Solving the Partial Inverse Knapsack Problem

Eva Ley[0009−0000−7927−4757] and
Nikolas Lykourinos[0009−0003−8211−0544] and
Maximilian Merkert[0000−0002−7838−445X] and
Andreas M. Tillmann[0000−0002−7744−836X]

Abstract In this paper, we investigate the partial inverse knapsack problem, a bilevel optimization problem in which the follower solves a classical 0/1-knapsack problem with item profit values comprised of a fixed part and a modification determined by the leader. Specifically, the leader problem seeks a minimal change to given item profits such that there is a follower optimum selecting or discarding items according to two given disjoint subsets. Based on the bilinear value-function reformulation, we develop a single-level linear mixed-integer model for the problem and propose a branch-and-cut solution algorithm. Our computational experiments demonstrate the overall viability of our approach and that utilizing a new exponential-size class of valid inequalities for the leader allows to substantially reduce the overall solution times.

1 Introduction

A partial inverse combinatorial problem (PICP) is a bilevel problem in which the leader modifies the cost parameters of the combinatorial follower problem (FP) such that an optimal follower solution includes or excludes specified elements, respectively. Thus, in the partial inverse knapsack problem (PIKP), the leader modifies item profits (by as little as possible) in order to entice the follower to select or disregard certain items in an optimal solution of their knapsack problem. Generally, PICPs arise, e.g., in economics when a planner may adjust incentives so that rational agents choose a prescribed outcome [3], or to model parameter estimation tasks to explain partially observed decisions [9]. A complete inverse combinatorial problem (CICP) is a special case of a PICP

Eva Ley, Maximilian Merkert
TU Braunschweig, Institute for Mathematical Optimization, Universitätsplatz 2, 38106 Braunschweig, Germany; e-mail: {eva.ley,m.merkert}@tu-braunschweig.de

Nikolas Lykourinos (corresponding author), Andreas M. Tillmann
TU Clausthal, Institute of Mathematics, Erzstraße 1, 38678 Clausthal-Zellerfeld, Germany; e-mail: {nikolas.lykourinos,andreas.tillmann}@tu-clausthal.de

in which a unique follower solution has to be optimal. For polynomial-time solvable follower problems, CICPs are also solvable in polynomial time [1], whereas PICPs can be NP-hard, see [7] and references therein. The complete inverse knapsack problem (CIKP) is coNP-complete [2, 10], while PIKP can be shown to be Σ_2^p-complete, in a similar way as for other PICPs in [5]. Moreover, the bilinear follower objective in PIKP makes the problem unsuitable for general-purpose methods for bilevel mixed-integer linear problems (see, e.g., [6]).

In the following Sect. 2, we first formally state the PIKP as a bilevel mixed-integer problem with an integer follower problem and recast it as an exponential-size single-level mixed-integer linear program (MIP). We then develop problem-specific cutting planes, a corresponding separation problem, and our overall solution approach. In Sect. 3, we report on numerical experiments, and provide some concluding remarks in Sect. 4.

This work is based on and extends the second author's Master's thesis [8].

2 Problem Formulation and Solution Approach

Given an item set $E \coloneqq \{1, 2, \ldots, n\}$, weights $\boldsymbol{w} \in \mathbb{N}^n$, profits $\boldsymbol{c} \in \mathbb{N}^n$ and budget $b \in \mathbb{N}$, the PIKP with forbidden items $F \subseteq E$ and required items $R \subseteq E$ can be stated as

$$\min_{\boldsymbol{x}\in\mathbb{R}^n,\, \boldsymbol{y}\in\{0,1\}^n} \quad \|\boldsymbol{x}\|_1 \tag{1a}$$
$$\text{s.t.} \quad y_e = 0 \qquad \forall e \in F \tag{1b}$$
$$y_e = 1 \qquad \forall e \in R \tag{1c}$$
$$\boldsymbol{y} \in \arg\max_{\bar{\boldsymbol{y}}} \left\{ (\boldsymbol{c}+\boldsymbol{x})^\top \bar{\boldsymbol{y}} : \boldsymbol{w}^\top \bar{\boldsymbol{y}} \le b, \bar{\boldsymbol{y}} \in \{0,1\}^n \right\}. \tag{1d}$$

