

Research.NET Web: uma aplicação para análise e exibição de redes de colaboração acadêmica utilizando grafos dinâmicos

Érick Luiz Fonsêca Lopes¹, Eduardo N. Borges¹

¹Centro de Ciências Computacionais – Universidade Federal do Rio Grande (FURG)
Av. Italia, km 8, 96203-900. Rio Grande – RS

si.erickluiz@gmail.com, eduardoborges@furg.br

Resumo. *Este artigo descreve o desenvolvimento de uma solução Web que permite a visualização dinâmica de redes de colaboração acadêmica. A aplicação utiliza dados da plataforma Lattes para identificar e exibir a relação entre conjuntos de pesquisadores.*

Abstract. *This paper describes the development of a Web solution that allows the dynamic visualization of academic collaboration networks. The application uses data from the Lattes platform to identify and display the relationship between sets of researchers.*

1. Introdução

Atualmente não é incomum a necessidade de avaliar redes formadas a partir da colaboração entre autores de determinado grupo, tanto no meio acadêmico quanto fora dele. Para [MENA-CHALCO and CESAR-Jr 2009], estas avaliações podem ser utilizadas, por exemplo, para documentar a produção científica do grupo. Extrair valores reais de uma rede de colaboradores e exibí-los de maneira simples não é tarefa fácil em tempos em que a informação é produzida de forma constante e abrangente. Isto implica em se utilizar uma boa fonte de dados para a absorção do conhecimento requerido.

Referente à pesquisa, o Brasil possui uma particularidade excepcional: “a existência de um cadastro nacional de currículos de pesquisadores, a Plataforma Lattes”, onde se concentram informações referentes à vida no meio científico/acadêmico dos profissionais cadastrados [DIGIAMPIETRI 2015]. Sistemas que extraem informações da plataforma Lattes têm como uma das principais funcionalidades a análise bibliométrica de possíveis redes de colaboração geradas pelo conhecimento obtido [FARIAS and BORGES 2013] e [CORREA et al. 2017].

A bibliometria é o estudo da produção científica em determinado contexto. Dentre as preocupações que a análise bibliométrica possui, a mais significativa dentro do escopo deste artigo é a sua utilização para controle bibliográfico que permite conhecer o tamanho e a característica do acervo: número real de produções, por exemplo [ARAÚJO 2006].

Deste modo, Research.NET Web é uma solução computacional que auxilia na observação de relacionamentos de coautoria construídos a partir dos currículos dos pesquisadores envolvidos no estudo, sendo estes obtidos através da plataforma Lattes¹. Sua funcionalidade principal é identificar o número real de publicações da rede, percebendo as coautorias e exibindo de forma gráfica os relacionamentos obtidos. A principal

¹<http://lattes.cnpq.br/>

contribuição deste trabalho é proporcionar uma visualização dinâmica da rede sem a necessidade de reprocessamento para avaliar períodos distintos.

2. Trabalhos Relacionados

Ainda que existam diversas aplicações com o propósito de analisar redes formadas por colaboração em pesquisa como ArnetMiner² e CiênciaBrasil³, este trabalho tomou como base o software Research.Net, desenvolvido na Universidade Federal do Rio Grande.

O Research.Net é “[...] um sistema de informação que facilita a análise das redes de colaboração acadêmica entre membros de uma ou mais IES” [FARIAS and BORGES 2013]. Desenvolvido na linguagem de programação Java, O sistema se apresenta em uma versão *desktop* para execução local [FARIAS and BORGES 2013].

Para gerar suas visualizações a ferramenta utiliza o software scriptLattes, que extrai informações dos currículos dos pesquisadores da rede advindos da Plataforma Lattes [MENA-CHALCO and CESAR-Jr 2009]. A utilização da plataforma Lattes se dá pela dificuldade encontrada em avaliar as produções de uma Instituição de Ensino Superior (IES), tendo em vista o alto número de publicações e que ferramentas de indexação de trabalhos acadêmicos podem não abranger as conferências e revistas científicas que contêm estas publicações [FARIAS et al. 2012].

