A Report

“To critically discuss Stanley Milgram's experiment and assess the extent of the "small-world effect" in different network models using a road network dataset.”

[Question – 03]

BUSM132 – Complex Networks and Innovation
Group – 16

School of Business and Management
(Word count: 4321)

Prepared by:

Kaviya Thirumal [231052333]
Nava Bhargav [230270185]
Pon Ananth Veppalodai Senthurapandi [230289833]

Table of Contents

Analysis of "Small-world effect" in different network models using a road network dataset.

Abstract

The small-world experiment, conducted by Stanley Milgram suggested that it does not require more than a handful of links to connect any two people. It showed how people navigate the social space between themselves and a distant person (Schnettler,2009). In this report, we analysed the real-world road network data and its proximity in establishing small world phenomenon. Furthermore, we discussed the road network establishing Barabasi's scale free network structure and characteristics of a road network dataset to draw insights into network properties. The methodology and implications from the experiment are explained. We also focused on comparing the random Poisson network and the Barabasi-Albert scale-free network models. By examining key metrics such as degree distribution, clustering coefficient, average path length, and betweenness centrality, we aimed to understand how network structure influences information flow, connectivity, and communication in a road network dataset.

Keywords

Milgram's experiment, small world effect, Random Poisson network, Barabasi-Albert scale-free network, Watts-Strogatz small world networks, degree distribution, clustering coefficient, average path length, betweenness centrality, road network, degree, Weighted Degree, Eccentricity, Closeness centrality, Harmonic closeness centrality, Authority, hub, modularity class and triangles.

Introduction

The small world phenomenon can explain the occurrence of various social processes, and this can be analysed through contact channels of individuals acting as nodes and their networks as links, this was initiated **by Pool and Kochen in 1970's.** They worked on anecdotal experience of two strangers, meeting and finding out to share a common acquaintance work a small-world phenomenon (Schnettler,2009). This idea was further empirically developed by Stanley Milgram for tracing the connection between people.

Stanley Milgram's experiment on the small world phenomenon remains a cornerstone in understanding social networks and human connectivity. His study revealed the surprisingly short average path length between individuals, popularly known as "six degrees of separation." This paper aims to analyze the small world effect within the context of network science, focusing on three network models: Random Poisson, Barabasi-Albert scale-free and Watts-Strogatz small world networks. Most briefly on the random Poisson and scale-free.

Evolution of Network theory

Milgram's experiment in the 1960's involved participants sending packages to a target person via intermediaries, showcasing the interconnectedness of social networks. Despite its groundbreaking insights, the experiment faced ethical criticisms due to potential psychological harm inflicted on participants (Hong Zhang,2011).

Milgram utilised a letter-referral technique on first name basis. Milgram and his students asked the randomly selected starters to send messages to the pre-selected target person through letter referrals to closest acquaintance of the target and limited that letter not to be sent to the target directly unless the target is a direct acquaintance. Milgram found that letters reached the target in 5 or 6 steps, thus giving the concept of six degrees of separation in the world.

Later, **Watts and Strogatz's** pioneering work in 1999 came up with the concept of small-world networks as a transition between ordered and random networks. They proposed a model where a regular lattice, representing a highly ordered network, is gradually transformed into a random network through the rewiring of a small fraction of edges, controlled by a tenable parameter.

This transition between order and randomness revealed networks that exhibit both high clustering, akin to ordered networks, and short average geodesics, characteristic of random networks (Watts, 1999). Small-world networks, thus identified, manifest the small-world effect, where nodes are connected by relatively short paths despite the large size of the network.

In 1999, **Barabási and Albert** introduced a dynamic network model that operates based on two fundamental mechanisms: growth and preferential attachment. The networks generated by this dynamic process exhibit a distinctive property known as scale-free behaviour. In a scale-free network, a few nodes, known as "hubs," acquire a disproportionately large number of connections, while most nodes have only a few connections. In our road network data of

Chesapeake, we further discuss the scale-free properties like hierarchy of roads and preferential attachment.

Types of networks and Analysis of Road network Dataset

Before diving into the analysis of our road network dataset, it is essential to understand the characteristics of several types of networks. In this section, we discuss two prominent types of networks: random Poisson networks and Barabasi-Albert scale-free networks. The below are the types of networks:

1. Regular Networks: In a regular network, nodes are arranged in a regular pattern where each node is connected to its nearest neighbours. These networks typically have a high clustering coefficient but a long average path length. Here the nodes are connected to their nearest neighbours in a regular pattern.

2. Random Networks: Random Poisson networks are characterized by random connections between nodes. These networks exhibit a uniform distribution of connections, with each pair of nodes having an equal probability of being connected resulting in a short average path length but a low clustering coefficient. Here the connections between nodes are established randomly, resulting in a uniform distribution of connections.

3. Scale-free Networks (Barabasi-Albert scale-free network): Barabasi-Albert scale-free networks, on the other hand, are characterized by a power-law degree distribution, where a few nodes, known as hubs, have a significantly higher number of connections than most nodes. These networks arise from the principle of preferential attachment, where new nodes are more likely to connect to existing nodes with a higher degree. Despite their vulnerability to targeted attacks on highly connected nodes, Barabasi-Albert scale-free networks exhibit robustness to random failures and are prevalent in various real-world networks.

