

Prestígio em grafos de genealogia acadêmica: Uma proposta baseada em PageRank - short paper -

Arthur V. Kamienski, Rafael J. P. Damaceno, Jesús P. Mena-Chalco



Agenda

1. Objetivo
2. Grafos de genealogia acadêmica
3. PageRank e propostas de adaptação
4. Resultados
 - 4.1. Distribuição de valores
 - 4.2. Análise por idades acadêmicas
 - 4.3. Pesquisadores de prestígio da Ciência da Computação no Brasil

Objetivo

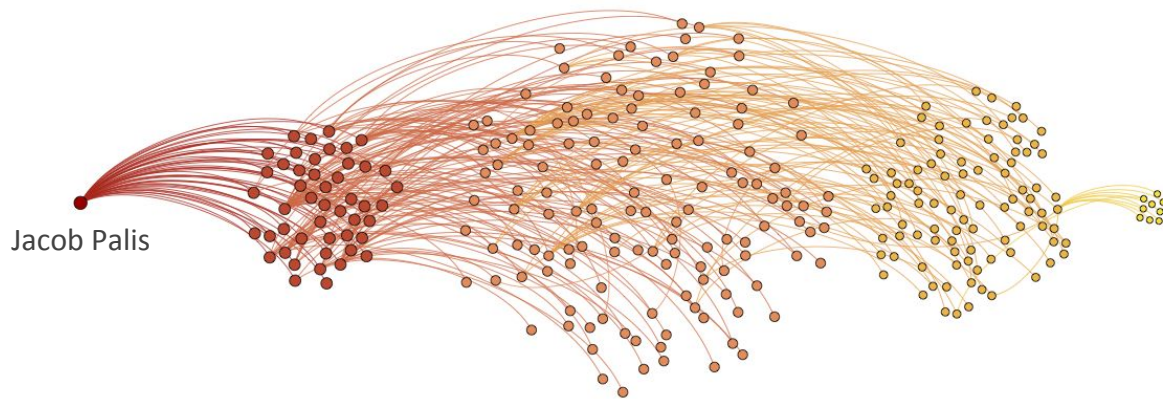
Objetivo

- Desenvolver uma medida capaz de classificar pesquisadores de acordo com sua genealogia acadêmica
 - Adaptar o algoritmo do PageRank para que se adeque às restrições estabelecidas pelos grafos utilizados

Grafos de genealogia acadêmica

Grafos de genealogia acadêmica

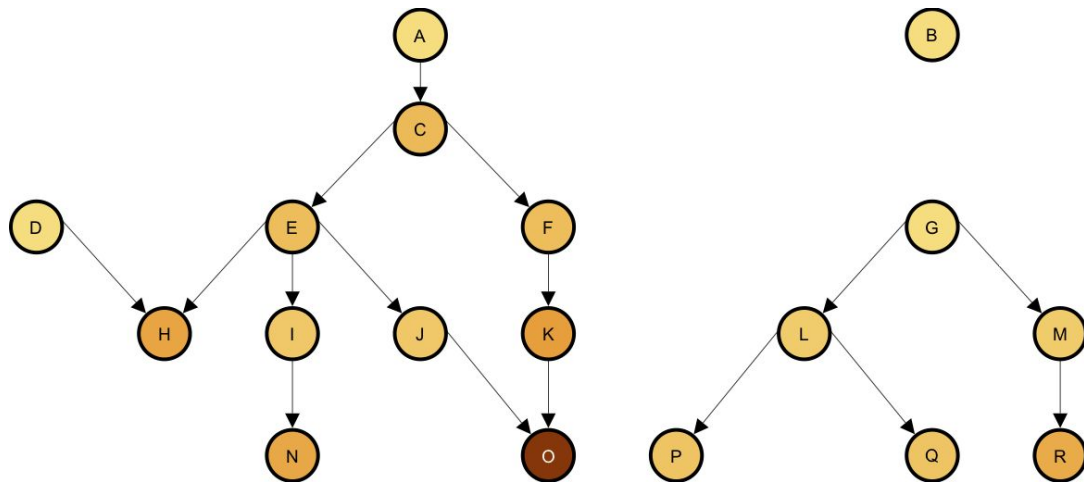
- Grafo no qual vértices representam pesquisadores e arestas representam relações de orientação acadêmica
- Grafo de genealogia acadêmica demonstrando a descendência de Jacob Palis:



