Advance Operating System

Project Report

5/4/2015

Shaomin (Samuel) Zhang
Suresh Vadlakonda

1. Abstract:

Peer to peer communication is process to share resources within set of users on internet without the help of central server. All peer-to-peer networks are overlay networks. An overlay network is built on top of an existing network, where the set of nodes is a subset of the set of nodes of the original network, and the edges correspond to paths between distinct nodes. Duncan watts and Steven Strogatz proposed random graph generation model base on small world properties Including average path length and high clustering. We here simulate the lattice ring using a model proposed by Duncan Watts and Steven Strogatz.

2. Introduction:

This simulation explores the formation of networks that result in the "small world" phenomenon. A popular example of the small world phenomenon is power grids to the neural networks of worms. This model illustrates some general, theoretical conditions under which small world networks between people or things might occur.

The aim of this project was to create small world network using Watts and Strogatz phenomenon, also to study changes in Diameter when probability of network changes.

The programming labors were divided as this: Suresh was responsible for GUI part. Samuel was responsible for logical model and the final integration.

The rest of the report is organized as follows:

- Background information
- Implementation and Analysis
- Conclusion and future work

3. Background information:

3.1. Small World Network:

A small-world network is a type of mathematical graph in which most nodes are not neighbors of one another, but most nodes can be reached from every other by a small number of hops or steps. Specifically, a small-world network is defined to be a network where the typical distance L between two randomly chosen nodes (the number of steps required) grows proportionally to the logarithm of the number of nodes N in the network.

3.1.1 Properties of small-world networks:

Small-world networks tend to contain cliques, and near-cliques, meaning sub-networks which have connections between almost any two nodes within them. This follows from the defining property of a high clustering coefficient. Secondly, most pairs of nodes will be connected by at least one short path. This follows from the defining property that the mean-shortest path length be small. Several other properties are often associated with small-world networks. Typically there is an over-abundance of hubs - nodes in the network with a high number of connections (known as high degree nodes). These hubs serve as the common connections mediating the short path lengths between other edges.

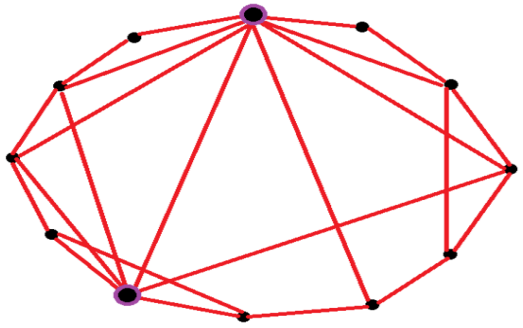


Figure 1

The Watts–Strogatz model is a random graph generation model that produces graphs with small-world properties, including Diameter and high clustering. It was proposed by Duncan J. Watts and Steven Strogatz. The model also became known as the (Watts) beta model after Watts used to formulate it in his popular science book Six Degrees.

3.1.2 Diameter (shortest path):

For a ring lattice the Diameter is l(0) = N/2K >>1 where N is number of nodes and K is number of short range neighbor and scales linearly with the system size. In the limiting case of β -> 1 where(β is probability) the graph converges to a classical random graph with l(1) $\frac{In\ N}{In\ K}$. However, in the intermediate region 0< β<1 the Diameter falls very rapidly with increasing β, quickly approaching its limiting value.

3.1.3 Clustering coefficient:

Clustering is an important property of social networks: People tend to have friends who are also friends with each other, resulting in sets of people among which many edges exist, while a set made from randomly chosen people would have a much smaller number of edges between them. The clustering

coefficient is a real number between zero and one that is zero when there is no clustering, and one for maximal clustering, which happens when the network consists of disjoint cliques.

3.2. Floyd algorithm:

The Floyd–Warshall algorithm compares all possible paths through the graph between each pair of vertices. It is able to do this with $\Theta(|V|^3)$ comparisons in a graph. This is remarkable considering that there may be up to $\Omega(|V|^2)$ edges in the graph, and every combination of edges is tested. It does so by incrementally improving an estimate on the shortest path between two vertices, until the estimate is optimal.

Consider a graph *G* with vertices *V* numbered 1 through *N*. Further consider a function shortest Path(*i*, *j*, *k*) that returns the shortest possible path from *i* to *j* using vertices only from the set {1,2,...,*k*} as intermediate points along the way. Now, given this function, our goal is to find the shortest path from each *i* to each *j* using only vertices 1 to *k* + 1.

For each of these pairs of vertices, the true shortest path could be either (1) a path that only uses vertices in the set {1, ..., *k*} or (2) a path that goes from *i* to *k* + 1 and then from *k* + 1 to *j*. We know that the best path from *i* to *j* that only uses vertices 1 through *k* is defined by shortest Path(*i*, *j*, *k*), and it is clear that if there were a better path from *i* to *k* + 1 to*j*, then the length of this path would be the concatenation of the shortest path from *i* to *k* + 1 (using vertices in {1, ..., *k*}) and the shortest path from *k* + 1 to *j* (also using vertices in {1, ..., *k*}).

