哈尔滨工业大学深圳研究生院

复杂网络建模报告

Social Network Analysis of Our Class

姓名胡文馨学号17S151563报告日期20171030

目录

Experimental principle		3
Model thinking		
1.	The original model	
2.		
3.	The second improved model	
4.	The third improved model	9
Network characteristics		11
1.Node-degree distribution		11
2.Average shortest path-length		12
3.Clustering coefficient		13
4.Robustness against intentional attack		14
5.R	Random attack	18

Experimental principle

Given the desired number of nodes N, the mean degree K , and a special parameter P, satisfying $0 \le P \le 1$ and $N \gg K \gg \ln N \gg 1$, the model constructs an undirected graph with N nodes and $\frac{NK}{2}$ edges in the following way:

- 1. Construct a regular ring lattice, a graph with N nodes each connected to K neighbors, K/2 on each side. That is, if the nodes are labeled $n_0 \dots n_{N-1}$, there is an edge (n_i, n_j) if and only if $0 < |i-j| \mod (N-1-\frac{K}{2}) \le \frac{K}{2}$.
- 2. For every node n_i take every egde (n_i, n_j) with i < j, and rewire it with probability P. Rewiring is done by replacing (n_i, n_j) with (n_i, n_k) where K is chosen with uniform probability from all possible values that avoid self-loops $(K \ne i)$ and link duplication (there is no edge (n_i, n_{k2}) with K2 = K at this point in the algorithm).

Model thinking

1. The original model

According to the principle of the algorithm, I try my best to build the model.

First of all, using two layers for loop nesting to establish a coupled network and the matrix to express the edge.

```
9#WS小世界模型构建
10 import random
11 #import numpy
12 from numpy import *
13 import networkx as nx
14 import matplotlib.pyplot as plt
16#每个节点都与左右相邻的各K/2节点相连,K为偶数
17 def CreateNetwork(n,k,p,matrix):
18
      for i in range(n):
          for j in range( k // 2 + 1):
19
20
              if i-j >= 1 and i+j <= (n-2):
                  matrix[i][i-j] = matrix[i][i+j] = 1
21
22
              elif i-j < 1:
                  matrix[i][(n-1)+i-j] = matrix[i][i+j] = 1
23
24
              elif i+j > (n-1):
25
                  matrix[i][i+j-(n-1)] = matrix[i][i-j] = 1
26
27
      for i in range(n):
28
          for j in range(n):
29
              print(matrix[i][j])
          print("\n")
30
```

Second, reconnecting randomly each edge. At the same time, I excluding two conclusions, one is a node connect with itself, the other is two nodes have two even more edges to connect. Through printing the matrix, I find the matrix is symmetrical. And using the test module, we can check clearly the number of changed edges and the ratio.

```
33 def SmallWorld(n,k,p,matrix):
     #随机产生一个概率p_change < p, 重新连接边
35
     p_change = 0.0
36
     edge_change = 0
37
     for i in range(n):
38
39
         \#t = int(k/2)
40
         for j in range( k // 2 + 1):
41
             #需要重新连接的过
42
             p_change = (random.randint(0,n-1)) / (double)(n)
43
             #重新连接
44
             if p_change < p:</pre>
                #随机选择一个节点,排除自身连接和重边两种情况
45
46
                while(1):
                    node_NewConnect = (random.randint(0,n-1)) + 1
47
                    if matrix[i][node_NewConnect] == 0 and node_NewConnect != i:
48
49
50
                if (i+j) <= (n-1):
51
                    matrix[i][i+j] = matrix[i+j][i] = 0
52
53
                    matrix[i][i+j-(n-1)] = matrix[i+j-(n-1)][i] = 0
54
                matrix[i][node_NewConnect] = matrix[node_NewConnect][i] = 1
55
                edge_change += 1
56
57
             else:
58
                print("no change\n",i+j)
59
        #test
        print("small world network\n")
60
        for i in range(n):
61
62
             for j in range(n):
63
                  print(matrix[i][j])
64
             print("\n")
        print("edge_change = ",edge_change)
65
        print("ratio = ",(double)(edge_change)/(n*k/2))
66
67
```

Third, I writing the data to a file to check the matrix and process the file. At that time, I intended to use R language to deal with the graph. But I find the matrix can be translate to the graph in python later. In particularly, I meet too many complex difficults due to I write the algorithm by myself and the style is different with the source code in python networkx.