Mathematically, the **power law** is expressed using an equation.

$$\mathbf{p(d) \sim d\ (-\gamma\ )} \text{ or } \mathbf{p(d) \sim (1/d)}$$

Where the fraction **p(d)** (%) of the nodes with a certain number of connections will decrease at power rate **a** to the number **d** of the connections (hence, the power law).

In a Barabási-Albert scale-free network, new nodes are added to the network one by one, and they preferentially attach to existing nodes with high degrees, resulting in a "rich-get-richer" phenomenon. This leads to the formation of hubs, nodes with very high degrees. In such networks, the clustering coefficient can vary, but it tends to be higher compared to random networks. This is because nodes with high degrees tend to form clusters with their neighbours, increasing the local clustering coefficient in the network.

4.Small-world networks (Watts and Strogatz's): This is a combination of regular and random networks, exhibiting both local clustering and short average path lengths. Small-world networks exhibit the small-world effect, where most nodes can be reached from any other node through a relatively small number of steps. Initially, nodes are connected to nearby neighbours, forming clusters Then, with a probability p, edges are randomly rewired, creating shortcuts. This maintains local clustering while shortening path lengths, creating a "small world" effect where most nodes are reachable in a few steps. Tuning p adjusts the balance between local clustering and global connectivity. These networks model real-world systems like social networks and power grids.

The small-world effect is a phenomenon observed in small-world networks, which are characterized by high local clustering and short average path lengths. These networks exhibit both the local connectivity found in regular networks and the global efficiency seen in random networks (Goodrich and Ozhel,2022).

Degree Distribution: In small-world networks, degree distribution plays a crucial role. It describes how the number of connections (degree) of nodes is distributed throughout the network (Newman ,2018). Small-world networks typically have a degree distribution that follows a power-law distribution, meaning that a few nodes have a very high degree of connections while most nodes have relatively few connections.

Clustering co-efficient: Small-world networks have a high clustering coefficient, indicating that nodes in the network tend to cluster together. This means that nodes are more likely to be connected to each other's neighbours, forming tightly knit clusters or communities within the network. This high clustering facilitates local communication and interaction between nodes (Newman ,2018).

Average path length: Despite the high clustering, small-world networks also have a short average path length. This means that even though nodes are clustered together, most nodes can be reached from any other node through a relatively small number of steps or connections. This characteristic of small-world networks enables efficient global communication and connectivity, contributing to the small-world effect.

In summary, the small-world effect emerges from the combination of high local clustering and short average path lengths in small-world networks. These networks strike a balance between local connectivity and global efficiency, allowing for efficient communication and navigation despite the complex network structure.

Collective Dynamics of "small world" networks

Understanding the properties of these network models provides valuable insights into the underlying principles governing network structure and connectivity. Now, let's apply this knowledge to the analysis of our road network dataset. Let n be vertices and k edges, edges rewired with p as random probability. The average minimum path in the network is measured by the variable L(p), while the average neighbourhood's connectivity is measured by the variable C(p) (Watts,1998). Clustering(C(p)) is a measurement of the percentage of neighbouring nodes that are connected to one another, while L(p) is the minimum number of connections required to connect one node to another, averaged across the whole network. It is possible to conceptualize L(p) as the network's global property and C(p) as its local property (Watts,1998). .

Regular lattice --- p = 0 (densely clustered)

L>> when n>> i.e. ($L \propto n$)

Random network ---p=1(sparsely clustered)

L>>only when log n >>

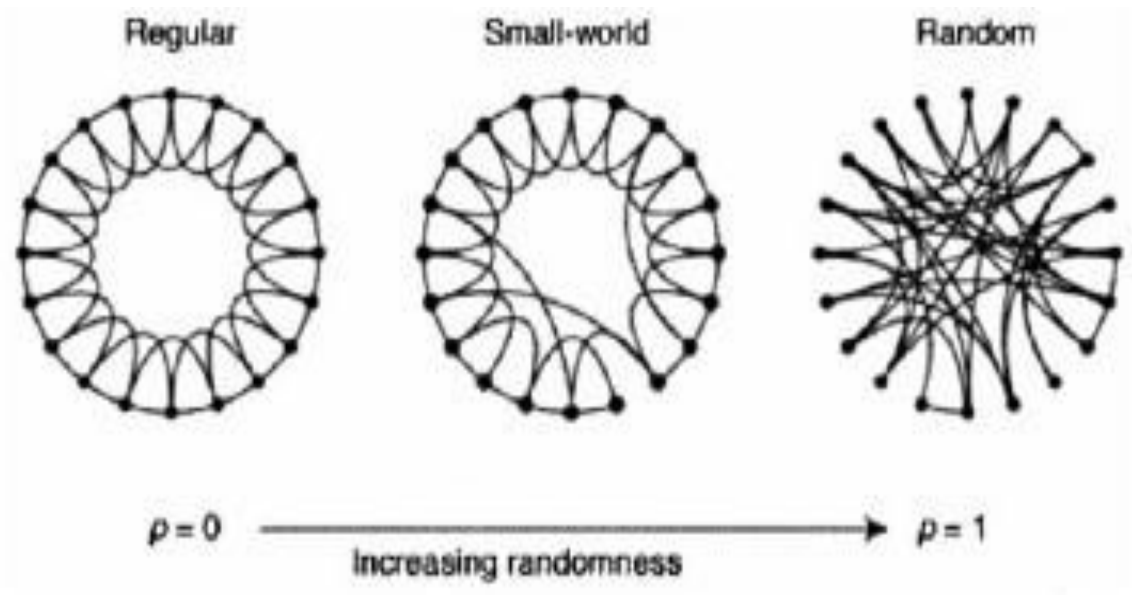

Figure: Represents increase in randomness in regular, small world and random networks

