Graph Coloring (1994, 1998; Karger, Motwani, Sudan)

Michael Langberg, The Open University of Israel, www.openu.ac.il
entry editor: Chandra Chekuri

INDEX TERMS: Graph coloring, approximation algorithms, semidefinite programming.

SYNONYMS: Clique cover.

1 PROBLEM DEFINITION

An independent set in an undirected graph $G = (V, E)$ is a set of vertices that induce a subgraph which does not contain any edges. The size of the maximum independent set in G is denoted by $\alpha(G)$. For an integer k, a k-coloring of G is a function $\sigma : V \to [1 \dots k]$ which assigns colors to the vertices of G. A valid k-coloring of G is a coloring in which each color class is an independent set. The chromatic number $\chi(G)$ of G is the smallest k for which there exists a valid k-coloring of G. Finding $\chi(G)$ is a fundamental NP-hard problem. Hence, when limited to polynomial time algorithms, one turns to the question of estimating the value of $\chi(G)$ or to the closely related problem of *approximate coloring*.

Problem 1 (Approximate coloring)**.**
INPUT*: Undirected graph $G = (V, E)$.*
OUTPUT*: A valid coloring of G with $r \cdot \chi(G)$ colors, for some approximation ratio $r \geq 1$.*
OBJECTIVE*: Minimize r.*

Let G be a graph of size n. The approximate coloring of G can be solved efficiently within an approximation ratio of $r = O\left(\frac{n(\log\log n)^2}{\log^3 n}\right)$ [12]. This holds also for the approximation of $\alpha(G)$ [8]. These results may seem rather weak, however it is NP-hard to approximate $\alpha(G)$ and $\chi(G)$ within a ratio of $n^{1-\varepsilon}$ for any constant $\varepsilon > 0$ [14, 9, 21]. Under stronger complexity assumptions, there is some constant $0 < \delta < 1$ such that neither problem can be approximated within a ratio of $n/2^{\log^\delta n}$ [17, 21]. This chapter will concentrate on the problem of coloring graphs G for which $\chi(G)$ is *small*. As will be seen, in this case the approximation ratio achievable significantly improves.

1.1 Vector coloring of graphs

The algorithms achieving the best ratios for approximate coloring when $\chi(G)$ is small [15, 3, 13, 1] are all based on the idea of *vector coloring*, introduced by Karger, Motwani and Sudan [15] [1].

Definition 1. *A vector k-coloring of a graph is an assignment of unit vectors to its vertices, such that for every edge, the inner product of the vectors assigned to its endpoints is at most (in the sense that it can only be more negative) $-1/(k-1)$.*

The *vector chromatic number* $\overrightarrow{\chi}(G)$ of G is the smallest k for which there exists a vector k-coloring of G. The vector chromatic number can be formulated as follows:

[1] Vector coloring as presented in [15] is closely related to the Lovász θ function [19]. This connection will be discussed shortly.

$$\vec{\chi}(G) \qquad \begin{array}{lll} \text{Minimize} & k & \\ \text{subject to:} & \langle v_i, v_j\rangle \le -\frac{1}{k-1} & \forall (i,j)\in E \\ & \langle v_i, v_i\rangle = 1 & \forall i \in V \end{array}$$

Here, assume that $V = [1, \ldots, n]$ and that the vectors $\{v_i\}_{i=1}^n$ are in R^n. Every k-colorable graph is also vector k-colorable. This can be seen by identifying each color class with one vertex of a perfect $(k-1)$-dimensional simplex centered at the origin. Moreover, unlike the chromatic number, a vector k-coloring (when it exists) can be found in polynomial time using semidefinite programming (up to an arbitrarily small error in the inner products).

Claim 1 (Complexity of vector coloring [15])**.** *Let $\varepsilon > 0$. If a graph G has a vector k-coloring then a vector $(k+\varepsilon)$-coloring of the graph can be constructed in time polynomial in n and $\log(1/\varepsilon)$.*

One can strengthen Definition 1 to obtain a different notion of vector coloring and the vector chromatic number.

$$\vec{\chi}_2(G) \qquad \begin{array}{lll} \text{Minimize} & k & \\ \text{subject to:} & \langle v_i, v_j\rangle = -\frac{1}{k-1} & \forall (i,j)\in E \\ & \langle v_i, v_i\rangle = 1 & \forall i \in V \end{array}$$

$$\vec{\chi}_3(G) \qquad \begin{array}{lll} \text{Minimize} & k & \\ \text{subject to:} & \langle v_i, v_j\rangle = -\frac{1}{k-1} & \forall (i,j)\in E \\ & \langle v_i, v_j\rangle \ge -\frac{1}{k-1} & \forall i,j \in V \\ & \langle v_i, v_i\rangle = 1 & \forall i \in V \end{array}$$

The function $\vec{\chi}_2(G)$ is referred to as the *strict* vector chromatic number of G and is equal to the Lovász θ function on $\bar{G}$ [19, 15], where $\bar{G}$ is the *complement* graph of G. The function $\vec{\chi}_3(G)$ is referred to as the *strong* vector chromatic number. An analog to Claim 1 holds for both $\vec{\chi}_2(G)$ and $\vec{\chi}_3(G)$. Let $\omega(G)$ denote the size of the maximum clique in G, it holds that: $\omega(G) \le \vec{\chi}(G) \le \vec{\chi}_2(G) \le \vec{\chi}_3(G) \le \chi(G)$.