PageRank e propostas de adaptação

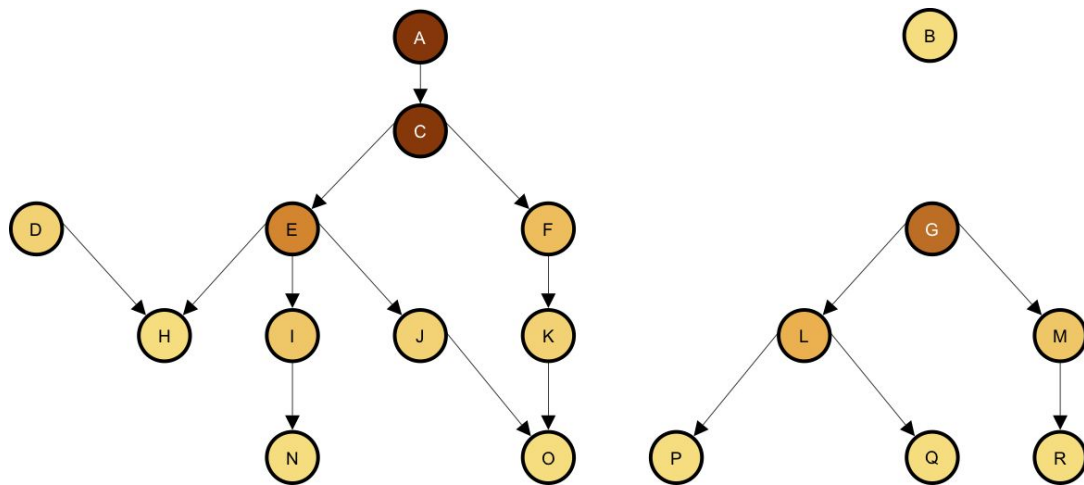
PageRank

- PageRank de um vértice é a soma de parcelas dos PageRanks dos vértices que o apontam
- Vértices folha tendem a acumular PageRank



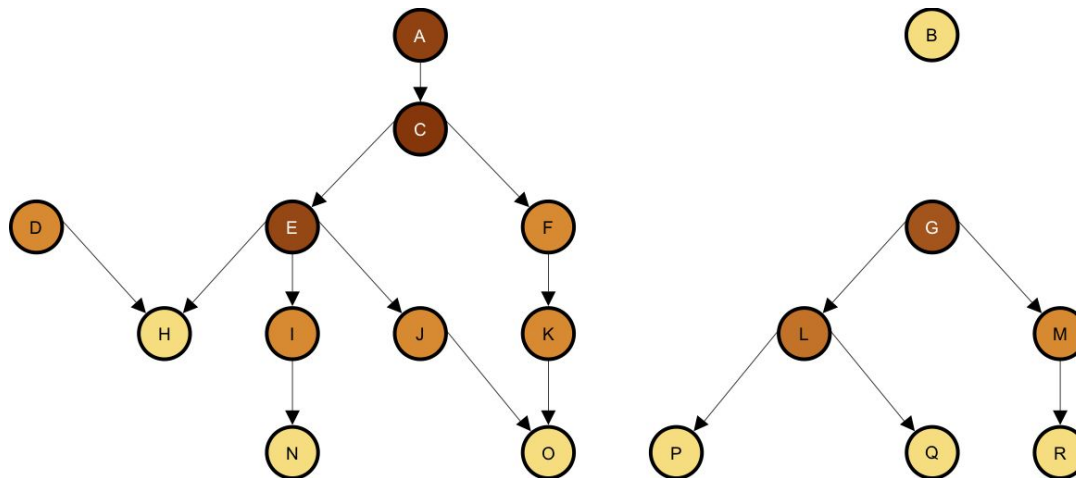
PageRank Invertido (PRI)

- Inversão do fluxo de PageRank
- Pesquisadores próximos às raízes (orientadores) são exaltados



PageRank Invertido Local Normalizado (PRILN)

- PageRank sobre um sub-grafo de k gerações a partir do vértice
- Normalização para a comparação entre sub-grafos de tamanhos diferentes



Definição formal do PRILN

- Inversão dos valores de PRIL mantendo ordenação para sub-grafos de mesmo tamanho

