



Bounding the Price-of-Fair-Sharing Using Knapsack-Cover Constraints to Guide Near-Optimal Cost-Recovery Algorithms

Sander Aarts^(✉), Jacob Dentes, Manxi Wu, and David B. Shmoys

School of Operations Research and Information Engineering, Cornell University,
Ithaca, NY, USA
sea78@cornell.edu

Abstract. We consider the problem of fairly allocating the cost of providing a service among a set of users, where the service cost is formulated by an NP-hard *covering integer program (CIP)*. The central issue is to determine a cost allocation to each user that, in total, recovers as much as possible of the actual cost while satisfying a stabilizing condition known as the *core property*. The ratio between the total service cost and the cost recovered from users has been studied previously, with seminal papers of Deng, Ibaraki, & Nagamochi and Goemans & Skutella linking this *price-of-fair-sharing* to the integrality gap of an associated LP relaxation. Motivated by an application of cost allocation for network design for LPWANs, an emerging IoT technology, we investigate a general class of CIPs and give the first non-trivial price-of-fair-sharing bounds by using the natural LP relaxation strengthened with knapsack-cover inequalities. Furthermore, we demonstrate that these LP-based methods outperform previously known methods on an LPWAN-derived CIP data set. We also obtain analogous results for a more general setting in which the service provider also gets to select the subset of users, and the mechanism to elicit users' private utilities should be group-strategyproof. The key to obtaining this result is a simplified and improved analysis for a cross-monotone cost-allocation mechanism.

Keywords: Cost sharing · Covering integer programs · LoRaWAN

1 Introduction

Allocating costs fairly when providing some collective service to a set of users has been a mainstream topic within cooperative game theory in which the notion of the core, and the associated *core property*, plays a central role. Specifically, the core property requires that the payment from each user is such that each subset of users pays, in total, no more than the cost of providing service just to them. LP duality has long been known to play a fundamental role in such allocations, and the seminal work of Deng, Ibaraki & Nagamochi [7] and Goemans & Skutella

[12] proved the link between the integrality gap (or lack thereof) for several NP-hard discrete optimization settings and the fraction of the service cost that can be recovered by an allocation satisfying the core property. We term the ratio between optimal cost and the cost recoverable satisfying the core property as the *price-of-fair-sharing*.

For a natural subclass of *covering integer programs* (CIPs), known as multiset-multi-covering problems, the strongest known price-of-fair-sharing was not obtained using a dual LP-based framework [22]. We show that an LP-based approach can yield improved bounds and simplified analyses, not just for this special case considered by [22], but for general CIPs. More specifically, our main result is that by strengthening the natural LP relaxation for CIPs with knapsack-cover inequalities, the resulting dual LP does have the property that any feasible solution yields a cost-allocation that satisfies the core property. This reduces the problem of finding a cost-allocation to that of finding a strengthened-LP-relative approximation algorithm, allowing us to leverage approximation algorithms already known for these settings (e.g., [3–5, 18]). Not only do these existing algorithms yield bounds on the price-of-fair sharing for general CIPs, they also exploit existing bounds for families of sparse CIP instances. These existing algorithms provide improved cost-recovery ratios for families of sparse CIP instances.

We also obtain analogous results in a more general setting where the service provider selects a subset of users to receive the service based on their elicited private valuations of the service. The mechanism used to elicit users' private valuations needs to be group-strategyproof so that no subset of users have an incentive to misreport their valuations. The key to obtaining this result is a simplified and improved analysis for a cross-monotone cost-allocation mechanism. These results also rely on our knapsack-cover-strengthened LP relaxation.

Our investigation into these questions was directly motivated by challenges in the development of an emerging technology, LPWANs, a popular Internet of Things solution for wirelessly connecting devices to the internet [28]. CIPs emerge from the appropriate model for determining the location of gateways to provide LPWAN coverage to a collection of users, and our result is useful to determine how the cost of LPWANs should be shared among users in an incentive compatible manner. We complement our algorithmic and structural results with an empirical investigation into the effectiveness of our methods on relatively large-scale LPWAN data. Our empirical study demonstrates that, in comparison to several natural heuristic alternatives, our theoretically-motivated methods provide far superior results.

2 Optimizing Coverage for LPWAN Networks and IoT

The growth of the Internet of Things (IoT) is creating both new opportunities and challenges in wireless coverage provision and pricing. Low-power wide-area networks (LPWANs) are a ubiquitous technology for connecting devices, or things, to the internet. These networks use radio communication to transmit

signals over long distances. Demodulating, and transferring these signals to the internet requires networks of physical wireless receivers. Many such networks are already deployed. In 2023 alone, LPWANs reportedly served over 1.3 billion IoT connections [1]. LoRaWAN is among the most popular LPWAN solutions, representing roughly 40% of all connections made outside of China. Globally, over 150 different network operators provide LoRaWAN coverage [29].

Sharing LoRaWAN coverage among multiple users is an effective yet challenging approach to better utilize wireless infrastructure and lower costs. Two key features make sharing LoRaWAN coverage especially beneficial. Firstly, the receivers, or *gateways*, can process orders of magnitude more wireless traffic than that produced by a typical single IoT application (see, e.g., the application in [2]). Secondly, the quality of coverage improves as the number of gateways increases. Transmitted LoRaWAN packets tend to fail at random, and coverage quality is measured by the *packet reception rate*. However, transmissions are *association free*; a device broadcasts to all nearby gateways, many of which may process the same packet [28]. A packet is lost if and only if all nearby gateways simultaneously fail to receive it. Thus, having multiple gateways to cover the transmitted signal can significantly improve the reception rate. This multi-coverage feature is a key aspect of our model and raises interesting computational challenges.