(The choice of the ℓ_1-norm objective is somewhat arbitrary, cf. Sect. 4.) We first rewrite this bilevel problem by means of an exact linearization of the bilinear value-function reformulation (see, e.g., [6]), utilizing a standard variable split and implicit variable bounds. To that end, let $U \coloneqq E \setminus (F \cup R)$ be the set of unrestricted items and denote the feasible set of the follower problem variables $\bar{\boldsymbol{y}}$ by $Y \coloneqq \left\{\boldsymbol{y} \in \{0,1\}^n : \boldsymbol{w}^\top \boldsymbol{y} \le b\right\}$. Moreover, let $k \coloneqq \max\left\{\boldsymbol{c}^\top \bar{\boldsymbol{y}} : \bar{\boldsymbol{y}} \in Y\right\}$ and $k' \coloneqq \max\left\{\boldsymbol{c}^\top \bar{\boldsymbol{y}} : \bar{y}_e = 0\ \forall e \in F,\ \bar{y}_e = 1\ \forall e \in R, \bar{\boldsymbol{y}} \in Y\right\}$ denote the optimal values for the unrestricted and restricted follower problem without profit modifications (i.e., $\boldsymbol{x} = \boldsymbol{0}$), respectively.

Theorem 1 *The PIKP* (1) *is equivalent to the following single-level MIP* (2)*:*

$$\min_{\boldsymbol{x}^\pm\in\mathbb{R}^n_{\ge 0},\, \boldsymbol{y}\in\{0,1\}^n} \quad \sum_{e\in E\setminus F} x_e^+ + \sum_{e\in E\setminus R} x_e^- \tag{2a}$$
$$\text{s.t.} \quad y_e = 0 \qquad \forall e \in F \tag{2b}$$
$$y_e = 1 \qquad \forall e \in R \tag{2c}$$
$$\boldsymbol{w}^\top \boldsymbol{y} \le b \tag{2d}$$
$$\boldsymbol{c}^\top \boldsymbol{y} + \sum_{e\in E\setminus F} x_e^+ \ge (\boldsymbol{c} + \boldsymbol{x}^+ - \boldsymbol{x}^-)^\top \bar{\boldsymbol{y}} \qquad \forall \bar{\boldsymbol{y}} \in Y \tag{2e}$$

$$x_e^+ = 0 \qquad \forall e \in F \qquad (2f)$$
$$x_e^- = 0 \qquad \forall e \in R \qquad (2g)$$
$$x_e^+ \le (k - k')y_e \qquad \forall e \in U \qquad (2h)$$
$$x_e^- \le c_e(1 - y_e) \qquad \forall e \in U \qquad (2i)$$

Proof. In the value-function reformulation of (1), we replace (1d) by $\boldsymbol{w}^\top \boldsymbol{y} \le b$ and

$$(\boldsymbol{c} + \boldsymbol{x})^\top \boldsymbol{y} \ge (\boldsymbol{c} + \boldsymbol{x})^\top \bar{\boldsymbol{y}} \qquad \forall \bar{\boldsymbol{y}} \in Y,$$

which constitutes an (exponentially large) family of bilinear constraints. Replacing $\boldsymbol{x} = \boldsymbol{x}^+ - \boldsymbol{x}^-$ with $\boldsymbol{x}^\pm \ge \boldsymbol{0}$, we apply [7, Lemma 2.5] to infer that in any optimal solution, it holds that $x_e^+ = 0$ (i.e., $x_e \le 0$) for all $e \in E$ with $y_e = 0$, and $x_e^- = 0$ (i.e., $x_e \ge 0$) for all $e \in E$ with $y_e = 1$. Hence, we have $y_e x_e = 0 = x_e^+$ if $y_e = 0$, and $y_e x_e = x_e^+$ if $y_e = 1$. In order to linearize the model, we need to explicitly enforce the implications $y_e = 0 \Rightarrow x_e^+ = 0$ and $y_e = 1 \Rightarrow x_e^- = 0$. To that end, we can exploit the restrictions (1b) and (1c) directly, along with big-M-type constraints enforcing, for $e \in U$, the implicit upper bounds $k - k'$ for x_e^+ and c_e for x_e^-, respectively, yielding (2f)–(2i). To see the validity of the upper bounds (2h) and (2i), consider some $e \in U$ with $y(\boldsymbol{x})_e = 1$ in an optimum $\boldsymbol{y}(\boldsymbol{x})$ of the FP w.r.t. $\boldsymbol{x}$, F and R. Increasing x_e^+ to at most $k - k'$ will lead to $y_e = 1$ in the selected (optimistic) follower optimum without restrictions (2c), (2d). Similarly, setting $x_e^+ = 0$ and $x_e^- \ge c_e$ for any $e \in U$ makes item e unprofitable, so $y_e = 0$ in any unrestricted follower optimum $\boldsymbol{y}$. Putting it all together, we arrive at (2). □

Since the MIP (2) has exponentially many inequalities (2e), we start with $Y' = \emptyset$ instead of Y and dynamically append optimal follower problem (FP) solutions; we refer to the correspondingly relaxed problem as the master problem (MP). Solving the FP for some $\boldsymbol{x}$ amounts to generating a *follower cut* (inequality of type (2e)), which cuts off the current MP solution unless it is already optimal for the PIKP. Moreover, the polyhedral description underlying (2) can be improved by virtue of the following results.

