Covering Analysis of the Greedy Algorithm for Partial Cover

Tapio Elomaa and Jussi Kujala

Department of Software Systems
Tampere University of Technology
P.O. Box 553, FI-33101 Tampere, Finland
elomaa@cs.tut.fi, jussi.kujala@iki.fi

Abstract. The greedy algorithm is known to have a guaranteed approximation performance in many variations of the well-known minimum set cover problem. We analyze the number of elements covered by the greedy algorithm for the minimum set cover problem, when executed for k rounds. This analysis quite easily yields in the p-partial cover problem over a ground set of m elements the harmonic approximation guarantee $H(\lceil pm \rceil)$ for the number of required covering sets. Thus, we tie together the coverage analysis of the greedy algorithm for minimum set cover and its dual problem partial cover.

1 Introduction

Minimum Set Cover is a fundamental combinatorial optimization problem with many practical applications. It is one of the oldest problems known to be NP-complete [1,2]. The goal in Minimum Set Cover is to cover all elements of the ground set by using as few subsets as possible from a given collection.

What also makes this problem very interesting is the fact that it can be approximated efficiently within guaranteed performance by the straightforward greedy algorithm [3,4,5]. Greedy approximation of Minimum Set Cover underlies approximation algorithms in many application fields; e.g., in machine learning [6,7], combinatorial pattern matching [8,9], and bioinformatics [10,11,12]. Thus, Minimum Set Cover has received a lot of analytical attention over the years. The endmost approximation possibilities of the problem and the performance of the greedy algorithm are well understood topics today.

Partial Cover [6] is a generalization of Minimum Set Cover in which one asks how many subsets are required to cover at least a fraction p, $0 < p \leq 1$, of the elements of the ground set. The greedy algorithm can be used also to approximate this problem, but it has to be changed in order to cope with Partial Cover. The required modifications, though, are small.

In this paper we draw a connection that has not been explicit before. We show that directly by analyzing the element covering performance of the greedy algorithm for Minimum Set Cover during its execution, one can obtain reasonably tight performance bounds for the p-Partial Cover problem. The bound that we obtain for a ground set of m elements is the harmonic bound $H(\lceil pm \rceil)$, which

T. Elomaa et al. (Eds.): Ukkonen Festschrift, LNCS 6060, pp. 102–113, 2010.

is the best known performance guarantee for the weighted PARTIAL COVER problem [13]. A somewhat tighter bound is known to hold for the unweighted version of the problem [14].

Our analysis asks how large portion of the elements in the ground set can be covered by using at most k subsets. We analyze the relation between the number of covered elements when the subsets are selected greedily and that when the subsets are chosen optimally. We then apply this relationship to the PARTIAL COVER problem to obtain the harmonic bound.

The remainder of this paper is organized as follows. In Section 2 we briefly review work on MINIMUM SET COVER together with its variants and recapitulate the greedy algorithm for PARTIAL COVER. The element covering analysis for the greedy algorithm is presented in Section 3 and its application to partial covers is the topic of Section 4. We consider possibilities to extend this approach further in Section 5. A brief survey of related work is given in Section 6 before concluding this paper in Section 7.

2 Minimum Set Cover and the Greedy Algorithm

A collection $S = \{ S_1, \ldots, S_n \}$ of subsets of some finite set U is a *cover* of U if $\bigcup_{i=1}^{n} S_i = U$. Moreover, $S' \subseteq S$ is a *subcover* of U if S' itself is a cover of U. In the classical MINIMUM SET COVER problem one is given as an instance a finite set U and a cover $S = \{ S_1, \ldots, S_n \}$ of U and is requested to find a subcover $S' \subseteq S$ of U of minimum cardinality. To put this more exactly, in terms of the approximation setting MINIMUM SET COVER problem is as follows:

Instance: A cover $S = \{ S_1, \ldots, S_n \}$ of U.
Solution: A subcover $S' \subseteq S$ of U.
Measure: Cardinality of the subcover, $|S'|$.

In the decision version of this problem one asks whether there exists a subcover of cardinality at most K. This problem was shown to be NP-complete by Karp [1] through a polynomial-time reduction from the VERTEX COVER problem. Throughout this paper we denote the cardinality of the ground set U by m.

