COMP 572 Bioinformatics: Network Analysis
Homework Assignment #3
Fall 2009
Due Date: Dec 3, 2009 (in class).

1 Betweenness of a set of nodes

We saw in class the definition of betweenness of a single node. How would you define the betweenness of a set of nodes but still capture the essence of this measure? Graphically show the effectiveness of your definition (in other words, draw some graphs to illustrate why the betweenness measure you propose satisfies the properties you require). In particular, discuss the relationship between the betweenness of a set X of nodes, and the betweenness of every element in X.

2 The Discovery of Scaling

In class, we studied the Barabási-Albert model for constructing random scale-free network. The idea is network growth and preferntial attachment. Same idea can be captured using *deterministic fractal* algorithms.

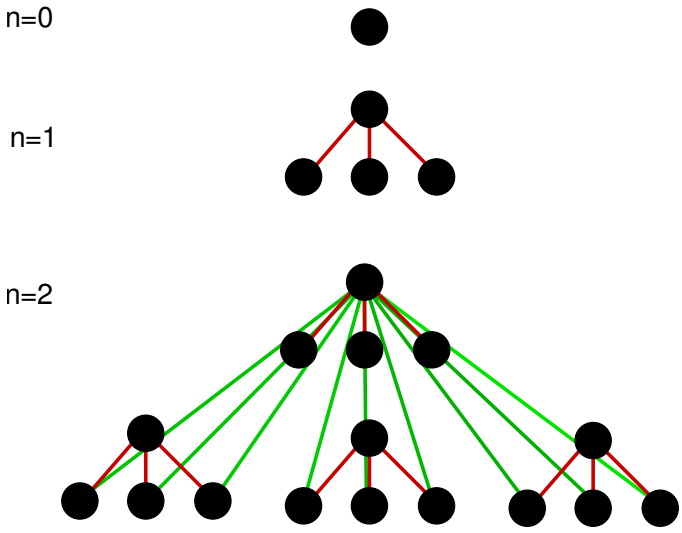
Now implement a simple program which does the following:

Initialization Start from a single node (called root) and add three nodes and connect all of them to the root. We call this network G_1

Iteration Add 3 copies of G_n and connect each of the 3^n bottom nodes of these two units to the root of the network. We call this updated network G_{n+1}.

End Until the total number of nodes exceeds certain threshold N

The following diagram shows the first 2 steps of this algorithm.



(a) What is the number of nodes of this model as a function of n? What is the number of edges? (Hint: Try $n = 1, 2, \ldots$ and see how many nodes do you get).

(b) Plot degree distribution for large $N = 10000$. How does the tail of the degree distribution scale (logarithmic, linear or exponential) ? If it is logarithmic, determine its base. If it is linear determine its slope. If it is exponential determine its power.

(c) (extra credit) Analytically prove your observation.

3 Evolution of Networks

Read the following two papers

> T. Yamado and P. Bork, "Evolution of biomolecular networks—lessons from metabolic and protein interactions", Nature Reviews Molecular Cell Biology, 10:791-803, 2009

and

> M. Lynch, "The evolution of genetic networks by non-adaptive processes", Nature Reviews Genetics, 8:803-813, 2007.

(1) For each of the papers, write down four questions you have (approach, data, conclusions, assumptions, etc.), and (2) compare the two papers in terms of the ideas they propose, and comment on whether you see them as complimentary, contradictory, identical, or completely unrelated! Discuss how you reached your conclusions (based on issues raised in the two papers).