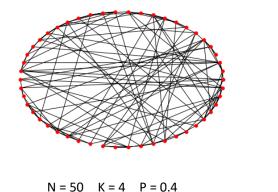
```
69#将matrix与人文件
70 def DataFile(n,k,p,matrix):
71
      # 打开一个文件
      f = open("C:/Onetwork/data.txt", "w")
72
73
      for i in range(n):
74
          for j in range(n):
              netdata = ','.join(str(matrix[i][j]))
75
76
              f.write(netdata)
77
          f.write('\n')
78
      # 关闭打开的文件
79
      f.close()
80
      print('end')
81
82
```

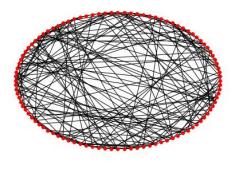
Fourth, I add the nodes and edges and export the picture.

```
84# 圃图
85 def Drawmap(n,matrix,G):
86#添加n个节点
     for i in range(n):
87
          G.add_node(i)
89
90#添加边
      for i in range(n):
91
92
         for j in range(n):
93
              if matrix[i][j] == 1:
94
                 G.add_edge(i,j)
95
     #定义一个布局,采用circular布局方式
      pos = nx.circular_layout(G)
97
98
      #绘制图形
99
100
      nx.draw(G,pos,with_labels=False,node_size = 30)
      #输出方式1: 将图像存为一个png格式的图片文件
101
      plt.savefig("WS-Network-byme.png")
102
      #输出方式2:在窗口中显示这幅图像
103
104
      plt.show()
105
```

Finally, I using the main function to achieve. In particular, it is very important to force conversion type for N,K,P and initialize matrix to zeros.

```
106#主函数
107 if __name__==" main ":
       print("main")
108
109
       #输入三个参数: 节点数N, 参数K, 概率P
       n = input("请输入节点数 n = ",)
110
       k = input("请输入参数(偶数) k = ",)
111
       p = input("请输入概率 p = ",)
112
113
       n=int(n)
114
       k=int(k)
115
       p=float(p)
       matrix = zeros((n,n),int)
116
117
       G = nx.Graph()
118
119
       value = [n,k,p]
120
121
       CreateNetwork(n,k,p,matrix)
122
       SmallWorld(n,k,p,matrix)
123
124
       print(matrix)
125
       #导出到一个文件中
       DataFile(n,k,p,matrix)
126
       #画图
127
       Drawmap(n,matrix,G)
128
```





N = 100 K = 4 P = 0.4

2. The first improved model

The first improved model is to change the initial connecting way. Hypothesis you take part in a birthday party, everyone must know the host and they maybe do not know other people. Extremly everone only know the host, so the host is the social center.

```
14

15 #假设参加聚会,每个人只认识一个主角

16 def CreateNetwork(n,k,p,matrix):

17          i = 1

18          for j in range(n):

19          matrix[1][j] = 1

20          matrix[j][1] = 1

21

N = 50     K = 4     P = 0.4

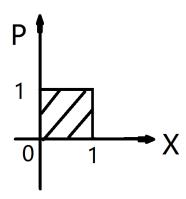
Edge change = 65

change ratio = 0.65
```

As the pictures show, if every node connects to a node except itself, the other node will have many chances to know more than one person after random reconnection.

3. The second improved model

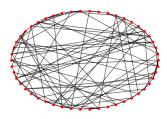
The second improved model is to change the fixed probability. If p=0, then the network is completely regular. If p=1, then the network is completely random. However, in the reality, p is a complex and fluctuating between the random and the regular. So I define p is the number of linear changes P=A*X+B. Setting x in the range of [0,1], p also in the range of [0,1]. I conclude that A determines the change speed in randomness and B determines the size of the initial randomness.



According to the principle, I simulate two situations. The first situation is on the basis of the original randomized reconnection, the second situation is randomized reconnection three times respectively.

