



**11ª Jornada Científica e
Tecnológica do IFSULDEMINAS**
& **8º Simpósio de
Pós-Graduação**

Proposta de Ambiente Interativo para a Execução Automática de Funções de Identificação de Comunidades em Grafos

Lucas P. PAIVA¹; Hugo RESENDE²; Taffarel BRANT-RIBEIRO³

RESUMO

O problema de detecção de comunidades em grafos tem sido amplamente estudado pela comunidade científica e figura como um tema de pesquisa relevante. Geralmente, analisar comunidades é uma tarefa complexa devido à atomicidade de tais comunidades. O uso do framework Igraph contribui para a implementação dos diversos algoritmos de detecção de comunidade presentes na literatura, sendo um dos frameworks mais usados na resolução do problema. Entretanto, a operacionalização de tal torna o trabalho oneroso e dispendioso por seu uso ser derivado de linhas de comando em terminais específicos. Sendo assim, planejando melhorar a manipulação de algoritmos para detecção de comunidades de grafos presentes no Igraph, o presente trabalho pretende apresentar a criação de interfaces gráficas interativas para possibilitar a execução automática de funções de detecção de comunidades em grafos. Espera-se propiciar aos pesquisadores que investigam o problema exposto uma maneira mais prática de lidar com algoritmos relacionados à temática.

Palavras-chave:

Igraph; Clusterização; Algoritmos de Clusterização; Interfaces Gráficas.

1. INTRODUÇÃO

A Teoria dos Grafos pode ser compreendida como uma das áreas com maior número de estudos dentro da programação matemática, contendo uma ampla quantidade de problemas reais que podem ser resolvidos ou representados em formas básicas de grafos, além de também estar presente em aplicações nas mais diversas áreas da engenharia (AHUJA *et al.*, 1993, BAZARAA *et al.*, 1990, HARARY, 1972). Dentro da Teoria dos Grafos, um tema bastante pesquisado se refere à detecção de comunidades, as quais são caracterizadas por serem grupos de vértices que compartilham propriedades comuns e possuem mais arestas entre si do que com vértices de outras comunidades. Aplicações, como a recuperação de informações, a segmentação de imagens, entre outras, usam ou são operacionalizadas por meio da detecção de conjuntos de vértices bem relacionados em grafos (MACHADO, 2011).

As comunidades em grafos ou agrupamentos estão presentes em diferentes tipos de estudos, para os mais variados propósitos (FORTUNATO *et al.*, 2007; RESENDE, 2014). De modo a

¹Discente, IFSULDEMINAS – Campus Passos. E-mail: luccasppaiva196@gmail.com.

²Orientador, IFSULDEMINAS – Campus Passos. E-mail: hugo.resende@ifsuldeminas.edu.br.

³Coorientador, IFSULDEMINAS – Campus Passos. E-mail: brant.ribeiro@ifsuldeminas.edu.br.

detectar grupos de vértices bem relacionados, diversos algoritmos exatos, heurísticos e hibridizados foram propostos. Vários deles utilizam de distintas técnicas, as quais são tradicionais no cenário computacional. No entanto, existe certa complexidade em se desenvolver soluções para problemas que envolvem grafos, inclusive o problema de detecção de comunidades. Por essa razão, faz-se comumente o uso de frameworks que são projetados exclusivamente para este fim.

O *Igraph* é um *framework* composto por uma coleção de bibliotecas que permite a criação e a manipulação de grafos e de diversos algoritmos relacionados à Teoria dos Grafos. Tal *framework* é escrito em linguagem de programação C e a sua versão mais estável possui pacotes em linguagens *Python* e *R*. (CSÁRDI; NEPUSZ, 2006).

A viabilidade de utilização do *Igraph* o caracterizou como um dos softwares mais utilizados em pesquisas acadêmicas relacionadas à Teoria dos Grafos (FERREZ *et al.*, 2006; HAN *et al.*, 2010). Mesmo sendo usada em diversos trabalhos e pesquisas, a biblioteca *Igraph* apresenta uma operacionalização trabalhosa. O seu uso é derivado de linhas de comando em terminais, o que torna o trabalho geralmente dispendioso.

Visando melhorar a manipulação de algoritmos para detecção de comunidades de grafos presentes no *Igraph*, o presente trabalho propõe a criação de um ambiente interativo para a execução automática de funções de detecção de comunidades em grafos que implementam esses algoritmos, tendo em vista o exposto no parágrafo anterior. Com base no que será desenvolvido, espera-se que estudiosos da área encontrem uma maior flexibilidade na manipulação de algoritmos relacionados à detecção de comunidades.

2. MATERIAL E MÉTODOS

Uma vez entendidos os principais conceitos relacionados à Teoria dos Grafos e, especificamente, ao problema de detecção de comunidades, o *framework Igraph* foi instalado e alguns algoritmos relacionados ao problema mencionado foram experimentados via linha de comando. Nesse sentido, com base no modelo de operacionalização do *framework*, as interfaces gráficas que comporão o ambiente interativo serão modeladas e alguns protótipos serão elaborados. Devido à disponibilidade dos algoritmos de detecção de comunidades no *Igraph*, serão implementados alguns dos algoritmos mais conhecidos e utilizados segundo a literatura computacional. Tais algoritmos são conhecidos como *spinglass*, *walktrap*, *infomap*, *fastgreedy*, *multilevel*, *label propagation* e *edgebetweenness*.