The clustering coefficient C gauges the extent to which nodes in a network form clusters. The clustering coefficient and average shortest path length of small-world networks generated by the Watts-Strogatz model as follows:

$$C = 3\,(k-2)\,/\,4\,(k-1)$$

$$L \approx n/2\,(\log n\,/\,\log k)$$

Analysis of road network dataset

In this section, we delve into the specifics of our road network dataset and explore its relevance within the context of different network models. Understanding the small-world effect in road networks is crucial for optimizing transportation efficiency, predicting traffic flow dynamics, and designing resilient infrastructure. By examining real-world road networks, we seek to uncover patterns of connectivity that contribute to both local cohesion and global efficiency. This dataset used is obtained from networkreprositery.com (Road network of Chesapeake) and we have used open-source network software Gephi in graphical analysis.

The table below provides us with some stats from the road network.

Network data statistics	
Nodes	39
Edges	169
Graph type	Undirected
Density	0.22942
Maximum degree	33
Minimum degree	3
Average degree	8.667
Avg. weighted degree	8.667
Graph density	0.228
Assortative	-0.375783
Number of triangles	582
Average number of triangles	14
Maximum number of triangles	71
Avg. clustering coefficient	0.450237
Avg. path length	1.838
Fraction of closed triangles	0.28418
Maximum k-core	7
Lower bound of maximum clique	5

Table: Represents network data statistics from the road dataset.

Insights from the dataset

The dataset comprises of 39 nodes which represent the point of interest and 169 edges which represent the roads connecting the nodes. The road dataset is unidirectional, indicating roads are bidirectional. The density of the graph is 0.22942, suggesting that approximately 22.94% of all possible road connections exist in the network. This indicates a relatively dense road network. The road network in our case has different degree for the nodes some have fewer connections, and some have more connections, which implies uniformity in the strength of connections across the network.

The road network is dense with an overall density of 0.22492. The negative assortative (-0.375783) suggests that nodes with high degrees tend to connect with nodes with lower degrees, and vice versa. The clustering in this road network is high and can be justified by the number of triangles present in a network. In our network we have 582 triangles. In our dataset node 39 has the highest degree with 33 edges which is the biggest hub/ intersection which connects to most of the nodes. The maximum number of triangles that a node is part of is 71, indicating some highly interconnected areas within the network.

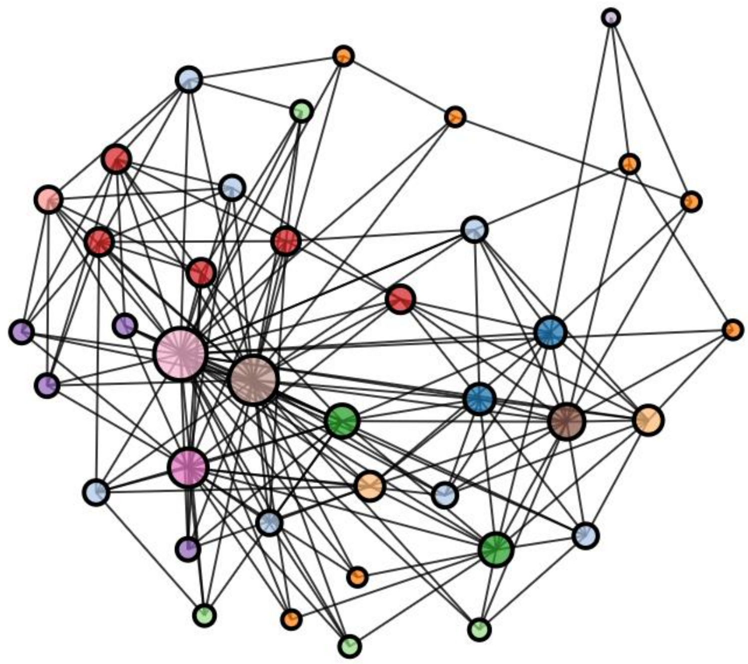

Figure: Interactive visualization of road-Chesapeake's graph structure.

The average clustering coefficient, which measures the degree to which nodes in a graph tend to cluster together, is 0.450237. This suggests a relatively high level of clustering in the network. The average path length in the network is 1.838, indicating that on average, it takes approximately 1.838 steps to travel from one node to another. This analysis provides insights into the structural properties and connectivity patterns of the road network dataset, which can be valuable for various applications such as traffic optimization, route planning, and urban planning.