2 KEY RESULTS

In what follows, assume that G has n vertices and maximal degree Δ. The $\tilde{O}(\cdot)$ and $\tilde{\Omega}(\cdot)$ notation are used to suppress polylogarithmic factors. We now state the key result of Karger, Motwani and Sudan [15]:

Theorem 1 ([15])**.** *If $\vec{\chi}(G) = k$ then G can be colored in polynomial time using* $\min\left\{\tilde{O}(\Delta^{1-2/k}), \tilde{O}(n^{1-3/(k+1)})\right\}$ *colors.*

As mentioned above, the use of vector coloring in the context of approximate coloring was initiated in [15]. Roughly speaking, once given a vector coloring of G, the heart of the algorithm in [15] finds a large independent set in G. In a nutshell, this independent set corresponds to a set of vectors in the vector coloring which are *close* to one another (and thus by definition cannot share an edge). Combining this with the ideas of Wigderson [20] mentioned below yields Theorem 1.

We proceed to describe related work. The first two theorems below appeared prior to the work of Karger, Motwani and Sudan [15].

Theorem 2 ([20])**.** *If $\chi(G) = k$ then G can be colored in polynomial time using $O(kn^{1-1/(k-1)})$ colors.*

Theorem 3 ([2])**.** *If $\chi(G) = 3$ then G can be colored in polynomial time using $\tilde{O}(n^{3/8})$ colors. If $\chi(G) = k \ge 4$ then G can be colored in polynomial time using at most $\tilde{O}(n^{1-1/(k-3/2)})$ colors.*

Combining the techniques of [15] and [2] the following results were obtained for graphs G with $\chi(G) = 3, 4$ (these results were also extended for higher values of $\chi(G)$).

Theorem 4 ([3])**.** *If $\chi(G) = 3$ then G can be colored in polynomial time using $\tilde{O}(n^{3/14})$ colors.*

Theorem 5 ([13])**.** *If $\chi(G) = 4$ then G can be colored in polynomial time using $\tilde{O}(n^{7/19})$ colors.*

The currently best known result for coloring a 3-colorable graph is presented in [1]. In their algorithm, [1] use the strict vector coloring relaxation (*i.e.* $\vec{\chi}_2$) enhanced with certain *odd cycle* constraints.

Theorem 6 ([1])**.** *If $\chi(G) = 3$ then G can be colored in polynomial time using $O(n^{0.2111})$ colors.*

To put the above theorems in perspective, it is NP-hard to color a 3-colorable graph G with 4 colors [16, 11] and a k-colorable graph (for sufficiently large k) with $k^{\frac{\log k}{25}}$ colors [17]. Under stronger complexity assumptions (related to the Unique Games Conjecture [18]) for any constant k it is hard to color a k-colorable graph with any constant number of colors [6]. The wide gap between these hardness results and the approximation ratios presented in this section has been a major initiative in the study of approximate coloring.

Finally, the limitations of vector coloring are addressed. Namely, are there graphs for which $\vec{\chi}(G)$ is a poor estimate of $\chi(G)$? One would expect the answer to be "yes" as estimating $\chi(G)$ beyond a factor of $n^{1-\varepsilon}$ is a hard problem. As will be stated below, this is indeed the case (even when $\vec{\chi}(G)$ is small). Some of the results that follow are stated in terms of the maximum independent set $\alpha(G)$ in G. As $\chi(G) \geq n/\alpha(G)$, these results imply a lower bound on $\chi(G)$. Theorem 7 (i) states that the original analysis of [15] is essentially tight. Theorem 7 (ii) presents bounds for the case of $\vec{\chi}(G) = 3$. Theorem 7 (iii) and Theorem 8 present graphs G in which there is an extremely large gap between $\chi(G)$ and the relaxations $\vec{\chi}(G)$ and $\vec{\chi}_2(G)$.