Alguns dos algoritmos mencionados no parágrafo anterior são baseados na maximização de uma medida de qualidade que avalia comunidades conhecida como **modularidade**. Com base no exposto, pretende-se possibilitar ao usuário o acesso à informações específicas resultantes desses algoritmos. Por exemplo, no caso do algoritmo *multilevel*, serão apresentadas ao usuário, informações referentes ao valor de modularidade, as comunidades detectadas e a qual comunidade

cada um dos vértices do grafo de entrada pertence após a execução do algoritmo. Cabe ressaltar que cada algoritmo modelado possui particularidades na sua saída, as quais serão reportadas aos usuários.

A operacionalização do ambiente consistirá no processo de o usuário inserir a entrada (grafo) por meio de um arquivo, e solicitar a execução de algum algoritmo de agrupamento. De acordo com os testes e resultados gerados antes da proposição do ambiente interativo, por linhas de comando, teremos um padrão de aceitabilidade das interfaces, no que diz respeito à corretude das funcionalidades.

Para a validação do vigente trabalho, determinadas experiências serão realizadas por meio de testes de funcionalidade com grupos de usuários, os quais são alunos ou professores da área da computação, que entendem o básico de Teoria dos Grafos. Nesse momento, caso haja quaisquer comportamentos inadequados, como erros de sincronização ou mau funcionamento, serão feitas correções no ambiente interativo, de modo que ele consiga apresentar todo o aspecto evolutivo esperado do trabalho.

Por fim, os resultados serão gerados automaticamente por meio de relatórios em *.pdf*, sendo analisados e discutidos junto à literatura correlata, de modo a compreender os pontos de melhoria, em comparação a estudos similares e à execução do *framework* no terminal através de linhas de comando. A partir dessas verificações e análises, os resultados obtidos serão submetidos para veículos de comunicação científica.

3. RESULTADOS ESPERADOS E CONSIDERAÇÕES FINAIS

Com base no que foi apresentado no decorrer deste documento, percebe-se que o *framework Igraph*, embora bastante utilizado para lidar com tarefas relacionados à Teoria dos Grafos, apresenta obstáculos na sua manipulação, realizada via linhas de código em terminais de sistemas operacionais. Particularmente, para o problema de detecção de comunidades em grafos, os algoritmos mencionados neste trabalho demandam um estudo aprofundado acerca de seus parâmetros e a sua forma de construção de soluções (comunidades).

Uma vez conhecidas as limitações do *Igraph*, espera-se, por meio de um ambiente interativo, composto por interfaces gráficas, propiciar uma melhor forma de operacionalização de algoritmos relacionados ao problemas de detecção de comunidades em grafos. Espera-se como contribuições possibilitar uma maior familiaridade de pesquisadores/estudantes em investigar o problema. Nesse sentido, os algoritmos que serão automatizados poderão ter os seus parâmetros, na maioria das vezes com um entendimento de difícil compreensão, melhor entendidos por esses interessados.

Como uma outra contribuição, este trabalho poderá servir como modelo para outros projetos que objetivam propor ambientes interativos para a execução automática de algoritmos presentes em

outros *frameworks* (para lidar com outros problemas, por exemplo) passíveis de execução somente via terminal. Como trabalhos futuros, o ambiente proposto no presente trabalho poderá ser estendido para mapear outras funções da Teoria dos Grafos disponibilizadas no *Igraph*, por exemplo, algoritmos de detecção de caminhos mínimos, algoritmos de detecção de árvores geradoras de custo mínimo.

REFERÊNCIAS

AHUJA, R., MAGNATI, T. e ORLIN, J. **Network flows: theory, algorithms, and applications**, Prentice-Hall, 1993.

BAZARAA, M., JARVIS, J. e SHERALI, H. F. **Linear programming and network flows**, John Wiley, 1990.

CSÁRDI, G.; NEPUSZ, T.: **The igraph software package for complex network research**. InterJournal Complex Systems, 2006.

FERRES, L., PARUSH, A., ROBERTS, S., LINDGAARD, G. **Helping People with Visual Impairments Gain Access to Graphical Information Through Natural Language: The iGraph System**. In: Miesenberger K., Klaus J., Zagler W.L., Karshmer A.I. (eds) Computers Helping People with Special Needs. ICCHP 2006. Lecture Notes in Computer Science, vol 4061. Springer, Berlin, Heidelberg, 2006.

FORTUNATO, S.; BARTHELEMY, M. **Resolution limit in community detection**. Proceedings of the National Academy of Sciences, 2007.

HAN, W., LEE, J., PHAM, M., XU YU, Jeffrey. **IGraph: a framework for comparisons of disk-based graph indexing techniques**, Proceedings of the VLDB Endowment, v.3 n.1-2, September 2010.

HARARY, F. **Graph Theory**, Addison-Wesley, 1972.

MACHADO, F. M. **Detecção de comunidades em grafos multicamada muito grandes**. 2011. Dissertação (Mestrado) - Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, 2011.

RESENDE, H. **Algoritmos paralelos para determinar agrupamentos em grafos com única ou múltiplas camadas**. Dissertação (Mestrado) - Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo (UNIFESP), São José dos Campos, 2014.