**Web and Social Computing**
**Lab Assignment 2**

**Submitted to : Dr. Sowmya Kamath**
**Submitted by : Shaikh Sahil Ahmed**
**Roll number : 192IT022**

**Dataset 1:** wiki-Vote
**Dataset 2:** p2p-Gnutella04
**Dataset 3:** p2p-Gnutella06

Based on the earlier experiments using Dataset 1, 2 and 3 we calculated various parameters of graphs like geodesic, average clustering coefficient, strongly connected components, diameter, sparseness, k-connectedness and degree distribution. Now in this assignment, the task is to implement Erdos Renyi, Watts Strogatz and Barabasi Albert model with their properties.

**A. Erdos Renyi**
1. It is the simplest possible model to explain the random complex networks
2. Network is undirected
- Start with all isolated nodes that are having no edges. Now add edges between pair of nodes one at a time randomly.
- Two possible choices for adding edges randomly:
  - Randomize edge presence
  - Randomize node pairs

So, I have implemented the model using randomized node pairs.
Two parameters:
1. Number of nodes: n
2. Number of edges: m

Randomized node pair says that pick a pair of nodes at random among the n nodes and add an edge between them if already not present and repeat until exactly m edges have been added. This model is also represented by G(n,m).

Factors like degree distribution, average clustering coefficient and average shortest path length has been calculated. The function ‘gnm_random_graph(n, m)’ is used to generate random graph with given the value of n and m.

**<u>Dataset 1:</u>**
n = 7115
m = 103689

a) Degree distribution:

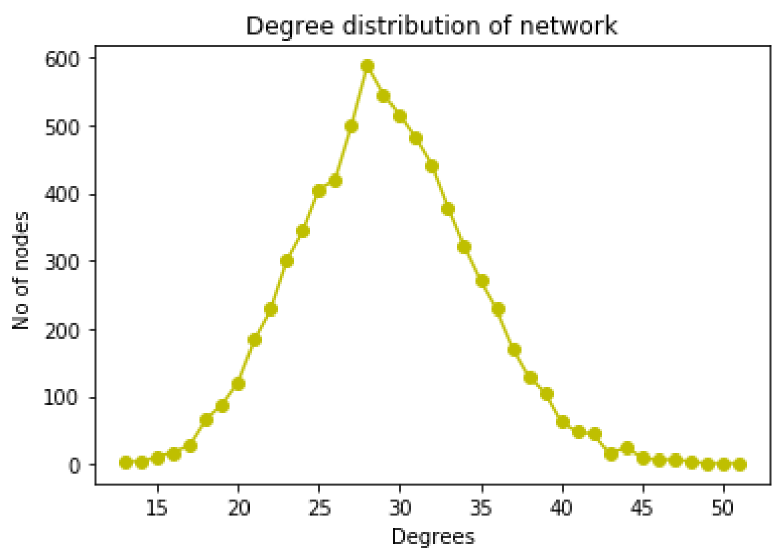


Fig 1. Degree distribution of Erdos Renyi model for dataset 1.

b) Average clustering coefficient = 0.004168010043203969
c) Average shortest path length = 2.919553280566207

**Dataset 2:**
n = 10876
m = 39994
a) Degree distribution:

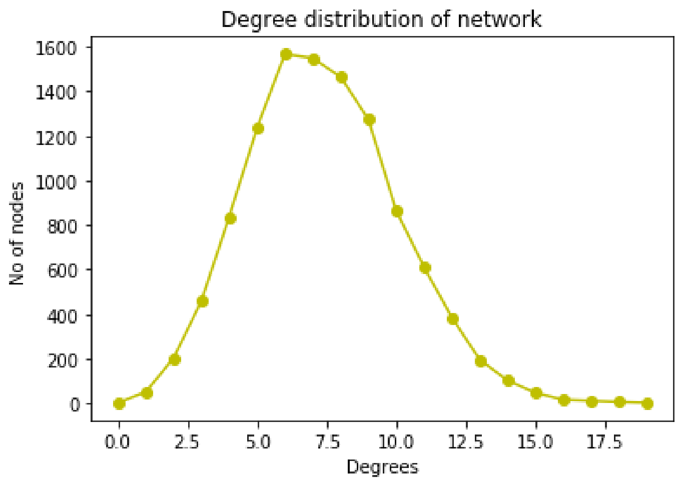


Fig 2. Degree distribution of Erdos Renyi model for dataset 2.