Theorem 7 ([10])**.** *(i) For every constant $\varepsilon > 0$ and constant $k > 2$, there are infinitely many graphs G with $\vec{\chi}(G) = k$ and $\alpha(G) \leq n/\Delta^{1-\frac{2}{k}-\varepsilon}$ (here $\Delta > n^\delta$ for some constant $\delta > 0$). (ii) There are infinitely many graphs G with $\vec{\chi}(G) = 3$ and $\alpha(G) \leq n^{0.843}$. (iii) For some constant c, there are infinitely many graphs G with $\vec{\chi}(G) = O(\frac{\log n}{\log\log n})$ and $\alpha(G) \leq \log^c n$.*

Theorem 8 ([7])**.** *For some constant c, there are infinitely many graphs G with $\vec{\chi}_2(G) \leq 2^{\sqrt{\log n}}$ and $\chi(G) \geq n/2^{c\sqrt{\log n}}$.*

Vector colorings, including the Lovász θ function and its variants, have been extensively studied in the context of approximation algorithms for problems other than Problem 1. These include approximating $\alpha(G)$, approximating the Minimum Vertex Cover problem, and combinatorial optimization in the context of random graphs.

3 APPLICATIONS

Besides its theoretical significance, graph coloring has several concrete applications that fall under the model of *conflict free* allocation of resources (see for example [5, 4]).

4 OPEN PROBLEMS

By far the major open problem in the context of approximate coloring addresses the wide gap between what is known to be hard and what can be obtained in polynomial time. The case of constant $\chi(G)$ is especially intriguing, as the best known upper bounds (on the approximation ratio) are polynomial while the lower bounds are of constant nature. Regarding the vector coloring paradigm, a majority of the results stated in Section 2 use the weakest from of vector coloring $\overrightarrow{\chi}(G)$ in their proof, while stronger relaxations such as $\overrightarrow{\chi}_2(G)$ and $\overrightarrow{\chi}_3(G)$ may also be considered. It would be very interesting to improve upon the algorithmic results stated above using stronger relaxations, as would a matching analysis of the limitations of these relaxations.

5 EXPERIMENTAL RESULTS

None is reported.

6 DATA SETS

None is reported.

7 URL to CODE

None is reported.

8 CROSS REFERENCES

Exact Algorithms for Coloring, Max Cut, Randomized Rounding, and Sparsest Cut.

9 RECOMMENDED READING

References

[1] S. Arora, E. Chlamtac, and M. Charikar. New approximation guarentee for chromatic number. In *Proceedings of the 38th annual ACM Symposium on Theory of Computing*, pages 215–224, 2006.

[2] A. Blum. New approximations for graph coloring. *Journal of the ACM*, 41(3):470–516, 1994.

[3] A. Blum and D. Karger. An $\tilde{O}(n^{3/14})$-coloring for 3-colorable graphs. *Information Processing Letters*, 61(6):49–53, 1997.

[4] G. J. Chaitin. Register allocation & spilling via graph coloring. In *Proceedings of the 1982 SIGPLAN Symposium on Compiler Construction*, pages 98–105, 1982.

[5] G.J. Chaitin, M.A. Auslander, A.K. Chandra, J. Cocke, M.E. Hopkins, and P.W. Markstein. Register allocation via coloring. *Computer Languages*, 6:47–57, 1981.

[6] I. Dinur, E. Mossel, and O. Regev. Conditional hardness for approximate coloring. In *Proceedings of the 38th annual ACM Symposium on Theory of Computing*, pages 344–353, 2006.

[7] U. Feige. Randomized graph products, chromatic numbers, and the Lovász theta function. *Combinatorica*, 17(1):79–90, 1997.

[8] U. Feige. Approximating maximum clique by removing subgraphs. *SIAM J. Discrete Math*, 18(2):219–225, 2004.

[9] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57:187–199, 1998.

[10] U. Feige, M. Langberg, and G. Schechtman. Graphs with tiny vector chromatic numbers and huge chromatic numbers. *SIAM Journal on Computing*, 33(6):1338–1368, 2004.

[11] V. Guruswami and S. Khanna. On the hardness of 4-coloring a 3-colorable graph. In *Proceedings of the 15th annual IEEE Conference on Computational Complexity*, pages 188–197, 2000.

[12] M. Halldorsson. A still better performance guarantee for approximate graph coloring. *Information Processing Letters*, 45:19–23, 1993.

[13] E. Halperin, R. Nathaniel, and U. Zwick. Coloring k-colorable graphs using smaller palettes. *Journal of Algorithms*, 45:72–90, 2002.

[14] J. Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Math*, 182(1):105–142, 1999.

[15] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of ACM*, 45(2):246–265, 1998.

[16] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20:393–415, 2000.

[17] S. Khot. Improved inapproximability results for max clique, chromatic number and approximate graph coloring. In *Proceedings of the 42nd annual IEEE Symposium on Foundations of Computer Science*, pages 600–609, 2001.

[18] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th annual ACM symposium on Theory of Computing*, pages 767–775, 2002.

[19] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:2–13, 1979.

[20] A. Wigderson. Improving the performance guarantee for approximate graph coloring. *Journal of the ACM*, 30(4):729 – 735, 1983.

[21] D. Zuckerman. Linear degree extractors and the inapproximability of max clique and chromatic number. In *Proceedings of the 38th annual ACM symposium on Theory of Computing*, pages 681–690, 2006.