

O PARADOXO DA AMIZADE EM REDES DE CONVITES DE BANCAS DE DISSERTAÇÕES E TESES

Victor Alexandre Ploeger Mansueli

Mestre em Gestão e Tecnologia em Sistemas Produtivos
Universidade Federal do ABC (UFABC)
E-mail: victor.mansueli@ufabc.edu.br

Diogo Fornaziero Segura Ramos

Mestre em Ciência da Computação
Universidade Federal do ABC (UFABC)
E-mail: diogo.ramos@ufabc.edu.br

Jesús Pascual Mena-Chalco

Professor Doutor
Universidade Federal do ABC (UFABC)
E-mail: jesus.mena@ufabc.edu.br

INTRODUÇÃO

O Paradoxo da Amizade é um fenômeno ilustrado por Feld (1991), que afirma que a maioria das pessoas tem menos amigos do que a média de amigos dos seus amigos. O autor verificou o paradoxo ao considerar uma rede de estudantes. Considerando-se a estrutura de uma rede de amigos em sua totalidade, não faria sentido essa afirmação por todos os participantes simultaneamente. Tal fato implica que o fenômeno é um contrassenso.

Em um estudo cientométrico tendo como tema a colaboração científica, Eom e Jo (2014) generalizaram o Paradoxo da Amizade. Os autores analisaram redes (complexas) de coautoria para a extração de indicadores visando medir a probabilidade de se manter o paradoxo individual e coletivo. Consideraram também, além dos relacionamentos existentes, características dos pesquisadores como, por exemplo, números de citações e de publicações. Já Kong e demais autores (2019) realizaram um levantamento sobre redes sociais acadêmicas, no que concerne à modelagem, análise, mineração e aplicações. Os autores compararam estudos destas redes complexas, formadas por entidades acadêmicas e seus relacionamentos.

Assim como no trabalho de Kong e demais autores (2019), esta pesquisa se insere no âmbito acadêmico, porém retratando o processo de formação de bancas de defesa para obtenção do título de mestre ou doutor. Uma banca de defesa é um grupo pontual, com-

posto por pessoas, com o intuito de avaliar trabalhos de conclusão de alunos. Embora não haja uma regra definida quanto ao número máximo de participantes de uma banca, ela contém, usualmente, de um a cinco pesquisadores (além do orientador, que também pode ser um avaliador).

O orientador, seguindo critérios guiados pelas coordenações de pós-graduação, realiza convites formais para que outros pesquisadores possam ler previamente o trabalho sendo avaliado, e em dia, hora e local estipulados, possam comparecer à defesa, arguir o aluno após a sua apresentação e decidir, em conjunto, pela aprovação ou reprovação do trabalho de conclusão. Bancas de defesa podem ser multidisciplinares, nas quais os avaliadores são de áreas distintas do tema tratado no trabalho de conclusão. Os membros podem, da mesma forma, pertencer à própria instituição do orientador, ou a instituições externas (incluindo outros países). Nesse aspecto, a oferta de convites configura uma modalidade de colaboração científica entre os pesquisadores que orientam alunos de pós-graduação.

Redes de convites foram formalizadas no trabalho de Ramos (2018) e representam relacionamentos duais no qual um convite parte de um ator que o realiza (convitador) para o ator que o recebe (convidado). Entre outras situações, podem ser aplicadas para modelar a composição de bancas de pós-graduação, como proposto nesta pesquisa, na qual um convite foi tratado como uma oferta efetivamente aceita por um pesquisador, oficializando-o como membro de uma banca.

O presente trabalho contribui para um entendimento mais profundo da dinâmica que envolve a oferta de convites entre pesquisadores acadêmicos e na composição de bancas de defesa no Brasil. Pode ser considerado, da mesma forma que o trabalho de Eom e Jo (2014), como uma generalização inédita do Paradoxo da Amizade, na medida em que não considera somente a estrutura topológica da rede, mas também a caracterização heterogênea dessa em dois papéis exercidos pelos atores envolvidos (convidados e convitadores), e a quantificação dos relacionamentos na forma de convites recebidos e/ou realizados.

