Several Problems about 3-Coloring on Random Graphs $G(n, p)$

Xiaoxu Guo Xuangui Huang Lin Li Chengyu Lin
Kuan Yang Lutian Zhao
Shanghai Jiao Tong University

January 18, 2013

Introduction

Given a random graph $G(n, p)$ with $p = \frac{d}{n}$, where d denotes the average degree of vertices in the graph, the threshold of 3-colorability had been found experimentally but not proved rigorously. Although several upper and lower bound results have been obtained, this problem remains open.

Introduction

Now we are focused on several interesting problems related to 3-coloring of random graphs. One is the propagation length of the color change. The other is the relationship between the propagation length and other properties of a graph, such as correlation decay.

Threshold

Main result

Theorem

Let $f_k(n,p)$ denote the probability that the random graph $G(n,p)$ is k-colorable. For every constant $k \geq 3$, there exists $d_k(n)$ such that for any $\epsilon > 0$,

$$\lim_{n\to\infty} f_k\left(n, \frac{d_k(n)-\epsilon}{n}\right) = 1 \quad \textit{and} \quad \lim_{n\to\infty} f_k\left(n, \frac{d_k(n)+\epsilon}{n}\right) = 0 \tag{1}$$

Remark

We believe that for all $k \geq 3$, $d_k(n)$ converges. However, we are not able to prove this and it remains an interesting open problem.

Threshold

Best known upper and lower bound

For $k = 3$, analytic (non-rigorous) verification of the previous experimental results by the cavity and replica symmetry breaking methods of Statistical Physics suggest that

$$d_3 \simeq 4.69 \tag{2}$$

Dubois, Mandler (03) The best known upper bound to 3-Col threshold

$$d_3 \leq 4.854 \tag{3}$$

Achlioptas, Moore (03) The best known lower bound to 3-Col threshold

$$d_3 \geq 4.03 \tag{4}$$

Propagation Length

Definition

Now let $dis(u, v)$ be the distance between vertices u and v, Σ be the set of all proper 3-colorings π on a graph.

Definition

The propagation length of a proper 3-coloring π of a 3-colorable graph is

$$\max_{u, c \neq \pi(u)} \{ \min_{\pi' \in \Sigma, \pi'(u) = c} \{ k \in \mathbb{N} | \forall v, dis(u, v) > k \implies \pi(v) = \pi'(v) \} \}$$

Propagation Length

Conjecture

Formally, if we first pick a 3-colorable graph from $G(n, p)$ at random, then pick a proper coloring of this graph uniformly at random, we want to know the expected propagation length of this proper coloring of the graph (Latter we'll call it PL for short).

Conjecture

$PL = O(\log n)$ on $G(n, \frac{3}{n})$ model.

Random coloring

Generation of random coloring

The following algorithm was used to find random coloring of given graph.

```
shuffle (v_1, v_2, ..., v_n)

def search(i)
    if i <= n then
        for x in [RED, GREEN, BLUE] do
            c_{v_i} = x
            search(i + 1)
    else
        if (c_1, c_2, ..., c_n) is a proper coloring then
            report (c_1, c_2, ..., c_n)
    end
end
```

