Optimization Letters manuscript No.
(will be inserted by the editor)

Taming The Set Covering Problem: The Value of Dual Information

Belma Yelbay, Ş. İlker Birbil and Kerem Bülbül

Received: date / Accepted: date

Abstract We offer a fresh perspective on solving the set covering problem to near optimality with off-the-shelf methods. We formulate minimizing the gap of a generic primal-dual heuristic for the set covering problem as an integer program and analyze its performance. The empirical insights from this analysis lead to a simple and powerful primal-dual approach for solving the set covering problem to near optimality with a state-of-the-art standard solver. We demonstrate the effectiveness of this approach on a rich set of benchmark instances compiled from the literature. We conclude that set covering problems of various characteristics and sizes may reliably solved to near optimality without resorting to custom solution methods.

Keywords set covering · exact method · primal-dual heuristic · LP relaxation · dual information · empirical study

1 Introduction

With the boost in computing technology and the striking advances in linear programming (LP) solvers, many large-scale combinatorial optimization problems can now be solved in a reasonable time. Although the performance of integer programming (IP) solvers is not comparable to that of LP solvers, many moderate-size hard IP problems in academic and industrial contexts are being solved with an increasing success every passing day. Consider for instance the famous state-of-the-art `CPLEX` solver, which has just become free for academic use. As of 2010, the latest version of its mixed integer linear programming solver is now on average 50% faster than its earlier releases in

Belma Yelbay, Ş. İlker Birbil and Kerem Bülbül
Sabancı University, Manufacturing Systems and Industrial Engineering, Orhanlı-Tuzla, 34956 Istanbul, Turkey
E-mail: byelbay@su.sabanciuniv.edu, sibirbil@sabanciuniv.edu, bulbul@sabanciuniv.edu

the last ten years. Even more impressive, this latest version has improved the performance of the previous version by 10% within merely six months [22].

Realizing these promising developments with the exact methods, we revisit the set covering problem (SCP) and conduct an empirical study on a set of problems that appeared in the literature over the last three decades. This famous problem can be defined as follows.

Definition 1 Given a collection $\mathcal{S}$ of sets over a finite universe $\mathcal{U}$, a set cover $J \subseteq \mathcal{S}$ is a sub-collection of these sets, whose union is $\mathcal{U}$. When each set in the collection has an associated cost, then the set covering problem is about finding a set cover J such that the total cost is minimized.

Our motivation for selecting SCP is two-fold: First, SCP has a wide popularity among researchers and practitioners because a wide range of applications from scheduling to routing, and from manufacturing to telecommunications can be cast as set covering problems, possibly with side constraints. Second, this wide interest allows us to review a large body of work from the literature as well as access many acknowledged and frequently studied set of problems. To obtain a fairly representative problem set, we have strived to compile from the literature not only the research problems but also some actual problems that arise in practice.

We need to emphasize that the need for fast and efficient heuristic methods persists especially for large-scale combinatorial problems. On this account, SCP is no different. It is fair to say that unless $\mathcal{NP} = \mathcal{P}$, there will always be hard SCP instances that are intractable with the exact methods. On one hand the competition between heuristic and exact methods becomes fiercer and fiercer. On the other hand, it is also known that exact and heuristic methods can be complementary to each other. Leveraging on this idea, we also propose such a complementary approach by considering the LP-IP relationship, particularly, through the dual information. Based on our comprehensive empirical study, we shall indeed infer that the dual information is a significant source for designing new heuristics and approximation methods with excellent empirical performance. While primal-dual heuristics for the SCP have been thoroughly investigated in the context of approximation algorithms, we believe that their full empirical potential is yet to be realized. We take a first step in this direction.