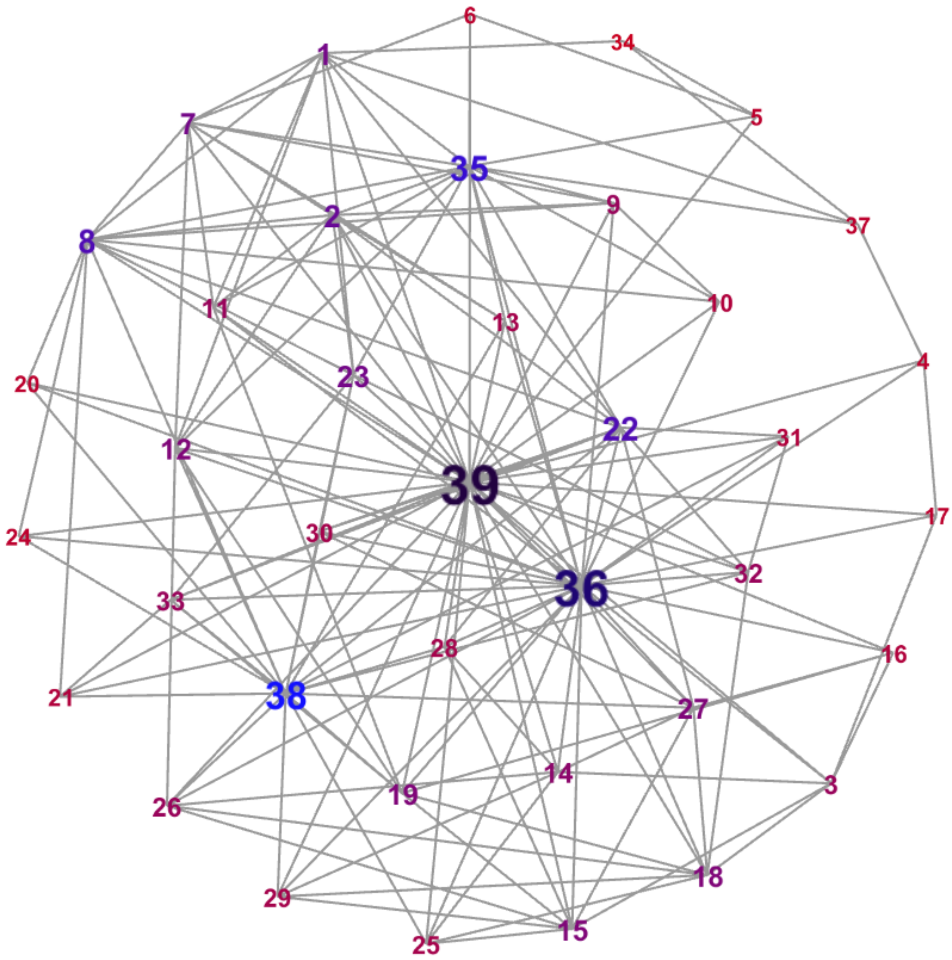

Figure: A network showing nodes, hubs, and their connectivity.

Conclusions from the data

__Random Poisson network__:

1. **Degree distribution**: The degree of distribution is more uniform, and the values range from 3 (minimum degree) to 33 (maximum degree). Thus, it is in line with the characteristics of the random Poisson network.

2. **Clustering coefficient**: The average clustering coefficient is 0.45 which is unusually high for a random network. This implies that there is a presence of localized clusters which increases the clustering coefficient. Random Poisson network has no impact with clustering coefficient.
3. **Average path length**: In a random Poisson network, with an average path length of 1.838, this value is remarkably low. Typically, random networks have longer average path lengths, closer to the diameter of the network. The diameter of a random Poisson network tends to be logarithmic in the number of nodes, so an average path length close to 2 is quite unusual.

The average path length is 1.838 which tends to be longer when compared to small world networks. It indicates less distance between the nodes which indicates that the network is highly interconnected. For our road network example, this path length indicates high level of connectivity and efficiency in terms of communication and road transport.

Barabasi Scale-free network:

1. **Degree Distribution**: The degree distribution appears to follow the power-law distribution with few highly connected nodes (hubs) and many few connected nodes.

