Chromatic Number in Preferential Growth Models

Vishnu Sampathkumar
National Public School,
Bengaluru, Karnataka, India
vishnu.sampathkumar@cfrce.org

October 21, 2021

Abstract
Barabási and Albert proposed a preferential-attachment based graph construction in 1999, where the temporal addition of nodes to the 'seed' graph was probabilistically driven based on the degree of the existing nodes. In this paper, we attempt to characterize the chromatic number of Barabási-Albert graph by comparing extant theorems that model chromatic number as well as provide a case for comparatively more accurate bounds based on cliques. In addition, we also direct the reader and future investigator towards highly accurate lower bounds, particularly the Lovasz number of the concerned graph's complement. Adopting a considerably stronger experimental component, our research makes use of a preferential-attachment simulator (code in the appendix) to test and characterize various conditions and situations on the graph.

1 Introduction

In 1999, the preferential-attachment model was proposed by Barabási and Albert [6] to accurately represent growing networks such as the internet. Barabási-Albert model (which will, going forward, be referred to as the BA model for brevity), is similar to the Erdős–Rényi and Watts–Strogatz models [4], [9], in the sense that all of them generate random graphs as well as describe network evolution.

However, the starkest contrast is that the BA model has a temporal dimension as well, starting out with a fixed set of nodes, with further nodes added stochastically added over a period of time. Due to this temporal dimension, BA graphs have other distinct characteristics. First, they are scale-free, that is, their degree distribution follows a power law. Further, these graphs are characterized by network 'hubs' - nodes in the network which successively accumulate a significantly higher degree than other networks.

Beyond characterizing web-pages on the internet, the BA preferential attachment model has been considered to describe the evolution of many other real-world networks. One major application is in the neuroscience field of connectomics, where the BA model has been proposed as a possible descriptor of brain development, describing the growth of neural networks from neurogenesis to the creation of complex structural and functional networks. Outside of neuroscience, the BA model has also been proposed to characterize the growth of social networks. Once again,

the BA model's concept of hubs - 'rich get richer' - plays a strong role in both these applications.

There are a number of derivatives of the BA model, all based on the concept of preferential attachment - salient models include the Buckley-Osthus model [7] (which introduces a parameter called the 'attraction constant'), the Móri model [8] (a specialization towards trees), and many others. Irrespective, this paper will characterize the chromatic number in a generalized preferential growth model.

2 Technical Preliminaries

2.1 Mechanism of Preferential Attachment

We define a graph $G_m^{n_0}$ having n_0 vertices and mn_0 edges for some value m, which defines the number of new links to attach at a given time-step t. Then, the probability $\prod(k)$ of a new node attaching to the ith node of $G_m^{n_0}$ is given by

$$\prod(k) = \frac{k_i}{\sum_j k_j} \tag{1}$$

where k_i is the degree the ith node, which is divided by the sum of all j pre-existing nodes.

Then, based on the Linearized Chord Diagram (LCD) [5], for $m = 1$ we can take initial graph G_1^0. Then, for G_1^{t-1} we add a single edge to generate G_1^t, such that

$$P = \begin{cases} \dfrac{k_i}{2t-1} & 1 \leqslant i \leqslant t-1 \\ \dfrac{1}{2t-1} & i = t \end{cases} \tag{2}$$

where P is the probability of attachment.

2.2 Describing Network Degree

2.2.1 Network Degree Approximations

The node degree, k, can be approximated, represented with respect to t, a continuous, real variable. Following from (1), the instantaneous rate of attachment for degree k_i node can be taken as the product of m and $\prod(k_i)$, that is

$$\frac{dk_i}{dt} = m\prod(k_i). \tag{3}$$

Then, looping over all j nodes except ith node, we have

$$\sum_j k_j = m(2t-1) \tag{4}$$

From equation (3) we have

$$\frac{dk_i}{dt} = \frac{1}{2t-1}. \tag{5}$$

For large values of t we can neglect the constant in the denominator such that

$$\frac{dk_i}{dt} = \frac{1}{2t}. \tag{6}$$

Since at time t_i, a new node has m links with the existing graph, we can say that

$$k_i(t_i) = m. \tag{7}$$

Finally, integrating, we get

$$k_i = m\left(\frac{t}{t_i}\right)^{\beta} \tag{8}$$

where β is taken as the dynamical exponent having the value 0.5.