2 Literature Review

SCP is long known to be $\mathcal{NP}$-hard in the strong sense [16]. Therefore, many heuristic and enumerative algorithms have been developed to solve SCP effectively. Exact algorithms generally rely on the branch-and-bound method to obtain optimal solutions [3,5,7,15]. Since solving a large SCP with an exact method takes an excessively long time, sacrificing optimality but obtaining fairly good solutions within an acceptable time by means of a heuristic is a compromise fix. Many researchers list various heuristics and approximation algorithms, and they show that their empirical performance is quite good [11, 17, 18]. The simplest algorithms are the greedy algorithms, which can be used to solve large-scale set covering problems almost in no time. However, their myopic nature may easily yield solutions far from optimality. Haouari and Chaouachi [20] as well as Vasko and Wilson [32] introduce randomness and penalization into the greedy algorithms to improve the solution quality. Along this line, two recent local search heuristics appear in [24, 35]. Finger et al. conduct an analysis on some benchmark instances by measuring the correlation between the cost of a solution and the closeness to the optimal solution [14]. This study gives useful insights to understand the problem structure and to develop problem-specific local search algorithms. Several meta-heuristics have also been proposed. Among these, we can list simulated annealing [9, 23], genetic algorithms [1, 6, 25], tabu search [12, 28], ant colony optimization [30] and electromagnetism meta-heuristic [2]. In a recent work, Muter et al. give a generic framework that uses LP-relaxation information for promoting meta-heuristics to diversify or intensify while searching for the optimum [29].

Similar to our work in this paper, several studies design heuristics based on the Lagrangian relaxation or the LP-relaxation [10, 13, 21]. The primal-dual approach is commonly used for approximating $\mathcal{NP}$-hard optimization problems that can be modeled as IP problems such as the metric traveling salesman problem, the Steiner tree problem, the Steiner network problem, and the set covering problem [33]. Bar-Yehuda and Even [4] are the first researchers who consider a generic primal-dual approach to approximate the set covering problem. The primal-dual approach is based on finding only a feasible solution to the dual of the LP-relaxation of the IP formulation of SCP presented in the next section. Using this solution, an integral solution for the SCP is constructed. Although the worst case performance of the primal-dual algorithm of Bar-Yehuda and Even [4] is poor [19], its empirical performance turns out to be much more promising. Therefore, several studies that have sprung out of the primal-dual approach [8, 27, 34, 36] in the set covering literature.

3 Primal-Dual Methods Revisited

In this section, we discuss in-depth our motivation for using the relationship between the IP formulation of the set covering model and its LP-relaxation. In a nutshell, we gather dual information from the optimal solution of the LP-relaxation, and then reduce the problem size considerably so that the resulting SCP can be solved by an IP solver with much less computational effort.

Before delving into the details of this approach, we first give the mathematical model of the SCP. Using Definition 1, we obtain the integer programming model of the SCP as

$$\text{minimize} \quad \sum_{j \in \mathcal{S}} c_j x_j, \tag{1}$$

$$\text{subject to} \quad \sum_{j \in \mathcal{S}} a_{ij} x_j \geq 1, \; i \in \mathcal{U}, \tag{2}$$

$$x_j \in \{0, 1\}, \;\; j \in \mathcal{S}. \tag{3}$$

Here $c_j > 0$ is the coverage cost of the jth set, x_j is a binary variable, which is equal to 1, if $j \in J$, and a_{ij} is a binary parameter, which is equal to 1, if item i is covered by the jth set. The set of constraints (2) ensures that each item is covered by at least one set, and the constraints (3) impose integrality on the variables. If the cost of coverage is the same for each set; that is, $c_1 = c_2 = \cdots = c_n$ with $n = |\mathcal{S}|$, then the problem is called the unicost set covering problem. When we consider the LP-relaxation of the SCP, the integrality constraints (3) are replaced by simple bounds on the variables and a continuum of values is considered for the solution. The dual of the LP-relaxation of SCP is then given by:

$$\text{maximize} \sum_{i \in \mathcal{U}} y_i, \tag{4}$$

$$\text{subject to} \sum_{i \in \mathcal{U}} a_{ij} y_i \leq c_j, \qquad j \in \mathcal{S}, \tag{5}$$

$$y_i \geq 0, \qquad i \in \mathcal{U}, \tag{6}$$

where the dual variables y_i, $i \in \mathcal{U}$ correspond to the coverage constraints in the LP-relaxation of (1)-(3).

In the subsequent discussion, we shall describe the proposed approach and provide computational results. Let us first introduce the problem sets and the testing environment that we used in our empirical study. We classify the compiled problems into the following categories:

(a) Standard benchmark problems from the OR-library [7] (70 instances).
(b) Euclidean-type cost and coverage correlated problems [36] (320 instances).
(c) Airline problems [3] (16 instances).
(d) Railway problems [13] (7 instances).
(e) Hard cost and coverage correlated problems as defined in [31] (30 instances).
(f) Combinatorial [18] and Steiner triple [26] unicost problems (17 instances).