The problem of optimally providing wireless coverage to a group of users using multiple gateways can be modeled as a covering integer program. Let \mathcal{U} denote a set of m users, and \mathcal{F} a set of n potential gateway locations, or facilities. Each user $j \in \mathcal{U}$ has a service requirement $r_j > 0$, and each gateway has an opening cost $c_i > 0$. Requirements and costs are represented by vectors $\mathbf{r} = (r_1, r_2, \dots, r_m)$ and $\mathbf{c} = (c_1, \dots, c_n)$, respectively. Each facility-user pair is associated with a contribution parameter $a_{ij} \geq 0$ that specifies the quality of coverage that facility i provides user j if opened. Contributions are represented as an $n \times m$ matrix \mathbf{A} . The objective is to minimize cost, subject to providing sufficient coverage to each user.

$$c^* = \min\{\mathbf{c}^T \mathbf{x} : \mathbf{Ax} \geq \mathbf{r}, \mathbf{x} \in \{0, 1\}^n\} \quad (1)$$

Here, the binary decision variables $\mathbf{x} = (x_1, \dots, x_n)$ indicate whether each facility i is opened or not. Details on how CIPs capture LoRaWAN coverage provision are given in an extended online version. (A more general formulation of CIPs permits the purchase of up to d_i integer copies of facility i . Our results extend to this setting in a straightforward way, but we omit this extension for simplicity and clarity of exposition.)

3 Cost-Sharing Fundamentals

The main objective in this work is to develop a principled algorithmic framework for finding fair and stable ways to share the cost of CIP solutions among the set of users. We follow a standard approach in mechanism design and define “fair and stable” using the *core property* [16, 26]. An allocation satisfies the core property if no subset of users is allocated more cost than the minimum cost of

providing coverage for this subset alone. This property can also be viewed as an absence of cross-subsidies; no sub-group of users pays more than what it would cost them to serve themselves, thereby not subsidizing other users [10].

More formally, a cost allocation is a vector $\xi = (\xi_1, \dots, \xi_m)$, where $\xi_j \geq 0$ represents the dollar amount charged to user $j \in \mathcal{U}$. Allocations ξ_j are also called cost-shares. Following the standard definition, cost-shares ξ satisfy the core property if for each sub-group $J \subseteq \mathcal{U}$,

$$\sum_{j \in J} \xi_j \leq c_J^*, \quad \forall J \subseteq \mathcal{U}, \quad (2)$$

where c_J^* is the minimum cost of serving group J [16]. Cost allocations can vary in magnitude. We say cost allocation ξ is *budget-balanced* whenever the total amount paid by the served users, $\sum_{j \in \mathcal{U}} \xi_j$, equals the cost of serving them. The *core* is the set of budget-balanced cost allocations satisfying the core property [16]. In some problems, finding cost-shares in the core can be challenging.

Linear programming (LP) duality is the workhorse tool for finding cost-shares satisfying the core property, but it is not always possible to simultaneously satisfy both the core property and be budget-balanced. There are simple CIP instances that have no cost-allocations in the core (see, e.g., Li *et al.* [22]). Deng, Ibaraki, and Nagamochi [7] show that a generic covering problem has a non-empty core if and only if its natural LP-relaxation has no integrality gap. As such, it is common to maximize the cost recovered, subject to satisfying the core-property [12, 16]; a cost-allocation ξ is β -budget-balanced if it recovers a fraction β of the cost.

For many optimization problems, every dual feasible solution is a cost allocation that satisfies the core property. This connection has important algorithmic implications. First, one can leverage duality to find cost-shares satisfying the core property by using (approximation) algorithms that produce feasible dual solutions as well as feasible integer solutions to the primal. If the integer solution costs at most α times the dual, we see that the corresponding cost allocation is $1/\alpha$ -budget balanced. We shall say then that α is the *price-of-fair-sharing* (or equivalently, $1/\alpha$ has also been termed the cost-recovery ratio). This relationship has been frequently used to obtain strong results in many settings [7, 8, 12, 13, 19, 21, 23, 25].

On the other hand, seminal work by Deng, Ibaraki, & Nagamochi [7] and Goemans & Skutella [12] shows that all cost allocations satisfying the core property are dual feasible solutions in the set cover problem and the facility location problem. Since the cost-allocation mechanism purchases an integer solution, but only allocates costs as a fractional solution, duality implies that price-of-fair-sharing is lower-bounded by the integrality gap of the problem. It is folklore that the price-of-fair-sharing is α whenever the integrality gap of the natural LP relaxation is α [16, 23].

The relationship between the price-of-fair-sharing and the integrality gap of the natural IP/LP formulation does not, however, appear to hold for CIPs. The multi-cover constraints in CIPs render the natural linear programming relaxation

ill-suited for cost-sharing. Even with just one user $|U| = 1$ and integral inputs, the integrality gap of the CIP is unbounded [4]. Naively, this would imply that the price-of-fair-sharing in CIPs is unbounded. However, Li, Sun, Wang, and Lou [22] show that a bound can be attained for the special case of CIPs with integer-valued \mathbf{A} and \mathbf{r} . Their analysis, however, does not make explicit use of duality. Moreover, our IoT application yields CIPs that do not satisfy the assumption of integer-valued \mathbf{A} and \mathbf{r} . This begs two questions: How do cost-sharing and duality relate in CIPs, and can there be an effective cost-recovery at all when data are non-integer-valued?

