European Journal of Operational Research 237 (2014) 62–70

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Discrete Optimization

A fast approximation algorithm for solving the complete set packing problem

Tri-Dung Nguyen *

Centre of Operational Research, Management Sciences and Information Systems, University of Southampton, Southampton SO173TJ, UK

ARTICLE INFO

Article history:
Received 19 June 2013
Accepted 12 January 2014
Available online 20 January 2014

Keywords:
Set packing problem
Combinatorial auctions
Winner determination problem
Coalition structure generation
Large-scale MILP

ABSTRACT

We study the complete set packing problem (CSPP) where the family of feasible subsets may include all possible combinations of objects. This setting arises in applications such as combinatorial auctions (for selecting optimal bids) and cooperative game theory (for finding optimal coalition structures). Although the set packing problem has been well-studied in the literature, where exact and approximation algorithms can solve very large instances with up to hundreds of objects and thousands of feasible subsets, these methods are not extendable to the CSPP since the number of feasible subsets is exponentially large. Formulating the CSPP as an MILP and solving it directly, using CPLEX for example, is impossible for problems with more than 20 objects. We propose a new mathematical formulation for the CSPP that directly leads to an efficient algorithm for finding feasible set packings (upper bounds). We also propose a new formulation for finding tighter lower bounds compared to LP relaxation and develop an efficient method for solving the corresponding large-scale MILP. We test the algorithm with the winner determination problem in spectrum auctions, the coalition structure generation problem in coalitional skill games, and a number of other simulated problems that appear in the literature.

1. Introduction and literature review

1.1. The Set Packing Problem (SPP)

The set packing problem and its variants (the set covering and set partitioning problems) are among the most well-studied problems in combinatorial optimisation thanks to their wide ranges of applications, their elegant mathematical formulation, and their special structural properties. In the SPP, there are n objects which can be packed into a number of subgroups among m predefined feasible subsets labelled as $\mathcal{S}_1, \ldots, \mathcal{S}_m$. Each subset $\mathcal{S}_j$ has a payoff value of v_j. The SPP aims to divide these n objects into non-overlapping subgroups such that their total payoff is maximised. The SPP can be formulated as an MILP as follows:

$$\begin{aligned} \mathbf{SPP}(\boldsymbol{A}, \boldsymbol{v}) := \max_{\boldsymbol{x}} \ & \boldsymbol{v}^t \boldsymbol{x} \\ \text{s.t. } & \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{e}, \\ & \boldsymbol{x} \in \{0,1\}^m, \end{aligned} \tag{1}$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)$ is a vector of binary decision variables with x_j indicating whether subset $\mathcal{S}_j$ is selected in the packing, $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ is a matrix with element a_{ij} in row i and column j indicating whether subset $\mathcal{S}_j$ contains object i, $\boldsymbol{v} \in \mathbb{R}^m$ is a vector of payoffs, and $\boldsymbol{e} \in \mathbb{R}^n$ is a vector with all elements being equal to one.

The SPP has many applications such as for routing and scheduling trains at intersections in railway operations Zwaneveld, Kroon, and Van Hoesel (2001), for selecting winning bids in combinatorial auctions (De Vries & Vohra, 2003), for surgical operations scheduling Velásquez and Melo (2006), and for packets scheduling and transmission in communication networks Emek et al. (2012). In the context of combinatorial auctions, the winner determination problem (WDP) is essentially a set packing problem. Sandholm (2002) develops an algorithm that utilises the graphical representation of the coalition structure search space for solving the WDP. Fujishima, Leyton-Brown, and Shoham (1999) develop an exact algorithm, where caching and pruning are used to speed up the search, and a heuristic algorithm for solving the WDP.

