CSE 6010 - Assignment 3

Part1 By Xun Cao
Part2 By Qi Zheng

Graph generation Description:

To represent my graph. I used a linked list. For each vertex i, there is a linked list.
The first node of the linked list for vertex i is one node which vertex i is connected with. And it is represented by an ***edge structure*** including node_id, next pointer. The next pointer points to next node which is also connected to vertex i.

The head pointer for vertex I's linked list is stored as a member of ***node structure*** in an array. The node structure includes first pointer which points to the first node (stored as an ***edge structure***) in vertex i's linked list, last pointer which points to the last node in vertex i's linked list and also the vertex i's degree. The array size is the size of vertices from 0 to n-1.

For initialization process, I set up 3 nodes – 0, 1, 2. And they form a fully connected graph.

In order to get a N nodes graph, I insert nodes using a loop from 3 to N-1.
In the insert process. First I choose which node should be selected to connect newly added node. This follows the probability P(i) = d(i) / total degree, where i is the selected node. To implement this, I generate a random value T between 1 and total degree. Then, I use a loop to scan existing k nodes from 0 to k-1, in each step, I compare the d(i) with random value T. If d(i) >= T, add the new node to i. If not, subtract d(i) from T and repeat the process for next node.

After deciding which node i to add to, add the new node to node i 's linked list and increase node i 's degree by one. Also, create the new node's linked list and add i to it, initialize its degree as 1.

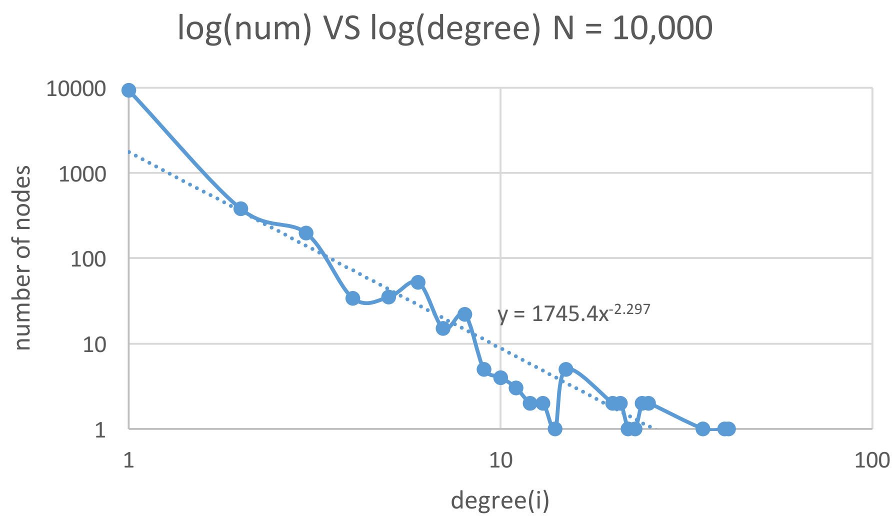
Thus, we will get a N nodes graph.

Format of Graph generation Output:

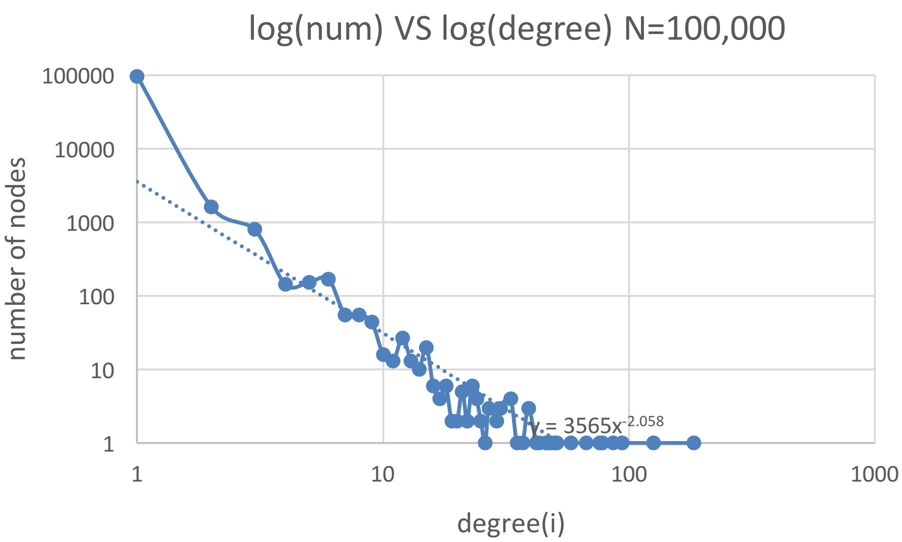
So the output file is a representation of linked list. The very first line is the total number of nodes (N). Each line represents a node's linked list from node 0 to node n-1. On each line, if first number the degree of node i, and the second number is the node id i. The left numbers are those nodes connected to node i.

Topology analysis:

The topology generated is scale-free.
As shown in the graph, the distribution of node degree follows power law.
10000 nodes are used for the simulation.
Plot the degree i on x axis, and number of nodes with that degree on y axis.
We can get the correlation formula $y = 1745.4x^{-2.297}$
When I create log(x) VS log(y), the following plot produce approximately a straight line.



Then, I try different values. The plot for N = 100,000:



Part_2. Study of the growth of the path length between vertices as the size of the scale-free network grows.