Em redes de convites, embora não ocorra, necessariamente, uma reciprocidade entre pesquisadores como ocorre em redes de amigos (nem todo pesquisador que recebe um convite realiza um convite para o pesquisador que o convidou), pode prevalecer a “noção” – individual – de que mais convites são recebidos ou realizados pelos outros participantes. Nesse contexto, esta pesquisa testa duas hipóteses do Paradoxo da Amizade no âmbito de pesquisadores acadêmicos que recebem e/ou realizam convites para bancas de defesa no Brasil:

- **Paradoxo do Convidado:** *“Eu recebo menos convites do que a média de convites recebidos pelos pesquisadores que me convidam”.*
- **Paradoxo do Convitador:** *“Eu realizo menos convites do que a média de convites realizados pelos pesquisadores que eu convido”.*

MÉTODO

Os dados utilizados na pesquisa foram submetidos a um pré-processamento consistindo de normalização, remoção de acentos, padronização em letras minúsculas e supressão de espaços em branco, duplicados. Esse tratamento assegurou a unicidade dos pesquisadores na rede, isto é, onde cada vértice representa uma pessoa de forma única.

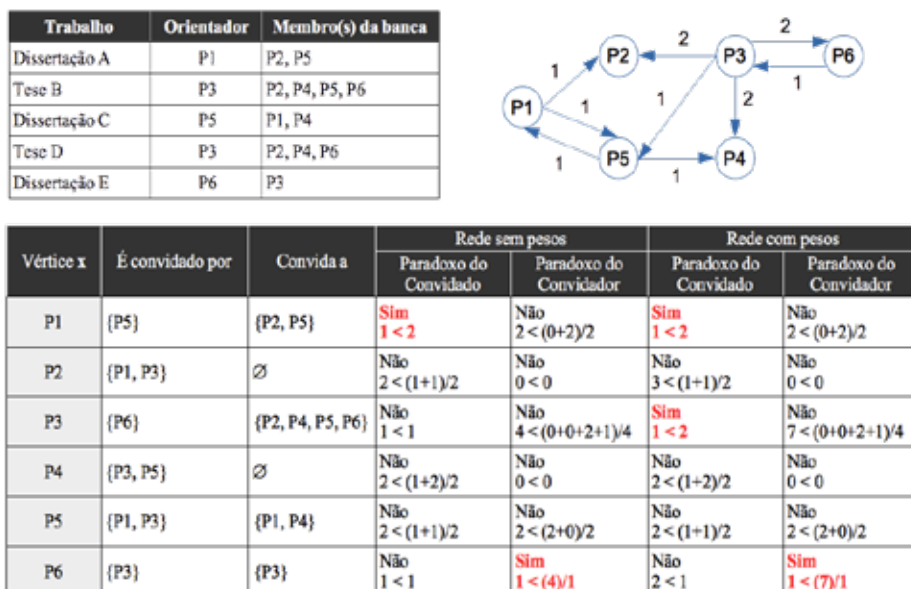
Em seguida, relações foram estabelecidas entre os pesquisadores conforme os convites recebidos e/ou realizados entre cada um deles seguindo as composições de bancas. Uma dupla nomeada foi criada para cada ocorrência de um convite e a “direção” do convite determinada pelos papéis do pesquisador como convidado e/ou convidador.

Duas redes de convites foram então estruturadas. A primeira delas sem peso, ou seja, desconsiderando a eventual multiplicidade de convites entre dois pesquisadores quaisquer (isto é, um relacionamento indica apenas uma oferta de convite). A segunda delas como uma rede ponderada (considerando-se a quantidade de convites entre os mesmos pesquisadores). Importante ressaltar que a soma de convites recebidos ao longo de um período não apresenta simetria com o número de membros que compõem as bancas de defesa (tamanhos das bancas).