Random coloring

Generation of propagation

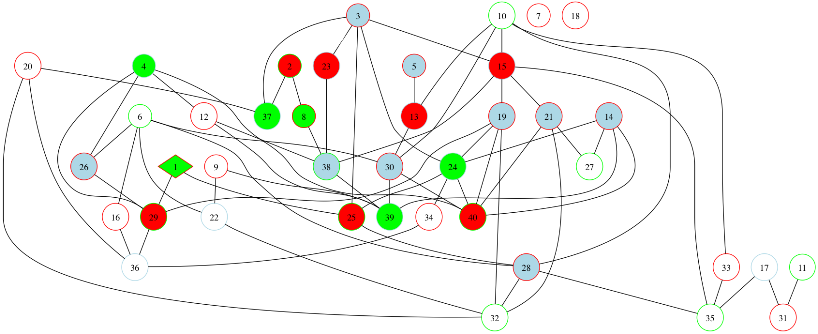
```
sort (v_2, ..., v_n) so that d_{v_2} >= d_{v_3} >= ... >= d_{v_n}
where d_v denotes the distance from v to v_1

def search(i)
    if i <= n then
        search(i + 1)
        for x in [RED, GREEN, BLUE] do
            if x != c_{v_i}
                c_{v_i} = x
                search(i + 1)
    else
        if (c_1, c_2, ..., c_n) is a proper coloring then
            report (c_1, c_2, ..., c_n)
    end
end
```

Experiment and result

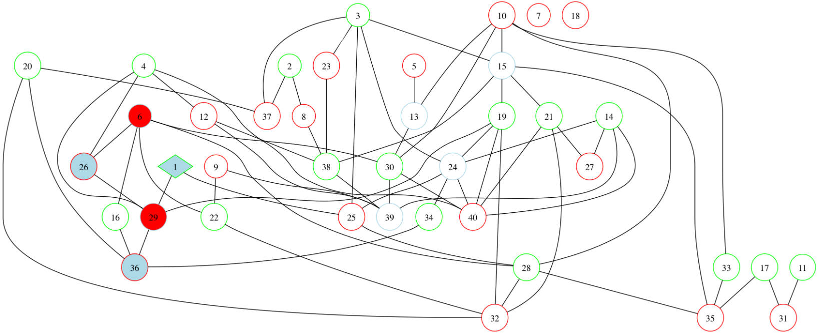
Fixed $n = 40$ and $d = 30$, random graph G was generated firstly.
After that, 100 random colorings were generated, and their propagation length were calculated.
The following figure shows the most common situation, which the propagation length falls between 3 and 4.

Experiment and result



Experiment and result

Degenerated cases will be one of the most interesting phenomenon.



Propagation length

The second set was designed to illustrate that the propagtion length of random graphs (with same d).
Fixed $d = 3$, for each n (the number of vertices), 100 random graphs with some of their coloring were generated, and the average propagation length of them were calculated.

Propagation length

n	$\bar{p}$
30	2.60
40	2.42
50	3.50
60	3.20
70	3.38

$\bar{p}$ doesn't increase as desired. One of the possible reasons is the assumption goes a bit wrong. There exists some coloring with very large propagation length, but rarely occurs.

The tree case

The simplied problem is, what is the expected propagation length of a infinite binary tree.

The tree case

Probability

Let p_n denote the probability that the propagation length is not more thant n. For simplicity, let $p_0 = \frac{1}{2}$.
There are totaly 4 cases, which of probability $\frac{1}{4}$:

- the left child conflicts with the root
- the right child conflicts
- both of the children conflict
- neither of the children conflicts

Therefore, for $n \geq 1$,

$$p_n = \frac{1}{4} + \frac{2}{4} p_{n-1} + \frac{1}{4} p_{n-1}^2$$

The tree case

Expected value

By definition, the expected value

$$E = \sum_{i=1}^{\infty} i \cdot (p_i - p_{i-1}) = \frac{1}{2} + \sum_{i=0}^{\infty} (1 - p_i)$$

The tree case

Expected value

The following table shows the result by computer simulation. $(E_n = \sum_{i=0}^{n}(1 - p_i))$

$\log_2 n$	E_n
1	0.158203125
2	0.395759789
3	0.915307650
4	1.867514073
5	3.322648169
6	5.233578079
7	7.484970289
8	9.915307650
9	12.56571487

The tree case

Expected value

One can conclude from the result that $\lim_{n\to\infty} E_n$ diverge. The reason is, the expected number of children changing its color is exactly $\frac{1}{4} \cdot 2 + \frac{1}{2} \cdot 1 = 1$. Thus, the propagation won't terminate.