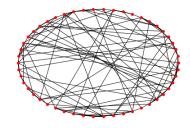
```
27#第一次随机化重连
28 def SmallWorld_1(n,k,x,matrix):
     #随机产生一个概率p change, 如果p change < p, 重新连接边
30
     A = 1.0
31
     B = 0.0
32
     p = A * x + B
33
 55#第二次随机化重连
 56 def SmallWorld_2(n,k,x,matrix):
      #随机产生一个概率p_change,如果p_change < p,重新连接边
 57
      A = 1.5
 58
 59
      B = 0.0
      p = A * x + B
 60
 83#第三次随机化重连
 84 def SmallWorld_3(n,k,x,matrix):
      #随机产生一个概率p change, 如果p change < p, 重新连接边
 85
      A = 1.5
 86
      B = 0.01
 87
      p = A * x + B
 88
```

The first situation: on the basis of the original randomized reconnection



A = 1.0

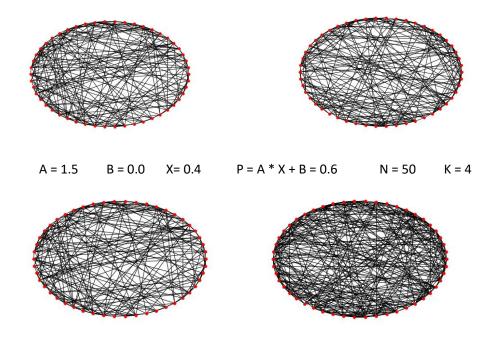




N = 50

K = 4

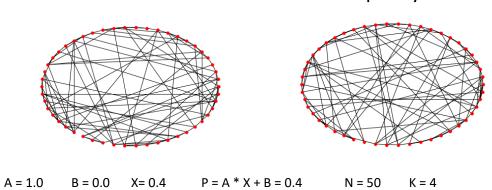
$$B = 0.0$$
 $X = 0.4$ $P = A * X + B = 0.4$



The second situation: randomized reconnection three times respectively

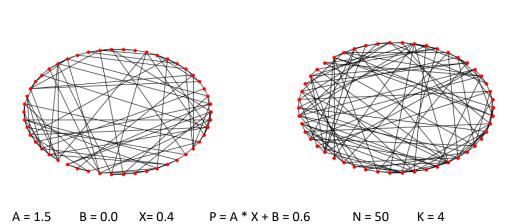
B = 0.2 X = 0.4

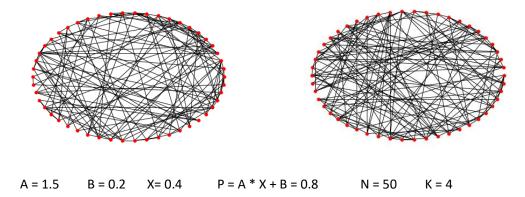
A = 1.5



P = A * X + B = 0.8 N = 50

K = 4





As the pictures show, if we random reconnect on the basis of the original, the connection will grow more and more with fast speed. If we random reconnection three times respectively, it means that we choose different p in different time. And the p is affected by A and B, A is the speed of change, B is the initial random value.

4. The third improved model

One of the constraints of the WS small world model is closed. In reality, the scale of network is constantly expanding. For example, hypothesis graduate sophomore can form a network. Graduate freshmen are newly added discrete nodes. And graduate sophomore and graduate freshmen form a new network now. At this time, the old network will affect the new network. How to deal with the hierarchical structure. I decide to change the WS model to grow. I set the half of the nodes form the old network, and the all of nodes form the new network.

```
17#选择一半的节点与左右相邻的各K/2节点相连,K为偶数
18 def CreateNetwork(n,k,p,matrix):
19
     for i in range(n//2):
20
         for j in range( k // 2 + 1):
             if i-j >= 1 and i+j <= (n//2-2):
21
                 matrix[i][i-j] = matrix[i][i+j] = 1
22
23
             elif i-j < 1:
                 matrix[i][(n//2-1)+i-j] = matrix[i][i+j] = 1
24
25
             elif i+j > (n//2-1):
                 matrix[i][i+j-(n//2-1)] = matrix[i][i-j] = 1
26
27
28#第一次随机化重连
29 def SmallWorld_1(n,k,p,matrix):
30
       n=n//2
       #随机产生一个概率p change,如果p change < p, 重新连接边
31
32
       p_{change} = 0.0
33
       edge_change = 0
       for i in range(n):
34
           for j in range( k // 2 + 1):
35
```

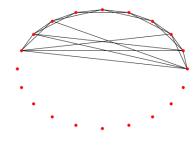
```
59#第二次随机化重连
60def SmallWorld_2(n,k,p,matrix):
61 print('samll net2 n = ',n)
```

First of all, the nodes are divided into two layers. And then according to the principle, I simulate two situations.

