Information Sciences 426 (2018) 101–116

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

An order-based algorithm for minimum dominating set with application in graph mining

David Chalupa [a,b]

[a] Operations Research Group, Department of Materials and Production, Aalborg University, Fibigerstræde 16, Aalborg 9220, Denmark
[b] Computer Science, School of Engineering and Computer Science, University of Hull, Cottingham Road, Hull HU6 7RX, United Kingdom

ARTICLE INFO

Article history:
Received 20 May 2017
Revised 12 October 2017
Accepted 13 October 2017
Available online 16 October 2017

Keywords:
Randomised local search
Order-based representation
Minimum dominating set
Complex networks
Heuristics

ABSTRACT

Dominating set is a set of vertices of a graph such that all other vertices have a neighbour in the dominating set. We propose a new order-based randomised local search (RLS_o) algorithm to solve minimum dominating set problem in large graphs. Experimental evaluation is presented for multiple types of problem instances. These instances include unit disk graphs, which represent a model of wireless networks, random scale-free networks, as well as samples from two social networks and real-world graphs studied in network science. Our experiments indicate that RLS_o performs better than both a classical greedy approximation algorithm and two metaheuristic algorithms based on ant colony optimisation and local search. The order-based algorithm is able to find small dominating sets for graphs with tens of thousands of vertices. In addition, we propose a multi-start variant of RLS_o that is suitable for solving the minimum weight dominating set problem. The application of RLS_o in graph mining is also briefly demonstrated.

1. Introduction

Dominating set of a graph is a set of its vertices such that each vertex is in the dominating set or has a neighbour in the dominating set. Dominating sets and their variants have applications in several diverse areas, including routing in wireless ad-hoc networks [11], multi-document summarisation [45] or modelling and studying of positive influence in social networks [12]. The problem of finding the minimum dominating set (MDS) is widely known to be NP-hard [19,32,37].

Let $G = [V, E]$ be an undirected graph and let $S \subseteq V$. Then, S is a *dominating set* if each vertex $v \in V$ is in S or is adjacent to a vertex in S. Dominating set with the lowest cardinality is called *minimum dominating set*, its cardinality is called *domination number* and is denoted by γ.

Fig. 1 presents an illustration of the dominating set for a sample of a social graph from Google+ with 200 vertices. Both drawings in the figure represent the same data. On the left hand side, the vertex with maximum degree is placed in the middle and other vertices are arranged into levels, based on their distance from the central vertex. On the right hand side, a dominance drawing of the network is presented, in which the dominating set is used to visually organise the network [2]. Vertices of the dominating set are highlighted in red and other vertices are grouped into clusters, effectively using the fact that each of these vertices has a neighbour in the dominating set. Dominating set can be used as a set of hubs of the network to form fine-grained clusters, since all other vertices have a neighbour in the dominating set. There is a large body of literature on dominating sets and their variations. The main focus in MDS literature is on theoretical aspects and