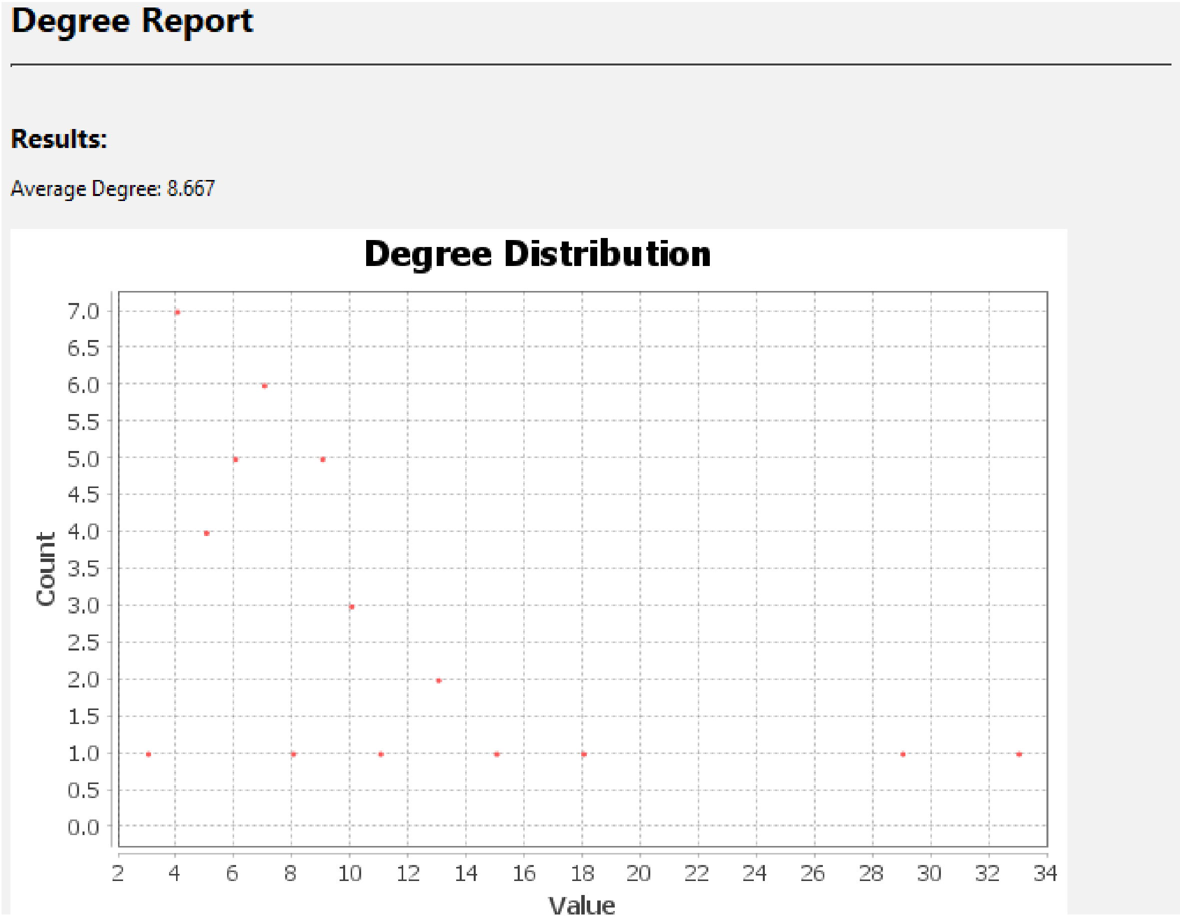

Figure: Representing Degree distribution which represents power-law

2. **Clustering coefficient**: While 0.45 isn't a high clustering coefficient in a scale free network, it is still significant. This level of clustering might indicate the presence of modules or communities within the network, where nodes with high degrees tend to cluster together. The presence of localized clusters also contributes to resulting in higher clusters.

