European Journal of Operational Research 000 (2016) 1–15

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Discrete Optimization

The generalized independent set problem: Polyhedral analysis and solution approaches

Marco Colombi[a,*], Renata Mansini[a], Martin Savelsbergh[b]

[a] Department of Information Engineering, University of Brescia, Via Branze 38 - 25123 Brescia, Italy
[b] H Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 755 Ferst Drive, NW, Atlanta, GA 30332, USA

ARTICLE INFO

Article history:
Received 21 October 2015
Accepted 28 November 2016
Available online xxx

Keywords:
Combinatorial optimization
Generalized independent set
Facets
0–1 quadratic programming
GRASP tabu search

ABSTRACT

In the Generalized Independent Set Problem, we are given a graph, a revenue for each vertex, and a set of *removable* edges with associated removal costs, and we seek to find an independent set that maximizes the net benefit, i.e., the difference between the revenues collected for the vertices in the independent set and the costs incurred for any removal of edges with both endpoints in the independent set. We study the polyhedron associated with a 0–1 linear programming formulation of the Generalized Independent Set Problem, deriving a number of facet-inducing inequalities, and we develop linear programming based heuristics to obtain high-quality solutions in a short amount of time. We also develop a heuristic method based on an unconstrained 0–1 quadratic programming formulation of the Generalized Independent Set Problem. In an extensive computational study, we assess the performance of these heuristics in terms of quality and efficiency. The best heuristic is then used to produce an initial solution for a branch-and-cut algorithm which uses some of the proposed facet-inducing inequalities.

1. Introduction

Let $G = (V, E)$ be a graph with vertex set V and edge set $E = E_1 \cup E_2$, where E_1 is a set of non-removable edges and E_2 is a set of removable edges. A revenue $w_i > 0$ is associated with each vertex $i \in V$, and a cost $c_{ij} > 0$ is associated with each removable edge $(i, j) \in E_2$. The goal is to find an independent set, i.e., a set of vertices such that no two vertices in the set are adjacent, that maximizes the difference between the total revenue associated with the vertices in the set and the total cost associated with the removal of edges with both endpoints in the set.

By assigning a zero cost to non-existing edges and by assigning an appropriately chosen large cost to non-removable edges, the problem can be seen to be equivalent to the Generalized Independent Set Problem (GISP) introduced by Hochbaum and Pathria (1997) in the context of forest management and harvesting. The authors consider a setting in which a forest is partitioned into cells and the problem is to decide which cells to harvest and which cells to leave unharvested. More specifically, a benefit is associated with harvesting a cell and a penalty is associated with harvesting adjacent cells. For a setting in which the cells of the forest can be represented with a bipartite graph, the authors present an efficient solution algorithm based on network flow methods. Additional results on GISP (as well as on a number of other combinatorial optimization problems) can be found in Hochbaum (2004).

More recently, Kochenberger, Alidaee, Glover, and Wang (2007) have proposed an unconstrained binary quadratic programming (UBQP) formulation for the GISP. The formulation has the advantage that it does not require (binary) variables associated with the removable edges, but only (binary) variables associated with the vertices. As the number of vertices is typically much smaller than the number of removable edges, the resulting formulation is relatively small, but, of course, comes at a price: a quadratic rather than linear objective function. The authors propose a tabu search algorithm that exploits the fact that all 0–1 vectors represent feasible solutions (although some with very high cost - representing the inclusion of vertices in the independent set that are connected by a non-removable edge) and demonstrate its effectiveness with a set of computational experiments using instances of varying size and density.

To the best of our knowledge, these are the only papers discussing GISP. Of course, a vast literature exists on the finding a maximum weight independent set in a graph, which is a special case of GISP. We refer interested readers to Mannino and Sassano (1994), Rebennack (2009), Warrier, Wilhelm, Warren, and Hicks (2005), Xiao and Nagamochi (2013) for exact algorithms, and to Bonomo, Durán, Lin, and Szwarcfiter (2006),

* Corresponding author.
E-mail addresses: m.colombi006@unibs.it (M. Colombi), renata.mansini@unibs.it (R. Mansini), martin.savelsbergh@isye.gatech.edu (M. Savelsbergh).

http://dx.doi.org/10.1016/j.ejor.2016.11.050

Burer, Monteiro, and Zhang (2002), Cogis and Thierry (2005), Feo, Resende, and Smith (1994), Lehmann, Kaufmann, Steigele, and Nieselt (2006), Marchiori (2002), Pelillo (2009) for heuristic methods and polynomial-time algorithms for special classes of graphs.

In this paper, we provide the first polyhedral analysis of GISP. We show that the polyhedron associated with a 0–1 linear programming formulation of GISP is full-dimensional, and study several classes of valid inequalities, all natural generalizations of valid inequalities for the traditional maximum weight independent set problem. In many cases, we identify conditions under which these inequalities are facet-inducing. In addition to the polyhedral analysis, we develop linear programming (LP) based heuristics in which the LP relaxation is strengthened by adding some of the valid inequalities derived. We also develop a meta-heuristic approach exploiting the binary quadratic programming formulation of GISP. That method is an adaptation of the probabilistic GRASP - tabu search algorithm proposed in Wang, Lü, Glover, and Hao (2013). We analyze the performance of all the heuristic methods on a large set of randomly generated instances by comparing their solutions to the one obtained by an integer programming (IP) solver (CPLEX 12.5.1) within one hour using a formulation enhanced with some randomly generated valid inequalities.

We solve instances with up to 400 vertices, with more than 70,000 edges, and with different fractions of removable edges. The computational results demonstrate that all the proposed heuristics are effective in producing high-quality solutions, but that the heuristic based on the unconstrained binary quadratic programming formulation is by far the most efficient. As part of our computational investigation, we also obtain a better understanding of the problem characteristics that impact the difficulty of GISP instances.

