

EE 232E
Graphs and Network Flows
Homework 2
Winter 2016

Liqiang Yu, Rongjing Bai, Yunwen Zhu
904592975, 204587519, 104593417

04-20-2016

Contents

1	Problem1	3
1.1	Part a	3
1.2	Part b	3
1.3	Part c	4
1.4	Part d	5
1.5	Part e	7
2	Problem2	9
2.1	Part a	9
2.2	Part b	9
2.3	Part c	10
2.4	Part d	10
2.5	Part e	13

1 Problem1

In this part, we do random walk on random network and try to analyze the relationship between random walk and network structure.

1.1 Part a

We create undirected random networks with 1000 nodes, the probability of the existence of a edge between any two nodes p is 0.01 by using the `random.graph.game()` function easily.

1.2 Part b

We simulate the random walk process by assigning random number to neighborhood of certain node. And we choose 1000 random walkers, and 100 as the maximum steps in order to get a better simulation result. For each one random walkers, the start node is randomly choosen. In order to get the distance from each starting point at step t , we find the shortest paths between them and measure it. The average distance $\langle s(t) \rangle$, and standard deviation $\sigma^2(t) = \langle (s(t) - \langle s(t) \rangle)^2 \rangle$ v.s. t is plotted in Figure 1 and 2 respectively.

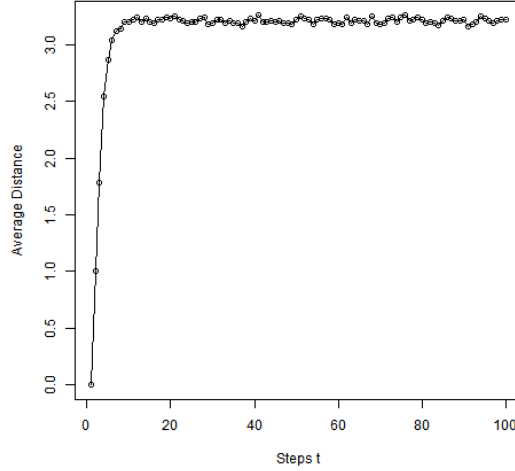


Figure 1: The average distance for random network with 1000 nodes

From the figures above, we can see that the average distance is around 3 and the

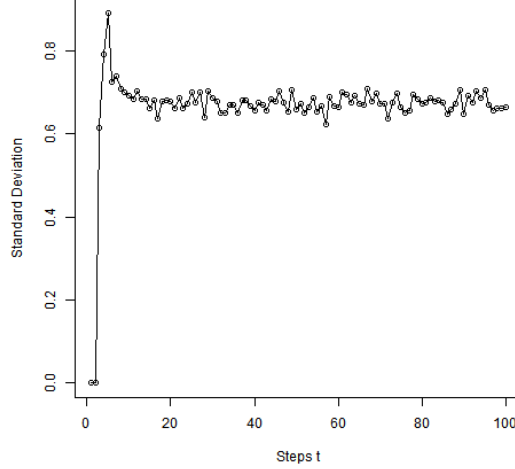


Figure 2: The standard deviation for random network with 1000 nodes

standard deviation is about 0.7. As step t increases, these two values converge to the fixed value.

1.3 Part c

In theory, values of a random walker in d dimensional are $\langle s(t) \rangle = 0$ and $\sqrt{\langle s^2(t) \rangle} = \alpha \sqrt{t}$. However, the results in our random network are totally different from the theoretical values. The standard deviation does not show any linear relationship with the number of steps t and average distance is not equal to zero. In fact, what we get in the random network, is that both the standard deviation and average distance converge to a constant.

As for the different result of average distance, this is due to the mechanism of our random walk algorithm: the distance of first step is adding one, after that, the minimum value of distance that the walker can ever achieve is zero (due to the fact that we cannot have signed distance). As the distance is always nonnegative, the average value cannot reach zero.

As for the different result of standard deviation, this is due to the structure of random network. We can see that the diameter is rather small in our random network which possible could make the nodes' distribution relatively separate. The limited number of edges can not guarantee the walker can randomly walk in a widely distributed pattern. The result we get seems that at first $\sqrt{\langle s^2(t) \rangle} = \alpha \sqrt{t}$, but later the random walker fell into a cluster and began walking in the local region.

1.4 Part d

When the number of nodes n is set to be 100, $\sigma^2(t) = \langle (s(t) - \langle s(t) \rangle)^2 \rangle$ v.s. t is plotted in Figure 3 and 4 respectively. And we choose 100 random walkers, and 100 as the maximum steps in order to get a better simulation result.

