Six degrees of separation

report

Sönke Beier

April 5, 2024

1 Introduction

The idea of six degrees of seperation says that everyone in the world is connected through a chain of six or fewer people. This means people are more closely connected than we might think.
In this project we translate the idea into a relation network, where different numbers of nodes are connected by partly randomly choosen edges. This type of networks can be found as "small world networks" in the literature (see [WS98]). The tasks of the report are adopted from [Set22].

2 Results and Discussion

2.1 Constructing a small world network

The task was to construct a network with L nodes, where each node is connected with Z neighbournodes. In addition, there are $\lfloor pLZ/2 \rfloor$ randomly selected shortcuts. If $pLZ/2$ is an integer p give us the ratio between the number of random shortcuts and the number of edges to neighbournodes.
In figure 2.1 we can see a network which was created by this construction rules.

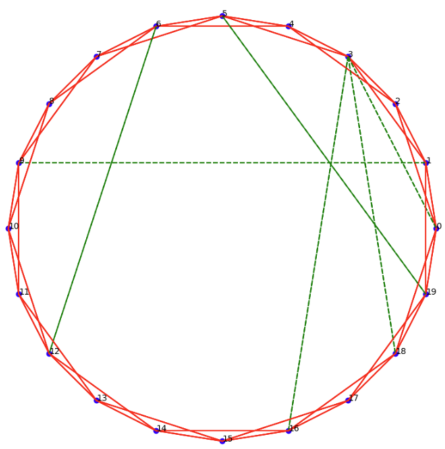
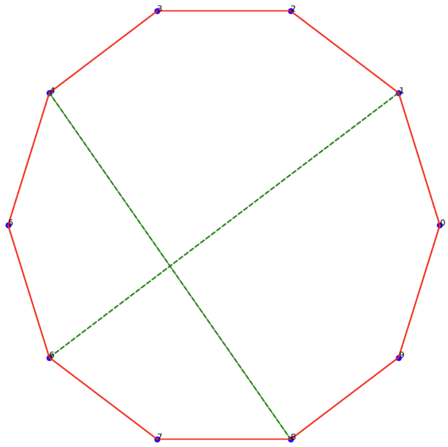


Figure 1: A small world network with $L = 20$, $Z = 4$, $p = 0.2$. There are 8 random shortcuts (green) and 40 edges to neighbournodes (red).

2.2 Measuring the minimum distances between nodes

An often considered property of small world problems is the distribution of the shortest paths between the nodes. To create a list of shortest paths from one node to other nodes we used the breadth-first algorithm. An example result of the code can be seen in figure 2. Summarized, the breadth first algorithm works as follows:

- Choose the starting node and put it into a queue and mark it as visited.
- Look at the first entry of the queue. For each unvisited neighbornode, mark it as visited and enqueue it. Save the number of required steps.
- Repeat the last step until the queue is empty.



(a) A small world network with $L = 10$, $Z = 2$, $p = 0.2$. Numbering starts counter-clockwise from the most right node.

	node1	node2	path_length	shortest_path
0	0	0	0.0	[0]
1	0	1	1.0	[0, 1]
2	0	2	2.0	[0, 1, 2]
3	0	3	3.0	[0, 1, 2, 3]
4	0	4	3.0	[0, 9, 8, 4]
5	0	5	3.0	[0, 1, 6, 5]
6	0	6	2.0	[0, 1, 6]
7	0	7	3.0	[0, 1, 6, 7]
8	0	8	2.0	[0, 9, 8]
9	0	9	1.0	[0, 9]

(b) Table of all shortest paths from node1 = 0 to all other nodes for the given example.

Figure 2: Example for computing the shortest paths.

In figure 3 the path lengths $l(p)$ at $p = 0$ are plottet to test the „FindAllPathLengths“method. It is constant for $0 < l < L/Z$ as expected and indicates that the method is functioning correctly.