2.2.2 Degree Distribution

Degree distribution in preferential attachment models follows an approximate power law. Based on the Linearized Chord Model, the power law, p_k, can be understood as,

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)} \tag{9}$$

which reduces to

$$p_k = k^{-3} \tag{10}$$

when $k >>> m$, a result which is inevitable with progressive time-steps, considering the concept of 'rich get richer' high-degree hubs.

3 Simulation

The Barabási-Albert Graphs were simulated according to the pseudo-code schema presented below.

Algorithm 1 Barabási-Albert Graph Construction

```
1: procedure PREFATTACH
2:     n_i ← nodes at given timestep
3:     if m < 1 then return false
4:     while n_i < n do
5:         n_i ← m links
6:         m links ← terminal node.
7:         goto loop.
8:         close;
9:     goto top.
```

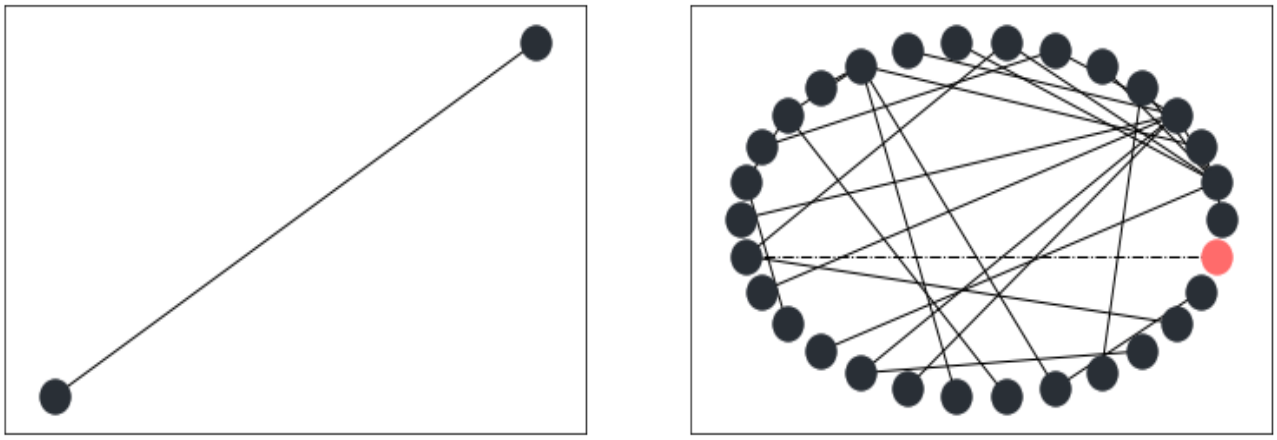

Figure 1: BA graph at two stages of construction: $m = 2$ at $t = 0$ (left); fully constructed graph at $t = 30$ where colored node is the most recently added node (right)