The tractability of the SPP depends on the structure of the underlying IP formulation. Specifically, Müller (2006) and Rothkopf, Pekeč, and Harstad (1998) summarise special cases where the corresponding LP relaxation solutions satisfy the integrality constraints and hence are also solutions of the SPP. These are, however, very restrictive cases and it is generally very difficult to solve the SPP. In fact, Karp (1972) shows that the SPP problem is NP-complete while Sandholm (2002) shows the inapproximability of the problem for general cases. Many methods, both exact and approximation, have been proposed for solving the SPP. Padberg (1973) and Cánovas, Landete, and Marın (2000) show different sets of facets of the set packing polyhedron which can be used to strengthen the LP relaxation solutions. Landete, Monge, and Rodríguez-Chía (2012) present an alternative formulation for the

* Tel.: +44 (0)23 8059 7759.
E-mail address: T.D.Nguyen@soton.ac.uk

0377-2217/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2014.01.024

SPP in a higher-dimensional space where a set of facets can be identified.

Methods for solving the SPP often start with solving the corresponding LP relaxation problem. De Vries and Vohra (2003) survey different methods such as a constraint generation method for solving the LP relaxation problem and a sub-gradient method for solving the Lagrangian relaxation. The authors also provide interesting insights on how the numerical algorithm is interpreted in the auctioning process. These methods have actually been well-studied in the context of the set covering problem (SCP), a variant of the SPP where the objective is to minimise the total cost of covering all the objects (see Beasley (1987) and Beasley & Jörnsten (1992) for examples). Caprara, Toth, and Fischetti (2000) survey methods to solve the SCP and compare their numerical performance on test problems that appear in the literature.

Kochenberger, Glover, Alidaee, and Rego (2004) provide an unified framework for solving combinatorial optimisation problems by transforming them into unconstrained quadratic binary optimisation problems (UQBP). The authors then suggest the use of Tabu search, a metaheuristic method that employs local search and a 'tabu'-list to keep track of the searched space, for solving the UQBP. Alidaee, Kochenberger, Lewis, Lewis, and Wang (2008) and Lewis, Kochenberger, and Alidaee (2008) apply these methods to the set packing and set partitioning problems and show that the algorithms outperform CPLEX (on the original MILP formulations) on many moderate-sized instances (with up to $n = 5000$ objects and $m = 15,000$ feasible subsets for the SPP case).

There are also many other heuristic methods for solving the SPP. In fact, Hoffman and Padberg (2001) state that "virtually every heuristic approach for solving general integer programming problems has been applied to the set-covering, packing and partitioning problems." Delorme, Gandibleux, and Rodriguez (2004) develops GRASP, a greedy randomised algorithm for solving the set packing problem. Beasley and Chu (1996) develop a genetic algorithm for solving the set covering problem and this method can be adapted to solve the SPP.

1.2. The Complete Set Packing Problem (CSPP)

In this paper, we aim to solve the SPP for cases when $m = 2^n$, i.e. any subset of objects can be grouped together in a packing, or when m is relatively large compared to n. This setting arises in applications such as combinatorial auctions where bidders submit their bids in the form of value functions on the objects selected. This bidding mechanism is favourable to auction designers and to bidders because the information can be communicated in a more compact way. Another application area is in multi-agent systems, e.g. in a sensor network (Dang, Dash, Rogers, & Jennings, 2006), where players are grouped into coalitions to maximise their total utility. As the number of possible subsets can grow exponentially, existing methods (such as Beasley & Jörnsten, 1992; Caprara et al., 2000; Fujishima et al., 1999; Sandholm, 2002) are not applicable due to the large number of binary decision variables involved.