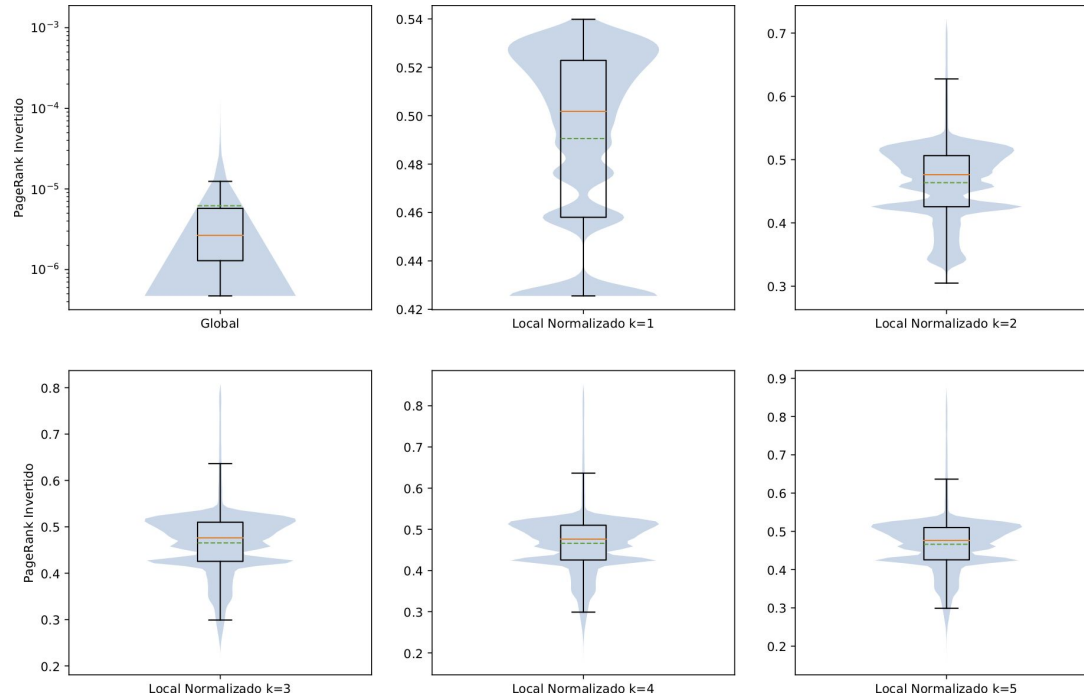
$$PRILN(v) = \frac{PRIL(v)}{PRIL(w)} \times (1 - PRIL(w)) = \frac{PRIL(v)}{PRIL(w)} - PRIL(v)$$

w = representante do conjunto de vértices de sub-grafo de mesmo tamanho

Resultados

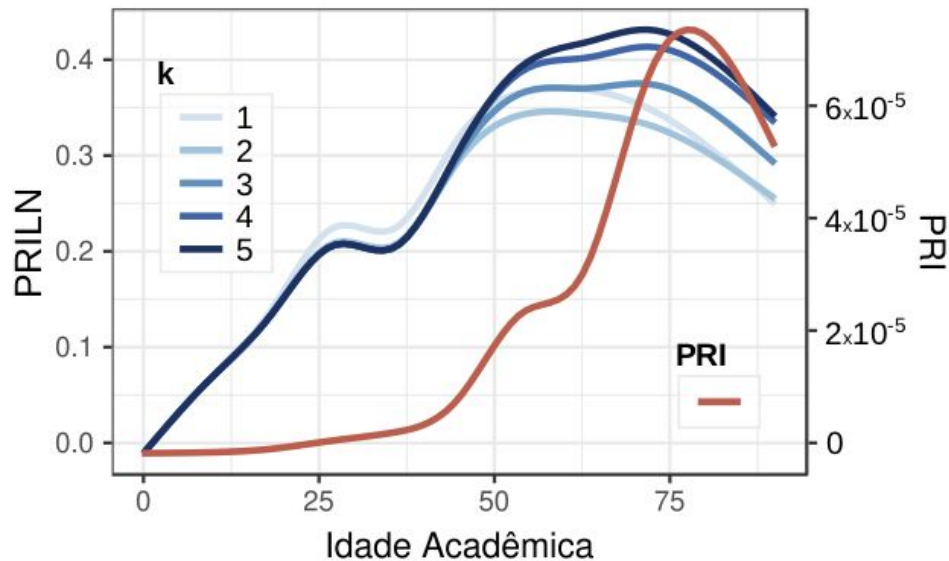
Distribuição de valores de PageRank

- PRILN está melhor distribuído que o PRI



Análise por idades académicas

- Pesquisadores mais jovens têm valores de PRILN mais altos que de PRI



Pesquisadores de prestígio

- Exaltação de pesquisadores já renomados e com bolsas de produtividade

Rank	PRI	PRILN k=1	PRILN k=2	PRILN k=3
1	CJP Lucena*	SRL Meira	N Maculan Filho*	CJP Lucena*
2	AV Staa	CJP Lucena*	SRL Meira	SBT Mendes
3	SRL Meira	AF Abreu	LP Magalhães	AV Staa
4	N Maculan Filho*	JM Souza*	CJP Lucena*	WP Paula Filho
5	JM Souza*	ARC Rocha	AV Staa	N Maculan Filho*

'*' indica bolsistas de produtividade em pesquisa do CNPq (2018)

Prestígio em grafos de genealogia acadêmica: Uma proposta baseada em PageRank

Arthur V. Kamienski

a.kamienski@aluno.ufabc.edu.br