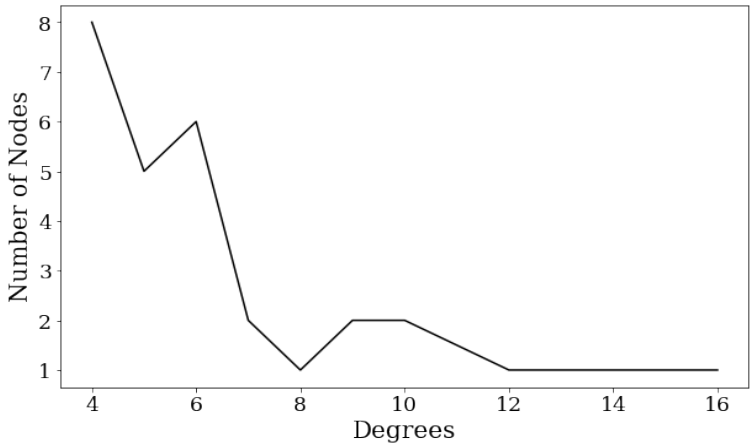

Figure 2: Degree distribution of above BA graph

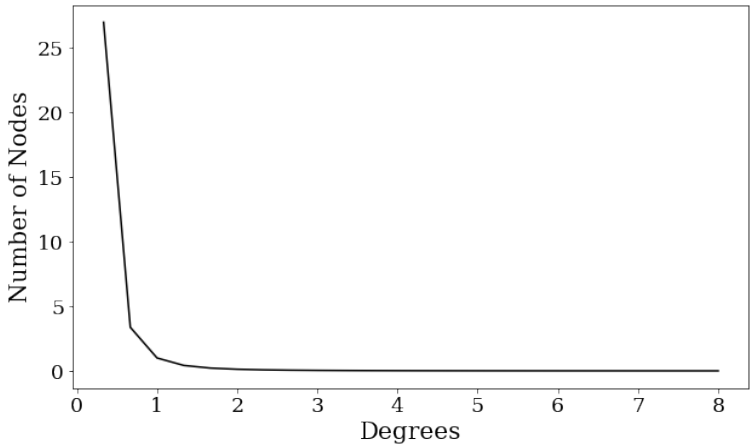

Figure 3: Predicted degree distribution using *eq.*(10)

4 Strange Behavior of Chromatic Number

Running the simulation, we observed a number of strange behavior of the chromatic number, χ, with relation to varying values of n and m. Namely:

- χ did not depend at all on the value of n. When n was varied from 5 to 800, keeping m constant, χ was also observed to remain constant.
- the value of χ incrementally increased (by approximately 4) when the value of m was incremented by 10, keeping n constant.
- When $m = n - 1$, G_m^n transforms to star S_m^n; since a star is a special case of a tree, we see that $\chi = 2$.

5 Characterizing Upper Bounds for χ

5.1 Theorized but Improbable Bounds

5.1.1 Brooks' Bound

A logical upper bound to chromatic number would be Brooks' [1] theorem, which states that the chromatic number, χ can be defined as:

$$\chi_b \leqslant \Delta(G) + 1, \tag{11}$$

where Δ(G) is the maximum node degree of the graph.

Alternatively, Brooks' theorem can be textually defined as $\chi(G) \leqslant \Delta(G)$ unless G contains a clique of size $\Delta(G) + 1$ or $\Delta(G) = 2$ and G contains an odd cycle. However, for the purpose of this paper, we will be adhering to the definition of *eq.*(11). However, due to the presence of high-degree hubs in the graph, $\Delta(G)$ is usually several times greater than the average node degree. For small values of n and m, $\Delta(G)$ was usually 25 to 30 times higher than the average node degree, while for $m \to n$ we observed that $\Delta(G) \to m$.

To judge the accuracy of defining χ with Brooks' theorem, we used greedy coloring in our simulation to approximate the actual chromatic number of the graph, and compared it to the predicted values, χ_b. To make an accurate comparison, we kept fixed value of n (as mentioned earlier, χ did not depend on n but rather only on m) and incrementally varied m from 1 (the minimum value of m for BA models) to $n - 1$ (the maximum value of m).

5.1.2 Kovalenko's Bound

As a comparison to Brooks' Theorem, we considered Kovalenko's [10] upper bound for χ. According to Kovalenko,

$$\chi_k \leqslant \frac{\log n - \log m}{\log\left(1 + \frac{1}{m}\right)} + m + 1 \tag{12}$$

We used the same comparative method for χ_k as we did for χ

We observed that in every case of m, both χ_b and χ_k had the least deviation from χ when $m = 1$, while maximum deviation for χ_k observed when $m \approx \frac{n}{3}$.

