LI Simpósio Brasileiro de Pesquisa Operacional
Limeira–SP, 02 a 06 de setembro de 2019

Um algoritmo genético para a detecção de comunidades com *overlapping* em redes de passagens aéreas

Igor Luppi
Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo
Avenida Cesare Mansueto Giulio Lattes, 1201 - Eugênio de Mello, CEP: 12247-014
igor.luppi@gmail.com

Camila P. S. Tautenhain
Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo
Avenida Cesare Mansueto Giulio Lattes, 1201 - Eugênio de Mello, CEP: 12247-014
santos.camila@unifesp.br

Mariá C. V. Nascimento
Instituto de Ciência e Tecnologia, Universidade Federal de São Paulo
Avenida Cesare Mansueto Giulio Lattes, 1201 - Eugênio de Mello, CEP: 12247-014
mcv.nascimento@unifesp.br

RESUMO

Empresas aéreas produzem um grande volume de dados de passagens. Esses dados podem ser modelados como dígrafos, nos quais os vértices representam os aeroportos e os arcos os voos entre os aeroportos. Agrupar os vértices de tais dígrafos pode auxiliar outras aplicações, como, por exemplo, na solução heurística de um problema de otimização que consiste em encontrar uma rota aérea que minimize custos de passagens, considerando tempos de estadia e de forma que o viajante visite todas as cidades. Nesse contexto, o agrupamento dos vértices desses dígrafos visa identificar grupos de aeroportos que possuam maior conectividade intra-grupo do que inter-grupos. Neste trabalho foi desenvolvido um algoritmo para o agrupamento de dígrafos em que os vértices podem pertencer a mais de uma comunidade (*overlapping*). Ele foi baseado nos métodos de propagação de rótulos *Speaker-Listener Label Propagation Algorithm* (SLPA) e *Genetic Algorithm - Label Propagation* (GA-LP) da literatura. O SLPA é uma extensão do algoritmo *Label Propagation Algorithm* (LPA) capaz de detectar comunidades com *overlapping*. O GA-LP, por sua vez, é um algoritmo genético para agrupar vértices em dígrafos com custo computacional linear. Como o GA-LP foi proposto inicialmente para lidar com comunidades disjuntas, no algoritmo proposto a sua função *fitness* foi substituída por uma medida da literatura baseada na de modularidade para avaliar comunidades com *overlapping*. Para avaliar os resultados do algoritmo proposto, foram realizados experimentos computacionais com dígrafos compilados a partir de uma base de dados que contém informações de cerca de 13 milhões de voos, dispostos em um período de 1 ano em 2400 aeroportos. Em particular, um experimento com um dígrafo que representa os voos ocorridos nos principais aeroportos do Brasil ao longo de um período de quatro bimestres mostrou que os voos tendem a se manter estáveis ao longo dos bimestres, com poucas modificações em suas rotas, provenientes de eventuais sazonalidades. Por exemplo, todos os aeroportos do período de novembro à dezembro foram classificados em comunidades com *overlapping*, indicando um maior número de conexões aéreas com aeroportos em diferentes comunidades na época de fim de ano. Além disso, os aeroportos tendem ser classificados em comunidades em comum com *hubs* aéreos, ou seja, aeroportos com muitas conexões. Em experimentos com dígrafos obtidos a partir da seleção aleatória de aeroportos da base de dados, os resultados obtidos pelo algoritmo foram próximos aos da literatura. Nesse sentido, trabalhos futuros envolvem a aplicação do algoritmo proposto para servir de apoio à problemas que utilizem esses dados de passagens aéreas como objeto de estudo.

https://proceedings.science/p/106963?lang=pt-br

PALAVRAS CHAVE. Detecção de comunidades em dígrafos; Comunidades com *overlapping*; Redes de passagens aéreas

TAG – Teoria e Algoritmos em Grafos; MH – Metaheurísticas