Theorem 2 *[8, Thm. 4.9] Let $\tilde{F} \subseteq F$ and $\tilde{R} \subseteq R$. If $\sum_{e \in \tilde{F}} w_e \le \sum_{e \in \tilde{R}} w_e$, then the inequality $\sum_{e \in \tilde{F}} (c_e + x_e) \le \sum_{e \in \tilde{R}} (c_e + x_e)$ is valid for any feasible solution of PIKP.*

Proof. Let $(\boldsymbol{x}, \boldsymbol{y})$ be feasible for the PIKP. Define $\hat{\boldsymbol{y}}$ by $\hat{y}_e = 1 - y_e$ for $e \in \tilde{R} \cup \tilde{F}$, $\hat{y}_e = y_e$ otherwise. Then, $\boldsymbol{w}^\top \hat{\boldsymbol{y}} = \boldsymbol{w}^\top \boldsymbol{y} + \sum_{e \in \tilde{F}} w_e - \sum_{e \in \tilde{R}} w_e \le \boldsymbol{w}^\top \boldsymbol{y} \le b$, i.e., $\hat{\boldsymbol{y}} \in Y$. Therefore, (1d) implies $0 \ge (\boldsymbol{c} + \boldsymbol{x})^\top (\hat{\boldsymbol{y}} - \boldsymbol{y}) = \sum_{e \in \tilde{F}} (c_e + x_e) - \sum_{e \in \tilde{R}} (c_e + x_e)$. □

As the family of inequalities from Theorem 2 is also exponentially large, we periodically solve the following separation problem to generate such *FR-cuts* dynamically:

$$\max_{\boldsymbol{u} \in \{0,1\}^F,\, \boldsymbol{v} \in \{0,1\}^R} \left\{ \sum_{e \in F} (c_e + x_e) u_e - \sum_{e \in R} (c_e + x_e) v_e \;:\; \sum_{e \in F} w_e u_e - \sum_{e \in R} w_e v_e \le 0 \right\} \quad (3)$$

Specifically, if and only if any *violated* FR-cut exists for a given $(\boldsymbol{x}, \boldsymbol{y})$, then (3) has a positive optimal value, achieved by some $\tilde{\boldsymbol{u}} \in \{0,1\}^F$ and $\tilde{\boldsymbol{v}} \in \{0,1\}^R$ (in fact, by integrality of $\boldsymbol{c}$, it then holds than $\sum_{e \in F} c_e \tilde{u}_e - \sum_{e \in R} c_e \tilde{v}_e \ge 1$). In this case, we add the corresponding cut

$$\sum_{e\in\tilde{F}:\tilde{u}_e=1}(c_e - x_e^-) \leq \sum_{e\in\tilde{R}:\tilde{v}_e=1}(c_e + x_e^+) \tag{4}$$

to the MP. In spite of the following intractability result (we omit its straightforward proof by reduction from the standard knapsack problem), problem (3) turned out to be solvable very quickly in all our experiments, as did the follower problems.

Proposition 1 *The FR-cut separation problem* (3) *is* NP*-hard, even for integral* $\boldsymbol{c}$*,* $\boldsymbol{x}$.

Finally, we also consider adding the following inequalities to our initial model:

Proposition 2 *[8, Cor. 4.28] Let* $s, t \in F$ *or* $s, t \in R$ *with* $s \neq t$*,* $w_s \leq w_t$ *and* $c_s \geq c_t$*. Then, the dominance inequality* $x_t \geq x_s$ *is valid for at least one optimum of PIKP.*