Let $\mathcal{N} = \{1, \ldots, n\}$ be the set of all objects and let $\boldsymbol{x} = (x_1, x_2, \ldots, x_{2^n})$ be a vector of binary variables with x_j indicating whether subset $\mathcal{S}_j$ is selected in the packing. The CSPP can be formulated as an MILP as follows:

$$\mathbf{CSPP}(\mathcal{N}, \boldsymbol{v}) := \max_{\boldsymbol{x}} \ \boldsymbol{v}^t \boldsymbol{x} \quad \text{s.t. } \boldsymbol{A}_{\mathcal{N}} \boldsymbol{x} \leqslant \boldsymbol{e}, \quad \boldsymbol{x} \in \{0,1\}^{2^n}, \tag{2}$$

where $\boldsymbol{A}_{\mathcal{N}} \in \mathbb{R}^{n \times 2^n}$ is a matrix with element a_{ij} in row i and column j indicating whether subset $\mathcal{S}_j$ contains object i. For convenience in notation, let $\boldsymbol{a}_j = (a_{1j}, \ldots, a_{nj})^t$ be a column vector of binary indicators for each $j \in \{1, \ldots, 2^n\}$. To avoid ambiguity in the ordering of $\boldsymbol{a}_j$, we assign $\boldsymbol{a}_j$ to the binary representation of $(j-1)$. Let us denote $v_j \equiv v(\boldsymbol{a}_j) \equiv v(\mathcal{S}_j)$ as the payoff of subset $\mathcal{S}_j$. For $n \leqslant 15$, problem $\mathbf{CSPP}(\mathcal{N}, \boldsymbol{v})$ can be solved efficiently by CPLEX through a classical branch and bound technique. However, the size of the MILP problem grows exponentially as the number of objects increases and it is impossible for CPLEX to solve instances with more than 20 objects. We aim to develop an approximation method for solving this MILP.

1.3. The winner determination problem in combinatorial auctions

Combinatorial auctions have been used in the procurement of London bus routes Cantillon and Pesendorfer (2006), radio spectrum Cramton (1997), and truckload transportation Caplice and Sheffi (2006), among many others. Combinatorial auctions arise in situations where bidders are interested in buying bundles of objects that inherit some level of synergies among themselves. One of the key problems in combinatorial auction is to find the best feasible combination of bids to maximise the total payoff. This problem is equivalent to a complete set packing problem where objects are those to be sold and the payoff of each subset is the maximum bid that the bidders offer. In the combinatorial auction literature, solution approaches such as Fujishima et al. (1999) and Sandholm (2002) often assume that the number of bids are relatively small compared to the number of objects, i.e. a few hundreds of objects and a few thousands of bids at most. However, in many real-life situations such as in spectrum auctions, bidders might be interested in buying any subset of their predefined frequencies. In this case, the bidders may express their interest through a compact value function that involves their objects of interest and their specific synergy parameters (Cramton, Ausubel, McAfee, & McMillan, 1997; De Vries & Vohra, 2003). Therefore, the set of feasible bids from all the bidders is an exponential function of the number of objects. We discuss about one such case in spectrum auctions in subSection 3.1.

1.4. The optimal coalition structure generation problem in cooperative game theory

Cooperative games with transferable utilities belong to a branch of game theory where groups of players can form coalitions in order to jointly achieve the groups' objectives. Cooperative game theory has many applications in economics and business (e.g. for setting insurance premiums (Lemaire, 1991), and for setting interchange fees for ATM bank networks (Gow & Thomas, 1998)), in law and political science (e.g. for computing voting power (Leech, 2003)), and in artificial intelligence (e.g. for coalition structure formation in multi-agent systems (Chalkiadakis, Elkind, & Wooldridge, 2011)), among many others. One of the key problems in coalitional games is to find a coalition structure, i.e. to divide the set of all players into disjoint subsets called coalitions, such that the total payoff of these coalitions is maximised. This problem is equivalent to a CSPP where players are viewed as objects, coalitions are viewed as subgroups, and a coalition structure is equivalent to a packing. Sandholm, Larson, Andersson, Shehory, and Tohmé (1999) present a coalition structure graph to visualise the set of all possible coalition structures. The authors then show interesting results about the guaranteed bound on the best coalition structure within certain parts of the graph. Since then, new exact methods have been introduced to exploit the special search space of the coalition structures. However, these existing methods are only applicable for games with less than 30 players Rahwan, Michalak, and Jennings (2012).