



ANÁLISE DA REDE DE COLABORAÇÃO DOS BOLSISTAS PRODUTIVIDADE EM PESQUISA: uma abordagem segmentada por níveis

Wellington Barbosa Rodrigues¹
Jesús Pascual Mena-Chalco¹

Resumo: Os bolsistas produtividade em pesquisa são pesquisadores reconhecidos pelo CNPq e são segmentados por níveis. Neste trabalho é analisada a rede de colaboração desses pesquisadores, tendo como objetivo geral apresentar as características da rede e da colaboração entre os pesquisadores considerando seus níveis. Foi possível evidenciar que o grupo de pesquisadores de nível 1A apresenta as maiores médias de centralidade dentro da rede, em contrapartida a maior média de colaboração ocorre com os pesquisadores de nível 2. Também foi possível identificar que, de maneira geral, os pesquisadores não possuem fortes relações com apenas um outro pesquisador, mas sim relações de maneira mais distribuída.

Palavras-Chave: Rede de colaboração. Bolsistas produtividade. Colaboração acadêmica.

1 INTRODUÇÃO

A colaboração entre pesquisadores é uma parte natural do processo de produção científica, pesquisadores trocam experiências e novas descobertas com a finalidade de aprimoramento do conhecimento produzido. As produções científicas vêm crescendo com o passar do tempo, assim como as colaborações entre os pesquisadores. Além da característica de crescimento, as produções científicas também apresentam diferentes formas de colaboração entre áreas distintas dos níveis dos pesquisadores envolvidos (DIAS; MOITA; DIAS, 2019).

Os pesquisadores produtividade em pesquisa (PQ's) são especialistas reconhecidos pelo CNPq entre seus pares. Atualmente, o nível da bolsa é dividido em três categorias (GUEDES; AZEVEDO; FERREIRA, 2015), sendo: Sênior; 1 (a qual é subdividida em A, B, C, D) e; 2.

O relacionamento entre pesquisadores pode ser visto como uma rede, onde os pesquisadores são os vértices e as arestas são as colaborações entre pesquisadores. Essa abordagem permite a obtenção de informações quantitativas e parâmetros referentes à forma de colaboração entre

¹ Universidade Federal do ABC (UFABC)

os pesquisadores (DING, 2011) de maneira a viabilizar o uso de estratégias de análises de redes sociais em redes de colaboração científica.

Este trabalho analisa a colaboração entre os PQ's, considerando os níveis de suas bolsas e também considera se a colaboração ocorre prioritariamente entre pesquisadores da mesma área ou de áreas diferentes. O objetivo geral deste trabalho é apresentar as características da rede e da colaboração entre os pesquisadores, considerando seus níveis, e como objetivos específicos temos: (i) apresentar medidas, segmentadas por níveis de PQ, relacionadas a topologia da rede; (ii) comparar a relação de colaboração entre os pesquisadores de diferentes níveis de bolsa e; (iii) analisar a diferença entre colaborações de pesquisadores da mesma área e de áreas distintas.

2 METODOLOGIA

Para a obtenção dos dados utilizados neste trabalho foram coletados os currículos da plataforma Lattes, através da ferramenta *scriptLattes* (MENA-CHALCO; CESAR JUNIOR, 2009), dos PQ's com bolsa vigente em 2021. Os pesquisadores e suas respectivas áreas foram identificados através do portal mapa de investimentos do próprio CNPq.

Os currículos foram analisados, foram excluídas publicações presentes em mais de um currículo e posteriormente criada uma rede de colaboração entre os PQ's através da coautoria de suas publicações, foram consideradas as publicações do tipo (i) Artigos em periódicos, (ii) Artigos em eventos, (iii) Livros, e (iv) Capítulos de livros. Esta rede de colaboração entre os pesquisadores (RCP) possui as informações referentes a todas as produções consideradas no período de 2011 a 2020, contemplando assim, um período de 10 anos das relações entre os pesquisadores, com 15.121 pesquisadores e 238.242 relações entre eles.

Para possibilitar identificar o nível de relação que cada pesquisador possui com os outros pesquisadores, ao criar a RCP para as relações foi utilizado peso nas arestas. A estratégia escolhida para atribuição de peso nas relações entre os pesquisadores foi o peso normalizado (LIU *et al.*, 2005), tal estratégia permite identificar o quanto da produção de um pesquisador é relacionada exclusivamente a outro pesquisador da rede.