Proof. Let s and t be given as specified, and let $(\boldsymbol{x}, \boldsymbol{y})$ be an optimal PIKP solution with $x_t < x_s$. Consider $(\hat{\boldsymbol{x}}, \boldsymbol{y})$ defined by $\hat{x}_e \coloneqq \frac{1}{2}(x_s + x_t)$ for $e \in \{s, t\}$, and $\hat{x}_e \coloneqq x_e$ otherwise. Clearly, $(\hat{\boldsymbol{x}}, \boldsymbol{y})$ satisfies $\hat{x}_t \geq \hat{x}_s$ and has the same objective value as $(\boldsymbol{x}, \boldsymbol{y})$, so it remains to show that $(\hat{\boldsymbol{x}}, \boldsymbol{y})$ fulfills (1d). As long as $c_s + x_s \geq c_t + x_t$ (which holds for both $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$), s dominates t, i.e., s is both more profitable and has smaller weight. Thus, the same selection $y_t = y_s$ as for $\boldsymbol{x}$ must still be optimal under $\hat{\boldsymbol{x}}$ (since packing both s and t changes nothing and only packing s has become less attractive). □

To solve PIKP, we generally proceed as follows: We start with the high-point-relaxation as the MP (i.e., MIP (2) with Y replaced by a possibly empty initial subset $Y' \subseteq Y$ in (2e)) and iterate, either proving optimality of the current solution or adding a "most violated" cutting plane that cuts it off. These cuts are generated by solving the FP w.r.t. the current modification $\boldsymbol{x}$ and/or the FR-cut separation problem (3). We have tested several variants of this overall algorithm, differing in linearization type, the order and frequency of cut separation, and the inclusion or omittance of dominance and/or an initial set of FR-inequalities (4); the details are given in the next section.

3 Computational Experiments

For each number $|E| \in \{50, 100, \ldots, 500\}$ of items, we created three 10-instance groups with $|U| \in \left\{\frac{|E|}{4}, \frac{|E|}{2}, \frac{3|E|}{4}\right\}$, and for each group, randomly generated $|R| \in \{1, \ldots, |E| - |U| - 1\}$. The sets U and $R \subset E \setminus U$ were chosen uniformly at random, and we set $F = E \setminus (R \cup U)$. For the budget b, we need $b \geq \sum_{e\in R} w_e$ to ensure PIKP feasibility, and further want $\sum_{e\in R} w_e < b < \sum_{e\in R\cup U} w_e$ to avoid implicit CICP instances. Thus, we generated $\alpha \in \{0.1, 0.2, \ldots, 0.8, 0.9\}$ uniformly at random and set $b = \sum_{e\in R} w_e + \alpha \sum_{e\in U} w_e$. The profit values and weights were drawn i.i.d. uniformally at random from the integer interval $\{1, \ldots, 100\}$. We ended up with 30 instances per item number, for a total of 300 instances.

We implemented our algorithm in Python 3.10.12, using Gurobi 12.0.2 [4] to solve the MIPs, with a time limit of 900 s per instance. All experiments were carried out on a 32-core Linux PC with 3.1 GHz Intel Core i9-14900 CPUs and 128 GB main memory.

We compare the following variants of the algorithm, all based on the MIP model (2) and using an initial $Y' = \emptyset$ (i.e., no inequalities of type (2e) in the first MP):

Table 1: Summary of computational experiments on all 300 instances

algo. variant	solved	time-out (fails)	runtime	nodes	gap %	follower cuts	FR-cuts
I	267	33 (33)	53.3	38,304.8	–	2,412.6	0
I(a)	248	52 (50)	104.9	4,843.5	0.2	6,468.0	0
I(b)	255	45 (44)	103.2	4,645.0	< 0.1	6,441.8	0
II	286	14 (11)	29.1	27,979.8	2.1	1,083.0	787.8
II(a)	298	2 (0)	29.8	1,754.0	< 0.1	2,061.0	1,600.9
II(b)	299	1 (0)	23.2	1,378.2	0.1	1,199.0	961.9

I : follower cut separation at each new MP incumbent,

II : FR- and follower cut separation at each new MP incumbent,

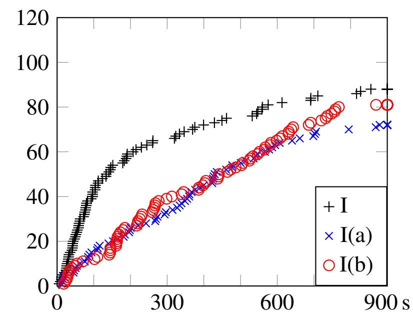
both in these basic versions as well as extended by (a): the respective cut separation also for every LP relaxation solution and (b): like (a) with additionally including the dominance inequalities from Prop. 2 in the initial model. (Further variants combining the above with initially adding all inequalities (4) for singleton sets $\tilde{F}$ and $\tilde{R}$, dominance inequalities, only separating follower cuts if no violated FR-cut was found, McCormick linearization of originally bilinear terms, and more were also tested, but inferior.)