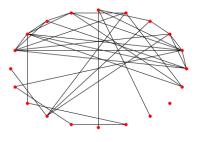
For the first situation, the first layer of the old network is a nearest neighbor coupled network. Then adding to the second layer of discrete nodes, all of nodes together with the random reconnected. So the new network is a small world network.

For the second situation, the first layer of the old network is a small world network which had been random re-connected. And then leaving the previous edges, adding to the second layer of discrete nodes and all of nodes together with the random re-connected. So the new network is also a small world network.

The first situation:

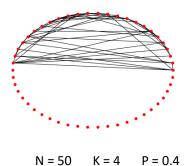


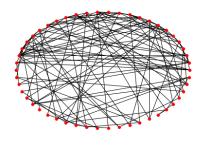
N = 25 K = 4 P = 0.4



N = 25 K = 4 P = 0.4

The second situation:





N = 50 K = 4 P = 0.4

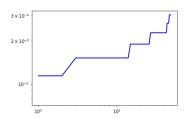
Network characteristics

1. Node-degree distribution

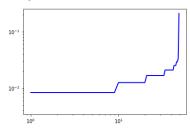
```
90#世点度分析
91def node_degree_distribution(n,matrix):
92 #求节点的度
                                                                         #生成x轴序列,从1到最大度
                                                              106
      degree = []
for i in range(n):
    sum = 0
    for j in range(n):
        sum += matrix[i][j]
    degree.append(sum)
                                                              107
                                                                         x = range(len(degree))
                                                                         #将频次转换为频率,列表内涵
                                                              108
96
97
98
99
100
101
102
103
                                                              109
                                                                         y = [z/sum_degree for z in degree]
                                                                         #在双对数坐标轴上绘制度分布曲线
                                                              110
      degree.sort()
print(degree)
                                                                         plt.loglog(x,y,color="blue",linewidth=2)
                                                              111
                                                              112
       for i in range(n):

sum_degree += degree[i]
                                                              113
                                                                         plt.show()
```

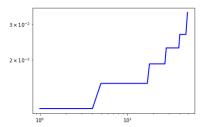
1. The original model:



2. The first improved model:

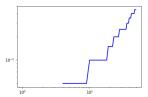


3. The second improved model:



4. The third improved model:

degree = [0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 7, 7, 7, 8, 8, 8, 8, 8, 9, 9, 9]



2. Average shortest path-length

```
54#FLoyd算法求最短路径
 55 def Ford(n, matrix):
       #出发点v,到达点w,中转点K
       #初始化新的邻接矩阵new_m,路径矩阵dis
 57
 58
       dis = zeros((n,n),int)
 59
       new_m = zeros((n,n),int)
 60
       for v in range(n):
 61
            for w in range(n):
 62
                dis[v][w] = w
 63
                if matrix[v][w] == 0:
 64
                    new_m[v][w] = 6666666
 65
                elif matrix[v][w] == 1:
 66
                     new_m[v][w] = 1
 67
                     dis[v][w] = 1
 68
 69
       for k in range(n):
 70
            for v in range(n):
                for w in range(n):
 71
 72
                     #如果经过中转点的路径比两点路径短
 73
                     if (\text{new}_m[v][k] + \text{new}_m[k][w]) < \text{new}_m[v][w]:
 74
                         new_m[v][w] = new_m[v][k] + new_m[k][w]
 75
                         \#dis[v][w] = dis[v][k]
76
                         dis[v][w] = 2
78
     #打印节点
79
     sum = 0.0
80
     for v in range(n):
81
            for w in range(v+1,n):
               print('v= ,',v,'w = ',w)
print('dis[v][w] = ',dis[v][w])
82
83
84
               sum = sum + dis[v][w]
85
               float(n)
86
     average_shortest_path_length = sum/(n*(n-1.0)/2)
88
     print('average_shortest_path_length = ',average_shortest_path_length)
1. The original model:
    average_shortest_path_length = 1.90530612245
2. The first improved model:
    average_shortest_path_length = 1.90285714286
3. The second improved model:
```