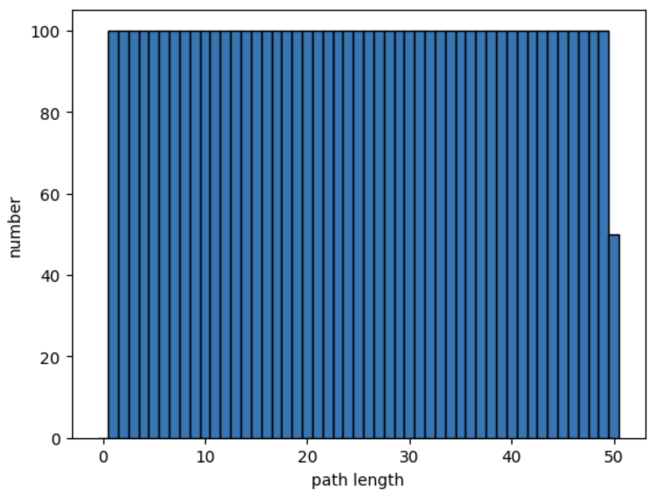
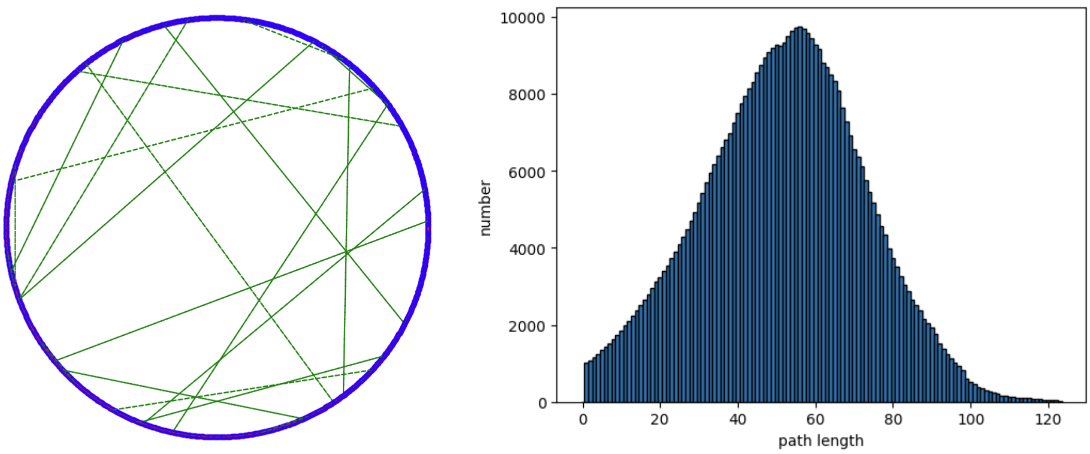


Figure 3: Testing the FindAllPathLengthts function: Histogram of path lengthts at $p = 0$, $Z = 2$ and $L = 100$. As expected it is constant for $0 < l < L/Z$.

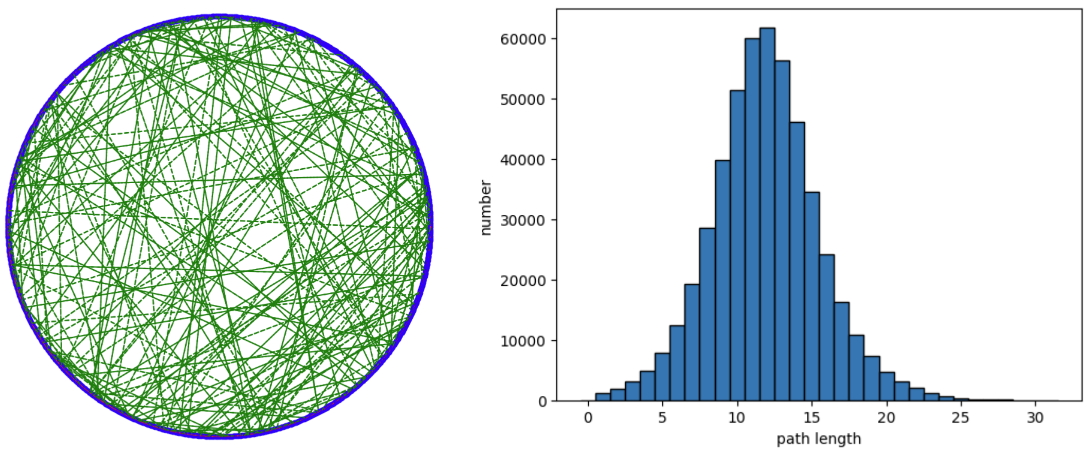
From the path length histograms we can see, that small worlds with only a few shortcuts have a broad distrution, however small worlds with many shortcuts have a narrow distribution.



(a)

(b)

Figure 4: Graph with $L = 1000$, $Z = 2$, $p = 0.02$ with the corresponding histogram of path lengths.