Para calcular o peso da aresta de saída é necessário identificar a exclusividade de um artigo. Um artigo que tenha apenas 2 autores terá a exclusividade 1 para cada autor. O segundo cálculo realizado é a frequência, que é dada pelo somatório de todas as exclusividades entre 2

autores. Por fim, o peso normalizado é calculado, onde se pondera a frequência pela quantidade total de artigos que um determinado pesquisador tem.

Exclusividade	Frequência	Peso normalizado
$g_{i,j,k} = \frac{1}{(f(a_k) - 1)}$	$c_{ij} = \sum_{k=1}^m g_{i,j,k}$	$w_{ij} = \frac{c_{ij}}{\sum_{k=1}^n c_{ik}}$

Para exclusividade os autores v_i e v_j são coautores em um artigo a_k e $f(a_k)$ é a quantidade total de autores no artigo, com isso $g_{i,j,k}$ representa o grau de exclusividade que cada autor tem com relação a esse artigo específico. A frequência, representada por c_{ij} , de colaboração entre dois autores é dada pelo somatório de todas as exclusividades das colaborações entre esses autores. Já w_{ij} é o peso normalizado na relação entre os dois autores, no qual é calculado através da exclusividade ponderada pela quantidade total de artigos de cada autor, assim, gerando um valor diferente para cada um, de maneira que o peso da relação de v_i para v_j é diferente do peso da relação de v_j para v_i .

Através da análise da RCP é possível obter resultados da relação que os PQ's possuem entre eles com base em uma análise da topologia da rede e as relações entre os níveis dos pesquisadores. Como medidas de topologia para análise da rede foram utilizadas as medidas propostas por Newman (2001) com o intuito de identificar como um pesquisador está relacionado dentro de uma rede de colaboração científica.

De maneira complementar é apresentado um comparativo entre as relações dos pesquisadores de maneira segmentada pelos níveis de PQ e as relações dos PQ's como um todo é apresentado de um diagrama de caixa elaborado com os pesos das relações de todos os PQ's. Essa é uma forma de permitir uma contextualização dos dados das relações entre os níveis com todo o conjunto de relações deste grupo de pesquisadores.

3 RESULTADOS

Com a RCP é possível analisar a forma de relacionamento entre os PQ's através da topologia da rede. Dentre os atributos que merecem destaque nessa rede, podemos citar que em média um pesquisador tem relação com outros 15,75 pesquisadores da rede (grau médio), a maior componente conexa é composta por 13.940 pesquisadores, indicando que há um caminho

entre cada um deles, onde a maior distância entre dois pesquisadores que tenham caminho entre eles é de 23.

Na **Tabela 1** pode-se conferir alguns dados referentes à RCP com divisão por níveis dos PQ's, para efeito comparativo. A primeira coluna apresenta a quantidade total de pesquisadores por nível, sendo que temos aproximadamente 40 vezes mais PQ's de nível 2 do que bolsistas SR, a segunda coluna é composta pelo cálculo da porcentagem que cada nível representa em relação ao todo da RCP: As quatro colunas seguintes foram calculadas com base na **média** das características de cada nível: **(i) grau médio**, indica em média com quantas pessoas da RCP um pesquisador de determinado nível teve colaboração no período analisado, em média um PQ-1A colaborou com quase 28 pessoas da rede, enquanto um PQ-2 em média colaborou com aproximadamente 13 pessoas diferentes; **(ii) Centralidade de grau**, é a fração de nós que um pesquisador está conectado, sendo a medida mais simples para identificar a centralidade de um pesquisador; **(iii) centralidade de intermediação**, é a fração de caminhos mais curtos entre todos os pesquisadores da rede a qual um determinado pesquisador faz parte, quanto mais alta esta centralidade indica que uma pessoa tem maior participação do fluxo de informações na rede e; **(iv) centralidade de proximidade**, é a inversa da distância média de caminho mais curto, indicando o quanto esse pesquisador tem de capacidade de propagar uma informação para toda a rede. As três formas calculadas de centralidade apresentam a mesma característica, seguem a seguinte ordem de maior centralidade média por nível para a menor: 1A, 1B, 1C, 1D, 2, SR.