The greedy algorithm for the set cover problem is one of the best-known polynomial-time approximation algorithms. It chooses at each step the unused set which covers the largest number of remaining elements. This algorithm was shown by Johnson [3] and Lovász [4] to have approximation ratio no worse than $H(m)$, where $H(m) = 1 + 1/2 + \cdots + 1/m$ is the mth harmonic number. Recall that $\ln m < H(m) \leq \ln m + 1$. Chvátal [5] extended the harmonic performance ratio also to the weighted version of MINIMUM SET COVER. This time the greedy selection picks at each step the covering set with the minimum cost per remaining element.

Feige [15] proved — using interactive proof techniques — that no polynomial time algorithm can approximate MINIMUM SET COVER within $(1 - \epsilon) \ln m$ for any $\epsilon > 0$, unless NP $\subseteq$ DTIME$(n^{\log \log n})$. Hence, under this plausible structural complexity assumption, the performance ratio of any polynomial time algorithm can improve on the harmonic bound of the greedy algorithm by at most $o(\ln m)$. The analysis of the greedy algorithm for Minimum Set Cover was essentially completed by Slavík [14,16] who proved a performance ratio of exactly $\ln m - \ln\ln m + \Theta(1)$ for the algorithm. More precisely

$$\ln m - \ln\ln m - 0.31 < |G|/|0| < \ln m - \ln\ln m + 0.78,$$

where G is the cover selected by the greedy algorithm and O is the optimal cover.

The proof technique of Slavík was to recursively define the "greedy numbers" $N(k, l)$, which correspond to the size of the smallest ground set U for which it is possible to have a cover of U with the optimal cardinality l and greedy cover of size k. The same technique can be adapted to also apply for fractional [4,17] and partial covers [6].

Two natural variations of Minimum Set Cover are its weighted version and d-Set Cover, where all members of the cover S have cardinality of at most d. In the weighted Minimum Set Cover all elements of S have a positive cost associated with them and the goal is to find a subcover of minimum total cost. Both variations, naturally, are NP-complete, since they contain the original problem as a special case. The greedy algorithm also approximates these problems within the harmonic bound in polynomial time [3,4,5].

A somewhat more general NP-hard set covering problem is Partial Cover [6]. We say that $S' \subseteq S$ is a *p-partial cover* of U if

$$\left| \bigcup_{S_j \in S'} S_j \right| \geq pm.$$

An instance of the Partial Cover problem consists of a finite set U, a finite cover $S = \{ S_1, \ldots, S_n \}$ of U, and a real p, $0 < p \leq 1$. The goal is to find a p-partial cover $S' \subseteq S$ of U of minimum cardinality:

Instance: A cover $S = \{ S_1, \ldots, S_n \}$ of U and a number p, $0 < p \leq 1$.
Solution: A p-partial cover $S' \subseteq S$ of U.
Measure: Cardinality of the p-partial cover, $|S'|$.

Table 1 shows the greedy algorithm for the weighted Partial Cover problem. In it one searches at each step for the unused subset that covers as many elements as possible — though, not excessive elements — with as low average cost per element as possible. When the required fraction of elements covered has been reached, the algorithm halts. Observe that in Step 5 of the algorithm the elements of the newly chosen covering subset are removed from the remaining subsets. Here it has to be done to keep the average cost per element of remaining subsets an informative measure. However, the same cleaning of remaining subsets can be carried out in the unweighted case as well without any harm. In the following we assume that such a cleaning operation is part of the greedy algorithm. The algorithm of Table 1 is very similar to the greedy algorithm for the weighted Minimum Set Cover problem [5,6].

Table 1. The greedy algorithm for the weighted PARTIAL COVER

Input: A cover $S = \{ S_1, \ldots, S_n \}$ of a finite set U, positive costs $c = \{ c_1, \ldots, c_n \}$ of the covering sets, and a number p.
Output: A p-partial cover $S' \subset S$ of U.