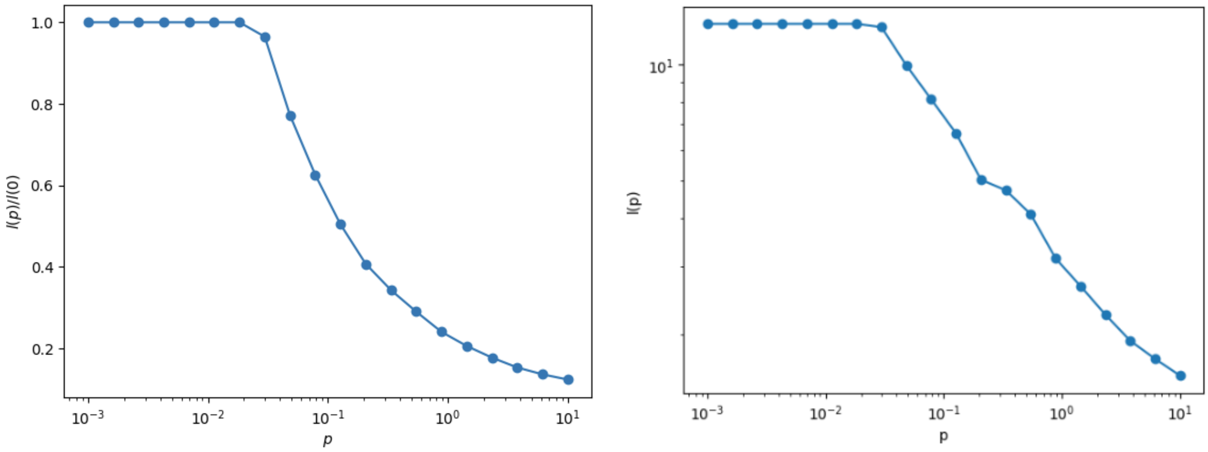


(a)

(b)

Figure 5: Graph with $L = 1000$, $Z = 2$, $p = 0.2$ with the corresponding histogram of path lengths.

In figure 6 we can see the average path length $l(p)$ divided by $l(0)$. It is fixet at one for small p, because p have to be big enought, that the number of random shortcuts $pLZ/2$ is bigger then one. Apart from this we can see from the log-log plot that it follows a power law.

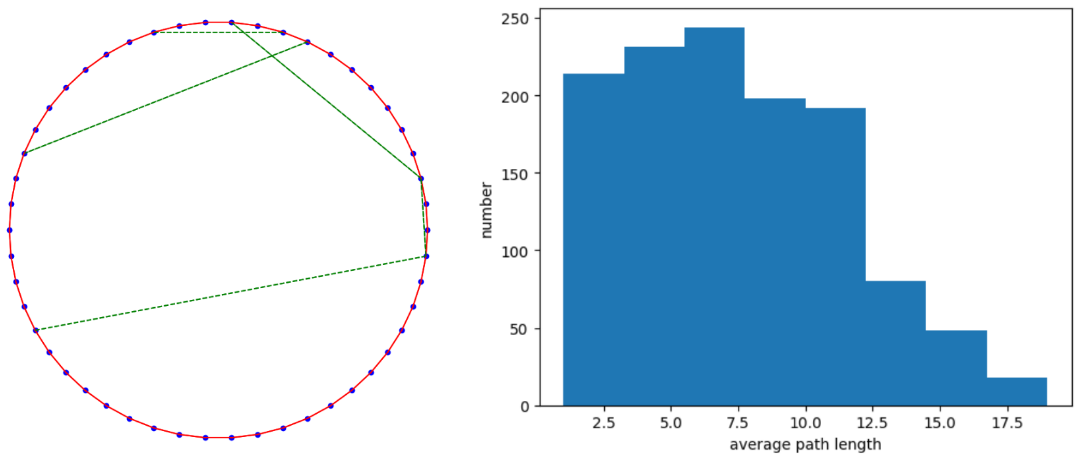


(a) (b)