Regular graphs

Regular graphs are graphs whose vertices have the same number of neighbors. It is also the simplest case of the propagation of color(other than trees). Here we analyze two cases, the cycle or the 3-regular graph.

Regular graphs

A cycle with n vertices

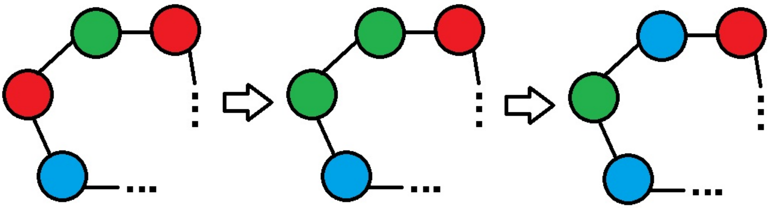


Figure: The propagation in a typical cycle, the left one is the original coloring, the middle one is the color change in a vertex, and the right one is the propagation ,it stops when turning the green one in the third graph into the blue one.

Regular graphs

A cycle with n vertices

Lemma

(Propagation in 2-regular graph) The propagation length in a cycle is at most 1.

Assumption

Consider a cycle, when it is infinitely large, we may assume that the left side color of a vertex does not affect that of the right side.

Regular graphs

A cycle with n vertices

Theorem

(Total Number of three coloring) The total number of 3-coloring of a n-nodes cycle $C(n)$ is given by $6J(n)$, where $J(n)$ is the Jacobsthal sequence, with $J(2) = J(3) = 1, J(n+1) = J(n) + 2J(n-1)$. Also, the total number of coloring, with $N_L \neq N_R$ with a fixed N, is given by $J(n-1)$, the same sequence as before.

Regular graphs

A cycle with n vertices

By the theorem before, we know that the propagation length in a $n-$cycle is given by

$$PL = \frac{\bar{C}(n)}{C(n)} + \frac{1}{2}\frac{C(n) - \bar{C}(n)}{C(n)} = \frac{3 \cdot 2^{n-1}}{2^n + 2 \cdot (-1)^n}$$

This calculation shows that the assumption 1 before is right.

Regular graphs

A cycle with n vertices

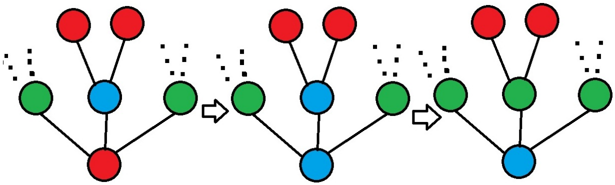
The Simpleness of the cycle enables us to extend it to a further result: the $k-$coloring.

$$\lim_{n\to\infty} PL_k(n) = \frac{1}{k-1}\left(1+\frac{\bar{C}(n)}{C(n)}\right) = \frac{k+(k-2)\sqrt{4k-3}}{2(k-1)^2}$$

From this formula, we know that when the number of colors grows larger, the propagation length is getting smaller, and eventually, it will converge to zero.

Regular graphs

A specific 3-regular graph



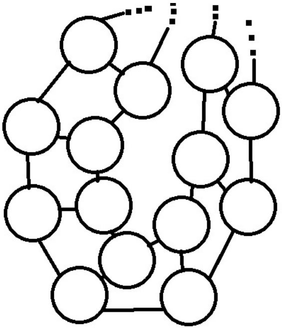
Assumption

When n is large, girth with length $o(n)$ in a $3-$regular graph will exists with low probability, and the three neighbors of a node will be colored nearly independently.