It is fair to state that some of the problems that we include in the compilation are not as standard as those used recurrently in the literature. However, this is inevitable because we believe that most of the standard benchmark problems solved in many past studies should be considered as relatively easy nowadays. For example, a group of instances from the highly cited OR-library (a) can be solved to optimality within less than a second on average by a standard IP solver. Therefore, we focus on hard instances (e-f) and large-scale practical problems (c-d). Using Euclidean-type distances, we also solve a set of randomly generated cost and coverage correlated instances (b), which commonly arise in location, telecommunication and routing problems.

The optimal LP and IP solutions in this study were obtained by `ILOG IBM CPLEX 12.1` running on a personal computer with an Intel Core i5 processor and 4 GB of RAM. In all problem instances, the upper limit on the solution time is set to 7200 seconds. The batch processing of the instances is carried out through simple C++ scripts. Our computational results are summarized in this paper. However, the entire set of results is available online[1].

3.1 Minding The Gap

As mentioned previously, our motivation is to use the LP information to obtain an integral solution for the SCP. A straightforward approach is to solve the LP relaxation and then use the dual information to identify the columns with zero reduced costs. These columns can be considered as promising ones that should likely appear in the IP optimal solution. Along this line, for instance, Hochbaum [21] solves the dual LP (4)- (6) and constructs a set cover composed of all primal variables with a zero reduced cost. Such approaches fall into the general category of primal-dual methods. Primal-dual methods find a feasible solution for the (primal) IP model (1)-(3) and a feasible solution for the dual LP model (4)-(6). In fact, the dual optimal solution can be obtained easily, since solving the LP model to optimality is not a major concern with the current status of the LP solvers. Using then elementary duality and the relation between the IP model and its LP relaxation, it is easy to see that the objective function values of a feasible IP solution and the optimal LP solution yield a pair of upper and lower bounds for the overall SCP, respectively. Therefore, the main drive behind the primal-dual methods is to find a way to minimize the gap between the objective function values of a feasible IP solution and that of an optimal or feasible dual LP solution.

This important relationship between the IP formulation and its LP relaxation has prompted us to concentrate on the best possible result that can be obtained by a primal-dual heuristic that only adds a set to the cover if the associated reduced cost is zero with respect to a feasible solution of (4)-(6). This consideration boils down to finding an optimal solution for the following mixed integer linear programming (MILP) model:

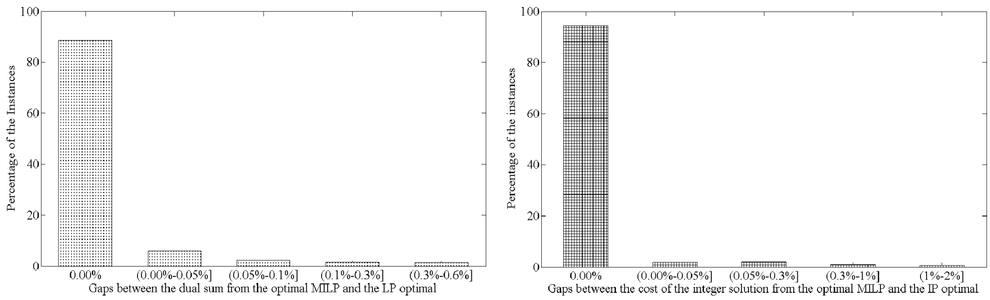
$$\begin{aligned}
\text{minimize} \quad & \sum_{j\in\mathcal{S}} c_j x_j - \sum_{i\in\mathcal{U}} y_i, && && (7)\\
\text{subject to} \quad & \sum_{i\in\mathcal{U}} a_{ij} y_i + z_j = c_j, && j \in \mathcal{S} && (8)\\
& 0 \le z_j \le (1 - x_j)c_j, && j \in \mathcal{S}, && (9)\\
& x_j \in \{0, 1\}, && j \in \mathcal{S}, && (10)\\
& y_i \ge 0, && i \in \mathcal{U}. && (11)
\end{aligned}$$

In this model the set of constraints (8) ensures the dual feasibility, while the constraints (9) prescribe that a primal variable set to 1 has a zero reduced cost.

[1] http://people.sabanciuniv.edu/sibirbil/scp/

First, we solved all test problems with known optimal solutions by the MILP model (7)-(11). Figure 1(a) shows the average percentage gap between the sum of the dual variables $\sum_{i \in \mathcal{U}} y_i$ in the optimal solution of the MILP model and the optimal objective function value of the dual LP (4)-(6). Similarly, Figure 1(b) depicts the average gap of the cost of the primal integer solution $\sum_{j \in \mathcal{S}} c_j x_j$ in the optimal solution of the MILP model to the optimal objective function value of (1)-(3).