Tabela 1 – Medidas da rede segmentada por níveis dos PQ's

Nível	Total Pesquisadores	Porcentagem do total	Grau Médio	Centralidade Grau	Centralidade Intermediação	Centralidade Proximidade
SR	215	1,4218	11,3535	0,0015	0,0001	0,1610
1A	1.166	7,7111	27,9949	0,0037	0,0005	0,1987
1B	1.220	8,0682	21,1385	0,0028	0,0003	0,1905
1C	1.432	9,4702	18,2730	0,0024	0,0003	0,1893
1D	2.474	16,3613	16,5081	0,0022	0,0002	0,1869
2	8.614	56,9671	12,8119	0,0017	0,0001	0,1810

Fonte: Elaborado pelos autores (2022).

Complementarmente na **Tabela 1** podemos ver que seja interna ou externamente, em média, as maiores colaborações, independentemente do nível, são estabelecidas com pesquisadores PQ-2 e PQ-1D. Por outro lado, os pesquisadores PQ-1A apresentam os maiores valores para as medidas de centralidade, com isso, estando em regiões de maior fluxo na RCP. Também é possível ver que embora os pesquisadores de nível PQ-2 possuam mais relacionamentos,

apresentam os valores de centralidade inferiores a todos os pesquisadores de nível 1, tendo centralidades maiores somente que os pesquisadores de nível SR.

A **Tabela 2** apresenta os valores médios dos relacionamentos entre os níveis dos PQ's, segmentando a média dos relacionamentos por interno e externo. Com o intuito de facilitar a visualização, a **Tabela 2** apresenta variação em suas cores indo do azul ao verde, sendo azul os relacionamentos com menor valor, verde os relacionamentos com maior valor e branco os intermediários.

Tabela 2 – Colaboração média entre os níveis de PQ's

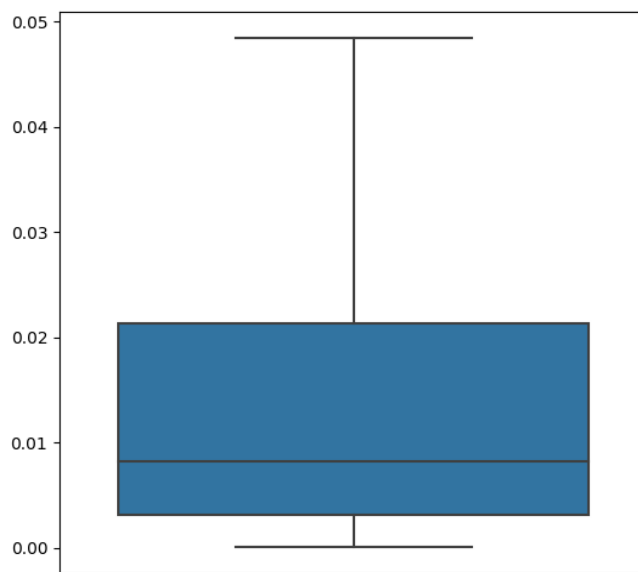
Nível	SR	1A	1B	1C	1D	2
SR-Interno	0,0744	0,1269	0,0959	0,0747	0,0756	0,0499
SR-Externo	0,1023	0,1672	0,1336	0,1222	0,0946	0,0751
1A-Interno	0,6884	1,7650	1,1705	1,0454	0,9507	0,6673
1A-Externo	0,9070	2,6192	1,9270	1,5656	1,3015	0,9692
1B-Interno	0,5442	1,2247	0,9131	0,7765	0,7296	0,5149
1B-Externo	0,7581	2,0163	1,5623	1,2465	1,0400	0,8126
1C-Interno	0,4977	1,2839	0,9115	0,8366	0,7603	0,5449
1C-Externo	0,8140	1,9228	1,4631	1,3142	1,0631	0,8085
1D-Interno	0,8698	2,0172	1,4795	1,3135	1,2959	0,9302
1D-Externo	1,0884	2,7616	2,1090	1,8366	1,6580	1,2350
2-Interno	2,0000	4,9297	3,6352	3,2779	3,2389	2,6564
2-Externo	3,0093	7,1604	5,7377	4,8631	4,2999	3,5479

Fonte: Elaborado pelos autores (2022).

Outro ponto a ser analisado é a intensidade da relação entre os pesquisadores, ou seja, quanto da produção de um determinado pesquisador está relacionada a apenas um outro pesquisador. Essa intensidade pode ser mensurada pelo peso normalizado das relações. Essa medida é composta pela quantidade total da produção de um determinado pesquisador e quanto dessa produção foi em colaboração com um outro determinado pesquisador, conforme descrito na **Seção 2**. Dessa maneira indicando o quão intensa é a relação direta de um pesquisador com outro.

