A Note on Coloring Sparse Random Graphs

Christian Sommer
sommer@nii.ac.jp

October 7, 2008

Abstract

Coja-Oghlan and Taraz (2004) presented a graph coloring algorithm that has expected linear running time for random graphs with edge probability p satisfying $np \leq 1.01$. In this work, we develop their analysis by exploiting generating function techniques. We show that, in fact, their algorithm colors $G_{n,p}$ with the minimal number of colors and has expected linear running time, provided that $np \leq 1.33$.

Keywords: graph algorithms, analysis of algorithms, combinatorial problems

1 Introduction

Deciding whether the chromatic number of a graph G is smaller than a given value ℓ is an **NP**-complete problem. Furthermore, Feige and Kilian [FK98] showed that, unless **ZPP** = **NP**, there is no polynomial-time algorithm that colors an input graph on n vertices, and has approximation ratio less than $n^{1-\epsilon}$, for all $\epsilon > 0$. Considering these worst-case results, a natural question arises: is there an algorithm that performs well on *random* instances?

In this work our focus is on the well-known and well-studied *binomial* model of random graphs [Gil59, ER60]. In this model, an edge is included in the resulting graph with probability p, independently of the presence or absence of other edges. We shall denote by $G_{n,p}$ a random graph drawn according to this distribution. Karp [Kar76] asked the following question: *Is there an algorithm that has expected polynomial running time for $G_{n,p}$ and always finds an optimal coloring?*

Coja-Oghlan and Taraz [COT04] affirmatively answered this question, provided that p is not too large, namely $np \leq 1.01$. However, they remarked that their analysis was not tight, and asked whether their result could hold for larger values of p. For $np \geq 3.35$, a random graph $G_{n,p}$ has a 3-core of linear size almost surely [PSW96], causing the algorithm of Coja-Oghlan and Taraz to run in exponential time only. In this work, we develop their analysis by exploiting generating function techniques, and show the following result.

Theorem 1. *There exists an algorithm that colors $G_{n,p}$ with the minimal number of colors and has expected* linear *running time, provided that $np \leq 1.33$.*

Lemma 1, which bounds the number of graphs with minimum degree k on ν nodes and μ edges, contributes the main improvement of [COT04] and is also of independent interest.

1.1 Algorithm

The algorithm of Beigel and Eppstein [BE05] decides whether a graph is 3-colorable in time $O(1.3289^n)$. Define $\tau_3 := 1.329$. The algorithm of Lawler [Law76] determines the chromatic number of a graph in time $O((1+\sqrt[3]{3})^n)$. Define $\tau_4 := 2.443$.

The k-core of a graph $G = (V, E)$ is the maximum subgraph with minimum degree k.

The algorithm of Coja-Oghlan and Taraz [COT04] optimally colors a graph as follows: the 3-core is built by repeatedly removing vertices with $\deg(v) < 3$ and putting them on a stack. Next, the algorithm of Beigel and Eppstein [BE05] is applied on the 3-core and if it is 3-colorable, the core is colored. Otherwise, it builds the 4-core by further removing vertices with $\deg(v) < 4$ and putting them on the stack. Then, the algorithm optimally colors the core using Lawler's algorithm [Law76]. Finally, the vertices from the stack are added one by one to the graph again and they are assigned the smallest color available. For an exposition of the algorithm we refer to [COT04].

2 Proof

Define $\mathcal{C}(k;\nu,\mu)$ to be the set of graphs having a k-core of size ν with μ edges and $\mathcal{G}(k;\nu,\mu)$ to be the set of graphs with ν nodes, μ edges and minimal degree k. Every $G \in \mathcal{C}(k;\nu,\mu)$ has a subgraph from $\mathcal{G}(k;\nu,\mu)$.

The expected running time (up to a constant factor and linear processing time to build the core) is, for all ν, μ, at most the probability that the graph has a k-core with ν vertices and μ edges multiplied with the exponential running time needed for the core, i.e.,

$$\sum_{k\in\{3,4\}} \sum_{\nu=k}^{n} \sum_{\mu=\frac{k}{2}\nu}^{\binom{\nu}{2}} \mathbf{Pr}[G_{n,p} \in \mathcal{C}(k;\nu,\mu)] \cdot (\tau_k)^{\nu}.$$

The probability that $G_{n,p}$, where $p = \frac{c}{n}$, has a k-core of size ν with μ edges is bounded as follows:

$$\begin{aligned}\mathbf{Pr}[G_{n,p} \in \mathcal{C}(k;\nu,\mu)] &\leq \binom{n}{\nu} \cdot \left(\frac{c}{n}\right)^{\mu} \cdot |\mathcal{G}(k;\nu,\mu)| \\ &\leq 2^{h(\nu/n)\cdot n} \cdot \left(\frac{c}{n}\right)^{\mu} \cdot |\mathcal{G}(k;\nu,\mu)|,\end{aligned}$$

where $h(p) := -p\log_2 p - (1-p)\log_2(1-p)$ denotes the binary entropy.