Figure 1 has a very important implication; the feasible IP solution obtained from the proposed MILP coincides with the optimal IP solution in almost all problem instances. Furthermore, the sum of the dual variables resulting from solving the proposed MILP is equal, or very close, to the optimal LP-relaxation solution. Of course, solving the proposed MILP is not simpler than solving the original SCP. However, the results obtained with the proposed MILP show that the dual optimal solution bears invaluable information that could help us select the optimal sets for the SCP. In the subsequent discussion, we concentrate on exploiting this dual information obtained by solving the LP relaxation.



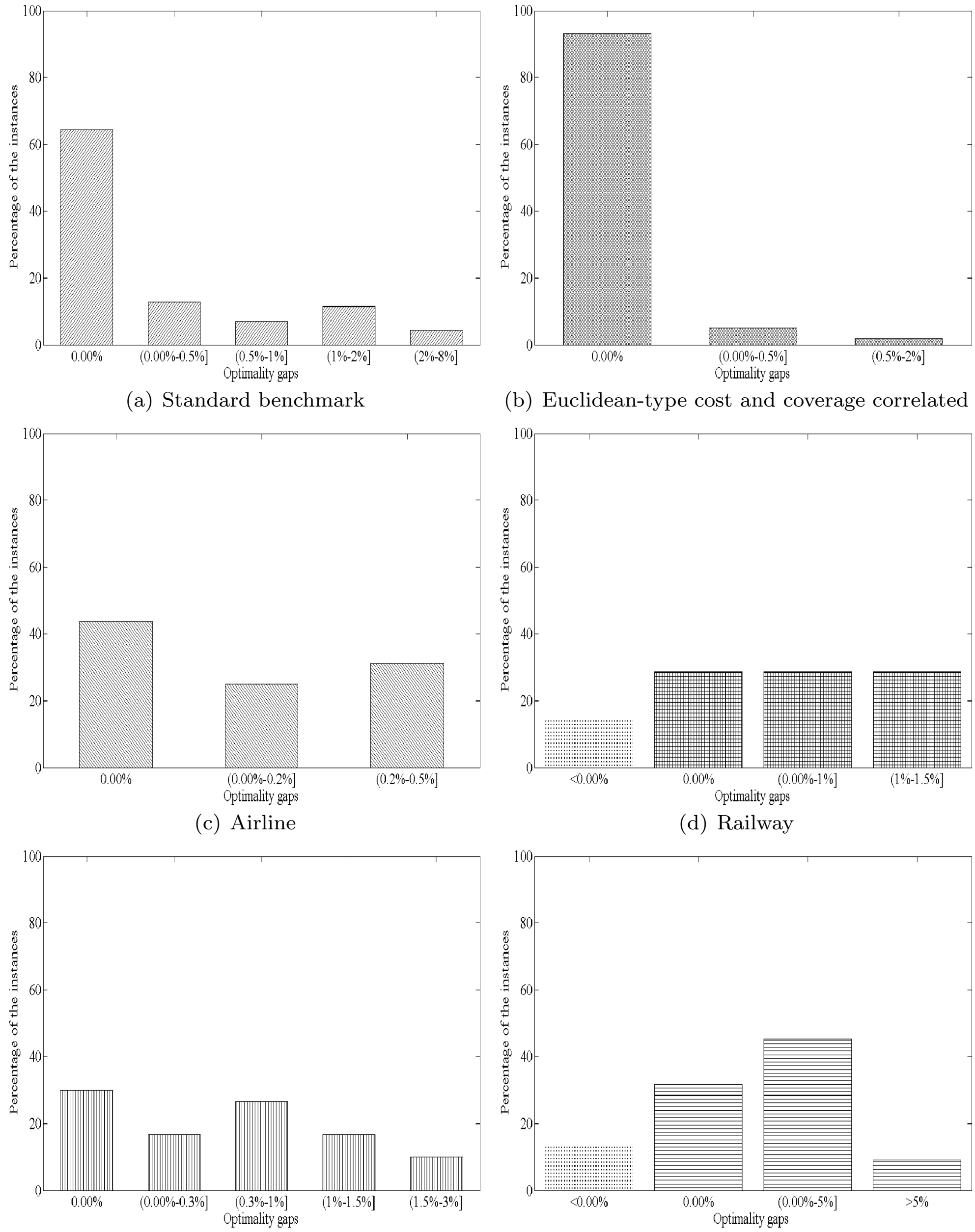
(a) LP Gap (b) IP Gap

Fig. 1 Average LP-IP percentage gaps: (a) between the optimal objective function value of the LP relaxation model and the sum of the duals from the MILP model, (b) between the optimal objective function value of the SCP model and the cost of the primal integer solution from the MILP model.

