European Journal of Operational Research 000 (2016) 1–11

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Discrete Optimization

An iterative pseudo-gap enumeration approach for the Multidimensional Multiple-choice Knapsack Problem

Chao Gao [a,c], Guanzhou Lu [a], Xin Yao [a,b], Jinlong Li [a,*]

[a] USTC-Birmingham Joint Research Institution in Intelligent Computation and Its Applications (UBRI), School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026, China
[b] The Centre of Excellence for Research in Computational Intelligence and Applications (CERCIA), School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[c] Department of Computing Science, University of Alberta, Edmonton, Canada

A R T I C L E I N F O

Article history:
Received 24 May 2016
Accepted 23 November 2016
Available online xxx

Keywords:
Integer programming
Heuristics
Multidimensional Multiple-choice Knapsack
Reduced cost constraint

A B S T R A C T

The Multidimensional Multiple-choice Knapsack Problem (MMKP) is an important NP-hard combinatorial optimization problem with many applications. We propose a new iterative *pseudo-gap* enumeration approach to solving MMKPs. The core of our algorithm is a family of additional cuts derived from the reduced costs constraint of the nonbasic variables by reference to a *pseudo-gap*. We then introduce a strategy to enumerate the *pseudo-gap* values. Joint with CPLEX, we evaluate our approach on two sets of benchmark instances and compare our results with the best solutions reported by other heuristics in the literature. It discovers 10 new better lower bounds on 37 well-known benchmark instances with a time limit of 1 hour for each instance. We further give direct comparison between our algorithm and one state-of-the-art "reduce and solve" approach on the same machine with the same CPLEX, experimental results show that our algorithm is very competitive, outperforming "reduce and solve" on 18 cases out of 37.

1. Introduction

The Multidimensional Multiple-choice Knapsack Problem (MMKP) is one of the hardest variants of the Knapsack Problem (Hifi & Sbihi, 2004). It has many real-world applications, such as logistics (Basnet & Wilson, 2005), running time resource management (Ykman-Couvreur, Nollet, Catthoor, & Corporaal, 2006), global routing of wiring in circuits (Shojaei, Wu, Davoodi, & Basten, 2010), web service composition (Yu, Zhang, & Lin, 2007) and capital budgeting (Pisinger, 2001), the strike force asset allocation problem (Li, Curry, & Boyd, 2004), etc.

Suppose there is a set of items N, which is divided into n disjoint subsets, where each item has an m dimensional cost and a profit value, the MMKP asks to select exactly one item from each subset such that the summed cost on each dimension will not exceed the given bound, while maximizing the summed profit. In the literature, the requirement of selecting exactly one item from each subset is commonly named as the *choice-constraint*, the subset of items is referred to as *group*.

More formally, let $\mathbf{x}$ be a zero-one vector where $x_j = 1$ means item with index j is selected, p_j and vector $\mathbf{v_j} = (v_j^1, v_j^2, \ldots, v_j^m)$ are respectively the profit value and cost vector associated with j. The resource bound is given by vector $\mathbf{b} = (b^1, b^2, .., b^m)$, and N_i is the set of items in *group* i. We can formulate the MMKP as a 0–1 Integer Linear Programming (ILP) problem:

$$\max \sum_{j \in N} p_j x_j \quad (1)$$

$$\text{subject to} \sum_{j \in N} v_j^k x_j \le b^k, \quad k = 1, ..., m, \quad (2)$$

$$\sum_{j \in N_i} x_j = 1, \quad i = 1, ..., n, \ \cup_{i=1..n} N_i = N, \quad (3)$$

$$x_j \in \{0, 1\}, j = 1, ..., |N|. \quad (4)$$

In this paper, we propose a new approach, namely Iterative Pseudo-Gap Enumeration, for solving MMKPs. Our algorithm starts by obtaining an upper bound from solving the Linear Programming (LP) relaxation, and then by reference to a *pseudo-gap* and a reduced cost constraint, we propose to derive a new family of pseudo cuts that constrain variables from different groups. Finally,