We split the proof into two cases for $\mu = \frac{k}{2}\nu + \lambda$. Set $\eta = 0.7$.

Case I: (Many Edges, $\mu \geq \frac{k}{2}\nu + \eta\nu$): The number of graphs with a core of size ν and at least μ edges $|\mathcal{G}(k;\nu,\mu)|$ is bounded by the number of ways to choose μ edges among the $\binom{\nu}{2}$ pairs of nodes in the core.

$$|\mathcal{G}(k;\nu,\mu)| \leq \binom{\binom{\nu}{2}}{\mu} \leq \left(\frac{e\nu^2}{2\mu}\right)^{\mu}$$

It remains to show that

$$2^{h(\nu/n)\cdot n} \cdot \left(\frac{c}{n}\right)^{\mu} \cdot |\mathcal{G}(k;\nu,\mu)| \leq (\tau_k)^{-\nu}.$$

We write $\nu = x_\nu n$ and $\lambda = x_\lambda x_\nu n$. ($x_\nu \in (0,1]$, $x_\lambda \geq \eta$)

$$\begin{aligned}2^{h(\nu/n)n} \cdot (\tau_k)^{\nu} \cdot \left(\frac{e\nu^2 c}{2\mu n}\right)^{\mu} &< 1 \\ \Leftrightarrow 2^{h(x_\nu)n} \cdot (\tau_k)^{x_\nu n} \cdot \left(\frac{e(x_\nu n)^2 c}{2(\frac{k}{2}x_\nu n + x_\lambda x_\nu n)n}\right)^{\frac{k}{2}x_\nu n + x_\lambda x_\nu n} &< 1 \\ \Leftarrow 2^{h(x_\nu)} \cdot (\tau_k)^{x_\nu} \cdot \left(\frac{e x_\nu c}{k+2\eta}\right)^{x_\nu(\frac{k}{2}+\eta)} &< 1\end{aligned}$$

The condition above holds for all $x_\nu \in (0,1]$.

Case II: (Few Edges, $\mu < \frac{k}{2}\nu + \eta\nu$): Let $e(\xi,k)$ denote the Euler series starting at k, i.e., $e(\xi,k) := e^{\xi} - \sum_{i=0}^{k-1} \xi^i/i! = \sum_{i\geq k} \xi^i/i!$.

We bound the number of graphs with minimal degree k. This lemma is also of independent interest.

Lemma 1 (Number of graphs with minimal degree k)**.** *The number of graphs with minimal degree k on ν vertices and μ edges is bounded as follows: There is a constant $C > 0$ such that for any $\xi > 0$*

$$|\mathcal{G}(k;\nu,\mu)| \leq C \cdot \left(\frac{2\mu}{\xi^2 e}\right)^{\mu} \cdot e(\xi,k)^{\nu}.$$

Proof. Given a degree sequence $\vec{d} = (d_1, d_2, \dots d_\nu)$ with $\sum_{i=1}^{\nu} d_i = 2\mu$, according to Bollobás' configuration model [Bol80], there are at most

$$\frac{(2\mu)!}{\mu! \cdot 2^\mu \cdot \prod_{i=1}^{\nu} d_i!} \leq C \cdot \left(\frac{2\mu}{e}\right)^\mu \cdot \prod_{i=1}^{\nu} \frac{1}{d_i!}$$

labelled graphs on ν vertices and μ edges such that the i-th vertex has degree d_i. Define $\mathcal{D}(k; \nu, \mu)$ to be the set of degree sequences of ν nodes with μ edges and all degrees at least k, i.e.,

$$\mathcal{D}(k; \nu, \mu) := \{\vec{d} = (d_1, d_2, \dots d_\nu) : \sum_{i=1}^{\nu} d_i = 2\mu, \forall j : d_j \geq k\}.$$

To obtain the bound, we sum over all possible degree sequences, i.e., all $\vec{d} \in \mathcal{D}(k; \nu, \mu)$. The next step, essentially, performs the difference. Coja-Oghlan and Taraz [COT04] bound the sum by $(\frac{2}{k!})^\nu \cdot (\frac{2}{k+1})^{2\mu - \nu k}$. We give a bound utilising generating functions. For $f(\xi) = \sum_{i \geq 0} a_i \xi^i$ let $[\xi^i] f(\xi) := a_i$. We claim that

$$\sum_{\vec{d} \in \mathcal{D}(k;\nu,\mu)} \prod_{i=1}^{\nu} \frac{1}{d_i!} = [\xi^{2\mu}] e(\xi, k)^\nu.$$