The computational results are summarized in Table 1, where we report average running times (in seconds) and branch-and-bound node numbers as shifted geometric mean values (shifts 10 and 100, resp.). We also report the number of solved instances and failures (when the solver did not find a feasible solution for the full PIKP within the time limit) as well as the average (arithmetic mean) optimality gap on timed-out non-failure instances. All decimals are rounded to the first significant digit.

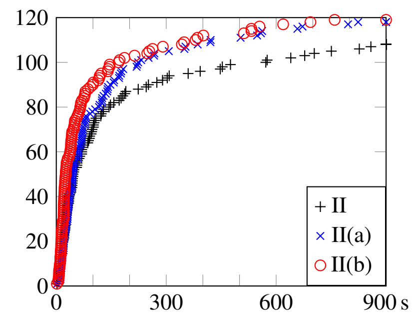
From Table 1, we see that any variant of II substantially outperforms all I-variants in terms of solved instance count and solver runtimes, showing that separation of FR-cuts significantly improves the basic version that only employs the (required) follower cuts. Moreover, enriching the initial MIP model (2) with the dominance inequalities can also help when combined with cut separation at every intermittent LP solution (compare variants (b) with the basic and (a) variants, resp.). Overall, variant II(b) stands out as the fastest and most reliable algorithm version. Fig. 1 shows that the above-described behavior persists on the larger instances with at least 350 items: II(b) solves more instances more quickly than all other variants, followed by II(a); these two are also the only variants that find feasible solutions for every instance within the given time limit, cf. Table 1. While almost all time-outs occur on this subset of 120 instances, the mean runtimes amount to, in particular, roughly 211 s for II, 75 s for II(a) and 51 s for II(b).

4 Concluding Remarks

We proposed a branch-and-cut algorithm for solving the partial inverse knapsack problem using a novel linearization of bilinear constraints and the dynamic separation of model inequalities as well as new cutting planes. Evaluating various variants of our



a variants of I



b variants of II

Fig. 1: Comparison of the algorithm variants on instances with $|E| \geq 350$; depicted are the numbers of instances (vertical axis) solved within a given runtime (horizontal axis)

algorithm on a large testset, the most effective one included dominance inequalities and separation of both types of cuts in all solving nodes. In the future, one could consider other leader objectives (e.g., $\|\boldsymbol{x}\|_\infty$ or the cardinality of $\boldsymbol{x}$, which simplifies the PIKP or makes it more difficult, resp.); note that follower and FR-cuts remain valid for those as well. Moreover, benchmarking and analysis w.r.t. different instance properties could lead to further algorithmic improvements. For instance, here, weights and profits were drawn i.i.d. uniformally at random, but using correlated values could yield more difficult PIKP instances given that it leads to harder knapsack problems, cf. [11].

References

1. Ahuja, R.K., Orlin, J.B.: Inverse optimization. Oper. Res. **49**(5):771–783 (2001)
2. Bulut, A., Ralphs, T.K.: On the complexity of inverse mixed integer linear optimization. SIAM J. Optim. **31**(4):3014–3043 (2021)
3. Dial, R.B.: Network-optimized road pricing: Part I: A parable and a model. Oper. Res. **47**(1):54–64 (1999)
4. Gurobi Optimization, LLC: Gurobi Optimizer Reference Manual. Available via `https://www.gurobi.com`. Accessed 10 July 2025
5. Johannes, B.: New Classes of Complete Problems for the Second Level of the Polynomial Hierarchy. PhD thesis, Technische Universität Berlin (2011)
6. Kleinert, T., Labbé, M., Ljubić, I., Schmidt, M.: A survey on mixed-integer programming techniques in bilevel optimization. EURO J. Comput. Optim. **9**:100007 (2021)
7. Ley, E., Merkert, M.: Solution methods for partial inverse combinatorial optimization problems in which weights can only be increased. Preprint at `https://optimization-online.org/?p=25609` (2024)
8. Lykourinos, N.: Schnittebenen für partiell inverse Knapsackprobleme. Master's Thesis, Technische Universität Braunschweig, Germany (2025). In German.
9. Moser, T.J.: Shortest path calculation of seismic rays. Geophys. **56**(1):59–67 (1991)
10. Roland, J., Figueira, J.R., De Smet, Y.: The inverse {0, 1}-knapsack problem: Theory, algorithms and computational experiments. Discret. Optim. **10**(2):181–192 (2013)
11. Kataoka, S., Yamada, T.: Upper and lower bounding procedures for the multiple knapsack assignment problem. Eur. J. Oper. Res. **237**(2):440–447 (2014)