Regular graphs

A specific 3-regular graph

Next step, we should do some experiment to show the property of some special $3-$regular graph.
First we consider the contamination of two cycles like this:



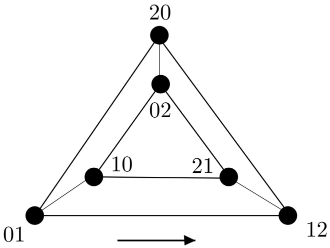
Regular graphs

A specific 3-regular graph

Theorem

(Total number of 3-coloring in the graph)
The total number of the proper coloring of the graph is given by

$$C(n) = 3^n + 2(-2)^n + 1$$



Regular graphs

A specific 3-regular graph

Lemma

(Upper bound of propagation) The propagation in this graph is at most 2

Regular graphs

A specific 3-regular graph

Neighbors	N	Pr(exist)	$L=0$	$L=1$	$L=2$
(10,10)	$\frac{1}{6}(3^{n-2}+2(-2)^{n-2}+1)$	$\frac{3^{n-2}+2(-2)^{n-2}+1}{3^n+2(-2)^n+1}$	$\frac{1}{2}$	$\frac{1}{2}$	0
(10,12)	$\frac{1}{6}(3^{n-2}+(-2)^{n-2}-1)$	$\frac{3^{n-2}+(-2)^{n-2}-1}{3^n+2(-2)^n+1}$	$\frac{1}{2}$	0	$\frac{1}{2}$
(10,20)	$\frac{1}{6}(3^{n-2}+(-2)^{n-2}-1)$	$\frac{3^{n-2}+(-2)^{n-2}-1}{3^n+2(-2)^n+1}$	0	1	0
(12,12)	$\frac{1}{6}(3^{n-2}+2(-2)^{n-2}+1)$	$\frac{3^{n-2}+2(-2)^{n-2}+1}{3^n+2(-2)^n+1}$	$\frac{1}{2}$	$\frac{1}{2}$	0
(12,10)	$\frac{1}{6}(3^{n-2}+(-2)^{n-2}-1)$	$\frac{3^{n-2}+(-2)^{n-2}-1}{3^n+2(-2)^n+1}$	$\frac{1}{2}$	0	$\frac{1}{2}$
(12,20)	$\frac{1}{6}(3^{n-2}-(-2)^{n-2}+1)$	$\frac{3^{n-2}-(-2)^{n-2}+1}{3^n+2(-2)^n+1}$	0	$\frac{1}{2}$	$\frac{1}{2}$
(20,20)	$\frac{1}{6}(3^{n-2}+2(-2)^{n-2}+1)$	$\frac{3^{n-2}+2(-2)^{n-2}+1}{3^n+2(-2)^n+1}$	0	1	0
(20,10)	$\frac{1}{6}(3^{n-2}+(-2)^{n-2}-1)$	$\frac{3^{n-2}+(-2)^{n-2}-1}{3^n+2(-2)^n+1}$	0	1	0
(20,12)	$\frac{1}{6}(3^{n-2}-(-2)^{n-2}+1)$	$\frac{3^{n-2}-(-2)^{n-2}+1}{3^n+2(-2)^n+1}$	0	$\frac{1}{2}$	$\frac{1}{2}$

Regular graphs

A specific 3-regular graph

So the propagation length, denoted by PL, is given by

$$PL = \sum \Pr(\text{exist})(\sum_i \Pr(L = i)) = \frac{3^n + \frac{5}{4}(-2)^n + 1}{3^n + 2(-2)^n + 1}$$

This time, we know that

$$\lim_{n \to \infty} PL = 1$$

for this 3−regular graph.

Calculation

Here we assume that the probability of a set of vertices is independent. That means if A_i denotes "vertex x_i is colored by color a_i" (All x_i are distinct), then $\Pr(\cap_{i=1}^{n} A_i) = \Pi_{i=1}^{n} \Pr(A_i)$.

Calculation

we let $pr(n, k)$ denote the probability of that the propagation is r in $G(n, p)$ model and $sum(n, k) = \sum_{j=0}^{k} pr(n, j)$ denote the probability of that the propagation is smaller than k in $G(n, p)$ model.
Also, we let $lim(k) = \lim_{n\to\infty} sum(n, k)$, we can assume that when n is large enough, $lim(k) = sum(n, k)$.