4. The third improved model:

average_shortest_path_length = 1.90612244898

3. Clustering coefficient

```
17#平均群聚系数
18 def average_clustering(n,matrix):
       number_three_tuple = 0.0
21
22
       Triangle = 0.0
23
24
       clustering_coefficient = 0.0
25
26
       for i in range(n):
27
            three_tuple = 0.0
28
           sum_edge = 0
29
          for j in range(n):
31
             if matrix[i][j] == 1 or matrix[j][i] == 1:
32
                     sum edge += 1
          float(sum_edge)
33
34
            three_tuple = int((sum_edge*(sum_edge-1.0))/2.0)
36
37
          #节点i的边组成列表myList,并且每次循环之前初始为空值
          myList = []
for j in range(i,n):
    if matrix[i][j] == 1 or matrix[j][i] == 1:
38
39
40
41
42
43
44
45
46
47
                   myList.append(j)
          #如果myList中的边(i, j)等于1,则形成三角形
          for k in range(len(myList)):
    for q in range(k,len(myList)):
                 if matrix[myList[k]][myList[q]] == 1 or matrix[myList[q]][myList[k]] == 1:
          Triangle += 1 if three_tuple != 0:
49
50
51
               clustering_coefficient += (Triangle/three_tuple)
      clustering_coefficient = clustering_coefficient/n
print('clustering_coefficient = ',clustering_coefficient)
```

1. The original model:

clustering_coefficient = 14.6541111111111

2. The first improved model:

clustering_coefficient = 36.61819727891157

3. The second improved model:

clustering_coefficient = 16.23684126984127

4. The third improved model:

clustering_coefficient = 17.105698412698413

4. Robustness against intentional attack

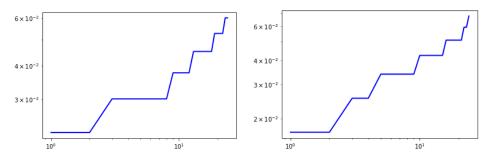
```
208 #动态行为
209#抗故意攻击 robustness against intentional attack
210 def node_robustness(n):
211
       #求出度最大的点
       degree = []
212
213
       for i in range(n):
           sum = 0
214
           for j in range(n):
215
216
               sum += matrix[i][j]
217
           degree.append(sum)
      #将度最大的点删除边
218
219
       node_flag = degree.index(max(degree))
       for i in range(n):
220
221
          matrix[node_flag][i] = 0
222
          matrix[i][node_flag] = 0
```

1. The original model:



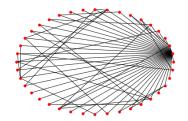


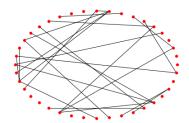
```
clustering_coefficient = 7.104952380952382
average_shortest_path_length = 1.80333333333
degree = [3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8]
```

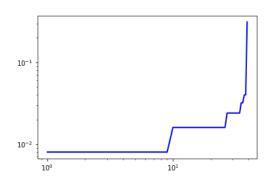


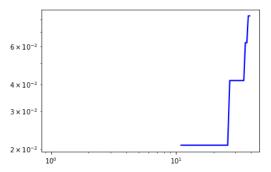
```
clustering_coefficient = 7.199238095238095
average_shortest_path_length = 2.68666666667
degree = [0, 2, 2, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 7, 7, 8]
```

2. The first improved model:





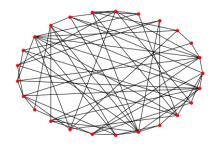


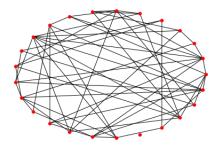


clustering_coefficient = 0.0
average_shortest_path_length = 23.1282051282

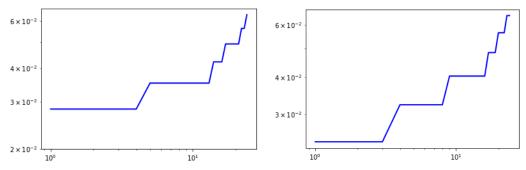
3. The second improved model:

The first situation:





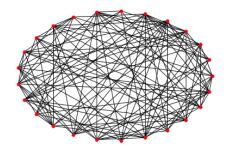
clustering_coefficient = 4.7184444444445 average_shortest_path_length = 1.77333333333 degree = [3, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 7, 7, 7, 7, 7, 8, 8, 9]

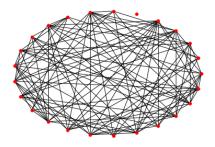


clustering_coefficient = 5.230476190476192 average_shortest_path_length = 3.27666666667

degree = [0, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 7, 7, 7, 8, 8]

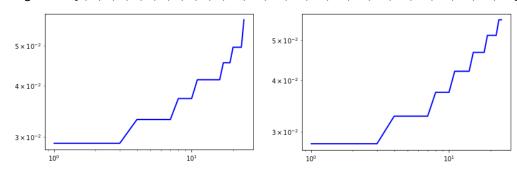
The second situation:





clustering_coefficient = 3.3337595737595738
average_shortest_path_length = 1.59666666667

degree = [7, 7, 7, 7, 8, 8, 8, 8, 9, 9, 9, 10, 10, 10, 10, 10, 11, 11, 11, 12, 12, 12, 12, 14]

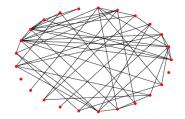


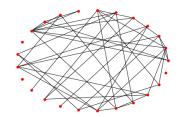
clustering_coefficient = 3.1082799422799416 average_shortest_path_length = 2.51666666667

degree = [0, 6, 6, 6, 7, 7, 7, 7, 8, 8, 8, 9, 9, 9, 10, 10, 10, 10, 11, 11, 11, 11, 12, 12]

4. The third improved model:

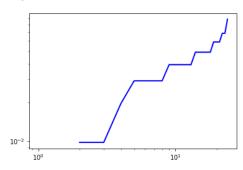
The first situation:

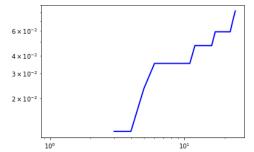




clustering_coefficient = 5.927873015873016 average_shortest_path_length = 4.66333333333

degree = [0, 0, 1, 1, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 7, 7, 9]





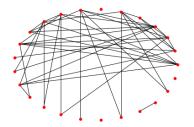
clustering_coefficient = 4.927301587301587

 $average_shortest_path_length = 5.6$

degree = [0, 0, 0, 1, 1, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 6, 7]

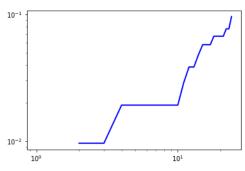
The second situation:

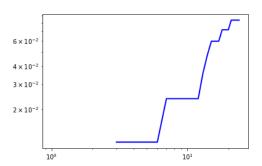




clustering_coefficient = 15.26015873015873
average_shortest_path_length = 4.92333333333

degree = [0, 0, 1, 1, 2, 2, 2, 2, 2, 2, 2, 3, 4, 4, 5, 6, 6, 6, 7, 7, 7, 7, 8, 8, 10]

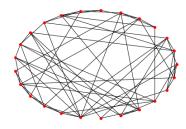




clustering_coefficient = 9.709142857142858 average_shortest_path_length = 8.32666666667

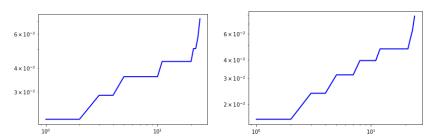
5.Random attack

1. The original model:



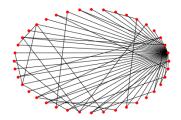


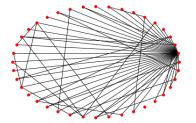
clustering_coefficient = 6.540825396825397 average_shortest_path_length = 1.78333333333 degree = [3, 3, 3, 4, 4, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 7, 7, 8, 10]