O scriptLattes é um *script* utilizado para compilar listas de produções e tratar as duplicatas⁴ identificadas. Ademais, possibilita criação de grafos de coautoria e mapa de geolocalização dos membros inseridos na análise. Possui a licença GNU-GPL e opera de forma semiautomática devido a validação de acesso - CAPTCHA - adotada pela plataforma Lattes, o que torna sua utilização mais complexa [MENA-CHALCO and CESAR-Jr 2009].

3. Research.NET Web

O Research.NET Web é um sistema de informação utilizado para facilitar a análise de redes de colaboração acadêmicas construídas a partir do conhecimento extraído do currículo Lattes dos pesquisadores da rede. Tendo como base o Research.Net, a aplicação foi implementada do zero e esta disponível⁵ na plataforma GitHub. Sua capacidade de verificar o número real de produções de um grupo e como os pesquisadores se relacionam por meio de seus trabalhos em eventos o tornam um sistema de análise bibliométrica. O funcionamento da aplicação ocorre por meio de três módulos: o recebimento de currículos, o construtor de redes e o visualizador de redes.

O módulo de recebimento de currículos recebe arquivos no formato XML, verifica sua existência e, caso o currículo ainda não tenha sido cadastrado, salva o currículo na aplicação. A construção da rede requer que os currículos estejam cadastrados, sendo assim, uma rede só pode ser construída após a submissão dos arquivos a este módulo.

O construtor de redes recebe os pesquisadores e constrói uma nova rede através dos arquivos do diretório de currículos. Após identificar as duplicatas e quantificar o

²<http://www.arnetminer.org/index.jsp>

³<http://www.cienciabrasil.org.br>

⁴Referências bibliográficas semanticamente equivalentes

⁵<https://github.com/researchnetWeb/Research.NET-WEB>

acervo da rede a aplicação salva um arquivo XML com a estrutura gráfica que a representa. Este deverá ser executado antes do visualizador de redes.

Por fim, a aplicação conta com o módulo visualizador de redes que recebe uma rede a ser lida e requisita ao servidor o arquivo da rede (previamente construído). Após a leitura do arquivo o módulo exibe a rede conforme demonstrado na Figura 1.

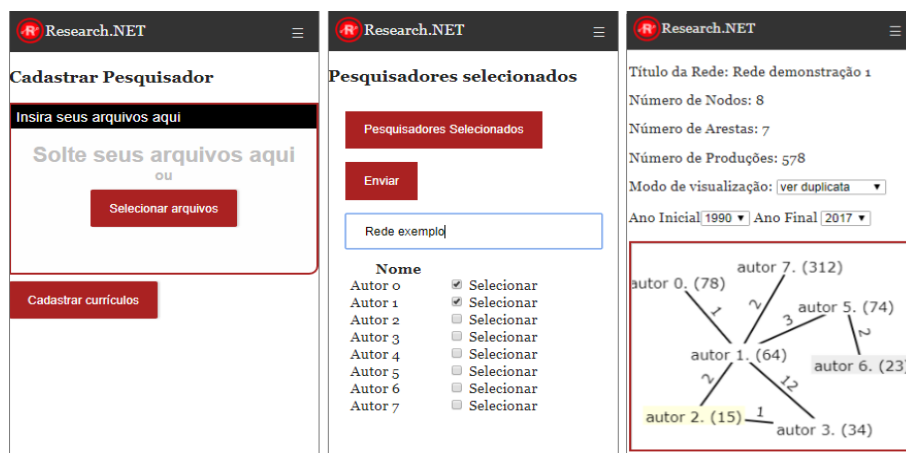


Figura 1. Interfaces gráficas das principais funcionalidades em versão mobile. A identidade dos autores foi preservada neste exemplo. Junto do nodo e aresta é possível identificar o número de produções do pesquisador e coautorias, respectivamente.

Seguindo o fluxo de funcionamento da aplicação de cadastrar os currículos, criar a rede e acessar a página de visualização, o usuário tem acesso a algumas informações sobre pesquisadores e redes cadastradas que auxiliam na análise, tais como: número de nodos ou vértices, número de arestas, número real de produções da rede (eliminando duplicatas), período de análise, relações e as produções que as constituem.