3.2 Columns With Zero Reduced Costs

Since solving the MILP model is too demanding, we attack it in two phases. In the first phase, we solve the LP relaxation of (1)-(3) and identify all columns with zero reduced costs. Then, in the second phase, we obtain an integer feasible solution by solving (1)-(3) over these columns only. We refer to this IP as the *restricted IP* or the *restricted SCP problem*.

Figure 2 represents the average optimality gap in percentages for *all* problem instances in the problem sets (a) to (f) obtained with this approach. The optimality gaps are computed with respect to the optimal or the best known solutions of these instances in the literature. These figures demonstrate that



(a) Standard benchmark (b) Euclidean-type cost and coverage correlated

(c) Airline (d) Railway

(e) Hard cost and coverage correlated (f) Unicost

Fig. 2 Average optimality gaps for the restricted SCP solved over columns with zero reduced costs.

extremely good results can be obtained when the IP formulation (1)-(3) is solved over the columns with zero reduced costs. Except for a very few problem instances, the optimality gap is always within 2%. Only for some of the instances in the sets (a), (e) and (f), the optimality gaps are slightly larger. This can be partly explained by the imposed restriction on the solution time (7200 seconds). For the unicost problems (f), all columns have a zero reduced cost after solving the LP relaxation. Thus, in this case we end up solving the original problem in the second phase. However, note that we have improved the best-known solutions in the literature for some of the instances in the sets (d) and (f) as implied by the negative percentages for the first bars of Figure 2(d) and Figure 2(f).

The next question is to investigate how much the results deteriorate, when they are compared against the optimal SCP solutions obtained with `CPLEX`. To this end, we have selected the problem sets (a-c) and (e) that can all be solved to optimality within our time limit. The stacked bar plot in Figure 3 summarizes these results. We report the average percentage time reduction when compared against the solution time of `CPLEX`, average percentage reduction in the problem size over the original size, and the average percentage optimality gap. We note that the reported times for our approach also include the time to solve the LP relaxation in the first phase. Figure 3 clearly shows that considerable reduction can be achieved both in solution time and problem size with a negligible increase in the optimality gap.

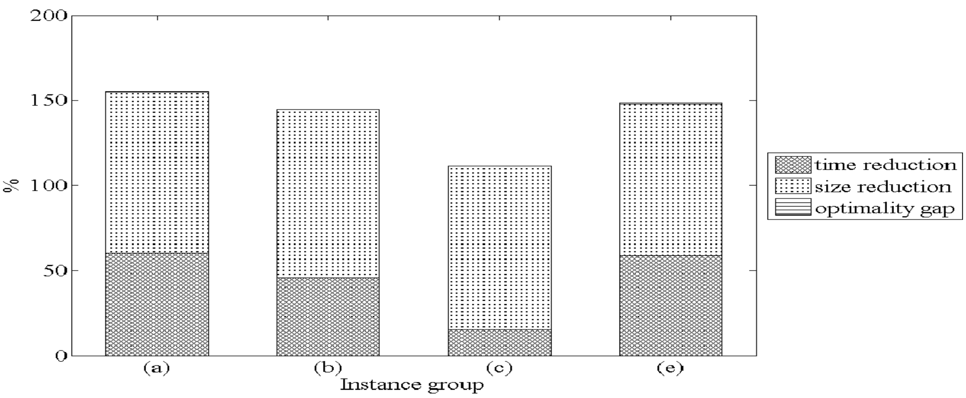


Fig. 3 Average percentage time and size reduction and the average optimality gap of the restricted IP solved over columns with zero reduced costs with respect to the original SCP formulation solved by `CPLEX` to optimality within the time limit.