b) Average shortest path length :
As the graph is having 3 component:
1st component average shortest path length : 4.886061947407882
2nd component average shortest path length : 0
3rd component average shortest path length : 0

c) Average clustering coefficient = 0.0008810223770849435

**Dataset 3:**
n = 8717
m = 31525

a) Degree distribution:

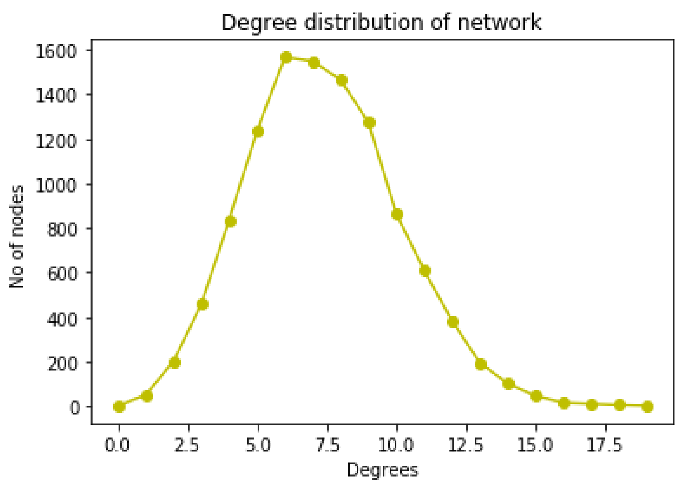


Fig 3. Degree distribution of Erdos Renyi model for dataset 3.

b)Average shortest path length :
As the graph is having 7 component:
$1^{st}$ component average shortest path length : 4.810906805699634
$2^{nd}$ component average shortest path length : 1.0
$3^{rd}$ component average shortest path length : 0
$4^{th}$ component average shortest path length : 0
$5^{th}$ component average shortest path length : 0
$6^{th}$ component average shortest path length : 0
$7^{th}$ component average shortest path length : 0

c) Average clustering coefficient = 0.0007510166958574055

From the above calculations, it denotes that dataset 3 is having relatively low average clustering coefficient as compared to other two datasets. Dataset 3 is having 7 components.

**B. Watts-Strogatz model**

1. It is the effort to extend the Erdos Renyi model to a better predictor of real network properties, that is it addresses the poor prediction of clustering.
2. Most nodes are just a few edges away on average.

The function 'watts_strogatz_graph(n,k,p,seed=None)' is used for building the graph where n is number of nodes, each node is connected to k nearest neighbors in ring topology and p is the probability of rewiring each edge.

**<u>Dataset 1:</u>**

n = 7115
k = 12
p = 0.65

a) Degree distribution:

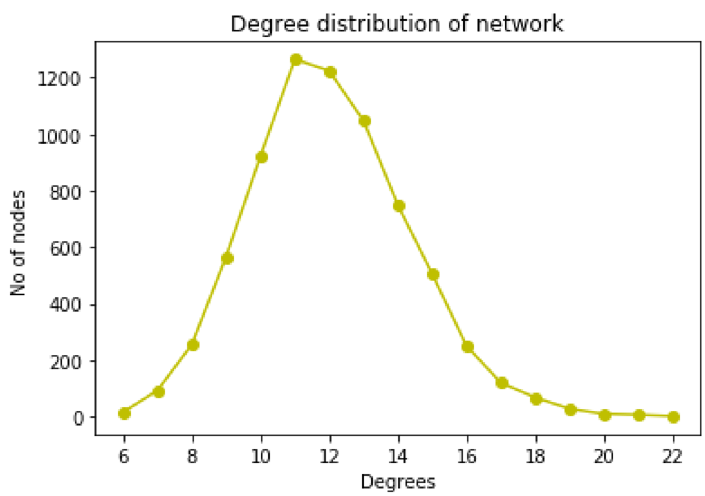


Fig 4. Degree distribution of Watts Strogatz model for dataset 1.

b) Average shortest path length : 3.8954970660526858
c) Average clustering coefficient = 0.03131629526512166

**<u>Dataset 2:</u>**
n = 10876
k = 12
p = 0.65

a) Degree distribution:

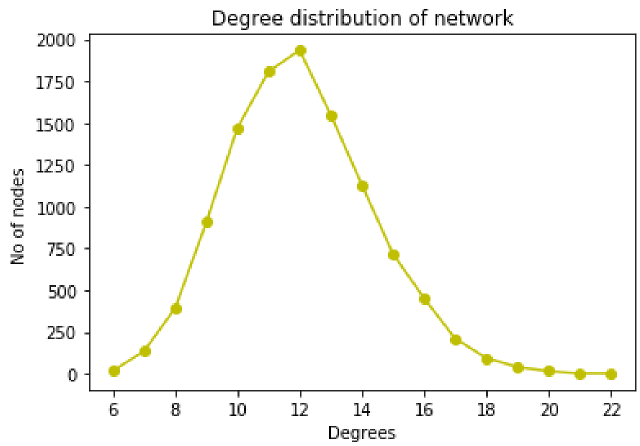


Fig 5. Degree distribution of Watts Strogatz model for dataset 2.

b) Average shortest path length : 4.081984333320651
c) Average clustering coefficient = 0.029920304015337122

**<u>Dataset 3:</u>**

n = 8717
k = 12
p = 0.65

a) Degree distribution:

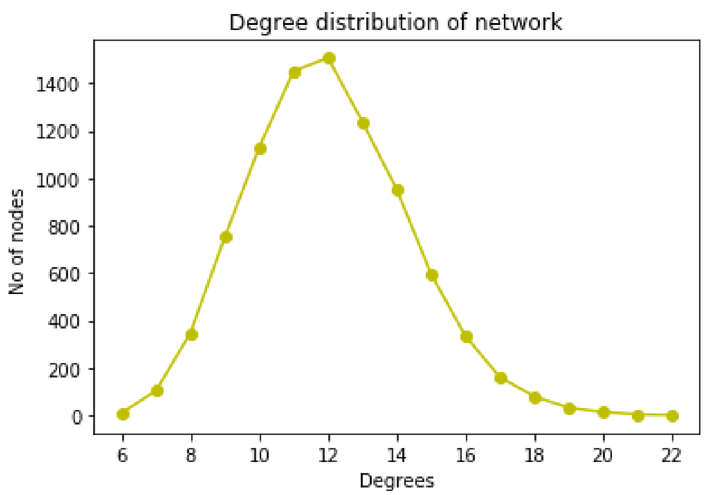


Fig 6. Degree distribution of Watts Strogatz model for dataset 3.

b) Average shortest path length : 3.9793208167294862
c) Average clustering coefficient = 0.02900195361959158

Average clustering coefficient is relatively low for dataset 3 as compared to dataset 1 and 2. Average shortest path length for dataset 1 is less as compared to dataset 2 and 3.

**C. Barabasi Albert**

Goal – Start with a small network of few nodes and understand how it grows by adding new nodes.

1. Tries to model “Prefential Attachment”
2. Likelihood of linking to a node is proportional to its current degree.
3. The greater the degree, the more edges it will get making its degree even greater.

4. Idea : Do not link new nodes uniformly and randomly but according to “Rich got Richer” scheme which is known to be heavy tailed distribution.

The function ‘barabasi_albert_graph(n,m,seed=None)’ is used to build the graph where n is number of nodes in a graph, m is number of edges to attach from a new node to existing nodes.

**Dataset 1:**
n = 7115
m = 10
a) Degree distribution:

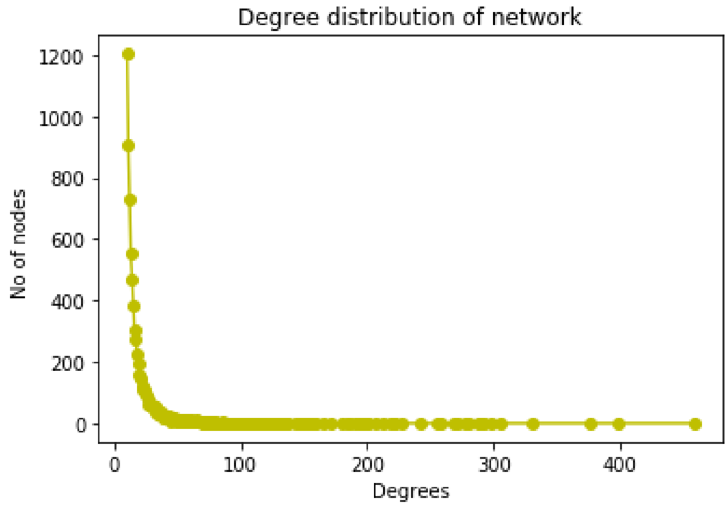


Fig 7. Degree distribution of Barabasi Albert model for dataset 1.