1. $S' \leftarrow \emptyset$.
2. Find out the number r of elements of U that still need to be covered in order to obtain a p-partial cover:
$$r \leftarrow \lceil pm \rceil - \left| \bigcup_{S_j \in S'} S_j \right| .$$
3. If $r \leq 0$, then return S'.
4. Find $S_i \in S \setminus S'$, $S_i \neq \emptyset$, that minimizes the quotient
$$\frac{c_i}{\min(r, |S_i|)} .$$
5. $S' \leftarrow S' \cup S_i$.
For each $S_j \in S \setminus S'$: $S_j \leftarrow S_j \setminus S_i$.
Go to step 2.

The straightforward analysis of the greedy method for PARTIAL COVER becomes quite complicated because the optimal solution may cover a different set of elements than those chosen by the greedy algorithm. Thus, the methods used by Johnson [3], Lovász [4], and Chvátal [5] to establish the harmonic bound in case of complete covers do not generalize directly to this problem.

Nevertheless, Kearns [6] managed to prove the weak harmonic performance guarantee of $2H(m) + 3$ for the greedy algorithm by bounding separately the weights of those elements that are covered by the greedy algorithm but do not belong to the optimal cover, and those that are members of both solutions. Using a completely different approach Slavík [13] proved that for the weighted p-partial cover problem a bound similar to the classical one holds: The performance ratio of the greedy algorithm for this problem is no worse than $H(\lceil pm \rceil)$. One can construct an example to show that this bound is also tight [13]. The bound contains, as special cases, the classical harmonic bounds for MINIMUM SET COVER.

This time the proof technique of Slavík was to contrast directly the weights of the optimal and greedy cover from iteration to iteration in the execution of the greedy algorithm. Slavík's [14] exact analysis of the MINIMUM SET COVER problem also holds for PARTIAL COVER when the subsets are unweighted. Thus, unweighted PARTIAL COVER can be approximated using the greedy algorithm with ratio $\ln\lceil pm \rceil - \ln\ln\lceil pm \rceil + \Theta(1)$.

Subsequent MINIMUM SET COVER approximation approaches — aiming to improve additive constants, which is the most one can hope for after Feige's [15] proof — include Srinivasan's [17] application of the randomized rounding technique [18] to obtain improved performance ratio in special cases. Another line of research has been the work of Halldórsson [19,20] who applied a *local improvements* modification to the greedy algorithm to obtain an improved upper bound of $H(m) - 0.43$. In this approach one applies optimization techniques to the subsets that are small enough. This approach was taken further by Duh and Fürer [21] in their *semi-local optimization* approach. This leads to the polynomial-time approximation algorithm with the best worst-case performance guarantee of $H(m) - 1/2$.

Slavík's [13] proof of the harmonic bound is based on an analysis of the cost per remaining relevant element of a subset chosen to the greedy partial cover. Unfortunately, this does not lead to an intuitive proof. In the following we show that the harmonic performance guarantee can be obtained directly through an analysis of the greedy algorithm for Minimum Set Cover.

The greedy algorithm works in rounds choosing the subsets to the evolving cover one by one. Hence, it is natural to consider how many elements are covered by the greedy algorithm after r rounds and compare it to the optimal covering in k rounds. Viewing the greedy algorithm as gradually covering more and more elements gives a concrete connection between its performance in the two problems. The analysis easily yields the harmonic bound for Partial Cover. This simplified proof is of interest because of the importance and wide use of Minimum Set Cover and its generalization.

3 Covering Analysis of the Greedy Algorithm

Let us analyze the greedy set covering algorithm from the point of view of its covering performance. We show that the number of elements covered by r greedily chosen subsets is not much less than the total number of elements in k, $k \leq r$, optimally chosen sets of S. Here the r greedily or k optimally chosen subsets do not have to constitute a cover for the whole of U. However, setting r and k large enough, will eventually yield a full cover of U.

Let g_i denote the size of the subset chosen by the greedy algorithm on the ith round and let $G_r = \sum_{i=1}^{r} g_i$. The maximum number of elements covered by k optimally chosen subsets is denoted by O_k.