* Corresponding author.
E-mail addresses: cgao3@ualberta.ca (C. Gao), lrlgz@ustc.edu.cn (G. Lu), x.yao@cs.bham.ac.uk (X. Yao), jlli@ustc.edu.cn (J. Li).

http://dx.doi.org/10.1016/j.ejor.2016.11.042
0377-2217/

we introduce a simple strategy to enumerate the *pseudo-gap* iteratively. Joint with CPLEX to solve the strengthened problem at each iteration, we test our approach on 37 instances from the literature. It updates 10 new lower bounds, given a run-time of 1 hour for each instance, outperforms the state-of-the-art approach in the literature when running on the same machine.

The rest of our paper is organized as follows. In Section 2, we review the related work. In Section 3, we explain our approach in detail. We then present our experimental studies in Section 4. Section 5 concludes this paper.

2. Related work

A number of algorithms have been proposed for tackling MMKP. Exact methods based on Branch and Bound (Ghasemi & Razzazi, 2011; Khan, 1998; Sbihi, 2007) are able to guarantee the obtained solution to be optimal after the algorithm terminates, however systematic search without heuristics usually requires intractable computation time to obtain high quality solutions for large-scale instances.

It is believed that the first heuristic results were due to Moser, Jokanovic, and Shiratori (1997), who proposed a heuristic algorithm based on Lagrangian Relaxation that starts from building an infeasible solution, then repeatedly permutes to reduce the infeasibility. Their method was improved by Akbar, Manning, Shoja, and Khan (2001). Khan, Li, Manning, and Akbar (2002) proposed a heuristic based on the aggressive resource usage, and they claimed that their heuristic performs better than Moser's. However, a guided local search and a reactive local search heuristic both proposed by Hifi and Sbihi (2004) and Hifi, Michrafy, and Sbihi (2006) were able to outperform Khan and Moser's heuristics. Then, a column generation method proposed by Cherfi and Hifi (2010) obtained better results on the benchmark instances used by the previous heuristics. Cherfi and Hifi (2009) later proposed a hybrid algorithm that combines local branching with column generation and a truncated branch-and-bound. Cherfi and Hifi's hybrid algorithm outperformed all previous approaches substantially.

In fact, due to the different real-world application requirements, the approaches for tackling MMKPs can be grouped into two categories. The first ones are fast heuristics that focus on finding feasible solutions at a small computation cost, particularly to meet the requirement of real-time applications. The methods proposed by Ykman-Couvreur et al. (2006), Htiouech, Bouamama, and Attia (2013), Parra-Hernandez and Dimopoulos (2005), Xia, Gao, and Li (2015) and Shojaei, Ghamarian, Basten, and Geilen (2009) belong to this route. The second ones pay more efforts on high quality solutions. The iterative relaxation based heuristic introduced by Hanafi, Mansi, and Wilbaut (2009), a family of iterative semi-continuous relaxation heuristics named ILPH, IMIPH, IIRH and ISCRH proposed by Crévits, Hanafi, Mansi, and Wilbaut (2012), and another hybrid heuristic by Mansi, Alves, Valério de Carvalho, and Hanafi (2013) that consists of a family of cuts to define a reduced problem and a reformulation procedure are all of this sort.

The most recent approach "reduce and solve" Chen and Hao (2014) adopts both group fixing and variable fixing to obtain reduced problems, and then solves the reduced problems by the Integer Linear Programming (ILP) solver CPLEX. Based on different enumerating methods, two variants namely PEGF and PERC are actually defined. The "reduce and solve" approach found most of the current best known results on the set of 27 standard benchmark instances and 10 new irregular structure instances introduced by a fully parameterized CPH heuristic based on pareto algebra (Shojaei, Basten, Geilen, & Davoodi, 2013). The comparison between the "reduce and solve" approaches and CPH (Chen & Hao, 2014) over these 37 instances demonstrates that the two variants PEGF and PERC of "reduce and solve" are overall better than CPH.