4 Knapsack-Cover Constraints to the Rescue

This paper presents a principled framework for finding cost-shares in CIPs using linear programming duality. Our approach makes use of a well-known strengthened LP formulation based on knapsack-cover inequalities [4]. Our main contribution is to show that every feasible dual solution in this strengthened LP, which we shall denote KC-LP, naturally induces cost-shares that satisfy the core property. This has significant algorithmic consequences. First, our results imply that any approximation algorithm that produces CIP solutions with cost at most α times the KC-LP optimum can be used to produce cost-shares that recover $1/\alpha$ of the cost; that is, a price-of-fair-sharing of α . There are many such algorithms [4–6, 18]. More generally, our framework can be used to find cost-shares for any integer CIP solution by solving the strengthened dual linear program.

These methods yield the first cost-sharing algorithm with bounded price-of-fair-sharing for general CIPs. Furthermore, we prove that the implied price-of-fair-sharing bounds are tight. In addition, we showcase the efficacy of our framework by recovering up to 93% of the cost in semi-stylized LoRaWAN covering problems at scale; this is over twice the recovery of the next-best method. We also use our KC-LP approach to obtain analogous results for a more general setting in which the service provider also selects a subset of users to receive the service based on the users' private valuations elicited from a mechanism that is group-strategyproof. This reinforces the central message of this paper, affirming the powerful connection between an effective cost sharing mechanism and KC-LP dual solutions.

The knapsack-cover inequalities are introduced in the seminal work by Carr, Fleischer, Leung, and Phillips [4] to strengthen the LP so as to bound its integrality gap. It is helpful to first understand how the integrality gap of the natural LP relaxation is unbounded. Carr *et al.* provide the following instance: let $R > 0$ be a large integer requirement for a single user, and let there be two facilities; facility a provides $R - 1$ units of coverage at near-zero cost, whereas facility b provides R units of coverage at a cost of 1. Clearly, a feasible integer solution must include item b , with a total cost of 1. A fractional solution, however, can select a full unit of item a , and only a $1/R$ fraction of item b , with a cost of $1/R$. The resulting integrality gap is R , which can be arbitrarily large.

Carr *et al.* [4] derive a strong bound on the integrality gap by adding exponentially-many valid *knapsack-cover* (KC) inequalities to produce a

strengthened LP relaxation. These inequalities represent residual coverage requirements at partial solutions. Suppose facilities $S \subseteq \mathcal{F}$ have been built. Now user j has a *residual requirement* $r_j^S = \max\{r_j - \sum_{i \in S} a_{ij}, 0\}$. To satisfy j 's residual requirement, an additional contribution of at least r_j^S is needed from the remaining unbuilt facilities $\mathcal{F} \setminus S$. In addition, there is no benefit to exceeding r_j^S , so the original contributions a_{ij} can be clipped to r_j^S if they exceed this value. We call $a_{ij}^S = \min\{a_{ij}, r_j^S\}$ the *residual contribution*, and set it to zero if i is in S . This defines the *knapsack-cover inequality*:

$$\sum_{i \in \mathcal{F} \setminus S} a_{ij}^S \geq r_j^S. \quad (3)$$

There are $|\mathcal{U}| \times |2^{\mathcal{F}}| = m2^n$ KC inequalities. The KC inequalities are valid; their presence does not change the set of feasible *integer* solutions, and reduces the feasible set of the LP relaxation. The strengthened linear program is as follows:

$$\begin{aligned} & \min_x \sum_{i \in \mathcal{F}} c_i x_i, \\ s.t. \quad & \sum_{i \in \mathcal{F} \setminus S} a_{ij}^S x_i \geq r_j^S, \quad \forall (j, S) \in \mathcal{U} \times 2^{\mathcal{F}}, \\ & x_i \geq 0, \quad \forall i \in \mathcal{F}. \end{aligned} \quad (\text{KC-LP})$$

Note that the constraints $x_i \leq 1$ have been dropped. (Multiplicity constraints are implicit, in that $a_{ij}^S = 0$ whenever $i \in S$.) Associated with this primal is the knapsack-cover dual program (KC-DP), referred to as the KC-LP dual. The dual has one constraint for each facility, and one variable for every KC inequality:

$$\begin{aligned} & \max_y \sum_{j \in \mathcal{U}} \sum_{S \subseteq \mathcal{F}} r_j^S y_j^S, \\ s.t. \quad & \sum_{j \in \mathcal{U}} \sum_{S \subseteq \mathcal{F} \setminus \{i\}} a_{ij}^S y_j^S \leq c_i, \quad \forall i \in \mathcal{F}, \\ & y_j^S \geq 0, \quad \forall (j, S) \in \mathcal{U} \times 2^{\mathcal{F}}. \end{aligned} \quad (\text{KC-DP})$$

There are known approximation algorithms for CIPs that yield integer solutions with costs within a multiplicative factor of the KC-LP optimum. The worst-case performance of these algorithms depend on the column and row sparsity of the contributions \mathbf{A} . Let $\Delta = \max_i \{\sum_{j \in \mathcal{U}} \mathbf{1}[a_{ij} > 0]\}$ denote the column sparsity, and let $\Gamma = \max_j \{\sum_{i \in \mathcal{F}} \mathbf{1}[a_{ij} > 0]\}$ denote the row sparsity. Carr *et al.* [4] propose a Γ -approximation algorithm based on rounding a KC-LP solution; Kolliopoulos and Young [18] design an alternative KC-LP rounding procedure to yield a $\mathcal{O}(\log(1 + \Delta))$ -approximation algorithm. Together, these provide upper bounds on the strengthened integrality gap of Γ and $\mathcal{O}(1 + \log \Delta)$. Some algorithms attain improved guarantees at the cost of small violations of the multiplicity constraints [5, 6]. Common to these LP-rounding rounding methods is that they first require obtaining an optimal solution to the KC-LP.