Lemma 1. *For $r \geq k \geq 1$ the following holds*

$$G_r \geq \left(1 - \left(1 - \frac{1}{k}\right)^r\right) O_k.$$

Proof. By the pigeonhole principle the largest subset g_1, which is chosen by the greedy algorithm on the first round, must contain at least as many elements as there on average are in the k maximally-covering sets, for any k; i.e., $g_1 \geq O_k/k$. The pigeonhole principle applies also on the second and subsequent rounds. However, g_2 can only be guaranteed to have size $(O_k - g_1)/k$. In general, on round $n+1$, $n \geq 1$, one must reduce the number of elements in the subsets already chosen by the greedy algorithm on previous rounds, G_n, from the maximum number of elements covered by k subsets and we have

$$g_{n+1} \geq \frac{O_k - G_n}{k}. \tag{1}$$

Let us, thus, consider the sequence

$$\begin{cases} x_1 = O_k/k; \\ x_{n+1} = x_n - (x_n/k) = (1 - 1/k)\, x_n, \end{cases}$$

for which it holds

$$x_n = \left(1 - \frac{1}{k}\right)^{n-1} \frac{O_k}{k}. \tag{2}$$

By induction we can show that $G_n \geq \sum_{i=1}^{n} x_i$. The base case was already stated above. Let us, then, assume that the claim holds for values less than n. Now, by inequality (1), the inductive hypothesis, and the definition of the sequence, we get

$$\begin{aligned} G_{n+1} &= G_n + g_{n+1} \\ &\geq G_n + \frac{O_k - G_n}{k} \\ &= \frac{O_k}{k} + \left(1 - \frac{1}{k}\right) G_n \\ &\geq \frac{O_k}{k} + \left(1 - \frac{1}{k}\right) \sum_{i=1}^{n} x_i \\ &= \frac{O_k}{k} + \sum_{i=1}^{n} \left(1 - \frac{1}{k}\right) x_i \\ &= \frac{O_k}{k} + \sum_{i=1}^{n} x_{i+1} \\ &= x_1 + \sum_{i=2}^{n+1} x_i \\ &= \sum_{i=1}^{n+1} x_i. \end{aligned}$$

Combining this with equality (2) gives the lower bound for G_r:

$$\begin{aligned} G_r &\geq \sum_{i=1}^{r} x_i \\ &= \sum_{i=1}^{r} \left(1 - \frac{1}{k}\right)^{i-1} \frac{O_k}{k} \\ &= \left(1 - \left(1 - \frac{1}{k}\right)^{r}\right) O_k, \end{aligned}$$

where the last equality is by the value of a geometric series.

Particularly interesting special case is $r = k$, which corresponds to asking what is the performance guarantee of the greedy algorithm with respect to the number of covered elements. This question has independently been studied by Hochbaum and Pahtria [22,23], who also obtained the following results.

By the above result, for example, two greedily chosen subsets are guaranteed to cover at least 3/4 of the largest number of elements that can, on the whole, be covered by two subsets. Asymptotically the lower bound behaves as follows when $r = k$. One can simplify the above result by recalling that for all x

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x$$

and observing that $(1-(1-1/k)^k)$ is decreasing, which gives the greedy algorithm the following approximation guarantee in the number of elements covered:

$$G_k \geq \left(1-\frac{1}{e}\right)O_k.$$

4 Application of the Analysis to Partial Cover

Let APP now denote the number of subsets chosen by the greedy algorithm in order to cover at least a fraction p of the m elements in the ground set. Respectively, OPT is the minimum number of subsets required to cover at least proportion $\lceil pm \rceil$ of the elements.

Theorem 1. *APP/OPT $\leq H(\lceil pm \rceil)$.*

Proof. Without loss of generality, we can assume that $O_{\text{OPT}} = \lceil pm \rceil$. This can be accomplished by removing some of the elements from the sets belonging to the optimal solution. The remaining elements still constitute a p-partial cover of U and, therefore, the value of the optimal solution does not change. On the other hand, this modification cannot improve the solution of the greedy algorithm.

Observe that the case OPT $= 1$ is not interesting, because the greedy algorithm will also output the one subset that covers a fraction p of U. Hence, in this case APP $=$ OPT. In the following we consider only partial covers for which OPT ≥ 2.