If w(i,j) is the weight of the edge between vertices *i* and *j*, we can define shortest Path(*i*, *j*, *k* + 1) in terms of the following recursive formula: the base case is shortest Path(I,j,0) = w(i,j) and the recursive case is shortest Path(I,j,k+1) = min(shortest Path(I,j,k), shortest Path(I,k+1,k)+ shortest Path(k+1,j,k)).

This formula is the heart of the Floyd–Warshall algorithm. The algorithm works by first computing shortestPath(*i*, *j*, *k*) for all (*i*, *j*) pairs for *k* = 1, then *k* = 2, etc. This process continues until *k* = *n*, and we have found the shortest path for all (*i*, *j*) pairs using any intermediate vertices. Pseudocode for this basic version follows:

4. Implementation and Analysis:

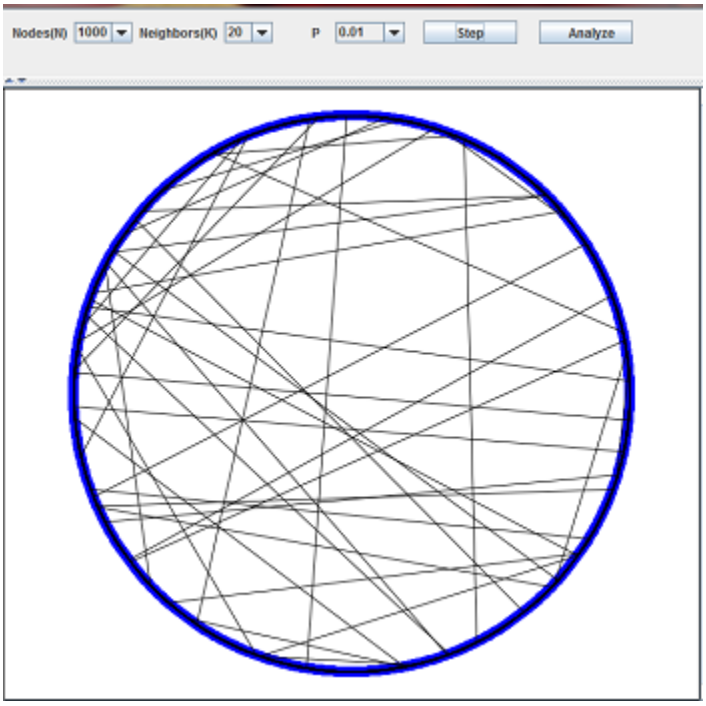
This simulation is an adaptation of a model proposed by Duncan Watts and Steve Strogatz (1998). It begins with a network where each node is connected to his or her two neighbors on either side. The network picks a random connection (or "edge") and rewires it. By rewiring, we mean changing one end of a connected pair of nodes, and keeping the other end the same.

To identify small worlds, the "Diameter" of the network are calculated and plotted. These two plots are separated because the x-axis is slightly different. The REWIRE-ONE x-axis is the fraction of edges rewired so far, whereas the REWIRE-ALL x-axis is the probability of rewiring. Networks with short Diameters are considered small world networks.

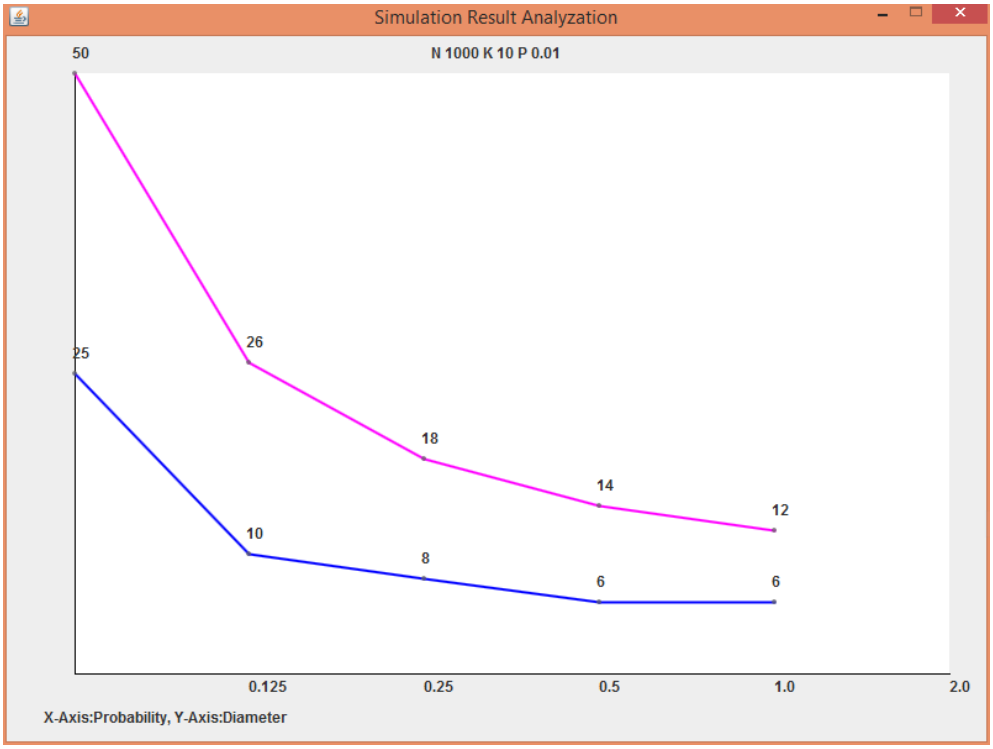
Diameter: Diameter is calculated by finding the shortest path between all pairs of nodes, adding them up, and then dividing by the total number of pairs. This shows us, on average, the number of steps it takes to get from one member of the network to another.

