

O paradoxo da amizade em grafos de genealogia acadêmica: análise experimental

Ana da Silva de Paula,¹ Jesús P. Mena-Chalco²
Centro de Matemática, Computação e Cognição, UFABC

1 Sobre o Paradoxo da Amizade e Genealogia Acadêmica

O paradoxo da amizade enuncia que, em média, uma pessoa tem menos amigos do que seus amigos [2]. É possível avaliá-lo em dois eixos: (i) O tipo de avaliação pode ser numérico, quando é atribuído a ele um valor através de uma expressão matemática, ou booleano, quando é atribuído um valor-verdade; (ii) A escala de avaliação pode ser local, quando é considerado o paradoxo individualmente, ou global, quando é observada a tendência do grafo. Sabe-se ainda que o paradoxo também é válido para características diferentes do grau, o que foi descrito como *paradoxo generalizado da amizade* [1]. Pode-se afirmar, portanto, que a maioria dos componentes de uma rede social está abaixo da média quando comparada a seus amigos, independentemente da característica observada.

Por outro lado, a genealogia acadêmica é o estudo da transmissão de conhecimento (herança intelectual) através de relações de orientação formal, permitindo, por exemplo, analisar a maneira como o conhecimento é passado a cada geração, assim como a identificação dos pesquisadores mais férteis, e previsões sobre o crescimento de áreas do conhecimento. Os grafos de genealogia são representações gráficas da genealogia acadêmica, nos quais os pesquisadores são representados por vértices e as relações de orientação são representadas por arestas direcionadas partindo do orientador e chegando ao orientando.

2 Método

Este trabalho tem o objetivo verificar a validade do paradoxo da amizade em grafos de genealogia. Nesse contexto, é efetuada a análise dos grafos utilizando como critério o *grau de saída* de cada vértice (i.e., sua quantidade de filhos) e comparando-o ao grau de saída médio de seus pais, filhos e irmãos — o que chamamos de *níveis de comparação* — conforme a fórmula $|F_v| < \frac{1}{|P_v|} \sum_{i \in P_v} |F_i|$, em que v é o vértice analisado, F_j é o conjunto dos filhos de j e P_j é o conjunto do nível de comparação no qual o paradoxo será avaliado. A análise local é booleana (o paradoxo é considerado *verdadeiro* se a fórmula for válida) e a análise global é numérica e baseada no percentual de vértices cujo paradoxo é verdadeiro.

¹ana.p@aluno.ufabc.edu.br

²jesus.mena@ufabc.edu.br

O valor do paradoxo para um vértice pode ser verdadeiro, falso ou indeterminado. Este último caso está presente porque nem sempre é possível avalia-lo, pois nem todos os vértices têm todos os familiares. Os vértices podem fazer parte de 12 categorias que estão relacionadas à existência de pais, filhos e irmãos.

3 Resultados Experimentais

Efetuamos análise de dois grafos de genealogia: (i) dos doutores cadastrados na Plataforma Lattes, com 372 218 vértices e 389 377 arestas, e (ii) de Johann Bernoulli, com 81 728 vértices e 88 921 arestas, obtido pelo *The Mathematics Genealogy Project* [3]. A genealogia de doutores da Plataforma Lattes reflete as características da prática da ciência no Brasil: é uma árvore rasa e recente (contendo, no máximo, 11 gerações e cujos registros mais antigos datam de aproximadamente 1940) e tem pesquisadores de diversas áreas do conhecimento. Já o grafo de Bernoulli é composto por 20 gerações, sua origem data de 1690 e os dados concentram-se nos pesquisadores doutores relacionados com matemática.

Na Tabela 1 apresentamos os resultados da aplicação do paradoxo de amizade nos grafos de genealogia considerando-se apenas os vértices nos quais era possível avalia-lo. As taxas de indeterminação nos níveis de pais, filhos e irmãos foram, respectivamente, 16.5%, 71%, 26.6% no grafo dos doutores da Plataforma Lattes, e 0, 1%, 80.7%, e 4.3% no grafo de Bernoulli.

Tabela 1: O paradoxo da amizade: doutores na P. Lattes (esq.), J. Bernoulli (dir.).

Nível	Falso	Verdadeiro	Nível	Falso	Verdadeiro
Pais	18602 (5,98%)	292261 (94.02%)	Pais	2926 (3,58%)	78841 (96.42%)
Filhos	94911 (87,91%)	13052 (12.09%)	Filhos	14898 (94,34%)	893 (5.66%)
Irmãos	121362 (44,44%)	151714 (55.56%)	Irmãos	34357 (43,92%)	43863 (56.08%)

4 Considerações Finais

Mesmo quando aplicados em grafos com características distintas (i.e., número diferente de gerações e vértices), o valor-verdade do paradoxo permanece o mesmo: *ele é verdadeiro no nível de pais, falso no nível de filhos e indeterminado no nível de irmãos, porém com uma leve tendência ao verdadeiro*. Este trabalho permitirá encontrar características comuns aos pesquisadores e tendências de diferentes áreas do conhecimento que se assemelham quando considerado o paradoxo, o que auxiliará na compreensão de como a Ciência acontece.

Referências

- [1] Y.-H. Eom and H.-H. Jo. Generalized friendship paradox in complex networks: The case of scientific collaboration. *Scientific Reports*, 4:4603, 2014.
- [2] S. L. Feld. Why your friends have more friends than you do. *American Journal of Sociology*, 96(6):1464–1477, 1991.
- [3] L. Rossi and J. P. Mena-Chalco. Caracterização de árvores de genealogia acadêmica por meio de métricas em grafos. In *BraSNAM*, pages 1–12, 2014.