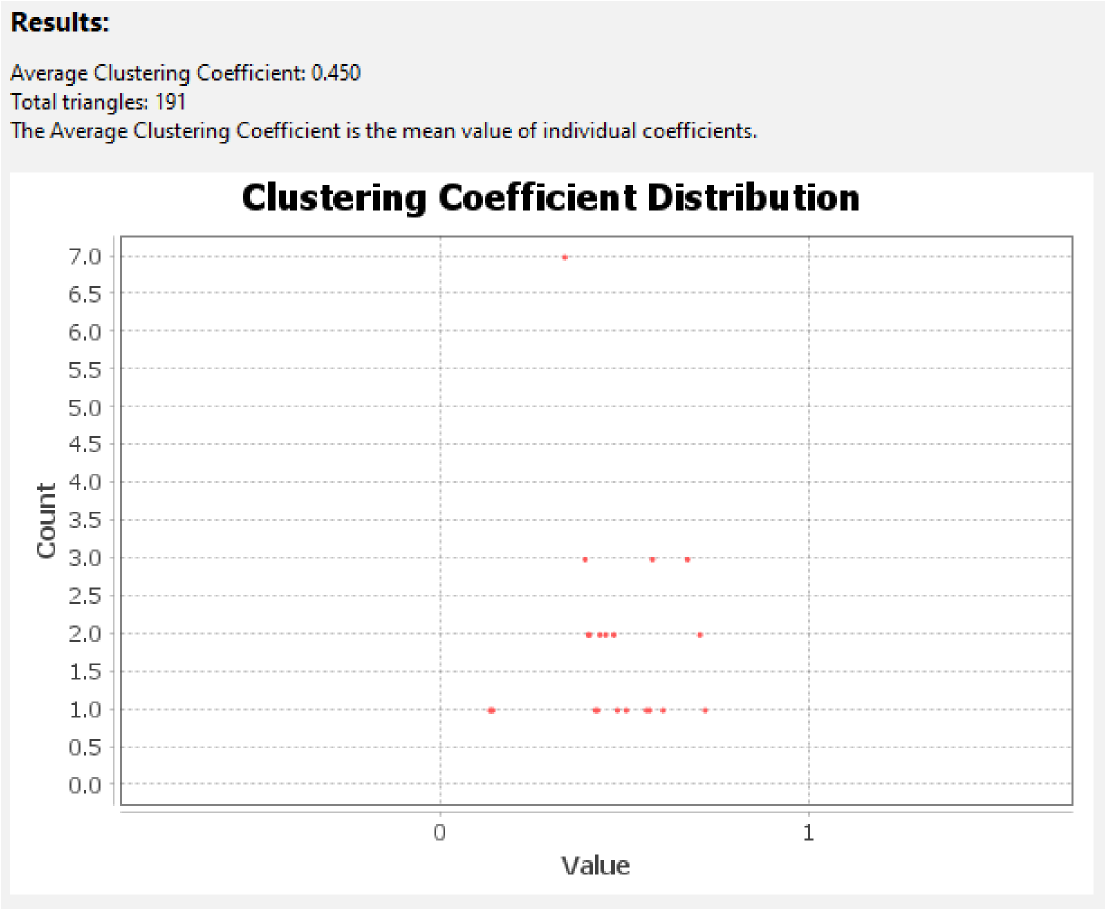

Figure: Represents the clustering coefficient distribution

3. **Average path length**: Scale-free networks often exhibit small-world properties, characterized by short average path lengths despite having hubs. In scale-free networks, the presence of hubs can significantly reduce the average path length, as most nodes are connected to a few highly connected hubs, allowing for efficient communication across the network. The average path length 1.838 exactly in line with the properties of a scale-free network.

   In our road dataset the centrality measures such as betweenness centrality distribution, closeness centrality distribution, harmonic closeness centrality distribution and eccentricity distribution emphasis small world properties, reinforcing their efficiency and robustness in information transmission and connectivity.

Comparison of Betweenness centrality and Clustering coefficient

Betweenness centrality and clustering coefficient provide complementary perspectives on network structure. For node 16 the betweenness centrality 0.606085 and closeness centrality 0.6 is similar which means the node 16 lies as connector between different networks and located in well-connected communities. Similar trend is seen in node 17 with 0.569 and 0.6667 as betweenness and clustering coefficient respectively.

In contrary, the node 39 has 240.481735 as betweenness value but the clustering coefficient is 0.132576 very less so it acts as a critical bridge or connector between different parts of the network while having fewer connections among its immediate neighbours. Likewise, the node 36 which is a prominent hub has betweenness value of 144.171958 but a very less clustering value of 0.137931.

In summary, the road network has a considerable number of nodes and edges which makes it easier to exhibit small world effect and any two nodes can be connected withing 7 connections. There are 169 edges which indicates that there are multiple routes between most pairs of locations, reducing travel time and improving accessibility. It also displays significant local clustering and short average path lengths, which are typical characteristics of real-world road networks. These properties facilitate efficient travel and navigation within the network while also reflecting the underlying geographical and functional organization of road systems.

The road dataset closely exhibits the 80/20 rule which is also known as the “rich getting rich” principle. The nodes 39,36,38 and 35 are the major hubs in our dataset. If any new road must be laid to connect maximum areas, there is a higher possibility that the new edge/road is connected via node 39 which has the highest edges/connects more places. Thus, reducing the average path length and increase efficiency. In terms of error and attack tolerance, the road network's redundancy, as evidenced by the multiple routes between most pairs of locations, enhances its stability against both random errors and targeted attacks. Random errors, such as road closures due to accidents or maintenance, are mitigated by alternative routes, ensuring continued connectivity. Similarly, from the nature of above road network being a scale free network it is error tolerance but not attack tolerance. Even though the presence of multiple routes provides alternative paths in the event of deliberate attacks on specific road segments, it might be serious if a node like 39 (hub) is being attacked. If it is not a hub then there will be a minimal impact on overall network functionality.

Limitations of Scale-free network

From the above we find the dataset depicts more of scale-free features, but we do find some limitations in real world with respect to this model. In most real-world networks, the degree of nodes is also a function of their intrinsic worth and there is no zero probability that a new node attaches to an isolated node. But scale free model doesn’t take this into account (Newman,2018). Also, scale-free network attachment is in proportion to sum of in and out degree but in real world different mechanisms can be responsible for this. The occurrence of removal of nodes and edges is not taken into consideration. Scale free network is undirected but in real world more directed networks exists and the research on directed road network data could unveil more advanced features to be focused.

Ethical Considerations

Since Stanley Milgram's landmark experiment in the 1960s, which raised profound ethical concerns regarding participant welfare, significant strides have been made in refining ethical standards in psychological research. Contemporary research now prioritizes informed consent as a cornerstone of ethical practice, ensuring that participants are fully aware of the study's objectives, potential risks, and their rights to withdraw at any time. Institutional review boards (IRBs) play a pivotal role in overseeing research protocols, safeguarding participant welfare, and ensuring compliance with ethical guidelines. Researchers are increasingly attuned to the potential psychological harm that studies may inflict on participants, implementing rigorous debriefing procedures and ethical training to minimize risks and address any adverse effects (American Psychological Association, 2017)

Furthermore, ethical frameworks such as those outlined in the Belmont Report underscore the importance of respecting participants' autonomy, promoting beneficence, and ensuring justice in research practices. These principles serve as guiding pillars for researchers in navigating complex ethical dilemmas and upholding the integrity of their studies. Additionally, the American Psychological Association (APA) provides comprehensive guidelines for ethical conduct in research, emphasizing the importance of obtaining informed consent, maintaining confidentiality, and mitigating harm to participants. By adhering to these ethical standards and frameworks, researchers can uphold the highest standards of integrity and promote the well-being of those involved in psychological research endeavours.

