UTILIZAÇÃO DE METAHEURÍSTICAS ILS E VND PARA RESOLVER O PROBLEMA DE CLUSTERIZAÇÃO EM GRAFOS ORIENTADOS

Viviane Cátia Köhler
Universidade Federal do Rio de Janeiro
COPPE, Programa de Engenharia de Sistemas e Computação,
Caixa Postal 68511,
21945-970,Rio de Janeiro, RJ
vivianeck@cos.ufrj.br

Márcia Helena Costa Fampa
Universidade Federal do Rio de Janeiro
COPPE, Programa de Engenharia de Sistemas e Computação,
Caixa Postal 68511,
21945-970,Rio de Janeiro, RJ
fampa@cos.ufrj.br

Luiz Satoru Ochi
Universidade Federal Fluminense
Instituto de Computação
Rua Passo da Pátria, 156, Bloco E, São Domingos
24210-240 Niterói, RJ
satoru@ic.uff.br

RESUMO

Os problemas de clusterização tem aplicação em engenharia de software, que muitas vezes trata com grandes sistemas que tendem a ter uma estrutura rica e complexa, que pode ser descrita a partir de um ou mais grafos dirigidos. Por exemplo, os nós e arestas de um grafo direcionado pode descrever os módulos (ou classes) de um sistema e suas inter-relações estáticas, respectivamente. Este grafo é denominado na literatura de grafo dependente modular (Module Dependency Graph, MDG). O MDG pode ser de grande dimensão e complexo. Uma maneira de torná-lo mais acessível consiste em encontrar partições de forma a separar os nós (módulos) em clusters (subsistemas).

Encontrar uma boa partição para um MDG é agrupar os vértices do grafo de tal forma que os grupos formados, ou clusters, tenham muitos arcos dentro do cluster e poucos arcos entre os cluster, o que caracteriza uma maior similaridade no intracluster e uma menor similaridade no intercluster, o qual é medido utilizando a função objetivo da qualidade da modularização (*Modularization Quality* -MQ), função é não-linear fracionária e inteira, o que dificulta resolver este problema de forma exata e devido a isso foi utilizado metaheurísticas.

As metaheurísticas propostas são baseados em Variable Neighborhood Descent (VND) e Iterated Local Search (ILS) com VND. O ILS+VND apresentou, em alguns casos, resultados melhores que simplesmente o VND, e na maior parte dos resultados ILS+VND e VND apresentaram resultados iguais. A maior contribuição deste trabalho é a qualidade da solução para instâncias grandes, uma vez que o algoritmos propostos sempre apresentaram soluções melhores ou iguais que os algoritmos evolutivos propostos na literatura, com um ganho de até 58%.

PALAVRAS CHAVE: clusterização em grafos, metaheurísticas, grafos.

ABSTRACT

The problems of clustering have applications in software engineering, which often deals with large systems tend to have a rich and complex structure, which can be described from one or more directed graphs. For example, nodes and edges of a directed graph can describe the module (or classes) of a system and their interrelationships static respectively. This graph is called in the literature of Module Dependency Graph (MDG). The MDGs can be large and complex. One way to make it more accessible is to find partitions to separate nodes (modules) into clusters (subsystems).

Finding a good partition for an MDG is to cluster the nodes of the graph so that the formed groups, or clusters, have many arcs in the cluster and few between the arches cluster, which features a greater similarity in the intracluster and a lower similarity in intercluster, which is measured using the objective function of the quality of modularization (Modularization Quality-MQ) function is non-linear fractional and integer, which makes much more difficult solving this problem exactly andfor this reason metaheuristics were used.

The proposed metaheuristics are based on Variable Neighborhood Descent (VND) and Iterated Local Search (ILS) with VND. ILS+VND presented in some cases perform better than simply VND, and most of the results ILS+VND and VND have similar results. The major contribution of this work is the quality of the solution for large instances, since the proposed algorithms always had equal or better solutions than the evolutionary algorithms proposed in the literature, with a gain of up to 58%.

KEYWORDS: graph clustering, metaheuristics, graphs.