3.3 Basic Columns

Even though we reduce the size of the problem by only considering columns with zero reduced costs, we may still end up solving a large restricted IP. This

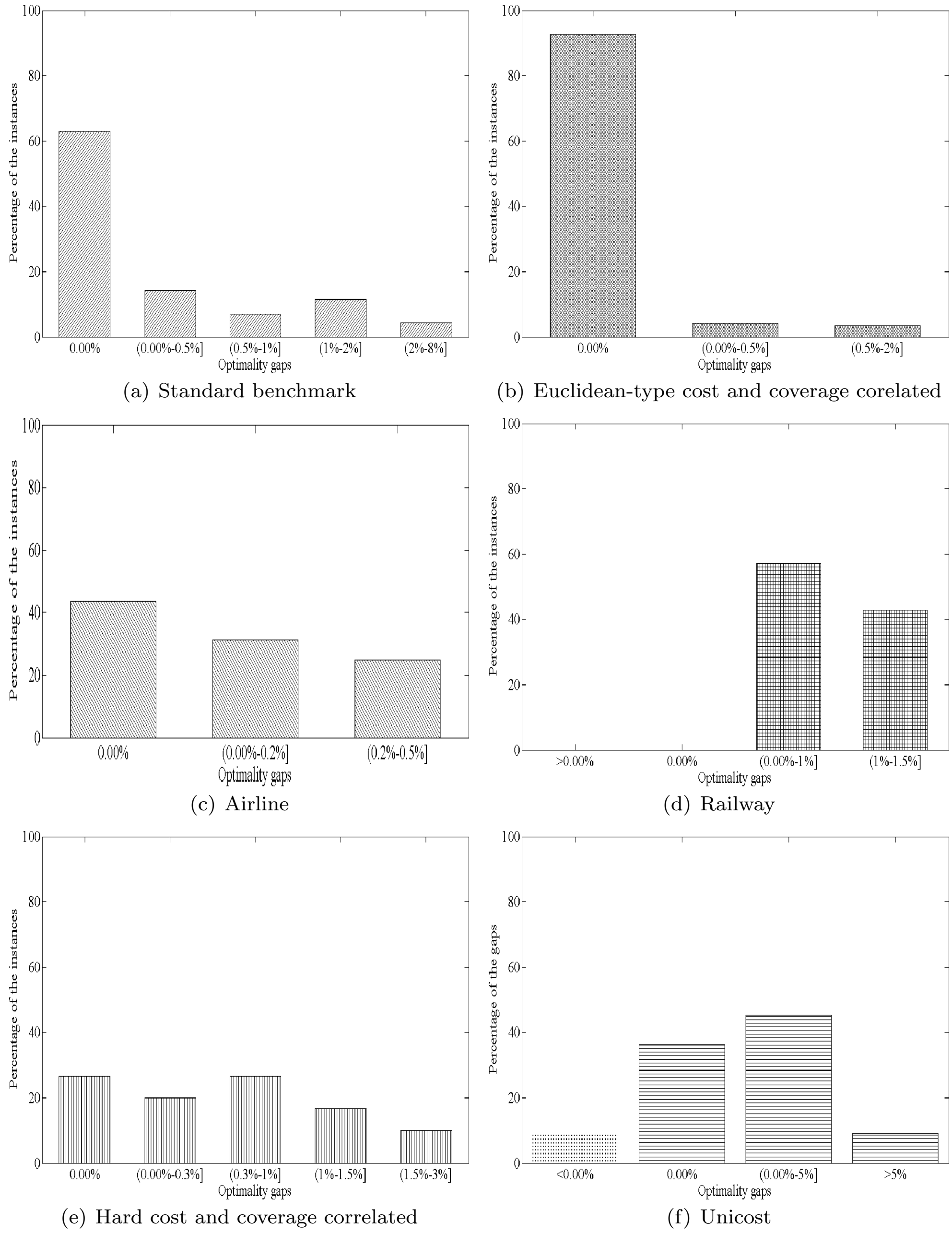


Fig. 4 Average optimality gaps for the restricted SCP solved over basic columns.

is indeed a valid concern as we observed with some of the unicost problems, for which the size of the restricted IP in the previous section is identical to that of the original problem. Therefore, as an alternate method, we propose to solve a restricted IP only over the columns that are basic at the optimal solution of the LP relaxation. There is a great benefit of this approach: we can both reduce the size of restricted IP and also know its exact size in advance.