E-mail address: dc@m-tech.aau.dk

https://doi.org/10.1016/j.ins.2017.10.033

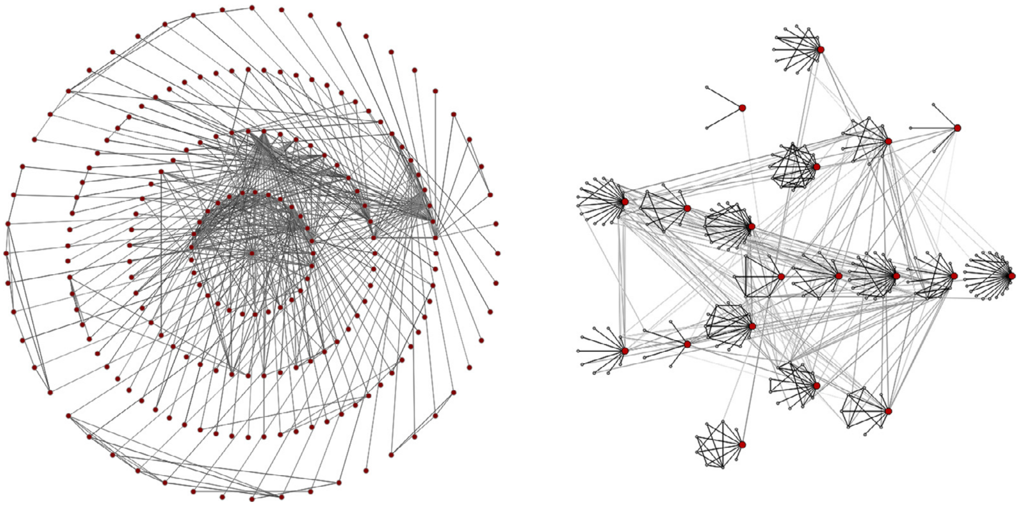

Fig. 1. Two visualisations of a sample social graph from Google+ with 200 vertices. On the left hand side, the vertex with maximum degree is taken as the central vertex and other vertices are visualised in a radial drawing, based on distance from the centre. On the right hand side, vertices are partitioned around vertices in the dominating set.

applications in wireless ad-hoc networks. NP-hardness and approximation characteristics of the problem suggest that works aimed at design of efficient heuristics and their scalability are of an interest. However, solving techniques for MDS are also of an increasing interest for large graphs due to their applications in social and information networks.

The current core of experimental literature is focused mainly on benchmarking and applications of algorithms for weighted dominating sets [5,24,28,41,48] and connected dominating set problems [11,23]. Design of scalable algorithms for large-scale instances of MDS and their real-world applications in social and information networks seem to be the areas to explore further. In this paper, we propose a new heuristic algorithm for MDS, which is highly scalable to large graphs and tends to perform better than a greedy approximation algorithm, as well as algorithms based on constructive approaches and ant colony optimisation with local search.

Our *order-based randomised local search* (RLS_o) algorithm tackles MDS indirectly by using a representation based on permutations of vertices, which are transformed into dominating sets using a greedy algorithm. The permutation of vertices is then optimised by repeated using of specific randomised *jump* moves.

Experiments were carried out using a wide collection of both synthetic and real-world graphs. We compare RLS_o to a classical greedy approximation algorithm, as well as a more recently proposed ant colony optimisation algorithm hybridised with local search (ACO-LS) [40] and its extension with preprocessing (ACO-PP-LS) [41]. ACO-PP-LS was originally designed for the minimum weight dominating set problem.

We first present results of RLS_o for unit disk graphs, since these graphs were used for evaluation in previous studies as models of wireless networks. Additionally, we provide results for scale-free networks generated by Barabási-Albert (BA) model. We generally obtain that RLS_o performs better than the greedy approximation algorithm, ACO-LS and ACO-PP-LS.

Results are also provided for an extensive collection of real-world graphs studied in network science, instances from DIMACS series, as well as samples of anonymised publicly available data from social networks *Google+* and *Pokec*. These results also confirm that RLS_o provides results of better quality than the greedy approximation algorithm, ACO-LS and ACO-PP-LS, while maintaining solid scalability for large graphs. In addition, the solutions found by RLS_o tend to be close to lower bounds, which have been computed as solutions to the linear programming relaxation of MDS. This relaxation represents the linear programming problem obtained from the formulation of MDS by assuming that decision variables can be any real values between 0 and 1, instead of binary variables.

RLS_o is also extended to a multi-start algorithm $MSRLS_o$ to solve the minimum weight dominating set (MWDS) problem. Finally, an application of RLS_o in graph mining is briefly discussed.

The structure of our paper is as follows. In Section 2, we review the background of the problem and related work. In Section 2.1, we describe our local search algorithm RLS_o for MDS. In Section 4, the multi-start local search algorithm $MSRLS_o$ for MWDS is presented. In Section 5, we present the experimental results and provide a short discussion. Finally, in Section 6, we summarise the contributions and identify open problems.

2. Background and related work

The problem of finding MDS remains NP-hard also for several very restricted graph classes. For example, NP-hardness of finding MDS for grids is known, the proof is attributed to Leighton [10]. For unit disk graphs, MDS problem is also NP-hard