Let us consider the least r such that $G_r \geq O_{\text{OPT}} - c$ for some constant c. Thus, by Lemma 1, we want to solve

$$\left(1-\left(1-\frac{1}{\text{OPT}}\right)^r\right)O_{\text{OPT}} = O_{\text{OPT}} - c \Leftrightarrow$$
$$-\left(1-\frac{1}{\text{OPT}}\right)^r = -\frac{c}{O_{\text{OPT}}}.$$

Taking natural logarithms of both sides gives

$$-r\ln\left(1-\frac{1}{\text{OPT}}\right) = \ln O_{\text{OPT}} - \ln c. \tag{3}$$

Recalling that for $x > -1 : \ln(1+x) \leq x$, where equality holds only for $x = 0$, leads to

$$\frac{r}{\text{OPT}} < \ln O_{\text{OPT}} - \ln c. \tag{4}$$

The r greedily selected subsets now cover $O_{\text{OPT}} - c$ elements. Thus, at most c further subsets are needed to cover in total at least O_{OPT} elements. Hence, $\text{APP} \leq \lceil r \rceil + c \leq r + 1 + c$, and by inequality (4) we have

$$\begin{aligned}\frac{\text{APP}}{\text{OPT}} &\leq \frac{r+c+1}{\text{OPT}} \\ &< \ln O_{\text{OPT}} - \ln c + \frac{c+1}{\text{OPT}}.\end{aligned}$$

The right-hand side obtains its minimum value when $c = \text{OPT}$, which further yields

$$\frac{\text{APP}}{\text{OPT}} \leq \ln O_{\text{OPT}} - \ln \text{OPT} + 1 + \frac{1}{\text{OPT}}.$$

Finally, because $\ln n > 1 + 1/n$ for all integers $n \geq 4$, we only need to check separately the cases $\text{OPT} = 3$ and $\text{OPT} = 2$ by substituting them and $c = 1$ to equation (3), to obtain the desired result

$$\frac{\text{APP}}{\text{OPT}} \leq \ln O_{\text{OPT}} < H(O_{\text{OPT}}) = H(\lceil pm \rceil).$$

The above derived harmonic bound $H(\lceil pm \rceil)$ is the tight bound for *weighted* Partial Cover [13], but Slavík's [14] greedy numbers technique yields the tight bound $\ln m - \ln \ln m + \Theta(1)$ for the *unweighted* problem. He obtains the tighter bound through a detailed analysis of the involved functions.

5 Further Application of the Analysis

Above we were eager to approximate the additive terms away in order to reach the harmonic bound for Partial Cover. Let us briefly consider what happens if they are not abstracted away.

Let

$$d = \ln\left(1 - \frac{1}{\text{OPT}}\right)^{-\text{OPT}}$$

and substitute it to equation (3) to obtain

$$\frac{rd}{\text{OPT}} = \ln O_{\text{OPT}} - \ln c.$$

Recalling that $\text{APP} \leq r + c + 1$ yields

$$\begin{aligned}\frac{\text{APP}}{\text{OPT}} &\leq \frac{r+c+1}{\text{OPT}} \\ &= \frac{\ln O_{\text{OPT}} - \ln c}{d} + \frac{c+1}{\text{OPT}}.\end{aligned}$$

The right-hand side of this inequality will obtain its minimum value when $c = \textsc{opt}/d$. Thus,

$$\frac{\textsc{app}}{\textsc{opt}} \le \frac{\ln O_{\textsc{opt}}}{d} + \frac{1 + \ln d - \ln \textsc{opt}}{d}. \tag{5}$$

This bound has the unfortunate property of not being independent of the value of OPT. Let us, thus, denote by $d_{\textsc{opt}}$ the value of d depending on the value of OPT. For instance, $d_2 = \ln 4 \approx 1.386$ and asymptotically $d_\infty = 1$.

How does the bound (5) behave in comparison to the harmonic one derived above? Assuming that $\ln O_{\textsc{opt}}$ is large, for small values of OPT bound (5) will be tighter than the harmonic one. However, $d_{\textsc{opt}}$ approaches one as the value of OPT increases and, thus, this bound eventually loses its advantage over the harmonic bound. Moreover, as the value of OPT is unknown, this performance guarantee is not a very practical one.

6 Related Work

The covering analysis of the greedy set covering algorithm (Lemma 1) has been settled in case $r = k$ by Hochbaum [22,23,24]. The result is relatively well known as Max Cover. The derivation of this result does not differ significantly from that given in this paper.