Clustering Coefficient: Another property of small world networks is that from one person's perspective it seems unlikely that they could be only a few steps away from anybody else in the world. The clustering coefficient for the entire network is the average of the clustering coefficients of all the nodes. A high clustering coefficient for a network is another indication of a small world.

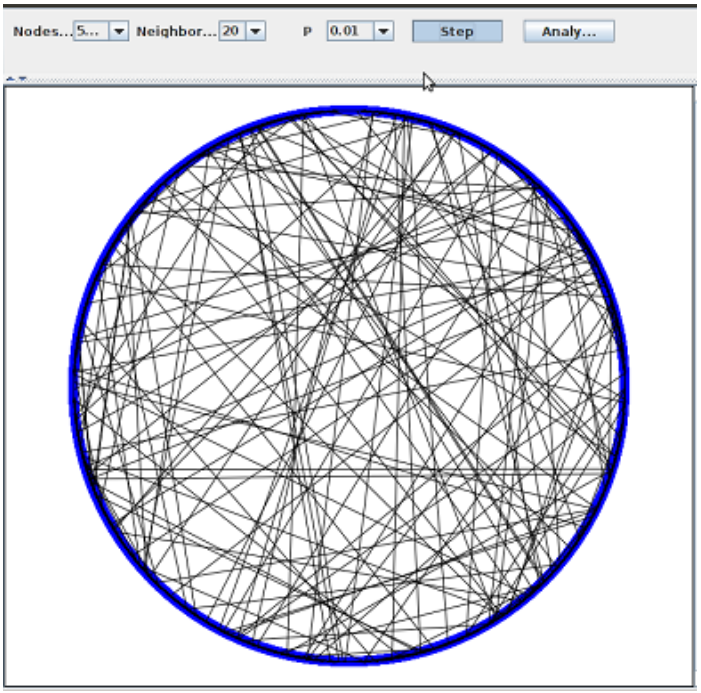
Example1: here we use 1000 node(N) for experiment with 20 short range neighbors(K). By using Floyd algorithm we calculated Diameter(Shortest path) and mean path.



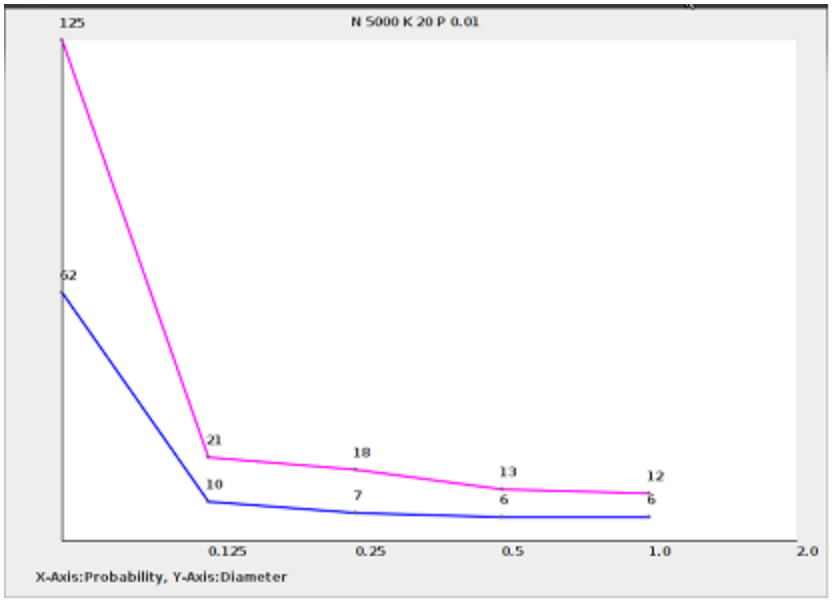
Analysis and result: Diameter is high on low β but in the intermediate region i.e. 0< β<1 the Diameter falls very rapidly with Increasing β, quickly approaching its limiting value. As shown in figure below.



Example2: here we use 5000 node(N) for experiment with 20 short range neighbors(K). By using Floyd algorithm we calculated Diameter(Shortest path) and mean path.



Analysis and result: Diameter is high on low β but in the intermediate region i.e. 0< β<1 the Diameter falls very rapidly with Increasing β, quickly approaching its limiting value. As shown in figure below.



5. Conclusion and future work:

In this simulation we use Floyd algorithm to accomplished shortest path between all pairs of nodes. The main drawback of this algorithm is it runs slowly for large number of nodes. That is because the time it takes for the Floyd algorithm (or other "all-pairs-shortest-path" algorithm) to run grows polynomially with the number of nodes.