

Enriquecendo bases de dados de currículos Lattes

Jamison José da Silva Lima, Luciano Antonio Digiampietri

Escola de Artes, Ciências e Humanidades, USP

jamison.lima@gmail.com, luciano.digiampietri@gmail.com

Objetivos

No Brasil existe uma base ímpar de dados sobre pesquisadores e suas pesquisas: a Plataforma Lattes. Nesta plataforma, diversas informações referentes a publicações, produções técnicas e projetos de pesquisa são encontradas e podem ser utilizadas para a identificação de redes de pesquisa, diversos tipos de relações entre pesquisadores, etc. O objetivo deste projeto é desenvolver ferramentas para enriquecer bancos de dados de currículos Lattes de forma a possibilitar uma gama de análises quantitativas e qualitativas sobre pesquisadores e grupos de pesquisa.

Métodos/Procedimentos

A metodologia foi iniciada com a realização de uma revisão bibliográfica sobre o tema: “análise de redes sociais”, com foco na identificação de métricas para a análise. A partir do estudo realizado, foram especificadas e implementadas diferentes ferramentas para o cálculo de métricas de análise. Além disso, também foram desenvolvidas ferramentas para obter informações complementares às disponíveis nos currículos Lattes (por exemplo citações de artigos utilizando-se o *Google Scholar*, ou informações de periódicos, como o *SJR*). Com base nas ferramentas desenvolvidas algumas análises foram realizadas.

Resultados

A Tabela 1 contém o nome e breve descrição de algumas métricas implementadas. Além disso, o banco de dados atual foi enriquecido com índices de periódicos científicos (JCR e SJR) e número de citações das publicações (extraídos dos sistemas *Google Scholar* e *Microsoft Academic Search*). As ferramentas desenvolvidas serviram de base para a escrita de dois artigos científicos que foram submetidos para publicação em um periódico brasileiro.

Tabela 1 - Algumas das Métricas Implementadas

Nome	Descrição
<i>Degree Assortativity</i>	Calcula globalmente a relação de similaridade de grau entre os nós vizinhos
<i>AuthorRank</i>	Função de prestígio baseada no PageRank
<i>Betweenness</i>	Calcula a quantidade de vezes que um nó está em algum caminho mínimo
<i>Closeness</i>	Métrica relacionada à distância de um nó a todos os outros no componente
<i>Clustering Coefficient</i>	Métrica global que calcula a proporção de subgrafos 3-completos no grafo
Densidade	Proporção de arestas no grafo ao máximo possível
Diâmetro	Calcula a maior distância de um componente conexo
Grau Médio	Calcula o grau médio dos nós do grafo
<i>PageRank</i>	Métrica de prestígio baseada na quantidade e qualidade das relações de um grafo

Conclusões

O desenvolvimento de diversas métricas de análise de redes sociais combinado com o enriquecimento de um banco de dados formado por currículos Lattes possibilita a realização de diferentes análises sobre a produtividade acadêmica brasileira. Como trabalho futuro pretendemos implementar outras ferramentas para facilitar este tipo de análise bem como estudar sobre diferentes aspectos algumas comunidades científicas brasileiras.

Referências Bibliográficas

DIGIAMPIETRI, L.; MENA-CHALCO, J.; ALCAZAR, J. J. P.; TUESTA, E. F.; DELGADO, K. V.; MUGNAINI, R.; SILVA, G. S., LIMA, J. J. S. (2013). Extração, caracterização e análises de dados de currículos Lattes. Revista Eletrônica de Sistemas de Informação (RESI), submetido, 2013.