Figure 6: Average path length for $Z = 2$, $L = 50$ and p from 0.001 to 1

2.3 Large numbers of nodes and statistical properties

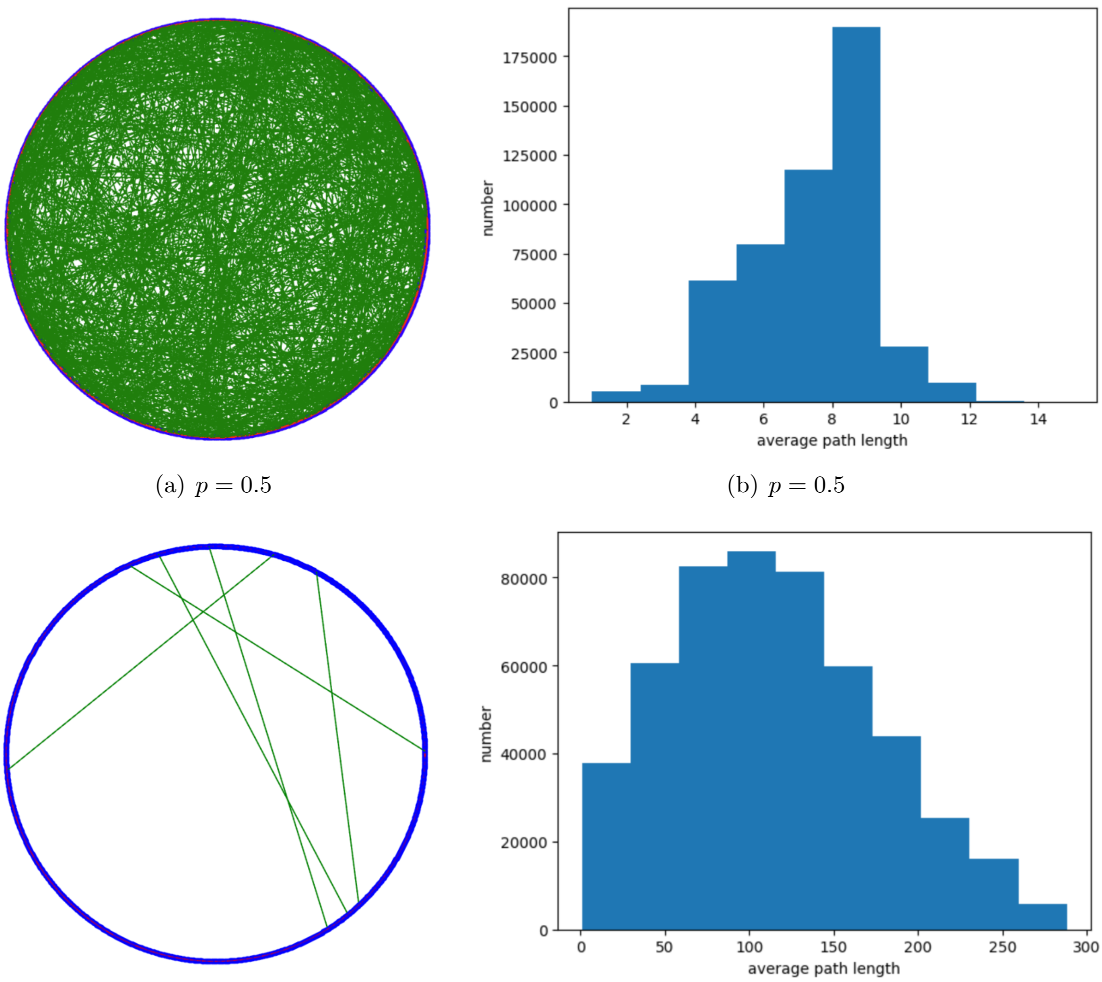
Now we want to compare the geometry of an example from figure 6 with $p = 0.1$ (see figure 7) with the the geometry Watts and Strogatz's geometry (see figure 8). [1]
With the histograms of the average path length we can compare the graphs in a statistical way. Statistically the graph of figure 7 looks more similar to the graph with $p = 0.005$ of figure 8. Both average path length distributions are much broader compared to the figure with $p = 0.5$. That similarity emerges because they have the same number of shortcuts.



(a) (b)

Figure 7: Graph and histogram of the average path length with $Z = 2$, $L = 50$ and $p = 0.1$

[1] Here we deviate slightly from the task as the calculation time was too long. We use a different number of neighbours and different p. The number of shortcuts stays the same.



(a) $p = 0.5$

(b) $p = 0.5$

(c) $p = 0.005$

(d) $p = 0.005$

Figure 8: Circle graphs of Watts and Strogatz's geomety with $Z = 2$, $L = 1000$ with the corresponding histograms of the average path length.

To compare the networks with different sizes, we can plot the rescaled average path length $\langle\theta\rangle = \pi Zl/L$ versus the total number of shortcuts $\mathbf{M} = pLZ/2$ in figure 9. The $\langle\theta\rangle$ shows a power law dependence with the total number of shortcuts. The first random shortcuts have a huge impact on the path length.