The strengthened linear program and its dual can be solved to near-optimality in polynomial time. Chekuri and Quanrud [5] develop a multiplicative weights method that returns approximately-optimal primal and dual solutions in near linear time, but with an $\mathcal{O}(1/\epsilon^5)$ dependence on the relative error ϵ . Their approach builds upon earlier work of Plotkin, Shmoys, and Tardos [27], and the solution method outlined by Carr *et al.* [4]. The returned dual solutions are feasible, lie within a $\frac{1}{1-\epsilon}$ -factor of the optimum, and have a polynomially-bounded number of non-zero variables [5]. In practice, the primal and dual problems can be solved exactly using column generation for quite large instances; the problem of finding a *most violated inequality* can be reduced to solving a sequence of pseudopolynomially-many minimum-cost knapsack problems, each of which admits a pseudopolynomial exact algorithm (e.g., [20]). In our case study, we solve the KC-LP dual to optimality using this approach on large-scale problem instances with thousands of users and thousands of facilities.

5 Cost-Shares and the Strengthened LP

Our main technical contribution is to show that any feasible solution to the strengthened dual (**KC-DP**) produces a cost-allocation that satisfies the core property. This reconciles the apparent inconsistency for the CIP with respect to the folk understanding cost-shares and duality. The cost-shares themselves are remarkably intuitive: each user pays the part of the dual objective associated with her share of service utilization.

Theorem 1. *Let $\mathbf{y} = (y_j^S)_{S \subseteq \mathcal{F}, j \in \mathcal{U}}$ be a KC-LP dual-feasible solution. Then, cost-shares*

$$\xi_j := \sum_{S \subseteq \mathcal{F}} r_j^S y_j^S, \quad \forall j \in \mathcal{U}, \quad (4)$$

satisfy the core property. We say the cost-shares in (4) are induced by \mathbf{y} .

Clearly, by summing the cost-shares over all users, we recover the dual objective in (**KC-DP**). The theorem is simple, but its consequences are profound. In particular, it can be used to exploit approximation algorithms that find integer solutions with costs bounded in terms of a corresponding dual-feasible solution.

Corollary 1. *Let $X \subseteq \mathcal{F}$ be a feasible solution to the CIP, and \mathbf{y} a KC-LP dual feasible solution. If X is a KC-LP relative α -approximationmate solution, i.e.,*

$$c(X) \equiv \sum_{i \in X} c_i \leq \alpha \sum_{j \in U} \sum_{S \subseteq \mathcal{F}} r_j^S y_j^S,$$

then the cost-shares induced by \mathbf{y} have cost-of-fair-sharing at most α .

This implies that the Carr *et al.* [4] algorithm yields cost-allocation with cost-of-fair-sharing at most $\Gamma \leq m$, and that the rounding approach of Koliopoulos and Young [18] produces a cost-of-fair-sharing of $\mathcal{O}(\ln \Delta)$, where $\Delta \leq n$.

Moreover, if small violations of the multiplicity bounds are tolerated, this is improved further on sparse instances [5, 6]. These are the first price-of-fair-sharing bounds for general CIPs. Furthermore, even on the more restricted multi-set multi-cover problem, i.e., CIPs with integer inputs, our work yields improvements. The Γ -bound is new, and the latter $\mathcal{O}(\ln \Delta)$ bound dominates the existing $\mathcal{O}(\max_i \sum_j a_{ij})$ bound of Li *et al.* [22]. Finally, our theorem implies that one can also find cost-shares directly by solving the KC-LP dual. This decouples the problem of finding an integer CIP solution from that of finding cost-shares. Our case study shows this can have great value in practice.

Proof of Theorem 1. The proof is relatively simple; it naturally uses dual feasibility, and a rearrangement of summations that is standard in the analysis of primal-dual schema (see, e.g., [31]). Fix a KC-LP dual feasible solution \mathbf{y} , and let ξ be the induced cost-shares. The main burden of proof is to show that no subset of users has incentive to act separately from the others. In other words, we need to show that for any $J \subseteq \mathcal{U}$, we have $c_J^* \geq \sum_{j \in J} \xi_j$, where c_J^* is the minimum cost of an integer solution serving group J only; this is the minimum cost to the CIP in which only constraints associated with users J are included.

We prove this by considering an optimal solution to the problem of serving the smaller set of users. Let $X_J^* \subseteq \mathcal{F}$ be a minimum-cost solution for serving group J . Using (KC-DP) feasibility of \mathbf{y} , we see that

$$c_J^* = \sum_{i \in X_J^*} c_i \geq \sum_{i \in X_J^*} \left(\sum_{j \in \mathcal{U}} \sum_{S \subseteq \mathcal{F} \setminus \{i\}} a_{ij}^S y_j^S \right) \geq \sum_{i \in X_J^*} \sum_{j \in J} \sum_{S \subseteq \mathcal{F} \setminus \{i\}} a_{ij}^S y_j^S,$$

where the last inequality is due to the fact that all variables are non-negative.

The last part can be shown using a standard change in the order of summation. The right-hand-side expression above counts, for every item $i \in X_J^*$, the sum over subsets $S \subseteq \mathcal{F}$ that do not contain i . Equivalently, one can sum, for each subset $S \subseteq \mathcal{F}$, every item in X_J^* not in S , i.e.,

$$\sum_{i \in X_J^*} \sum_{j \in J} \sum_{S \subseteq \mathcal{F} \setminus \{i\}} a_{ij}^S y_j^S = \sum_{j \in J} \sum_{S \subseteq \mathcal{F}} y_j^S \left(\sum_{i \in X_J^* \setminus S} a_{ij}^S \right). \quad (5)$$

Recall that X_J^* is a feasible solution to the sub-problem of serving users in set J only. The residual demand of each subset S is satisfied by X_J^* for all users in J :

$$\sum_{i \in X_J^* \setminus S} a_{ij}^S \geq r_J^S, \quad \forall j \in J. \quad (6)$$