Figure 4 illustrates the optimality gaps when the restricted SCP is solved over the basic columns. As our results demonstrate, the compromise solution of using only the basic columns is still comparable with the solutions obtained by using all columns with zero reduced costs. However, for the problem set (d) even this smaller restricted IP cannot be solved to optimality within the time limit, and it may be possible to further improve the integral solution by increasing the time limit. Figure 5 exhibits similar benefits to those observed in Figure 3.

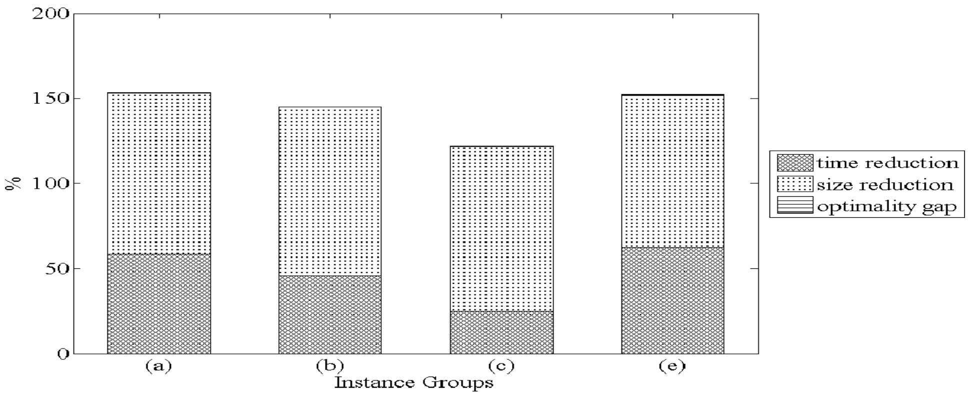


Fig. 5 Average percentage time and size reduction and the average optimality gap of the restricted IP solved over basic columns with respect to the original SCP formulation solved by `CPLEX` to optimality within the time limit.

4 Conclusions and Future Research Directions

Our computational study supports that the dual optimal solution of the LP-relaxation of the SCP supplies an important instrument for tackling this celebrated problem. Leveraging on this information, significant reductions in problem size and gains in solution quality can be achieved for large-scale instances which otherwise are out of reach for off-the-shelf solvers.

As our results demonstrate, there is a trade-off between incorporating all columns with zero reduced costs in the restricted SCP versus solving this IP over basic columns only. Clearly, the former yields integer solutions of higher quality and suggests that an algorithm may benefit from visiting alternate optimal solutions of the LP-relaxation of SCP. It is yet to be determined which of these multiple optimal solutions plays a more significant role in improving the IP solution. This may be an interesting path to explore for simple primal-dual heuristics as well as more sophisticated local search methods for the SCP.

It is well-known that in many practical applications, such as vehicle routing, scheduling and so on, a large-scale SCP is solved within a column generation or a branch-and-price setting. In such cases, embedding the partial dual information from the optimal solutions of the restricted master problems into the overall solution algorithm may prove useful for speeding up the convergence.

We also observed that a large class of standard benchmark instances for the SCP can be solved very efficiently by standard exact methods. There is a clear need for gathering new problem sets for benchmarking purposes. However, we emphasize that most of the frequently used unicost problem instances remain hard for off-the-shelf solvers.