A **Figura 1** é um diagrama de caixa que apresenta a distribuição do peso das relações entre os pesquisadores, o peso máximo possível de uma relação ao utilizar peso normalizado é 1. Observando o gráfico identifica-se que a mediana do peso das relações é inferior a 0,01, ou seja, a maioria das relações tem peso inferior a 1% do máximo possível, indicando que há grande diversidade nas colaborações e, de uma maneira geral, os PQ's não tem relações muito fortes com apenas um outro PQ.

Figura 1 - Diagrama de caixa do peso das relações da rede de colaboração



Fonte: Elaborado pelos autores (2022).

Com os resultados apresentados é possível ter um panorama sobre a relação dos PQ's, considerando a intensidade da relação que há entre eles, a centralidade média de cada nível em relação à rede como um todo e como os diferentes níveis se relacionam entre si.

4 CONCLUSÕES

O processo de colaboração acadêmica é fundamental para a construção da ciência e é essencial para complementação de conhecimentos dos pesquisadores de cada área. A produção científica vem crescendo nos últimos anos, junto com o crescimento da produção científica as colaborações também aumentaram. Tornando-se essencial um olhar para as formas como essas colaborações estão ocorrendo em grupos e como um todo.

Neste trabalho foi analisada a rede de colaboração composta pelo grupo de pesquisadores produtividade em pesquisa do CNPq, grupo escolhido por ser composto de pesquisadores reconhecidos por sua liderança acadêmica. Com base na análise das produções acadêmicas do PQ's foram evidenciados neste trabalho que os pesquisadores de nível 1A possuem grau médio de colaboração mais alto que os outros níveis de PQ's, além de também possuírem a média das centralidades calculadas maior que a dos outros níveis.

Uma outra visão obtida através da análise dos resultados foi que o grupo que em média recebe mais colaboração é o grupo composto pelos pesquisadores de nível 2, sendo que todos os níveis apresentam maior média de colaboração com esse grupo. Ao analisar as colaborações

segmentando por nível e por ambos os pesquisadores do relacionamento pertencerem à mesma área, nota-se que todos os níveis de PQ 's têm maior propensão de colaborar com pesquisadores externos à sua área.

Por fim, foi analisada o peso de todas as relações entre os PQ's, mostrando que, de uma maneira geral, os pesquisadores não possuem uma intensidade de colaboração muito alta com somente um pesquisador, tendo a tendência de colaborar pouco com muitas pessoas.

Como trabalhos futuros, pretendemos replicar as análises feitas neste trabalho, porém considerando a rede de colaboração de todos os doutores do Brasil com currículo cadastrado na plataforma Lattes. Tal análise permitirá identificar os grupos mais colaborativos com todos os pesquisadores, não somente entre os PQ's.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

REFERÊNCIAS

- DIAS, Thiago Magela Rodrigues; MOITA, Gray Farias; DIAS, Patricia Mascarenhas. Um estudo sobre a rede de colaboração científica dos pesquisadores brasileiros com currículos cadastrados na Plataforma Lattes. **Em Questão**, Porto Alegre, v. 25, n. 1, p. 63-86, 2019.
- DING, Ying. Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. **Journal of informetrics**, Amsterdam, v. 5, n. 1, p. 187-203, 2011.
- GUEDES, Moema de Castro; AZEVEDO, Nara; FERREIRA, Luiz Otávio. A produtividade científica tem sexo? Um estudo sobre bolsistas de produtividade do CNPq. **Cadernos Pagu**, Campinas, p. 367-399, 2015.
- LIU, Xiaoming *et al.* Co-authorship networks in the digital library research community. **Information processing & management**, London, v. 41, n. 6, p. 1462-1480, 2005.
- MENA-CHALCO, J.; CESAR JUNIOR, R. ScriptLattes: an open-source knowledge extraction system from the Lattes platform. **Journal of the Brazilian Computer Society**, Basingstoke, v. 15, n. 4, p. 31-39, 2009.
- NEWMAN, Mark EJ. Who is the best connected scientist? A study of scientific coauthorship networks. **Phys. Rev. E.**, [*s.l.*], v. 64, n. 016131, 2001.