Proof is by induction over ν. Base case ($\nu = 1$):

$$\sum_{\vec{d} \in \mathcal{D}(k;1,\mu)} \prod_{i=1}^{1} \frac{1}{d_i!} = \frac{1}{(2\mu)!} = [\xi^{2\mu}] e(\xi, k).$$

Inductive step ($\nu - 1 \to \nu$):

$$\begin{aligned}
[\xi^{2\mu}] e(\xi, k)^\nu &= \sum_{\delta \geq k}^{2\mu} [\xi^\delta] e(\xi, k) \cdot [\xi^{2\mu - \delta}] e(\xi, k)^{\nu - 1} \\
&= \sum_{\delta \geq k}^{2\mu} \frac{1}{\delta!} \cdot \sum_{\vec{v} \in \mathcal{D}(k;\nu-1,\mu-\delta/2)} \prod_{i=1}^{\nu-1} \frac{1}{d_i!} \\
&= \sum_{\vec{v} \in \mathcal{D}(k;\nu,\mu)} \prod_{i=1}^{\nu} \frac{1}{d_i!}.
\end{aligned}$$

Furthermore, for $f(\xi) = \sum_{i \geq 0} a_i \xi^i$ with $a_i \geq 0$

$$\frac{f(\xi)}{\xi^{2\mu}} = \sum_{i \geq 0} \frac{a_i \xi^i}{\xi^{2\mu}} \geq a_{2\mu} = [\xi^{2\mu}] f(\xi).$$

Therefore, since $a_i \geq 0$, for any $\xi > 0$,

$$[\xi^{2\mu}] e(\xi, k)^\nu \leq \frac{e(\xi, k)^\nu}{\xi^{2\mu}}. \qquad \square$$

Using the bound from Lemma 1 we can now continue the proof of the main theorem, case II (few edges). For any $\xi > 0$, $|\mathcal{G}(k;\nu,\mu)|$ is bounded as follows.

$$|\mathcal{G}(k;\nu,\mu)| \quad \leq \quad C \cdot \left(\frac{2\mu}{\xi^2 e}\right)^{\mu} \cdot e(\xi,k)^{\nu}$$

It remains to show that

$$2^{h(\nu/n)\cdot n} \cdot \left(\frac{c}{n}\right)^{\mu} \cdot |\mathcal{G}(k;\nu,\mu)| \leq (\tau_k)^{-\nu}.$$

We write $\nu = x_\nu n$ and $\lambda = x_\lambda x_\nu n$. ($x_\nu \in (0,1]$, $x_\lambda < \eta$)

$$\begin{aligned} C \cdot 2^{h(\nu/n)n} \cdot (\tau_k \cdot e(\xi,k))^{\nu} \cdot \left(\frac{2\mu c}{\xi^2 en}\right)^{\mu} &< 1 \\ \Leftarrow 2^{h(x_\nu)} \cdot (\tau_k \cdot e(\xi,k))^{x_\nu} \cdot \left(\frac{c x_\nu (k + 2x_\lambda)}{\xi^2 e}\right)^{x_\nu(\frac{k}{2}+x_\lambda)} &< 1. \end{aligned}$$

Set $\xi = 1.85$. The condition above holds for all $x_\nu \in (0,1]$ and $x_\lambda \in (0,\eta)$. $\square$

References

[BE05] Richard Beigel and David Eppstein. 3-coloring in time $O(1.3289^n)$. *Journal of Algorithms*, 54(2):168–204, 2005.

[Bol80] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[COT04] Amin Coja-Oghlan and Anusch Taraz. Exact and approximative algorithms for coloring $\mathbf{G}(n,p)$. *Random Structures and Algorithms*, 24(3):259–278, 2004.

[ER60] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[FK98] Uriel Feige and Joe Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57(2):187–199, 1998.

[Gil59] Edgar N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30:1141–1144, 1959.

[Kar76] Richard M. Karp. The probabilistic analysis of some combinatorial search algorithms. In *Algorithms and complexity*, pages 1–19. 1976.

[Law76] Eugene L. Lawler. A note on the complexity of the chromatic number problem. *Information Processing Letters*, 5(3):66–67, 1976.

[PSW96] Boris Pittel, Joel Spencer, and Nicholas C. Wormald. Sudden emergence of a giant k-core in a random graph. *Journal of Combinatorial Theory Series B*, 67(1):111–151, 1996.