Slavík was not the first author to show the tight harmonic approximation bound for partial cover. This result is already contained in the more general result of Wolsey from 1982 [25], although it is not a widespread fact.

The direct link between the covering analysis of the greedy set covering algorithm and its performance for partial cover (Theorem 1) is, to the best of our knowledge, an original contribution.

The cover of Minimum Set Cover $S = \{S_1, \ldots, S_n\}$ can be seen as a hypergraph over the vertices from the ground set U. Its dual problem is Minimum Vertex Cover (VC) over the dual of this hypergraph, which inverts the roles of hyperedges and vertices. The greedy approach and other algorithms for different variations of the partial VC problem have been studied extensively in recent years [26,27,28,29].

Several new variants of the Minimum Set Cover problem have been proposed and analyzed lately. Let us just mention a few of them. In the *red-blue* set cover [30] the ground set U contains red and blue elements and the aim is to cover all of the blue elements and as few as possible of the red elements. This is a strongly inapproximable problem [31]. A generalization of the red-blue set cover, *positive-negative* partial set cover was introduced by Miettinen [32]. In *multicover* problems the requirement is to cover each element a prescribed number of times. Also in this extension of Minimum Set Cover the greedy algorithm and its variants yield good approximation results [33,34,35].

7 Conclusion

We have shown that covering analysis of the greedy algorithm for the Minimum Set Cover problem quite easily yields the harmonic bound $H(\lceil pm \rceil)$ for the

p-Partial Cover. This makes the connection between the classical problem and its generalization explicit. The obtained bound is not the tightest one known to hold for the unweighted problem. Nevertheless, it is clearly better than the one that comes out of the analysis that bounds the sizes of the resulting sets.

As future work we leave studying whether the tighter performance guarantee for the unweighted Partial Cover could be reached by means of covering analysis. Also, the potential of this line of analysis for the weighted Partial Cover was not explored in this work.

Acknowledgements

This work has been supported by the Academy of Finland.

References

1. Karp, R.M.: Reducibility among combinatorial problems. In: Miller, R., Thatcher, J. (eds.) Complexity of Computer Computations, pp. 85–103. Plenum Press, New York (1972)
2. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman & Co, New York (1979)
3. Johnson, D.S.: Approximation algorithms for combinatorial problems. Journal of Computer and System Sciences 9(3), 256–278 (1974)
4. Lovász, L.: On the ratio of optimal integral and fractional covers. Discrete Mathematics 13, 383–390 (1975)
5. Chvátal, V.: A greedy heuristic for the set-covering problem. Mathematics of Operations Research 4(3), 233–235 (1979)
6. Kearns, M.J.: The Computational Complexity of Machine Learning. MIT Press, Cambridge (1990)
7. Kivinen, J., Mannila, H., Ukkonen, E.: Learning hierarchical rule sets. In: Proceedings of the Fifth Annual ACM Conference on Computational Learning Theory, pp. 37–44. ACM Press, New York (1992)
8. Brazma, A., Jonassen, I., Ukkonen, E., Vilo, J.: Discovering patterns and subfamilies in biosequences. In: States, D.J., Agarwal, P., Gaasterland, T., Hunter, L., Smith, R. (eds.) Proceedings of the Fourth International Conference on Intelligent Systems for Molecular Biology, pp. 34–43. AAAI Press, Menlo Park (1996)
9. Brazma, A., Ukkonen, E., Vilo, J.: Discovering unbounded unions of regular pattern languages from positive examples. In: Nagamochi, H., Suri, S., Igarashi, Y., Miyano, S., Asano, T. (eds.) ISAAC 1996. LNCS, vol. 1178, pp. 95–104. Springer, Heidelberg (1996)
10. Li, M.: Towards a DNA sequencing theory. In: Proceedings of the Thirty-First Annual Symposium on Foundations of Computer Science, pp. 125–134. IEEE Computer Society, Los Alamitos (1990)
11. Kivioja, T., Arvas, M., Saloheimo, M., Penttilä, M., Ukkonen, E.: Optimization of cDNA-AFLP experiments using genomic sequence data. Bioinformatics 21(11), 2573–2579 (2005)
12. Rantanen, A., Mielikäinen, T., Rousu, J., Maaheimo, H., Ukkonen, E.: Planning optimal measurements of isotopomer distributions for estimation of metabolic fluxes. Bioinformatics 22(10), 1198–1206 (2006)

