10):1–28.
- Iglič, H., Doreian, P., Kronegger, L., and Ferligoj, A. (2017). With whom do researchers collaborate and why? *Scientometrics*, 112(1):153–174.
- Kato, M. and Ando, A. (2013). The relationship between research performance and international collaboration in Chemistry. *Scientometrics*, 97(3):535–553.
- Lane, J. (2010). Let's make science metrics more scientific. *Nature*, 464(7288):488–489.
- Leão, J. C., Brandão, M. A., de Melo, P. O. V., and Laender, A. H. F. (2017). Classificação de Relações Sociais para Melhorar a Detecção de Comunidades. In *Proc. of the VI Brazilian Workshop on Soc. Netw. Anal. and Mining*, pages 647–657, São Paulo, Brazil.
- Mena-Chalco, J. P., Digiampietri, L. A., Lopes, F. M., and Junior, R. M. C. (2014). Brazilian bibliometric coauthorship networks. *JASIST*, 65(7):1424–1445.
- Milgram, S. (1967). The small-world problem. *Psychology Today*, 1(1).
- Mooney, H. A., Duraiappah, A., and Larigauderie, A. (2013). Evolution of natural and social science interactions in global change research programs. *PNAS*, 110(1):3665–3672.
- Silva, T. H. P., Laender, A. H. F., Davis Jr., C. A., da Silva, A. P. C., and Moro, M. M. (2016). The Impact of Academic Mobility on the Quality of Graduate Programs. *D-Lib Magazine*, 22(9/10).