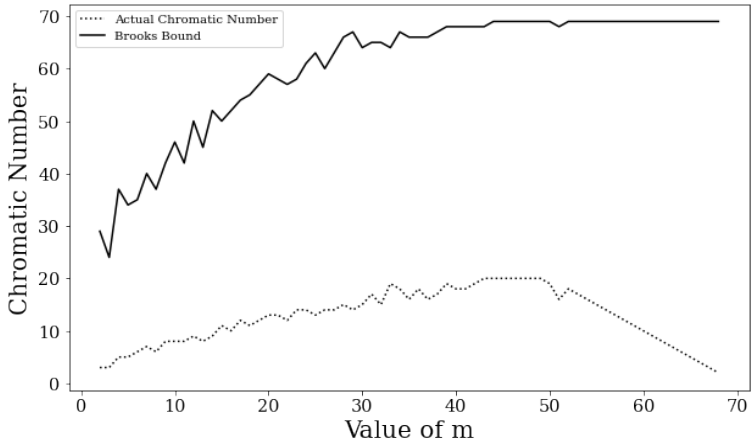

Figure 4: Actual vs Predicted: Distribution for Brooks' Upper Bound

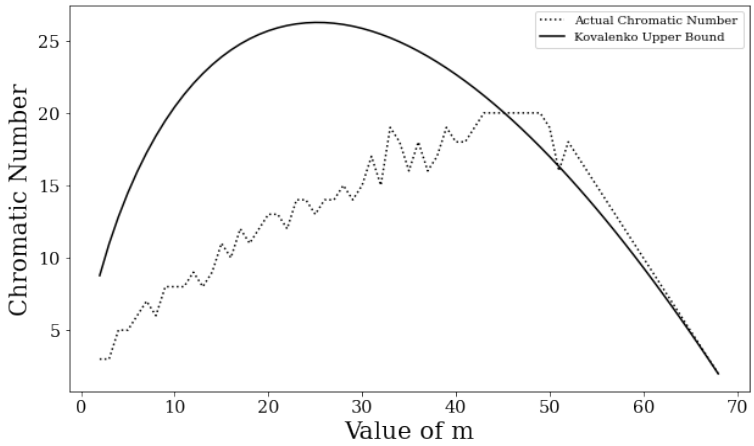

Figure 5: Actual vs Predicted: Distribution for Kovalenko's Bound

5.1.3 The Inapplicability of Brooks' and Kovalenko's Bounds

As observable from the above graphs, χ_b is not viable as a bound due to the high inaccuracy - since it is wholly dependent on $\Delta(G)$ it steadily deviates away from χ as m increases. While χ_k is significantly more accurate than χ_b, it was observed that $\chi_k - \chi$ reached negative values - this fundamentally contradicts the concept of an 'upper bound'. Analyzing the formulation of χ_k, it is clear that its foundational fallacy is due to its dependence on n, which, as observed from the simulation, did not have any effect on χ_k.

5.2 Improved Bound Based on Cliques

Considering the aforementioned inapplicability of χ_k and χ_b to form an upper bound for χ, we now turn our attention to incorporating the clique number $\omega(G)$ within the upper bound. In 1998, Reed [3] generalized an upper bound for chromatic number based on degree and clique number; considering that the process of preferential growth was biased towards generating complete induced subgraphs K_n (refer fig. 2).

Thus, we decided to repurpose Reed's upper bound for the chromatic number, χ_r to the preferential attachment model.

According to Reed, the chromatic number can be bounded such that:

$$\chi_r \leqslant \frac{\Delta(G) + 1 + \omega(G)}{2} \tag{13}$$

which can alternatively be reduced to

$$\chi_r \leqslant \frac{\chi_b + \omega(G)}{2}. \tag{14}$$

Based on the above relationships, we can see that

$$\chi_k \leqslant \chi_r \leqslant \chi^b \tag{15}$$

observing that $\chi_k = \chi_r = \chi_b$ only when $\Delta(G) = 1$. From the below graph, we can see that while not as accurate as χ_k, χ_r does not suffer from the fundamental fault of dipping below the actual chromatic number that the former does when $m > \frac{2n}{3}$

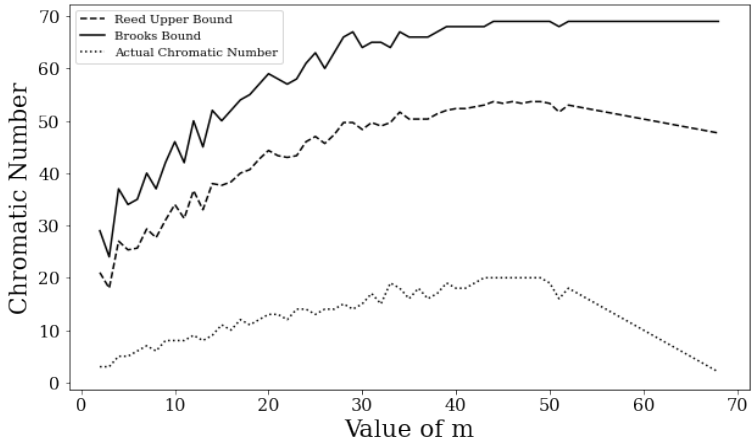

Figure 6: Three-way Comparison Between Brooks' Bound, Reed's Bound and Actual Chromatic Number