Para ilustrar o cálculo do Paradoxo do Convidado/Convidador, a Figura 1 apresenta a transformação de um conjunto de cinco bancas de defesa de mestrado ou doutorado (parte superior esquerda) em uma rede composta por seis pessoas (parte superior direita). Na rede, cada vértice representa uma pessoa, sendo que convites estão indicados pelas arestas, sempre partindo do convidador para o convidado. O número em cada aresta representa a quantidade de convites realizada pelo convidador. Na Tabela (parte inferior da figura) são apresentadas as avaliações dos paradoxos para cada vértice (“Sim” indica paradoxo válido, “Não” indica paradoxo inválido). Nas inequações, o termo que está na parte esquerda está associado ao vértice em evidência. O termo à direita corresponde à média dos graus de entrada (ou de saída) dos convidados (ou dos convidadores). No exemplo, o Paradoxo do Convidado é válido apenas para o vértice P1 na rede sem pesos, e nos vértices P1 e P3 na rede com pesos. Já o Paradoxo do Convidador é válido apenas para o vértice P6, em ambas as redes.

Nesse exemplo, a quantidade de convites não influenciou substancialmente o cálculo dos Paradoxos entre as duas redes (com e sem peso), excetuando-se o Paradoxo do Convidado para o vértice P3 (esse, embora tenha realizado 7 convites, recebeu somente 1). Nas seções seguintes apresentamos tanto o conjunto de dados utilizado, quanto a análise dos Paradoxos para defesas de dissertações e teses brasileiras, agrupadas por Grande Área.

Figura 1: Exemplo de construção da rede de convites a partir da composição de cinco bancas de defesa (parte superior), e testes dos paradoxos da amizade (parte inferior)



Fonte: elaborado pelos autores.

CONJUNTO DE DADOS UTILIZADO

A Tabela 1 sumariza o conjunto de dados coletado a partir do Catálogo de Teses e Dissertações da Capes, hospedado na Plataforma Sucupira, abrangendo o período de seis anos (2013 a 2018). O processo de coleta se estendeu durante os meses de outubro e novembro de 2019. Os totais estão agrupados por Grande Área do Conhecimento (GA) e por Grau Acadêmico, englobando mestrados e doutorados, tanto acadêmicos como profissionais.

As Grandes Áreas são nove agregações em segundo nível, definidas pela CAPES, com o intuito de facilitar o desenvolvimento das atividades de avaliação.¹ São distribuídas em: Ciências Agrárias (AGR), Ciências Biológicas (BIO), Ciências Exatas e da Terra (EXA), Engenharias (ENG), Ciências Humanas (HUM), Linguística, Letras e Artes (LIN), Multi-disciplinar (MUL), Ciências da Saúde (SAU) e Ciências Sociais Aplicadas (SOC).

Para o tratamento, exploração, análise e visualização dos dados utilizados nesta pesquisa, foram produzidos subconjuntos dos 472.474 trabalhos de conclusão, agrupados por GA, contendo os nomes dos pesquisadores e suas relações: convidador (orientador) e convidado(s) de cada uma das bancas de defesa de mestrado ou doutorado cadastradas. Os dados coletados evidenciaram que o número de membros registrados em bancas no

¹ Ver: <https://www.capes.gov.br/avaliacao/sobre-as-areas-de-avaliacao>

Brasil é variável, ocorrendo casos de defesas de doutorado, por exemplo, com até 8 avaliadores (sendo o orientador um deles).

Tabela 1: Número de Trabalhos de Conclusão. Período 2013-2018

Grande Área	Trabalhos de Conclusão		
	Dissertações (D)	Teses (T)	D+T
AGR	33.101	15.158	48.259
BIO	20.723	11.373	32.096
ENG	39.534	11.530	51.064
EXA	31.005	11.224	42.229
HUM	56.688	19.872	76.560
LIN	21.960	7.290	29.250
MUL	46.041	9.505	55.546
SAU	51.547	21.710	73.257
SOC	53.056	11.157	64.213
Total	353.655	118.819	472.474

Fonte: elaborado pelos autores.