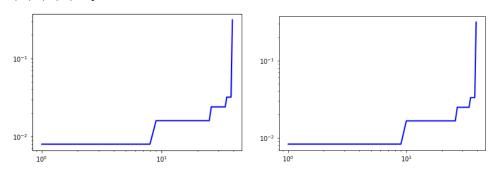


clustering_coefficient = 9.18336507936508 average_shortest_path_length = 2.64333333333 degree = [0, 2, 2, 3, 3, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 7, 8, 10]

2. The first improved model:





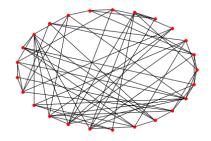


clustering_coefficient = 28.51876481744903

average_shortest_path_length = 3.54230769231

3. The second improved model:

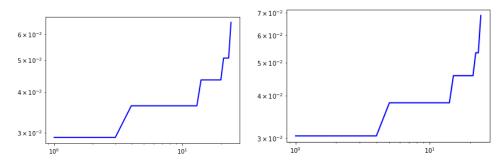
The first situation:





average_shortest_path_length = 1.78

degree = [4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 7, 9]

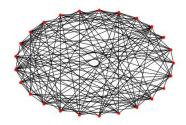


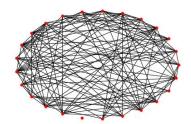
clustering_coefficient = 3.973904761904762

average_shortest_path_length = 2.64

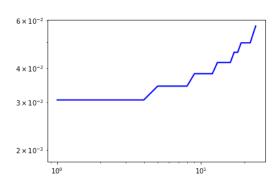
degree = [0, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6, 7, 7, 9]

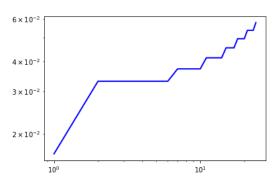
The second situation:





clustering_coefficient = 3.704933066933067 average_shortest_path_length = 1.56333333333 degree = [5, 8, 8, 8, 8, 9, 9, 9, 9, 10, 10, 10, 10, 11, 11, 11, 11, 12, 12, 13, 13, 13, 13, 14, 15]





clustering_coefficient = 3.5414398934398936 average_shortest_path_length = 2.94666666667 degree = [0, 4, 8, 8, 8, 8, 8, 9, 9, 9, 9, 10, 10, 10, 10, 11, 11, 11, 12, 12, 12, 13, 13, 13, 14]

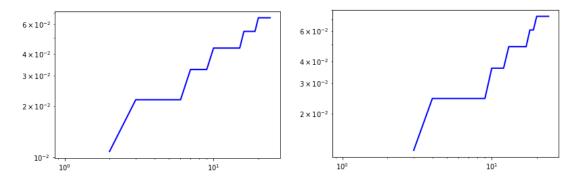
4. The third improved model:

The first situation:



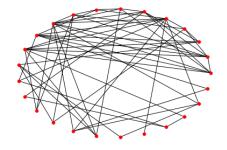


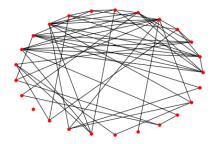
clustering_coefficient = 8.981142857142858 average_shortest_path_length = 4.63666666667 degree = [0, 0, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6]



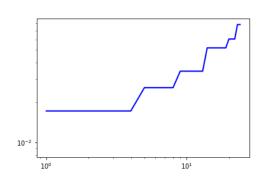
clustering_coefficient = 10.160571428571428 average_shortest_path_length = 5.60333333333 degree = [0, 0, 0, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 6, 6, 6, 6, 6]

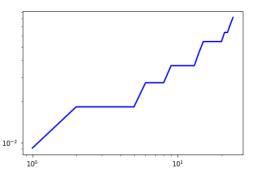
The second situation:





clustering_coefficient = 12.889428571428573 average_shortest_path_length = 1.80666666667 degree = [1, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 6, 6, 6, 6, 6, 6, 7, 7, 7, 9, 9]





clustering_coefficient = 12.09873015873016 average_shortest_path_length = 3.00666666667 degree = [0, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 6, 6, 6, 6, 6, 6, 7, 7, 8, 9]