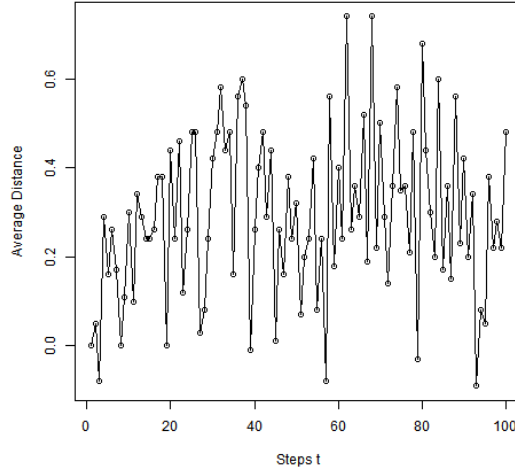


Figure 3: The average distance for random network with 100 nodes

When the number of nodes n is set to be 10000, $\sigma^2(t) = \langle (s(t) - \langle s(t) \rangle)^2 \rangle$ v.s. t is plotted in Figure 5 and 6 respectively. And we choose 1000 random walkers, and 100 as the maximum steps in order to get a better simulation result.

The diameter of these three network is shown in Table 1. Comparing with

Table 1: diameter of random network			
	n=100	n = 1000	n = 10000
diameter	9	6	3

Figure 1 and 2, we can see the trend of average distance and standard deviation in random network with 1000 and 10000 nodes are similar that both converge to constants. However, as the number of nodes increases (size of network increases), these two convergence values become smaller. This is because when the network becomes larger, the walker are freer to randomly walk in a relatively widely distributed pattern. Thus the average distance and the standard deviation will be more closer to the theoretical values.

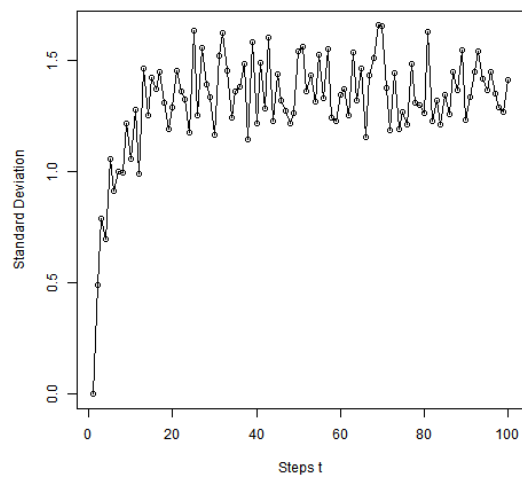


Figure 4: The standard deviation for random network with 100 nodes

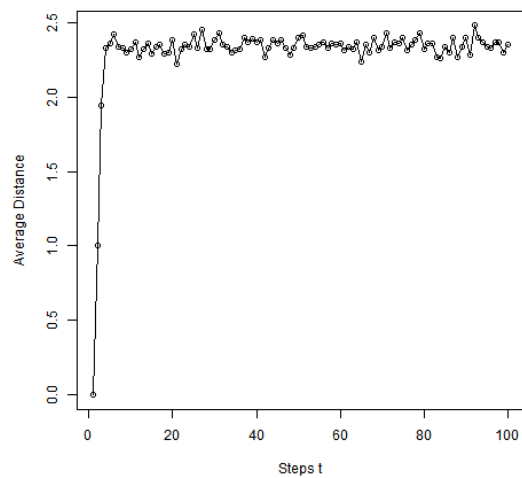


Figure 5: The average distance for random network with 10000 nodes

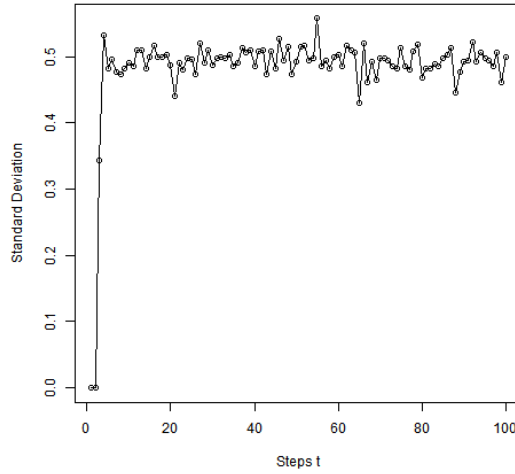


Figure 6: The standard deviation for random network with 10000 nodes

Note that the diameter value for these networks decreases with smaller average distance and standard deviation as the number of nodes increase. Furthermore, the average distance can never exceed the value of its diameter. Thus, to some extent, the diameter plays a role in the properties of network.