RESULTADOS

As redes de convites (com e sem pesos) foram criadas a partir dos conjuntos descritos na seção anterior. Foi utilizada a linguagem de programação Python (biblioteca *Networkx*) para contagem dos graus (ponderada e sem ponderação) de entrada e de saída, assim como para o cálculo das inequações sobre convidado e convidador, para cada vértice.

A Tabela 2 exibe os totais de vértices (pesquisadores), de arestas (um ou mais convites recebidos e/ou realizados), e os percentuais para os paradoxos do pesquisador como convidado e como convidador. Estes últimos dividem-se nos casos em que a quantidade de convites é desconsiderada (sem peso) e nos casos em que a quantidade de convites é rotulada na respectiva aresta e computada (com peso). A última linha contém os valores consolidados para todas as GAs. Em geral, na rede que envolve todas as áreas notamos que uma pessoa, em média, recebe ou realiza aproximadamente cinco convites (isto é, $845.270 / 171.184$).

Considerando-se um limiar de 80% (mínimo) para a rede de convidados, e de 50% (máximo) para a rede de convidadores, (por GA; global; sem e com peso), os números permitem aferir que: a) o Paradoxo do Convidado é *verdadeiro*; b) o Paradoxo do Convidador é *falso*. Ressalta-se, entretanto, que nesse último paradoxo o percentual ficou próximo do limiar no caso global com peso, com 48,02%.

Para analisarmos as assimetrias nos convites, examinamos a quantidade de convites enviados/recebidos. A Figura 2 apresenta dois diagramas de caixas (*boxplots*), de convites enviados e de convites recebidos, em função da quantidade de convites por Grande Área.

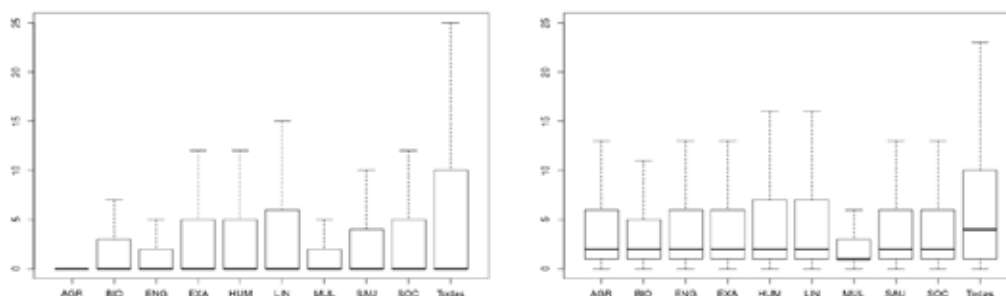
Tabela 2: Percentuais calculados para os paradoxos, agrupados por GA

GA	Vértices	Arestas	Rede de Convites sem peso		Rede de Convites com peso	
			Paradoxo do Convidado	Paradoxo do Convidador	Paradoxo do Convidado	Paradoxo do Convidador
AGR	27.985	84.362	84,31%	24,73%	83,74%	29,49%
BIO	21.977	67.568	82,95%	33,32%	82,29%	37,05%
ENG	28.506	87.898	87,15%	33,91%	86,49%	38,93%
EXA	24.301	79.798	82,26%	39,12%	81,96%	43,61%
HUM	36.782	138.697	87,52%	38,73%	87,55%	42,52%
LIN	13.440	49.877	85,70%	39,91%	85,54%	43,00%
MUL	46.739	102.313	80,28%	27,72%	80,71%	31,47%
SAU	41.692	142.566	84,48%	33,86%	84,22%	38,24%
SOC	28.368	104.459	86,71%	39,70%	87,07%	43,35%
Todas	171.184	845.270	84,53%	44,12%	84,40%	48,02%

Fonte: elaborado pelos autores.