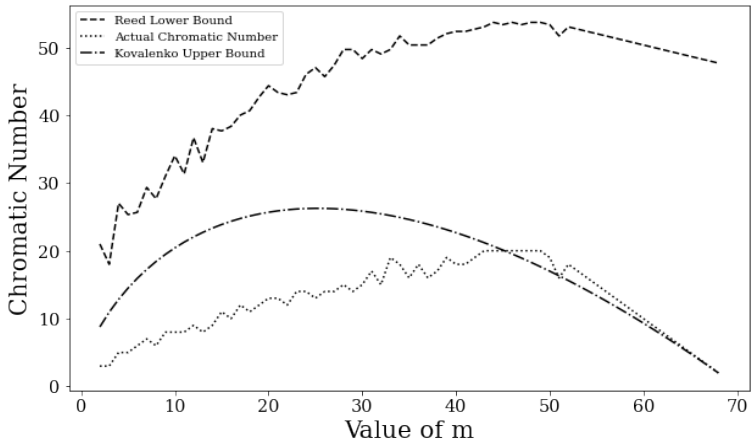

Figure 7: Three-way Comparison Between Kovalenko's Bound, Reed's Bound and Actual Chromatic Number

6 Characterizing Lower Bounds for χ

6.1 Kovalenko's Lower Bound

While there is a moderate amount of information available relating to characterizing the upper bound of chromatic numbers of preferential attachment models, significantly less information is available for characterizing the lower bound.

Kovalenko's lower bound is defined as:

$$\chi > \frac{m}{(4+\epsilon)\log m} \tag{16}$$

for large enough $m = m(\epsilon)$. However, it is uncertain how to determine the value of ϵ, and we assume that it is present to account for the rapid decrease in chromatic number seen when $m \to n$ (see section 4, bullet 2). Thus we can compare the values of actual and predicted chromatic number, as we had done with the upper bound as well.

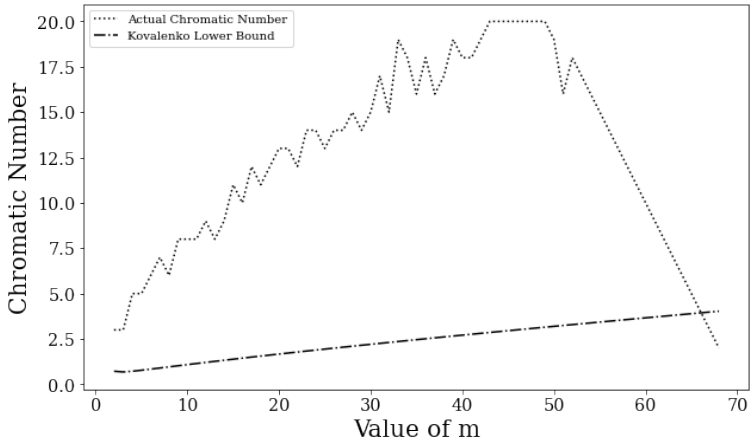

Figure 8: Lower bound Comparison

Once again, here, we can account for the dipping of chromatic number below Kovalenko's lower bound by considering the constant ϵ accordingly.

6.2 Lovasz Number

Finally we define the Lovasz number [2] as the most accurate predictor of chromatic number. The Lovasz number can be bounded by the relation

$$\alpha(G) \leqslant \Theta(G) \leqslant \vartheta(\overline{G}) \leqslant \chi, \tag{17}$$

where $\alpha(G)$, $\Theta(G)$, and $\vartheta(\overline{G})$ denote the Independence number, the Shannon capacity and the Lovasz number, respectively. While initially just a quantitative upper bound of the Shannon capacity (whose calculation is considered an NP-complete problem), we observed that in the context of the preferential attachment model, $\vartheta(\overline{G})$ could always be rounded to the nearest ones place, to yield the exact value of the chromatic number.