By applying the inequality (6) to (5), we arrive at the cost allocation to group J under the cost-shares ξ . In summary, we have shown that

$$c(X_J^*) \geq \sum_{j \in J} \sum_{S \subseteq \mathcal{F}} r_j^S y_j^S = \sum_j \xi_j. \quad (7)$$

This proves that each group $J \subseteq \mathcal{U}$ prefers to accept the cost-shares ξ over forming a group on their own. In other words, any (KC-DP) dual solution \mathbf{y} produces cost-shares satisfying the core property. \square

A natural question is whether there are cost-allocations, perhaps not derived from KC-LP dual variables, that have lower worst-case prices-of-fair-sharing. For CIP the answer is *no*; our bounds are provably tight, at least with respect to parameters Γ and Δ . In particular, in the special case in which all contributions and requirements are binary, the KC-LP and its dual reduce to the standard set cover linear programs. Here, the folk theorem applies, and the cost-recovery ratio is provably upper-bounded by the integrality gap [7]. For Set Cover, and hence CIP as well, the integrality gap can be as large as Γ and $\ln \Delta$ [11, 30]. Hence, the lower bounds of our cost-recovery ratios for CIP are tight.

6 A Case Study on LoRaWAN Coverage

We conduct a case study to evaluate the effectiveness of the KC-LP cost-sharing framework on a practical coverage-sharing problem for LoRaWAN. The study is based on a scenario in which the goal is to provide coverage over Brooklyn, NY, a relatively large and densely built urban area. The challenge is to find a gateway placement that provides sufficient coverage throughout the area at minimal cost, and to allocate the cost of the gateways across the users, subject to the core constraint, while recovering as much of the cost as possible. This is the main problem motivating our work. Our results are based on the average performance over a sequence of random problem instances. Our results suggest the KC-LP framework performs well in practice; by solving both the IP and KC-LP dual to optimality, we recover on average 93% of the cost of the gateways.

The set of problem instances is derived from a combination of real-world data and assumptions made in the absence of available data. Each instance is a CIP with random contribution matrix \mathbf{A} , requirements \mathbf{r} , and facilities \mathbf{f} . The demand points, or users \mathcal{U} , are defined as a regular grid of 7,808 points over the study-area. Each instance uses a sub-sample of size 2,000. Facility locations are derived from the corners of real building footprints, and sub-sampled down to 4,380 per instance. This mimics a practical constraint faced by operators who do not have access to every site due to the high access cost. Next, each facility i is associated with a cost c_i uniformly distributed between 0 and 1. For the contribution matrix \mathbf{A} , we use a distance-based Okumura-Hata model with normal noise to generate random connection qualities a_{ij} between each gateway-demand point pair (i, j) [14]. Finally, to ensure feasibility, each demand point has a requirement equal to the value of building all gateways, divided by a geometrically distributed random variable, sampled independently for each user. A total of 10 instances are used.

For each instance, we solve both the CIP and the KC-LP dual to optimality. The CIPs are solved using an off-the-shelf IP solver; the KC-LP dual is solved using the column generation routine, similar to [4]. This produces the optimal cost-shares, in the sense that the price-of-fair-sharing matches the instance-

specific integrality gap exactly. To reiterate, the ability to find cost-shares by solving the KC-LP dual is a valuable consequence of our work. We are not aware of another equally tractable method for finding cost-shares of this quality in practice. We also consider additional benchmark algorithms.

The KC-LP solver is compared against two natural approximation algorithms that produce cost-shares: the PRIMALDUAL algorithm and the GREEDY algorithm. The PRIMALDUAL algorithm, or dual-ascent algorithm, incrementally grows both a feasible KC-LP dual solution, and an infeasible integer CIP solution. The algorithm terminates as soon as the integer solution is feasible. The main idea behind this algorithm is standard (see e.g. [31]), and analogous to that of [3] in a multi-user setting. Next, we also employ an existing GREEDY algorithm that also produces both an integer solution to the CIP, as well as cost-shares [22]. This algorithm produces cost-shares via dual fitting; it greedily selects gateways to add, and amortizes the per-coverage cost of selected facilities into an infeasible KC-LP dual solution. Finally, the dual variables are scaled down by $\log(n)$ to ensure feasibility. We also introduce an improved variant of this algorithm, called GREEDY+. By viewing the greedy-generated cost-shares as KC-LP dual variables, these need to be scaled down by worst-case bound, but only the minimal amount to make them KC-LP dual feasible.

One caveat of both greedy algorithms is that they only seem to work well on CIPs with integer requirements \mathbf{r} and contributions \mathbf{A} [9]. As such, for these algorithms only, we modify the instances by multiplying the inputs \mathbf{A} and \mathbf{r} by a large constant, and then rounding the connections up, and the requirements down, to the nearest integer. This modification never compromises feasibility, and does not increase the minimum cost. However, solutions to the rounded CIPs need not be feasible for the original real-valued instance. This is a potential weakness of using greedy algorithms on real-valued CIPs.

The results are summarized in Fig. 1. Each set of bars along the x-axis represent one instance; the bar heights represent costs. The GREEDY algorithm performs remarkably well, with its solution cost exceeding the minimum by only 10% on average, whereas the primal-dual algorithm performs worse, averaging nearly 30% above optimal. More remarkably, the OPTIMAL method recovers 93% of the cost on average; over twice as much as the next best algorithm. GREEDY recovers relatively little cost, even after the improved dual-fitting GREEDY+; PRIMALDUAL recovers considerably more – 45% of the minimum cost.