It is worth noting that unconstrained binary quadratic programming formulations can be used to model a wide range of combinatorial optimization problems in different application domains (a review can be found in Kochenberger, Glover, Alidaee, & Rego (2004)). Various metaheuristics have been developed that can find near-optimal solutions to large UBQP instances in a short amount of time and, therefore, can provide a valid alternative to methods based on 0–1 linear programming formulations. On the other hand, existing exact solution methods for UBQP become computationally prohibitive even for relatively small instances and, therefore, exact methods based on 0–1 linear programming formulations still have a computational edge. Although the development of a full-blown exact method for GISP is out of the scope of the present paper, we conduct a small experiment with a branch-and-cut algorithm based on the 0–1 linear programming formulation and that uses some of the proposed valid inequalities, this time selected using a high-quality feasible solution produced by the GRASP - tabu search heuristic.

The remainder of the paper is organized as follows. In Section 2, we present a 0–1 linear programming formulation as well as an unconstrained binary quadratic programming formulation for GISP. In Section 3, we discuss the results of our polyhedral analysis. In Section 4, we introduce the LP-based heuristics, the GRASP - tabu search heuristic, and a branch-and-cut algorithm, whereas, in Section 6, we analyze their performance on a wide set of well-structured instances. Finally, in Section 7, we provide some concluding remarks and indicate possible directions for future research.

2. Problem formulations

To formulate GISP, we introduce binary variables x_i for $i \in V$, indicating whether vertex i is selected and the associated revenue is collected ($x_i = 1$) or not ($x_i = 0$), and binary variables y_{ij} for $(i, j) \in E_2$, indicating whether edge (i, j) is removed and the associated cost incurred ($y_{ij} = 1$) or not ($y_{ij} = 0$). With these variables, GISP can be formulated as follows:

$$\max \sum_{i \in V} w_i x_i - \sum_{(i,j) \in E_2} c_{ij} y_{ij} \tag{1}$$

$$\text{subject to } x_i + x_j \leq 1 \quad (i, j) \in E_1 \tag{2}$$

$$x_i + x_j - y_{ij} \leq 1 \quad (i, j) \in E_2 \tag{3}$$

$$x_i \in \{0, 1\} \quad i \in V \tag{4}$$

$$y_{ij} \in \{0, 1\} \quad (i, j) \in E_2. \tag{5}$$

The objective function (1) maximizes the net benefit, i.e., the difference between the sum of the revenues of the selected vertices and the sum of the costs of removed edges. Constraints (2) are the classical independent set constraints, capturing the restriction that vertices linked by a fixed edge cannot be in an independent set together, whereas constraints (3) capture the fact that vertices linked by a removable edge can be in an independent set together only if that edge is removed. Finally, constraints (4) and (5) define the integrality conditions.

To formulate GISP as an unconstrained binary quadratic program, we again use binary variables x_i for $i \in V$ to indicate whether vertex i is selected ($x_i = 1$) or not ($x_i = 0$), but capture the removal of an edge $(i, j) \in E$ by the product of the variables associated with the nodes at its end points, i.e., $x_i x_j$. More specifically,

$$\max \sum_{i \in V} w_i x_i - \sum_{(i,j) \in E} c_{ij} x_i x_j$$

$$\text{subject to} \qquad x_i \in \{0, 1\} \qquad i \in V,$$

where c_{ij} is set to a sufficiently large value for edges $(i, j) \in E_1$ to ensure that the vertices at the end points of these edges will never be selected together in an optimal solution.

3. Polyhedral analysis

Let P be the convex hull of all feasible solutions to the 0–1 linear programming formulation of GISP, i.e.,

$$P = conv\{(x, y) \in \mathbb{Z}_+^{|V|+|E_2|} \mid (x, y) \text{ satisfies } (2), (3), x_i \leq 1, i \in V, \text{ and } y_{ij} \leq 1, (i, j) \in E_2\}.$$

Proposition 1. *The GISP polyhedron P is full-dimensional, i.e.,* $dim(P) = |V| + |E_2|$.

Proof. The following $|V| + |E_2| + 1$ feasible points are trivially affinely independent:

- For each $i \in V$, the point with $x_i = 1$ and all other variables equal to zero.
- For each $(i, j) \in E_2$, the point with $y_{ij} = 1$ and all other variables equal to zero.
- The zero vector (i.e., all variables equal to zero). □

Let $S \subseteq V$ be a subset of vertices. We define $E_2(S) = \{(i, j) \in E_2 : i, j \in S\}$ and $\delta_2(S) = \{(i, j) \in E_2 : i \in S, j \notin S\}$ to be the set of removable edges with both endpoints in S and the cut set with only one endpoint in S, respectively. Let $\Gamma_1(i)$ be the set of vertices connected to i with a fixed edge and let $\Gamma_2(i)$ be the set of vertices connected to i with a removable edge, i.e., $\Gamma_1(i) = \{j \in V : (i, j) \in E_1\}$ and $\Gamma_2(i) = \{j \in V : (i, j) \in E_2\}$. Thus, $|\Gamma_1(i)|$ and $|\Gamma_2(i)|$ give the degree of vertex i with respect to edge set E_1 and E_2, respectively. The overall degree $|\Gamma(i)|$ of vertex i is $|\Gamma_1(i)| + |\Gamma_2(i)|$. Finally, we let $\bar{E}$ denote the set of edges in the complement graph $\bar{G}$, i.e., the graph with the same set of vertices, but with the edges that are not in E.

Please cite this article as: M. Colombi et al., The generalized independent set problem: Polyhedral analysis and solution approaches, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.11.050