A Figura 2 apresenta uma rede exemplificando as propriedades.

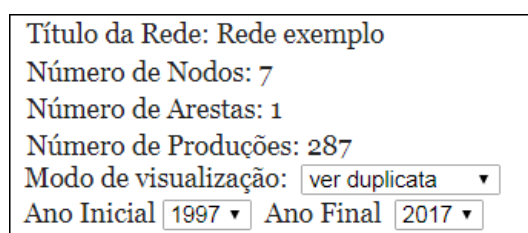


Figura 2. Informações apresentadas pelo módulo visualizador.

Ao manipular o grafo, o usuário ainda pode visualizar informações dos nodos, como nome completo do pesquisador, número de produções e grau para o período selecionado, e *link* do currículo Lattes. Além disso, consultando informações de cada aresta, são apresentadas as produções que lhe deram origem, ou seja, extraídas dos currículos dos pesquisadores e identificadas como duplicatas pelo Research.NET Web através de seu algoritmo de deduplicação que considera a distância levenshtein⁶ para comparar os títulos

⁶A distância levenshtein retorna o número de edições necessárias pra transformar um conjunto de caracteres em outro

das produções.

4. Conclusões

Este trabalho trouxe como contribuição uma solução computacional que cria redes a partir de currículos extraídos da plataforma Lattes, calcula suas métricas e exibe de forma dinâmica os grafos de saída. A análise da rede é feita para cada ano em que existam produções, isto permite visualizar a rede em períodos diferentes sem a necessidade de um novo processamento, apenas modificando o ano inicial ou final no descritor da mesma. Ademais, garante que o currículo de um pesquisador não será cadastrado duas vezes sem que o mesmo esteja com uma data de atualização mais recente.

Ainda limitada, a solução necessita de aprimoramento na inserção de métricas para um aproveitamento melhor do conhecimento obtido. Sobre a leitura de currículos, na versão atual a aplicação se limita em trabalhos em eventos e deve ser ampliada para utilizar também outros tipos de produções como artigos em periódicos, livros e capítulos de livros.

Por fim, entender uma rede e seu comportamento é um processo massivo que pode ser demorado quando não são utilizadas aplicações que o facilitem. Neste ponto, o trabalho soma ao ajudar na visualização, não apenas da rede, mas dos dados extraídos dela. Considera-se que uma nova fase do projeto irá proporcionar uma ampliação considerável no público alvo e no alcance da aplicação. Disponibilizá-lo online como um serviço da Universidade Federal do Rio Grande (FURG) pode auxiliar pesquisadores na construção e análise de redes de colaboração, principalmente em IES.

Referências

- ARAÚJO, C. A. (2006). Bibliometria: evolução histórica e questões atuais. *Em Questão*, 12(1):11–32.
- CORREA, T. S., SUZUKI, M. B., CINTRA, P. R., and COSTA, L. S. F. O. (2017). O fim do scriptlattes? uma análise de suas funcionalidades, alternativas para o presente e perspectivas para o futuro. *Revista do EDICC*, 3(3):138–148.
- DIGIAMPIETRI, L. A. (2015). Análise da rede social acadêmica brasileira. Livro docência, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo.
- FARIAS, L. R. and BORGES, E. M. (2013). Research.net: um sistema para análise de redes de colaboração baseado na plataforma lattes. Trabalho de conclusão de curso – monografia de graduação em engenharia de computação, Universidade Federal do Rio Grande, Rio Grande.
- FARIAS, L. R., VARGAS, A. P., and BORGES, E. N. (2012). Um sistema para análise de redes de pesquisa baseado na plataforma lattes. In *Escola Regional de Banco de Dados*. Sociedade Brasileira de Computação.
- MENA-CHALCO, J. P. and CESAR-Jr, R. M. (2009). scriptlattes: An open-source knowledge extraction system from the lattes platform. *Journal of the Brazilian Computer Society*, 15(4):31–39.