Overall, the KC-LP framework finds near perfect cost-shares when sharing the cost of LoRaWAN coverage. Using column generation to solve the KC-LP dual, one can recover nearly all of the cost of reasonably realistic coverage provision instances. If the instances are larger such that solving the IP and KC-LP to optimality is prohibitive, one can use both GREEDY and PRIMAL-DUAL; the former to solve the CIP, the latter to find cost-shares. Alternatively, one can solve the KC-LP to near-optimality using the algorithms methods of [4, 6]. The consistency of performance across this data set (along with other results that achieved analogous performance on variants in which the costs were Gaussian rather than uniform) demonstrate the effectiveness of this theoretically-inspired approach to deliver significant results in our application.

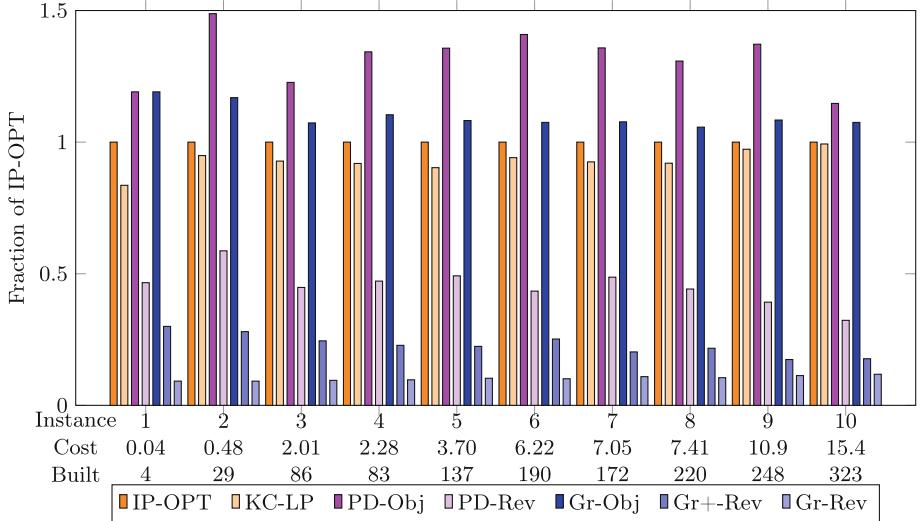


Fig. 1. Costs incurred and recovery for 10 instances. IP-OPT is the cost of the IP optimum, KC-LP the cost of the KC-LP optimum. Prefixes PD, Gr, and GR+, represent the PRIMALDUAL, GREEDY, and GREEDY+, respectively. Suffixes Obj and Rev represent the integer objective cost, and cost-share revenue, respectively.

7 Group-Strategyproof KC-LP Cost-Shares

So far it has been assumed that the group of users to be served is given; sometimes this choice also falls on the service provider. In an extended setting, the goal of the service provider is to elicit private user preferences and design a mechanism for choosing who to serve, in addition to allocating costs to those served. See Jain and Mahdian [16] for more details and definitions. Assume that each user $j \in \mathcal{U}$ has a private utility u_j for receiving service, and has the option of opting out for a utility of 0. When utilities are unknown, they must be elicited by a mechanism. This creates a challenge, because users and groups of users may be able to strategically misreport their utilities. A mechanism is said to be *strategyproof* if individual users cannot benefit from misreporting their utility, and *group-strategyproof* if no group of users can benefit by colluding to misreport their utilities. The goal of a mechanism in this setting is to elicit preferences in a group-strategyproof manner, decide who to serve, and allocate-costs in a way that maximizes the cost-recovery ratio.

To find a group-strategyproof mechanism, Moulin and Shenker [24] prove that it suffices to have cross-monotonic cost-allocation mechanism. Cost-shares $\xi : \mathcal{U} \times 2^{\mathcal{U}} \rightarrow \mathbf{R}^{|\mathcal{U}|}$ are *cross-monotonic* if the cost allocated to user j does not increase when more users are served, i.e., when for all $J \subseteq U$, $\xi_j(U) \leq \xi_j(J)$. Not all cost-shares that satisfy the core property are cross-monotonic. Pál and Tardos [25] provided a general approach for using primal-dual algorithms to find cross-monotonic cost-shares in the facility location problem and the

rent-or-buy problem. This approach has been widely used [17, 19, 21]. Interestingly, imposing cross-monotonicity usually lowers the best attainable cost-recovery ratio (or equivalently, increases the price-of-fair-sharing) [25]. Indeed, Immorlica, Mahdian, and Mirrokni [15] upper bounded the cost-recovery ratio for cross-monotonic cost allocations. In particular, a cross-monotonic cost allocation for set cover can recover at most $1/\Delta$ of the cost, where Δ is the size of the largest set. Moreover, even if all elements are covered by at most two sets, the cost recovery ratio is $\mathcal{O}((2+\epsilon)/n^{1/3})$. Li *et al.* [22] describe a cross-monotonic cost allocation for the CIP that recovers $1/(2n)$ of the cost without explicit use of duality; we simplify their proof using the strengthened dual, generalize the algorithm to general CIP, and improve this to $1/(2\Delta)$.

Our strengthened linear programming approach can be used to find cross-monotonic cost-shares. The approach uses a primal-dual algorithm, in line with the general framework of Pál and Tardos [25]. The algorithm is the same as in Li *et al.* [22], but our analysis is different, in that it uses the KC-LP dual.

Theorem 2. *Fix users $U \subseteq \mathcal{U}$. The primal-dual mechanism [22] produces a feasible solution X for users U , and cross-monotone cost-shares $\xi_j(J)$ that satisfy $\sum_{j \in J} \xi_j(J) \geq (\frac{1}{2\Delta}) \cdot c(X)$.*

The main algorithmic idea is to let each user independently select facilities in complete isolation from the other users. Cross-monotonicity is enforced by preventing any interactions between users' dual variables. Moreover, the problem faced by an individual user is a minimum cost-knapsack problem, each of which can be solved by a primal-dual algorithm [3]. This produces, for each user $j \in U$, a selection of facilities X_j , and a KC-LP dual solution \mathbf{y}_j that is feasible for the individual problem with constraints for user j only, and no variables corresponding to other users. Finally, our mechanism selects the union of all selected facilities $(X_j)_{j \in U}$, and scales down the individual dual variables \mathbf{y}_j by the column sparsity Δ . The procedure is summarized in Algorithm 1, in which `MinCostKnapsackPrimalDual` is given in [3, 22] (and our online version).