b) Average shortest path length : 2.986394173712678
c) Average clustering coefficient = 0.01411123151479387

**Dataset 2:**
n = 10876
m = 10
a) Degree distribution:

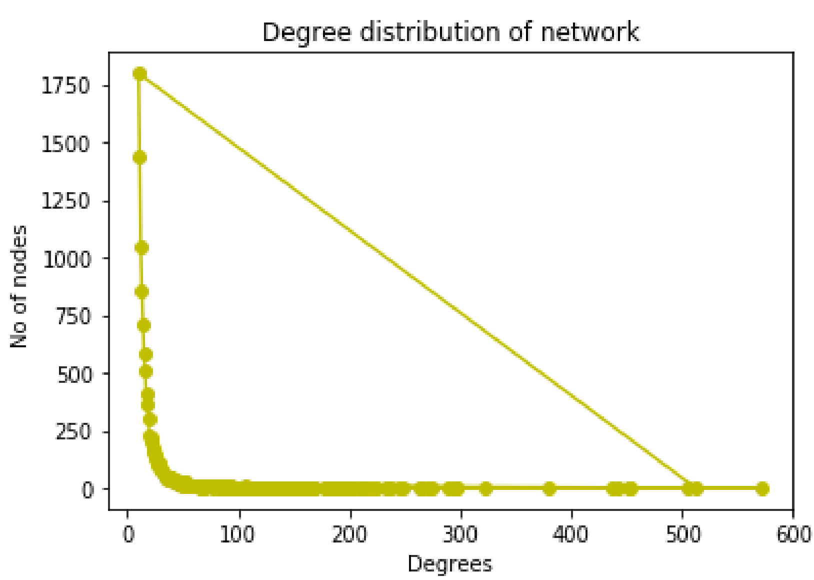


Fig 8. Degree distribution of Barabasi Albert model for dataset 2.

b) Average shortest path length : 3.0808283978643263
c) Average clustering coefficient = 0.0105220602424754

**Dataset 3:**
n = 8717
m = 10
a) Degree distribution:

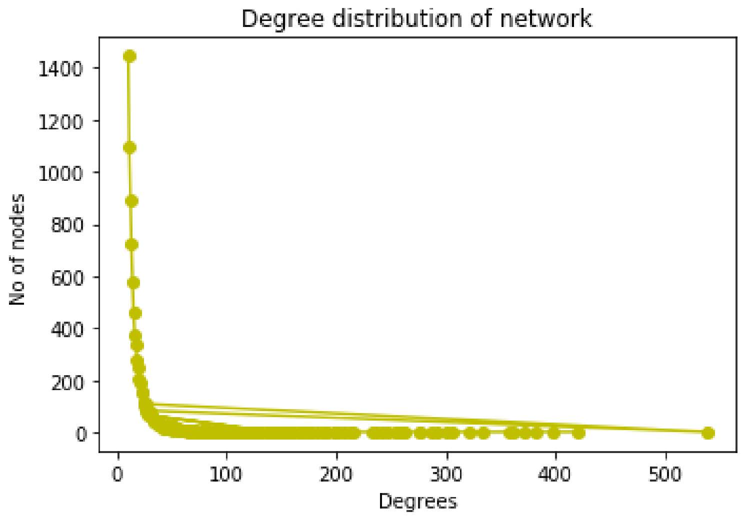


Fig 9. Degree distribution of Barabasi Albert model for dataset 3.

b) Average shortest path length : 3.0297813143629133
c) Average clustering coefficient : 0.012315596653276377

Average clustering coefficient of dataset 2 is relatively low as compared to other two dataset. Average shortest path length for dataset 1 is less as compared to dataset 2 and 3.

So, all the three models are observed with extracting the properties like average shortest path length, average clustering coefficient and degree distribution.