However, the random network with 100 nodes does not follow the same trend like the two other networks. This is possibly due to the fact that network with 100 nodes is disconnected and diameter is not applicable for disconnected graph.

1.5 Part e

The degree distribution at the end of random walk is shown in Figure 7 and the degree distribution of the graph is shown in Figure 8.

We can see from the figures above that the degree distribution at the end of the random walk and the degree distribution of graph are generally resembles each other. This similarity between two distributions means that nodes the walker went through are randomized, thus indicates the reliability of our random walk algorithm.

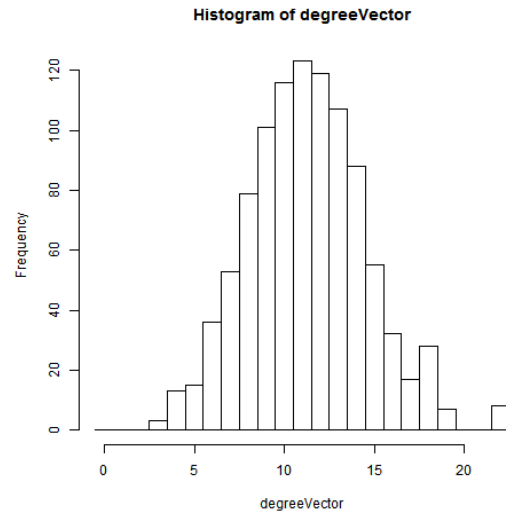


Figure 7: The degree distribution at the end of random walk

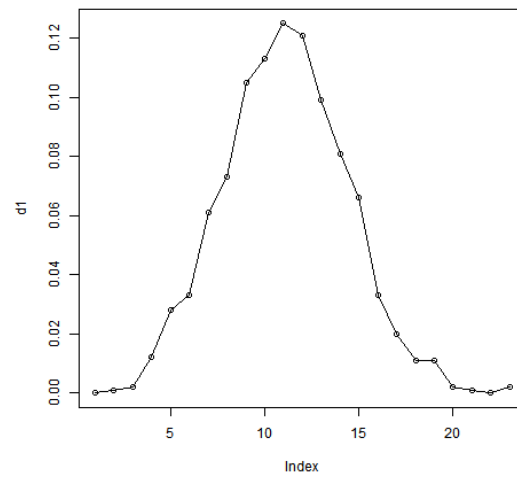


Figure 8: the degree distribution of the graph

2 Problem2

In this part, we do random walk on networks with fat-tailed degree distribution and try to analyze the relationship between random walk and network structure.

2.1 Part a

We create a fat-tailed network with 1000 nodes and degree distribution proportional to x^{-3} by using the `barabasi.game()` function easily.

2.2 Part b

Similarly, we simulate the random walk process by assigning random number to neighborhood of certain node. And we choose 1000 random walkers, and 100 as the maximum steps in order to get a better simulation result. For each one random walkers, the start node is randomly choosen. In order to get the distance from each starting point at step t , we find the shortest paths between them and measure it. The average distance $\langle s(t) \rangle$, and standard deviation $\sigma^2(t) = \langle (s(t) - \langle s(t) \rangle)^2 \rangle$ v.s. t is plotted in Figure 9 and 10 respectively.

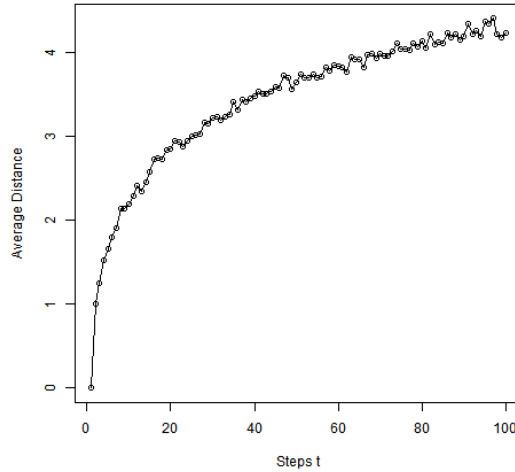


Figure 9: The average distance for fat-tailed network with 1000 nodes

From the figures above, we can see that the average distance and the standard deviation both grow proportional to the step t .