Algorithm 1: A cross-monotonic primal-dual

```

Input:  $(\mathcal{F}, \mathcal{U}, \mathbf{r}, \mathbf{c}, \mathbf{A})$ 
for all users  $j \in J$  independently do
     $X_j, \mathbf{y}'_j \leftarrow \text{MinCostKnapsackPrimalDual}(\mathcal{F}, r_j, c, a_j)$ 
     $\mathbf{y}_j \leftarrow \mathbf{y}'_j / \Delta$ 
     $X \leftarrow \bigcup_{j \in \mathcal{U}} X_j$ 
     $\xi_j(U) \leftarrow \frac{1}{2\Delta} \sum_{S \subseteq F} y_j^S$  for all  $j \in U$ 
return  $X, \xi$ 

```

Carnes and Shmoys [3] develop and analyze a primal-dual algorithm for the minimum-cost knapsack problem based on the KC-LP formulation. Fix a single user $j \in \mathcal{U}$ and let $\mathbf{a}_j = (a_{1j}, \dots, a_{nj})$ denote their contributions. The user starts with an all-zero dual solution \mathbf{y}_j , and an empty selection $X = \emptyset$. While the residual demand r_j^X is positive, they increase the dual variable y_j^X . Eventually, some constraint $\sum_{S \subseteq F - \{i\}} a_{ij}^S y_j^S \leq c_i$ becomes tight for a facility i . This facility

is added to the selection X , and the process repeats. The algorithm returns a selection of facilities X_j and a feasible KC-LP dual solution. Critically, the cost of X_j is at most twice the KC-LP dual objective under \mathbf{y}_j .

Theorem 3 (Carnes and Shmoys [3]). *Let $X_j \subseteq \mathcal{F}$ and \mathbf{y}_j be a selection of facilities, and the corresponding dual solution returned by the primal-dual algorithm for min-cost knapsack. These dual variables are feasible, and satisfy*

$$\sum_{i \in X_j} c_i \leq 2 \sum_{S \subseteq \mathcal{F}} r_j^S y_j^S.$$

This result is used in two parts of our proof. First, we use the dual feasibility of the individual dual variables \mathbf{y}_j to construct a feasible dual solution to the master CIP, which gives us the core-property via Theorem 1. Secondly, the approximation ratio is used to derive our cost-recovery ratio.

Proof of Theorem 2. To prove this result, we need to argue for cross-monotonicity, the core-property, and cost recovery. We argue for cross-monotonicity first. Clearly, the dual variables \mathbf{y}'_j of each user are independent of other users. Meanwhile, the maximum number of users in U served by any facility, Δ , is monotonically increasing in the size of U , so the dual variables $\mathbf{y}_j = \mathbf{y}'_j / \Delta$ are monotonically decreasing in U , as are the induced cost-shares.

The core property is easy to prove using dual feasibility and our Theorem 1. Consider some selected facility $i \in X$. Then,

$$\sum_{j \in U} \sum_{S \subseteq \mathcal{F} \setminus \{i\}} a_{ij}^S y_j^S = \frac{1}{\Delta} \sum_{\{j \in U : a_{ij} > 0\}} \left(\sum_{S \subseteq \mathcal{F} \setminus \{i\}} a_{ij}^S y'_j^S \right) \leq \frac{1}{\Delta} \sum_{\{j \in U : a_{ij} > 0\}} c_i \leq c_i.$$

The first equality follows from dropping users not served by facility i . The second equality uses the definition of $\mathbf{y}_j = \mathbf{y}'_j / \Delta$. The following inequality uses the fact that the dual variables are individually feasible to the min-cost knapsack LP of each user j . The final inequality follows from the definition of Δ . The above shows that dual variables $(\mathbf{y}_j)_{j \in U}$ are feasible for the CIP induced by users U , and thus Theorem 1 implies that the core-property is satisfied by the accompanying cost-shares $(\xi_j(U))_{j \in U}$.

Finally, the cost-recovery recovery ratio follows from the approximation ratio of the minimum-cost knapsack algorithm. In particular, observe that

$$\sum_{i \in X} c_i \leq \sum_{j \in U} \sum_{i \in X_j} c_i \leq 2 \sum_{j \in U} \sum_{S \subseteq \mathcal{F}} r_j^S y_j^S = 2\Delta \sum_{j \in U} \sum_{S \subseteq \mathcal{F}} r_j^S y_j^S.$$

The first inequality is obvious; the second follows from applying Theorem 3 to each user j in U individually. The equality follows from the definition of \mathbf{y}_j . This proves that the cost-of-fair-sharing is at most 2Δ , as claimed. \square

Finally, the proof suggests there is potential for improvement. In fact, whenever the contributions \mathbf{A} are binary, the min-cost knapsack algorithm is exact,

in which case the cost-recovery ratio is $1/\Delta$ [22]. Moreover, if we know that each selected facility X always selected by at least two users, it also follows that the cost-of-fair-sharing is a factor 2 smaller, i.e. Δ . On the other hand, no group-strategyproof mechanism can recover more than $1/\Delta$ of the cost in general [15]. Whether the 2Δ cost-of-fair-sharing is tight when contributions are non-binary, however, remains an open problem [22].