References

1. Aickelin, U.: An indirect genetic algorithm for set covering problems. Journal of the Operations Research **53**(10), 1118–1126 (2002)
2. Azimi, Z., Toth, P., Galli, L.: An electromagnetism metaheuristic for the unicost set covering problem. European Journal of Operational Research **205**, 290–300 (2010)
3. Balas, E., Carrera, M.: A dynamic subgradient-based branch-and-bound procedure for set covering. Operations Research **44**(6), 875–890 (1996)
4. Bar-Yehuda, R., Even, S.: A linear-time approximation algorithm for the weighted vertex cover problem. Journal of Algorithms **2**, 198–203 (1981)
5. Beasley, J.: An algorithm for set covering problem. European Journal of Operational Research **31**, 85–93 (1987)
6. Beasley, J., Chu, P.: A genetic algorithm for the set covering problem. European Journal of Operational Research **94**, 392–404 (1996)
7. Beasley, J., Jornsten, K.: Enhancing an algorithm for set covering problems. European Journal of Operational Research **58**, 293–300 (1992)
8. Bertsimas, D., Vohra, R.: Rounding algorithms for covering problems. Mathematical Programming **80**, 63–89 (1998)
9. Brusco, M., Jacobs, L., Thompson, G.: A morphing procedure to supplement a simulated annealing heuristic for cost- and coverage-correlated set-covering problems. Annals of Operations Research **86**, 611–627 (1999)
10. Caprara, A., Fischetti, M., Toth, P.: A heuristic method for the set covering problem. Operations Research **47**(5), 730–743 (1999)
11. Caprara, A., Toth, P., Fischetti, M.: Algorithms for the set covering problem. Annals of Operations Research **98**, 353–371 (2000)
12. Caserta, M.: Tabu search-based metaheuristic algorithm for large-scale set covering problems. Springer (2007)
13. Ceria, S., Nobili, P., Sassano, A.: A Lagrangian-based heuristic for large-scale set covering problems. Mathematical Programmimg **81**, 215–228 (1998)
14. Finger, M., Stützle, T., Lourenço, H.: Exploiting fitness distance correlation of set covering problems. Lecture Notes in Computer Science **2279**, 61–71 (2002)
15. Fisher, M., Kedia, P.: Optimal solution of set covering/partitioning problems using dual heuristics. Management Science **36**(6), 674–687 (1990)
16. Garey, M., Johnson, D.: Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman (1979)
17. Gomes, F., Meneses, C., Pardalos, P., Viana, G.: Experimental analysis of approximation algorithms for the vertex cover and set covering problems. Computers and Operations Research **33**, 3520–3534 (2006)

18. Grossman, T., Wool, A.: Computational experience with aproximation algorithms for the set covering problem. European Journal of Operational Research **101**, 81–92 (1997)
19. Hall, N., Vohra, R.: Pareto optimality and a class of set covering heuristics. Annals of Operations Research **43**, 279–284 (1993)
20. Haouari, M., Chaouachi, J.: A probabilistic greedy search algortihm for combinatorial optimization with application to the set covering problem. Journal of the Operational Research Society **53**, 792–799 (2002)
21. Hochbaum, D.: Approximation algorithms for the set covering and vertex cover problems. SIAM Journal on Computing **11**, 555–556 (1982)
22. IBM: (2010). CPLEX Optimizer performance benchmark details.
23. Jacobs, L., Brusco, M.: A local search heuristic for large set-covering problems. Naval Research Logistics **42**, 1129–1140 (1995)
24. Lan, G., DePuy, G., Whitehouse, G.: An effective and simple heuristic for the set covering problem. European Journal of Operational Research **176**, 1387–1403 (2007)
25. Lorena, L., Lopes, L.: Genetic algorithms applied to computationally difficult set covering problems. Journal of Operational Research Society **48**, 440–445 (1997)
26. Mannino, C., Sassano, A.: Solving hard set covering problems. Operations Research Letters **18**, 1–5 (1995)
27. Melkonian, V.: New primal-dual algorithms for steiner tree problems. Computers and Operations Research **34**, 2147–2167 (2007)
28. Musliu, N.: Local search algorithm for unicost set covering problem. Lecture Notes in Artificial Intelligence **4031**, 302–311 (2006)
29. Muter, I., Birbil, S., Sahin, G.: Combination of metaheuristic and exact algorithms for solving set covering-type optimization problems. INFORMS Journal on Computing **22**, 603–619 (2010)
30. Ren, Z., Feng, Z., Ke L.J., Z., Z.J.: New ideas for applying ant colony optimization to the set covering problem. Computers and Industrial Engineering **58**, 774–784 (2010)
31. Rushmeier, R., Nemhauser, G.: Experiments with parallel branch-and-bound algorithms for the set covering problem. Computers and Industrial Engineering **58**, 774–784 (2010)
32. Vasko, F., Wilson, G.: An efficient heuristic for large set covering problems. Naval Research Logistics Quarterly **31**, 163–171 (1984)
33. Vazirani, V.: Primal-dual schema based approximation algorithms. Springer (2002)
34. Williamson, D.: The primal-dual method for approximation algorithms. Mathematical Programming **91**, 447–478 (2002)
35. Yagiura, M., Kishida, M., Ibaraki, T.: A 3-flip neighborhood local search for the set cocering problem. European Journal of Operational Research **172**, 472–499 (2006)
36. Yelbay, B.: Primal-dual heuristics for solving the set covering problem. Master's thesis, Sabancı University, Istanbul,Turkey (2010)