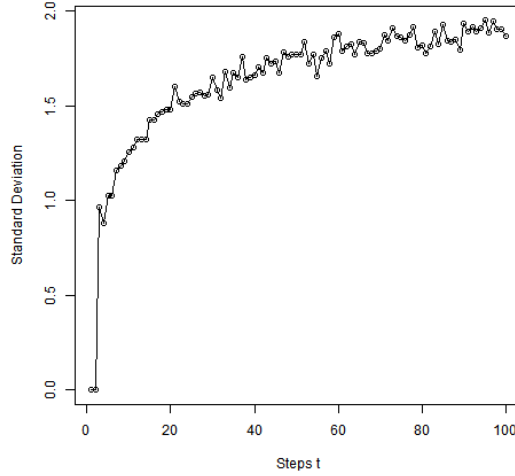


Figure 10: The standard deviation for fat-tailed network with 1000 nodes

2.3 Part c

The results of average distance and standard deviation is different from the random network in problem 1. In the fat-tailed network, the degree distribution of the nodes is more tightly bounded and densely located in a small range, so the degree of each nodes is more likely resemble. Thus, with the incrementation of the step t , the walker will be more likely to walk in a remote area. In this case, the average distance and the the standard deviation will be proportional to the t .

However, in the previous random network, the node is more sparsely distributed and the walk may be more likely to be trapped in a certain area, which could be the reason the the convergence of these two indicators.

The reason why the average distance still differs the theoretical value zero is due to the mechanism of our random walk algorithm. The minimum value of distance that the walker can ever achieve is zero (due to the fact that we cannot have signed distance). As the distance is always nonnegative, the average value cannot reach zero.

2.4 Part d

When the number of nodes n is set to be 100, $\sigma^2(t) = \langle (s(t) - \langle s(t) \rangle)^2 \rangle$ v.s. t is plotted in Figure 11 and 12 respectively. And we choose 100 random walkers, and 100 as the maximum steps in order to get a better simulation result.

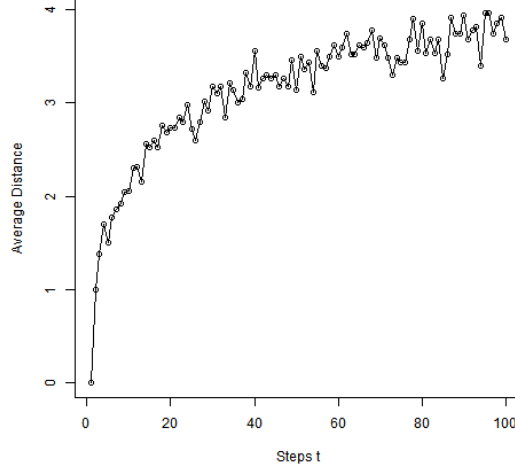


Figure 11: The average distance for fat-tailed network with 100 nodes

When the number of nodes n is set to be 10000, $\sigma^2(t) = \langle (s(t) - \langle s(t) \rangle)^2 \rangle$ v.s. t is plotted in Figure 13 and 14 respectively. And we choose 10000 random walkers, and 100 as the maximum steps in order to get a better simulation result.

The diameter of these three networks is shown in Table 2. Comparing with

Table 2: diameter of fat-tailed network

	n=100	n = 1000	n = 10000
diameter	11	19	27

Figure 9 and 10, we can see the trend of average distance and standard deviation in all three networks with 100, 1000 and 10000 nodes are similar that both proportional to t . Unlike 1(d), here 100 nodes fat-tailed network is connected, thus its result is similar to the other two and we can analyze these three results all together.

Since every network is connected, the random walker can travel to any node in the network, then diameter will not play a role. Thus, the average distances

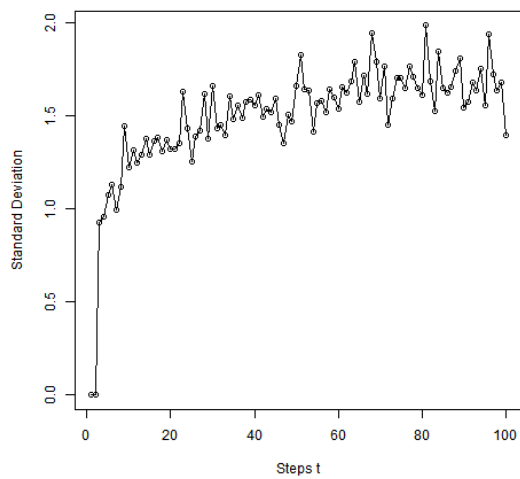


Figure 12: The standard deviation of average distance for fat-tailed with 100 nodes