Conclusion & Future work

To conclude, the analysis of the road network dataset reveals with any two nodes typically connected within a short number of steps, approximately seven connections exhibiting closeness to Milgram's finding on small world effect. The network demonstrates error tolerance due to its scale-free nature and exhibit vulnerability to targeted attacks, particularly if key hub nodes are compromised. Moreover, the road dataset displays prominent characteristics of real-world road networks, including significant local clustering and is decentralised because of many prominent hubs. This finding helps us to infer that the dataset we took exhibit **scale free model explicitly** more than the random model.

However, the limitations of the data set involve a non-dynamic nature and not a directed network as in many real cases must be considered. Incorporating dynamic models can help predict future network dynamics and inform long-term planning strategies. The analysis may be specific to the road network under study and may not be directly applicable to other networks with distinctive characteristics or contexts. **Future research could delve** deeper into understanding the dynamics of the scale-free road network, including its growth mechanisms, traffic flow patterns, and response to disruptions. Additionally, exploring methods to mitigate vulnerabilities and enhance the network's robustness could further improve its efficiency and reliability for transportation systems.

Appendix

The below table shows the values from Gephi for the road dataset for each node which contains details of label, degree, Weighted Degree, Eccentricity, Closeness centrality, Harmonic closeness centrality, Authority, Hub, Modularity class, Clustering coefficient and triangles.

Label	Degree	Eccentricity	closeness centrality	harmonic closeness centrality	betweenness centrality	Hub	clustering coefficient	triangles
1	10	2	0.575758	0.631579	33.554593	0.156557	0.422222	19
2	11	3	0.575758	0.640351	4.099939	0.203391	0.563636	31
3	7	3	0.527778	0.578947	2.411905	0.129795	0.47619	10
4	4	3	0.520548	0.548246	9.427381	0.093667	0.333333	2
5	4	3	0.520548	0.548246	10.336122	0.073595	0.333333	2
6	4	3	0.520548	0.548246	0.410256	0.082962	0.666667	4
7	10	2	0.575758	0.631579	5.310598	0.153881	0.555556	25
8	13	2	0.603175	0.671053	18.474803	0.194024	0.397436	31
9	7	3	0.542857	0.587719	0.840415	0.155219	0.714286	15
10	5	3	0.527778	0.561404	0.223748	0.129795	0.7	7
11	7	2	0.550725	0.592105	2.570213	0.156557	0.571429	12
12	9	2	0.567164	0.618421	7.160478	0.19001	0.444444	16
13	6	3	0.535211	0.574561	1.591791	0.155219	0.466667	7
14	8	3	0.542857	0.596491	2.91164	0.137824	0.428571	12
15	9	3	0.550725	0.609649	4.113492	0.149867	0.388889	14
16	5	3	0.520548	0.557018	0.606085	0.116414	0.6	6
17	4	3	0.513514	0.54386	0.569048	0.097681	0.666667	4
18	9	3	0.550725	0.609649	4.380159	0.151205	0.388889	14
19	9	3	0.550725	0.609649	6.356769	0.1686	0.388889	14
20	4	3	0.513514	0.54386	0.494467	0.124443	0.333333	2
21	5	3	0.520548	0.557018	0.694467	0.13381	0.5	5
22	13	2	0.603175	0.671053	17.233128	0.226138	0.397436	31
23	10	2	0.575758	0.631579	7.866461	0.199376	0.444444	20
24	4	3	0.513514	0.54386	0.494467	0.124443	0.333333	2
25	6	3	0.527778	0.570175	1.365102	0.141838	0.4	6
26	7	3	0.535211	0.583333	2.198435	0.153881	0.428571	9
27	9	3	0.550725	0.609649	4.950287	0.173953	0.416667	15
28	6	3	0.527778	0.570175	1.41206	0.147191	0.466667	7
29	6	3	0.527778	0.570175	1.365102	0.141838	0.4	6
30	6	3	0.527778	0.570175	0.733356	0.149867	0.666667	10
31	5	3	0.520548	0.557018	0.383356	0.13381	0.7	7
32	7	3	0.535211	0.583333	1.835208	0.156557	0.571429	12
33	7	3	0.535211	0.583333	1.633356	0.156557	0.571429	12
34	3	3	0.387755	0.429825	0.833333	0.024086	0.333333	1

35	15	2	0.622951	0.697368	35.753855	0.19001	0.333333	35
36	29	3	0.791667	0.877193	144.171958	0.281	0.137931	56
37	4	3	0.417582	0.469298	3.483766	0.042819	0.333333	2
38	18	3	0.584615	0.70614	38.266667	0.176629	0.130719	20
39	33	2	0.883721	0.934211	240.481735	0.315791	0.132576	70

Below is the graphical representation of betweenness centrality distribution, closeness centrality distribution, harmonic closeness centrality distribution and eccentricity distribution from Gephi.