It is worth noting that recent high solution quality aimed approaches (Chen & Hao, 2014; Crévits et al., 2012; Hanafi et al., 2009; Mansi et al., 2013) share the similar idea to reduce the problem by proposed pseudo cuts, and then the reduced problem is solved by an ILP solver, namely CPLEX. The key difference of these approaches is their proposed pseudo cuts and how they iteratively adjust their pseudo cuts.

In this paper, we present a new Iterative Pseudo Gap Enumeration (IPGE) approach for the MMKP. We introduce the concept of *pseudo-gap* which serves as a hypothesized gap between the upper bound and lower bound of the original problem. Based on the *pseudo-gap*, we show that a new family of cuts could be derived by the reduced cost constraints (Boussier, Vasquez, Vimont, Hanafi, & Michelon, 2010; Saunders & Schinzinger, 1970; Vimont, Boussier, & Vasquez, 2008). After applying these cuts, the strengthened problem is solved by the ILP solver CPLEX. We further introduce a strategy to enumerate the *pseudo-gap*, thereby realizing an iterative method that converges to an optimal solution after the *pseudo-gap* becomes valid.

To evaluate the effectiveness of IPGE, we conduct experimental studies on the 37 benchmark instances (Chen & Hao, 2014), among which 27 are with regular structures and 10 are with irregular structures, where regular or irregular structure indicates whether all groups of an instance have exactly the same number of items or not. The comparative experiments show that our algorithm competes favorably with the state-of-the-art "reduce and solve" approach. In particular, given a run time of 1 hour for each instance, IPGE is able to report 6 new better lower bounds on the 27 regular structure instances, and 4 on the 10 irregular structure instances, even though the best lower bounds from the literature have been regarded as very high.

3. An Iterative Pseudo Gap Enumeration approach to the MMKP

IPGE is essentially a two-step iterative procedure. In the first step, a family of pseudo cuts/constraints is derived from the reduced cost constraints with regarding to a *pseudo-gap*. Then the original problem with these pseudo cuts is solved by calling an ILP solver in the second step. In this section, we first show how to generate the pseudo cuts given there is a *pseudo-gap* at hand, after that we present how the *pseudo-gap* is initially defined and adjusted iteratively, finally we give our complete algorithm.

3.1. Definitions

For the convenience of understanding, we introduce some definitions that are consistently used in this paper.

- P is the given MMKP problem instance.
- $\mathbf{x}^*$ denotes an optimal solution to P.
- $LP(P)$ is the Linear Programming Relaxation of P.
- $\bar{\mathbf{x}}$ is the optimal solution to $LP(P)$.
- r_j denotes the reduced cost corresponding to $\bar{x}_j$, i.e., variable with index j.
- $\underline{v}(P)$ and $\bar{v}(P)$ are a lower bound and an upper bound for problem P, respectively.
- Strengthened problem $(P|C)$ denotes the problem instance P after applying a set of constrains (cuts) in C.

With regard to $\bar{\mathbf{x}}$, we further define the following sets:

- $J^0(\bar{\mathbf{x}}) = \{j | j \in N, \bar{x}_j \text{ is nonbasic and } \bar{x}_j = 0\}$ denotes the index set of nonbasic variables with value 0.
- $J^1(\bar{\mathbf{x}}) = \{j | j \in N, \bar{x}_j \text{ is nonbasic and } \bar{x}_j = 1\}$ denotes the index set of nonbasic variables with value 1.
- $J(\bar{\mathbf{x}}) = J^0(\bar{\mathbf{x}}) \cup J^1(\bar{\mathbf{x}})$ includes the indices of all nonbasic variables in $\bar{\mathbf{x}}$.