On Minimum Entropy Graph Colorings

Jean Cardinal, Samuel Fiorini and Gilles Van Assche
Université Libre de Bruxelles, B-1050 Brussels, Belgium
{jcardin, sfiorini, gvanassc}@ulb.ac.be

Abstract — **We study properties of graph colorings that minimize the quantity of color information with respect to a given probability distribution on the vertices. The minimum entropy of any coloring is the *chromatic entropy*. Applications of the chromatic entropy are found in coding with side information and digital image partition coding. We show that minimum entropy colorings are hard to compute even if a minimum cardinality coloring is given, the distribution is uniform, and the graph is planar. We also consider the minimum number of colors in a minimum entropy coloring, and show that this number can be arbitrarily larger than the chromatic number, even for restricted families of uniformly weighted graphs.**

I. Definitions

A *coloring* of a graph G is an assignment of colors to the vertices of G such that any two adjacent vertices have different colors. The well-known graph coloring problem is to find a coloring with as few colors as possible. The minimum number of colors in a coloring of G is the *chromatic number* of G, denoted by $\chi(G)$.

A *probabilistic graph* is a graph equipped with a probability distribution on its vertices. Let (G, P) be a probabilistic graph, and let X be any random variable over the vertex set $V(G)$ of G with distribution P. We define the *entropy* of a coloring ϕ as the entropy of the random variable $\phi(X)$. In other words, the entropy of ϕ is the sum over all colors i of $-p_i \log p_i$, where $p_i = \sum_{x:\phi(x)=i} P(x)$ is the probability that X has color i. The *chromatic entropy* $H_\chi(G, P)$ of the probabilistic graph (G, P) is the minimum entropy of any of its colorings. It was first defined by Alon and Orlistky [3] and gives the minimum quantity of color information contained in any coloring of a probabilistic graph.

Applications of this definition can be found in zero-error source coding with side information at the receiver [3, 5, 6], and compression of digital image partitions created by segmentation algorithms [1, 2]. In fact, minimum entropy colorings of a characteristic graph directly induce good codes for the two problems.

II. Hardness

This result improves on a recent hardness result from Zhao and Effros [6].

Theorem 1. *Computing a minimum entropy coloring of (G, P) is NP-hard, even if P is the uniform distribution, G is planar, and a minimum cardinality coloring of G is given.*

Proof. (sketch) Let F be a planar graph with $n \geq 1$ vertices, and let G be the graph obtained from F by attaching a triangle to each vertex and adding a vertex-disjoint 4-clique to the resulting graph. In other words, let G be the graph defined as $G = F' \cup K_4$, where $V(F') = \{v, v', v'' : v \in V(F)\}$, $E(F') = \{\{v, v'\}, \{v', v''\}, \{v'', v\} : v \in V(F)\}$, K_4 is the complete graph with four vertices and $V(F') \cap V(K_4) = \emptyset$. Since G is planar and has a clique of size 4, we have $\chi(G) = 4$. Moreover, a 4-coloring of G can be found in polynomial time.

Now consider the problem of finding a minimum entropy coloring of (G, U), where U is the uniform distribution. It can be shown that the chromatic entropy of (G, U) is at least the entropy, say H_0, of a coloring with three color classes of cardinality $n + 1$ and one class of cardinality 1, and that we have $H_\chi(G, U) = H_0$ if and only if F is 3-colorable.

Hence a polynomial algorithm for computing a minimum entropy coloring of (G, U) could be used to check the 3-colorability of any planar graph F, a problem known to be NP-complete. □

III. Number of colors

We define $\chi_H(G, P)$ as the minimum number of colors needed in any minimum entropy coloring of (G, P), and $\Delta(G)$ as the maximum degree of any vertex in G.

Theorem 2. $\chi_H(G, P) \leq \Delta(G) + 1$.

Theorem 3. *$\chi_H(G, P)$ is not bounded by any function of $\chi(G)$, even if P is uniform and G is a bipartite graph, a tree, or a planar graph.*

Proof. (sketch) The construction is as follows, starting with a graph with chromatic number at least 2. We attach a large set of independent vertices to each of its vertex. If these independent sets are sufficiently large, it can be shown that minimizing the entropy requires that they are colored with the same color. In this way, we increase χ_H by one unit. The chromatic number is not changed by this operation. The construction can be iterated to increase χ_H at will while keeping χ constant. □

Theorem 2 can be proved by contradiction.

A long version of this paper together with updated results is available [4]. The authors wish to thank Noga Alon, Ertem Tuncel and Nicolas Cerf for useful comments.

References

[1] M. Accame, F.G.B. De Natale and F. Granelli, Efficient Labeling Procedures for Image Partition Encoding, *Signal Processing*, 80(6), pp. 1127–1131, june 2000.

[2] S. Agarwal and S. Belongie, On the Non-Optimality of Four Color Coding of Image Partitions, *Proc. IEEE International Conference on Image Processing*, 2002.

[3] N. Alon and A. Orlitsky, Source Coding and Graph Entropies, *IEEE Transactions on Information Theory*, 42(5), pp. 1329–1339, sept. 1996.

[4] J. Cardinal, S. Fiorini and G. Van Assche, Technical Report 517, http://www.ulb.ac.be/di/publications/.

[5] P. Koulgi, E. Tuncel, S. Regunathan, and K. Rose, On Zero-error Source Coding with Decoder Side Information, *IEEE Transactions on Information Theory*, 49(1), pp. 99–111, jan. 2003.

[6] Q. Zhao and M. Effros, Low Complexity Code Design for Lossless and Near-Lossless Side Information Source Codes, *Proc. IEEE Data Compression Conference*, 2003.