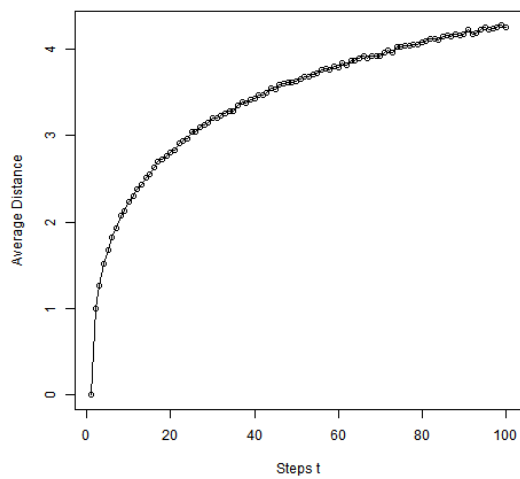


Figure 13: The average distance for fat-tailed network with 10000 nodes

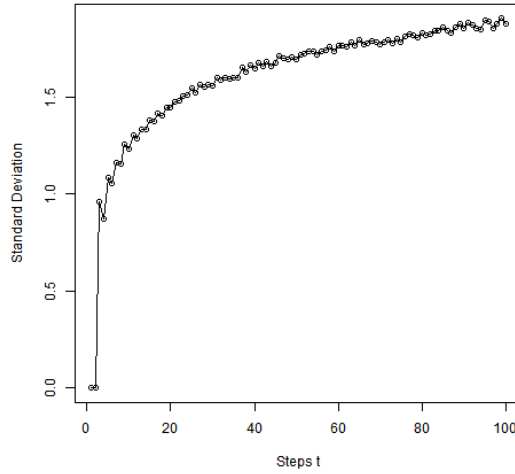


Figure 14: The standard deviation of average distance for fat-tailed network with 10000 nodes

among the three cases are nearly the same. However, the standard deviation in the case of 100 nodes is a bit larger than other two cases. This is possibly due to the limited size in that case. If the network size is too small, walker are relatively easier to get trapped in certain subregion, thus the walker behavior becomes more uncertain.

2.5 Part e

The degree distribution at the end of random walk is shown in Figure 15 and the degree distribution of the graph is shown in Figure 16.

We can see from the figures above that the degree distribution at the end of the random walk and the degree distribution of graph are generally resembles each other. This similarity between two distributions means that nodes the walker went through are randomized, thus indicates the reliability of our random walk algorithm.

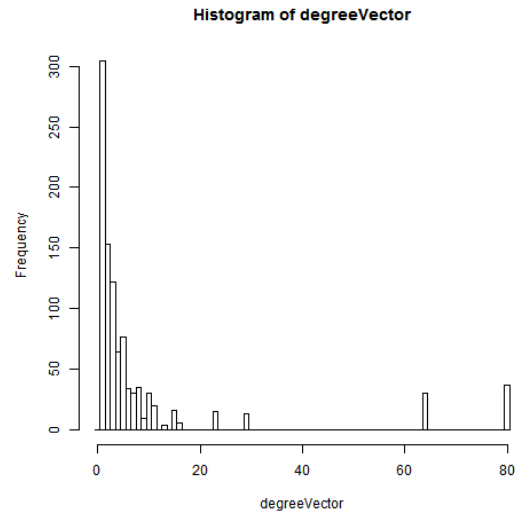


Figure 15: The degree distribution at the end of fat-tailed walk

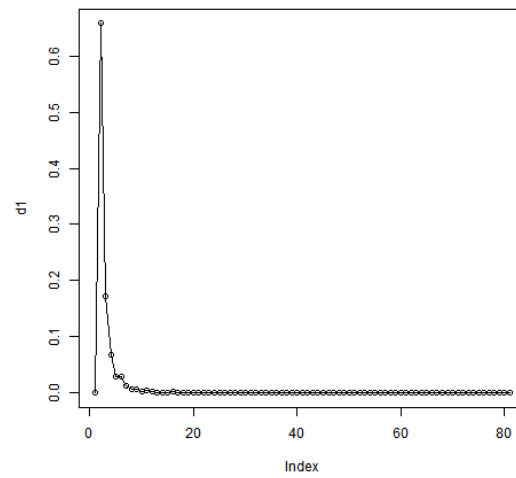


Figure 16: the degree distribution of the graph