Acknowledgements. This material is based on work supported by the NSF under Grant CNS-1952063.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. LPWAN Market. <https://iot-analytics.com/lpwlan-market/>. Accessed 29 Jun 2024
2. Aarts, S., Shmoys, D.B., Coy, A.: An interpretable determinantal choice model for subset selection (2023)
3. Carnes, T., Shmoys, D.: Primal-dual schema for capacitated covering problems. In: 13th International Conference on Integer Programming and Combinatorial Optimization, IPCO 2008, Bertinoro, Italy, 26–28 May 2008 Proceedings 13, pp. 288–302. Springer (2008)
4. Carr, R.D., Fleischer, L.K., Leung, V.J., Phillips, C.A.: Strengthening integrality gaps for capacitated network design and covering problems. In: Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '00, pp. 106–115. Society for Industrial and Applied Mathematics, USA (2000)
5. Chekuri, C., Quanrud, K.: On Approximating (Sparse) Covering Integer Programs. In: Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1596–1615. SIAM (2019)
6. Chen, A., Harris, D.G., Srinivasan, A.: Partial resampling to approximate covering integer programs. Random Struct. Algorithms **58**(1), 68–93 (2021)
7. Deng, X., Ibaraki, T., Nagamochi, H.: Algorithmic aspects of the core of combinatorial optimization games. Math. Oper. Res. **24**(3), 751–766 (1999)
8. Devanur, N.R., Mihail, M., Vazirani, V.V.: Strategyproof cost-sharing mechanisms for set cover and facility location games. In: Proceedings of the 4th ACM Conference on Electronic Commerce, pp. 108–114 (2003)
9. Dobson, G.: Worst-case analysis of greedy heuristics for integer programming with nonnegative data. Math. Oper. Res. **7**(4), 515–531 (1982)
10. Faulhaber, G.R.: Cross-subsidization: pricing in public enterprises. Am. Econ. Rev. **65**(5), 966–977 (1975)
11. Feige, U.: A threshold of $\ln n$ for approximating set cover. J. ACM (JACM) **45**(4), 634–652 (1998)
12. Goemans, M.X., Skutella, M.: Cooperative facility location games. J. Algorithms **50**(2), 194–214 (2004)
13. Gupta, A., Kumar, A., Pál, M., Roughgarden, T.: Approximation via cost-sharing: a simple approximation algorithm for the multicommodity rent-or-buy problem. In: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science, 2003, pp. 606–615. IEEE (2003)

14. Hata, M.: Empirical formula for propagation loss in land mobile radio services. *IEEE Trans. Veh. Technol.* **29**, 317–325 (1980)
15. Immorlica, N., Mahdian, M., Mirrokni, V.S.: Limitations of cross-monotonic cost-sharing schemes. *ACM Trans. Algorithms (TALG)* **4**(2), 1–25 (2008)
16. Jain, K., Mahdian, M.: Cost Sharing, pp. 385–410. Cambridge University Press (2007)
17. Jain, K., Vazirani, V.V.: Equitable cost allocations via primal-dual-type algorithms. *SIAM J. Comput.* **38**(1), 241–256 (2008)
18. Koliopoulos, S.G., Young, N.E.: Approximation algorithms for covering/packing integer programs. *J. Comput. Syst. Sci.* **71**(4), 495–505 (2005)
19. Könemann, J., Leonardi, S., Schäfer, G., van Zwam, S.H.: A group-strategyproof cost sharing mechanism for the Steiner forest game. *SIAM J. Comput.* **37**(5), 1319–1341 (2008)
20. Lawler, E.L.: Fast approximation algorithms for knapsack problems. In: 18th Annual Symposium on Foundations of Computer Science, SFCS 1977, pp. 206–213. IEEE (1977)
21. Leonardi, S., Schäfer, G.: Cross-monotonic cost sharing methods for connected facility location games. *Theor. Comput. Sci.* **326**(1), 431–442 (2004). <https://www.sciencedirect.com/science/article/pii/S0304397504004876>
22. Li, X.Y., Sun, Z., Wang, W., Lou, W.: Cost sharing and strategyproof mechanisms for set cover games. *J. Comb. Optim.* **20**(3), 259–284 (2010)
23. Meir, R., Bachrach, Y., Rosenschein, J.S.: Minimal subsidies in expense sharing games. In: Algorithmic Game Theory: Third International Symposium, SAGT 2010, Athens, Greece, 18–20 October 2010, Proceedings 3, pp. 347–358. Springer (2010)
24. Moulin, H., Shenker, S.: Strategyproof sharing of submodular costs: budget balance versus efficiency. *Econ. Theor.* **18**, 511–533 (2001)
25. Pál, M., Tardos, É.: Group strategyproof mechanisms via primal-dual algorithms. In: Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science, 2003, pp. 584–593. IEEE (2003)
26. Parsons, S.G.: Cross-subsidization in telecommunications. *J. Regul. Econ.* **13**, 157–182 (1998)
27. Plotkin, S.A., Shmoys, D.B., Tardos, É.: Fast approximation algorithms for fractional packing and covering problems. *Math. Oper. Res.* **20**(2), 257–301 (1995)
28. Semtech: LoRa® and LoRaWAN®: a technical overview. <https://lora-developers.semtech.com/documentation/tech-papers-and-guides/lora-and-lorawan/>. Accessed 07 Sept 2021
29. Semtech: Network Providers. <https://www.semtech.com/lora/ecosystem/networks>. Accessed 29 Jun 2024
30. Singh, M.: Integrality gap of the vertex cover linear programming relaxation. *Oper. Res. Lett.* **47**(4), 288–290 (2019)
31. Williamson, D.P., Shmoys, D.B.: The Design of Approximation Algorithms. Cambridge University Press (2011)