Calculation

We can calculate $pr(n,0)$ for each n because

$$
\begin{aligned}
pr(n,0) &= \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \left(\frac{1}{2}\right)^k \\
&= \left(1 - \frac{d/2}{n}\right)^{n-1}
\end{aligned}
$$

where $p = d/n$. Hence,

$$\lim_{n\to\infty} sum(n,0) = pr(n,0) = e^{-d/2}$$

Calculation

We can also use the value of $sum(n,k)$ to calculate the value of $sum(n,k+1)$.

$$\begin{aligned} sum(n,k+1) &= \sum_{i=0}^{n-1}\left(\binom{n-1}{k}p^k(1-p)^{n-1-k}\right. \\ &\quad \left. \times\left(\sum_{j=0}^{i}\binom{i}{j}\left(\frac{1}{2}\right)^j\left(\frac{1}{2}\right)^{i-j} sum(n,k)^j\right)\right) \\ &= \left(1-\frac{(1-sum(n,k))\times d/2}{n}\right)^{n-1} \end{aligned}$$

Calculation

Hence,

$$\lim_{n\to\infty} sum(n, k+1) = e^{-(1-sum(n,k))\times d/2}$$

Calculation

Finally, we calculate the expectation of propagation:

$$E(n) = \sum_{k=0}^{n} k \times pr(n,k) = \sum_{k=1}^{n} k \times (sum(n,k) - sum(n,k-1))$$

When n is large enough, we can assume that $E(n)$ equals to the limit of $E(n)$:

$$\begin{aligned} \lim_{n\to\infty} E(n) &= \lim_{n\to\infty} \sum_{k=1}^{n} k \times (sum(n,k) - sum(n,k-1)) \\ &= \sum_{k=1}^{\infty} k \times (lim(k) - lim(k-1)) \end{aligned}$$

Analysis

Lemma

$$0 < lim(k) < 1$$

for each $k \in \mathbb{N}$

Lemma

$$lim(k) < lim(k+1)$$

for each $k \in \mathbb{N}$

Analysis

We can conclude that when k goes to infinity, $lim(k)$ will converge to a number which is larger than 0 and not larger than 1. If d is small enough, $lim(k)$ will converge to 1.

Analysis

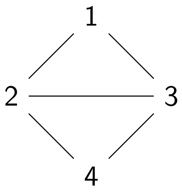
According to some programming experiment, $\lim_{n\to\infty} E(n) = \sum_{k=1}^{\infty} k \times (lim(k) - lim(k-1))$ is divergent. Also, we find that for fixed d, $E(n) = \Theta(\log n)$. It's an experimental result, but we cannot prove it.

Further Toughts

To prove the conjecture above is not a easy thing to do. The problem occurs while we want to calculate the marginal probability of a given set of vertices having a coloring configuration. It has a strange behavior.

Further Toughts

For example look at the figure:



If we consider the q-coloring problem on $G(n, \frac{d}{n})$ model, where q can be large, say $q = 2d$ or something else. We may feel better while making an analysis.

Further Toughts

Conjecture

If $PL = O(1)$, there exists an FPTAS to count the number of proper 3-colorings of $G(n,p)$ model with high probability.

Since $PL = O(1)$ may indicate the 'locality' for a random graph is extremely strong.

Further Toughts

We can just enumerate a small amount of vertices to get an estimation for the marginal probability $\Pr(x_v = c | x_1, x_2, \cdots, x_k)$ (x_i is the color for vertex i).
If we first have a proper coloring $\{x_i = c_i\}$, the number of proper 3-colorings is $\Pi_{i=1}^{n} \Pr(x_i = c_i | x_1 = c_1, x_2 = c_2, \cdots, x_{i-1} = c_{i-1})^{-1}$.

Thank you!