It must be noted that the Lovasz number is explicitly calculated for the complement of the graph, not for the graph itself. It is defined by the equation:

$$\vartheta(G) = \max_{d,V} \sum_{i \in V} (d^{\mathrm{T}} v_i)^2, \tag{18}$$

where d is a unit vector and v_i is each entry in the orthonormal representation of $\overline{G}$. While corrected versions of the Lovasz number like the Szegedy and Schriver number do exist, discussion regarding their ability to predict chromatic number is outside the scope of this paper. As with the simulation, the code regarding the computation of the Lovasz number is present in the appendix.

References

[1] R. L. Brooks, "On colouring the nodes of a network," in *Mathematical Proceedings of the Cambridge Philosophical Society*, Cambridge University Press, vol. 37, 1941, pp. 194–197.

[2] L. Lovász, "On the shannon capacity of a graph," *IEEE Transactions on Information theory*, vol. 25, no. 1, pp. 1–7, 1979.

[3] B. Reed, "ω, δ, and χ," *Journal of Graph Theory*, vol. 27, no. 4, pp. 177–212, 1998.

[4] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[5] B. Bollobás and O. Riordan, "Linearized chord diagrams and an upper bound for vassiliev invariants," *Journal of Knot Theory and its Ramifications*, vol. 9, no. 07, pp. 847–853, 2000.

[6] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[7] P. G. Buckley and D. Osthus, "Popularity based random graph models leading to a scale-free degree sequence," *Discrete Mathematics*, vol. 282, no. 1-3, pp. 53–68, 2004.

[8] T. F. Móri, "The maximum degree of the barabási–albert random tree," *Combinatorics, Probability and Computing*, vol. 14, no. 3, pp. 339–348, 2005.

[9] P. Erdös and A. Rényi, "On the evolution of random graphs," in *The structure and dynamics of networks*, Princeton University Press, 2011, pp. 38–82.

[10] K. Kovalenko, "On the independence number and the chromatic number of generalized preferential attachment models," *Discrete Applied Mathematics*, vol. 285, pp. 301–306, 2020.

7 Appendix

7.1 Simulation Code

```
import networkx as nx
def Barabasi_albert_graph(n, m, seed):
    if m < 1 or m >=n:
        raise nx.NetworkXError("BarabasiAlbert network must have m >= 1")
    if seed is not None:
        random.seed(seed) #randomize the graph construction process
    G=empty_graph(m) #begin with an empty graph
    G.name="Barabasi_albert_graph"
    tgt=list(range(m))
    r_nodes=[]
    src=m
    while src<n:
        G.add_edges_from(zip(*m,tgt))
        r_nodes.extend(tgt)
        r_nodes.extend(*m)
        tgt = _random_subset(r_nodes,m)
        src += 1
    return G
```

7.2 Lovasz Number

```
from __future__ import print_function
import numpy as np
import cvxopt.base
import cvxopt.solvers

def lovasz_theta(G, long_return=False, complement=False):

    (nv, edges, _) = parse_graph(G, complement)
    ne = len(edges)

    # This case needs to be handled specially.
    if nv == 1:
        return 1.0

    c = cvxopt.matrix([0.0]*ne + [1.0])
    G1 = cvxopt.spmatrix(0, [], [], (nv*nv, ne+1))
    for (k, (i, j)) in enumerate(edges):
```

```
    G1[i*nv+j, k] = 1
    G1[j*nv+i, k] = 1
for i in range(nv):
    G1[i*nv+i, ne] = 1

G1 = -G1
h1 = -cvxopt.matrix(1.0, (nv, nv))

sol = cvxopt.solvers.sdp(c, Gs=[G1], hs=[h1])

if long_return:
    theta = sol['x'][ne]
    Z = np.array(sol['ss'][0])
    B = np.array(sol['zs'][0])
    return { 'theta': theta, 'Z': Z, 'B': B }
else:
    return sol['x'][ne]
```