Todas as GAs possuem uma mediana de zero convites enviados por pessoas que participaram de bancas de defesa. A GA que possui o maior intervalo interquartil (diferença entre o quartil superior e o quartil inferior) é LIN, com seis convites enviados. O *boxplot* consolidado de convites enviados possui uma mediana de 0 e um intervalo interquartil de 10. AGR possui o quartil superior e inferior iguais a 0 devido ao baixo percentual de convidadores (24,9%). Todas as GAs possuem uma mediana de dois convites recebidos pelos pesquisadores que participaram de bancas, com exceção de MUL, que possui 1. HUM e LIN são GAs com maior intervalo interquartil, com 6 convites recebidos cada. O *boxplot* consolidado de convites recebidos possui uma mediana de 4 e um intervalo interquartil de 9.

Figura 2: *Boxplots* de convites enviados (esquerda) e de convites recebidos (direita)



Fonte: elaborado pelos autores.

CONCLUSÃO

A análise do Paradoxo da Amizade em redes de convites relacionadas a mais de 472 mil bancas de defesa de dissertações e teses permitiu confirmar o Paradoxo do Convidado e refutar o Paradoxo do Convidador. Embora tenham sido considerados dados oriundos do Brasil em um período de seis anos, acreditamos que os resultados são relevantes dada a possibilidade de se evidenciar o mesmo comportamento em outros contextos acadêmicos.

Este trabalho, de caráter empírico-descritivo, procurou explorar a dinâmica das redes de convites entre acadêmicos para composição de bancas de defesa no Brasil, não levando em consideração o grau acadêmico (mestrado ou doutorado). É importante ressaltar, contudo, que o período utilizado na coleta (de apenas seis anos) pode ser considerado uma limitação temporal do conjunto de dados, já que a Plataforma Sucupira não disponibiliza informações de bancas de defesa para trabalhos de conclusão anteriores à 2013.

Como possibilidades futuras de se explorar e complementar os dados coletados, prevemos estudos que analisem o paradoxo considerando não somente a topologia, mas correlacionando convites com outras propriedades de redes acadêmicas (por exemplo, grau acadêmico, área de concentração dos trabalhos de conclusão, Instituições de Ensino Superior (IES) dos pesquisadores, publicações e citações, interdisciplinaridade de tópicos e evolução temporal da produção acadêmica de autores). Adicionalmente, o método utilizado pode ser aplicado em conjuntos de dados coletados de fontes internacionais como, por exemplo, a NDLTD (Networked Digital Library of Thesis and Dissertations).

Além do fenômeno aqui apresentado, tratamento similar pode ser usado como insumo quantitativo para uma compreensão mais profunda das motivações que levam alguns pesquisadores a preferencialmente convidar – ou a serem convidados – pelas mesmas pessoas, ou mesmo estudar os diferentes níveis de experiência dos membros convidados para bancas de defesas de mestrado e doutorado. Por fim, endogenia e interdisciplinaridade também seriam aspectos a serem explorados.

REFERÊNCIAS

- EOM, Y. H.; JO, H. H. Generalized friendship paradox in complex networks: the case of scientific collaboration. *Scientific Reports*, London, v. 4, p. 4603, 2014.
- FELD, S. L. Why Your Friends Have More Friends than You Do. *The American Journal of Sociology*, Chicago, v. 96, n. 6, p. 1464-1477, 1991.
- KONG, X.; SHI, Y.; YU, S. et al. Academic social networks: modeling, analysis, mining and applications. *Journal of Network and Computer Applications*, London, n. 132, p. 86-103, 2019.
- RAMOS, D. F. S. *Território causal em grafos temporais de convites e coparticipações*. 2018. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal do ABC, Santo André, 2018.