13. Slavík, P.: Improved performance of the greedy algorithm for partial cover. Information Processing Letters 64(5), 251–254 (1997)
14. Slavík, P.: A tight analysis of the greedy algorithm for set cover. Journal of Algorithms 25(2), 237–254 (1997)
15. Feige, U.: A threshold of ln n for approximating set cover. Journal of the ACM 45(4), 634–652 (1998)
16. Slavík, P.: Approximation algorithms for set cover and related problems. PhD thesis, State University of New York at Buffalo, Department of Computer Science, Buffalo, NY (April 1998)
17. Srinivasan, A.: Improved approximation of packing and covering problems. In: Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing, pp. 268–276. ACM Press, New York (1995)
18. Motwani, R., Raghavan, P.: Randomized Algorithms. Cambridge University Press, Cambridge (1995)
19. Halldórsson, M.M.: Approximating set cover via local search. Technical Report IS-RR-95-0002F, Japan Advanced Institute of Science and Technology (1995)
20. Halldórsson, M.M.: Approximating discrete collections via local improvements. In: Proceedings of the Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 160–169. ACM Press, New York (1995)
21. Duh, R., Fürer, M.: Approximation of k-set cover by semi-local optimization. In: Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing, pp. 256–264. ACM Press, New York (1997)
22. Hochbaum, D.S.: Approximating covering and packing problems: set cover, vertex cover, independent set, and related problems. In: Hochbaum, D.S. (ed.) Approximation Algorithms for NP-hard Problems, pp. 94–143. PSW Publishing, Boston (1997)
23. Hochbaum, D.S., Pathria, A.: Analysis of the greedy approach of maximum k-coverage. Naval Research Quarterly 45, 615–627 (1998)
24. Hochbaum, D.S.: Approximation algorithms for the set covering and vertex cover problems. SIAM Journal on Computing 11(3), 555–556 (1982)
25. Wolsey, L.A.: An analysis of the greedy algorithm for the submodular set covering problem. Combinatorica 2(4), 385–393 (1982)
26. Hochbaum, D.S.: The t-vertex cover problem: Extending the half integrality framework with budget constraints. In: Jansen, K., Hochbaum, D.S. (eds.) APPROX 1998. LNCS, vol. 1444, pp. 111–122. Springer, Heidelberg (1998)
27. Bar-Yehuda, R.: Using homogeneous weights for approximating the partial cover problem. Journal of Algorithms 39(2), 137–144 (2001)
28. Gandhi, R., Khuller, S., Srinivasan, A.: Approximation algorithms for partial covering problems. Journal of Algorithms 53(1), 55–84 (2004)
29. Sümer, Ö.: Partial covering of hypergraphs. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 572–585. SIAM, Philadelphia (2005)
30. Carr, R.D., Doddi, S., Konjevod, G., Marathe, M.V.: On the red-blue set cover problem. In: Proceedings of the Eleventh ACM-SIAM Symposium on Discrete Algorithms, pp. 345–353. SIAM, Philadelphia (2000)
31. Peleg, D.: Approximation algorithms for the label-cover$_{\max}$ and red-blue set cover problems. J. Discrete Algorithms 5(1), 55–64 (2007)

32. Miettinen, P.: On the positive-negative partial set cover problem. Information Processing Letters 108(4), 219–221 (2008)
33. Fujito, T., Kurahashi, H.: A better-than-greedy algorithm for k-set multicover. In: Erlebach, T., Persinao, G. (eds.) WAOA 2005. LNCS, vol. 3879, pp. 176–189. Springer, Heidelberg (2005)
34. Berman, P., DasGupta, B., Sontag, E.D.: Randomized approximation algorithms for set multicover problems with applications to reverse engineering of protein and gene networks. Discrete Applied Mathematics 155(6-7), 733–749 (2007)
35. Chekuri, C., Clarkson, K.L., Har-Peled, S.: On the set multi-cover problem in geometric settings. In: Hershberger, J., Fogel, E. (eds.) Proceedings of the 25th ACM Symposium on Computational Geometry, pp. 341–350. ACM, New York (2009)