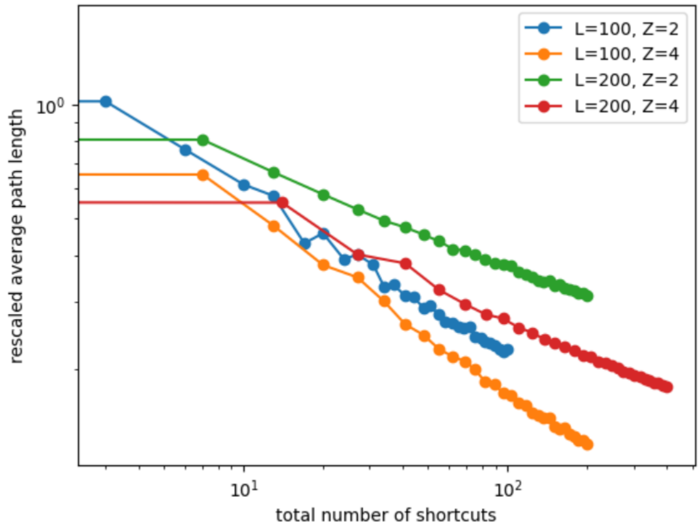
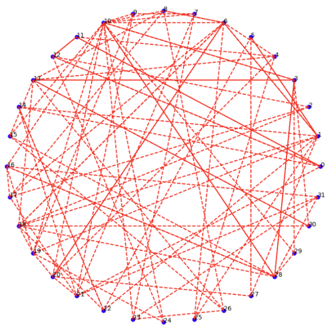


Figure 9: Rescaled average path length $\pi ZL/2$ to the number of shortcuts $pLZ/2$ for networks with different values for L and Z with p reaching von 0.001 to 1.

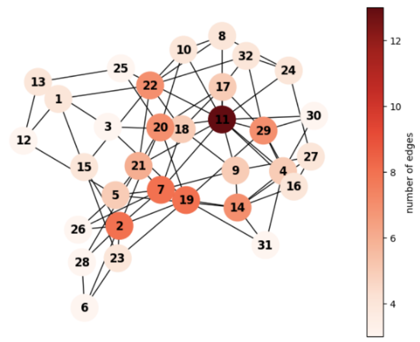
2.4 Real world network

Several real world networks are examined in [Mil+04]. These can be downloaded from https://www.weizmann.ac.il/mcb/UriAlon/download/collection-complex-networks. As a data set I used the "social network 3" mentioned there, which is based on the data from [Zel50]. Here college students were asked to choose which three students they wanted to have in a college committee. So we have a minimum of 3 edges/connections for each node/students.

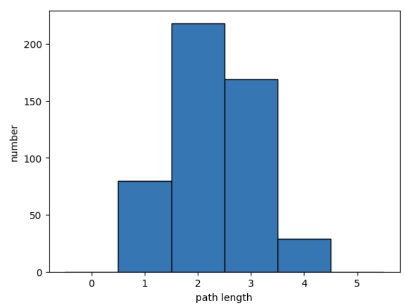
To analyze the network, I imported the network into my program structure and calculated the mean distance and the histogram of distances between nodes.



(a) Graph in circular form.



(b) Graph using a spring layout. The colourcode indicates the number of edges of one node.



(c) Histogram of distances between nodes from this example.

Figure 10: Data from [Zel50]. The nodes stand for college students. Edges indicates wether one student want another student in a college committee. The network has 32 nodes and 96 edges.

3 Code

The code for the project can be found on Github at https://github.com/SoenBeier/six_degrees_of_seperation

References

[Mil+04] Ron Milo et al. "Superfamilies of Evolved and Designed Networks". In: *Science* (2004). URL: https://www.science.org/doi/10.1126/science.1089167.

[Set22] James Sethna. *Entropy, Order Parameters, and Complexity*. Clarendon Press, 2022. URL: https://sethna.lassp.cornell.edu/StatMech/.

[WS98] Duncan Watts and Steven Strogatz. "Collectivedynamics of 'small-world' networks". In: *Nature* (1998). URL: https://www.ncbi.nlm.nih.gov/pubmed/9623998.

[Zel50] Leslie D. Zeleny. "Adaptation of Research Findings in Social Leadership to College Classroom Procedures". In: *Sociometry* 13.4 (1950), pp. 314–328. ISSN: 00380431. URL: http://www.jstor.org/stable/2785274 (visited on 12/20/2023).