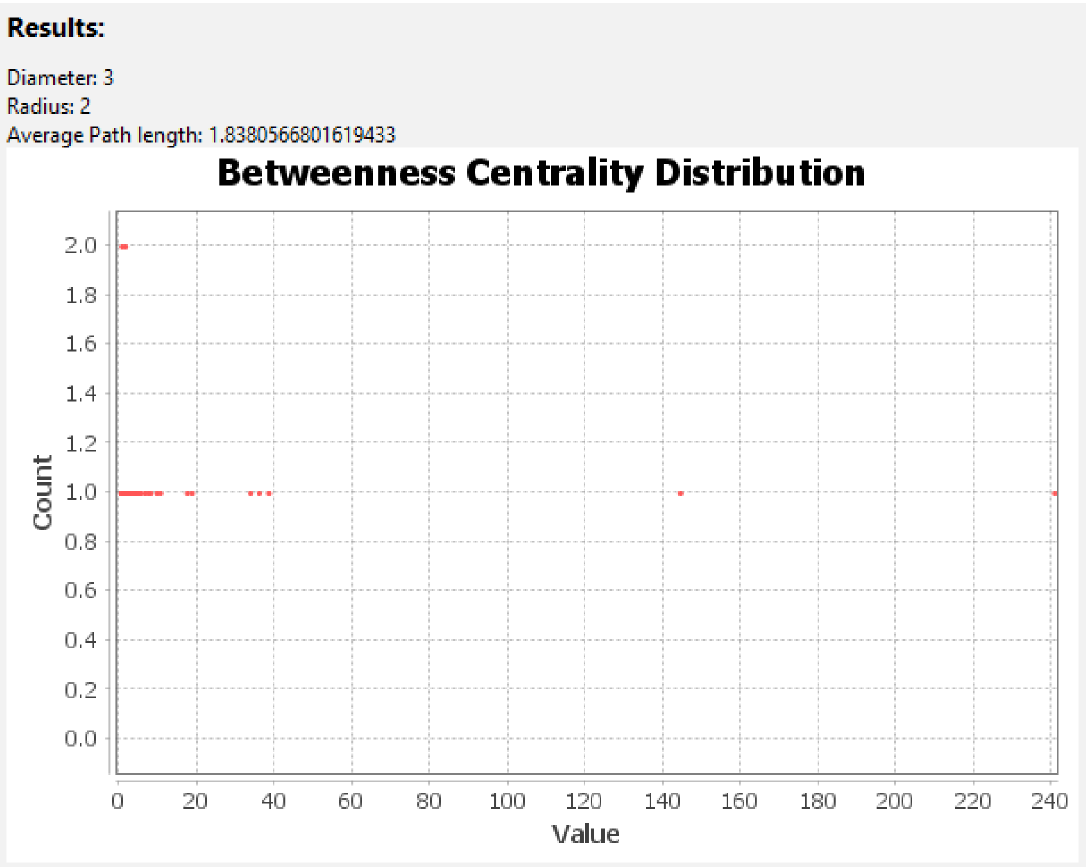

Figure: Represents the betweenness centrality distribution

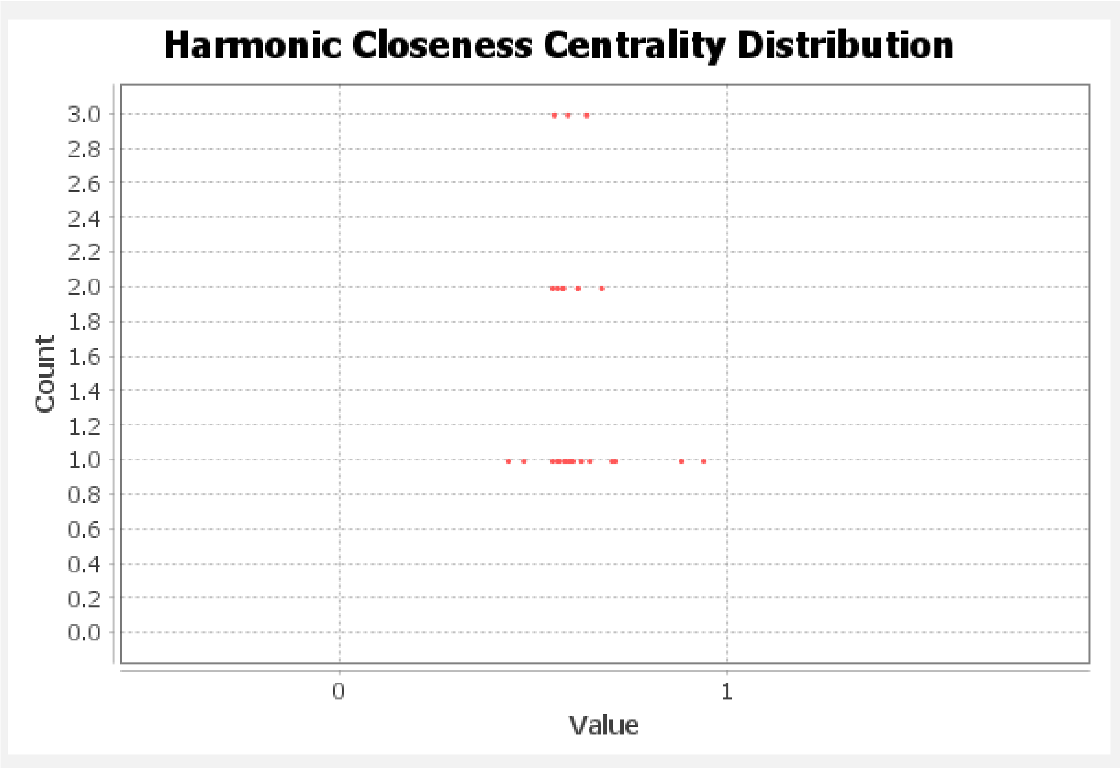

Figure: Represents Harmonic closeness centrality distribution

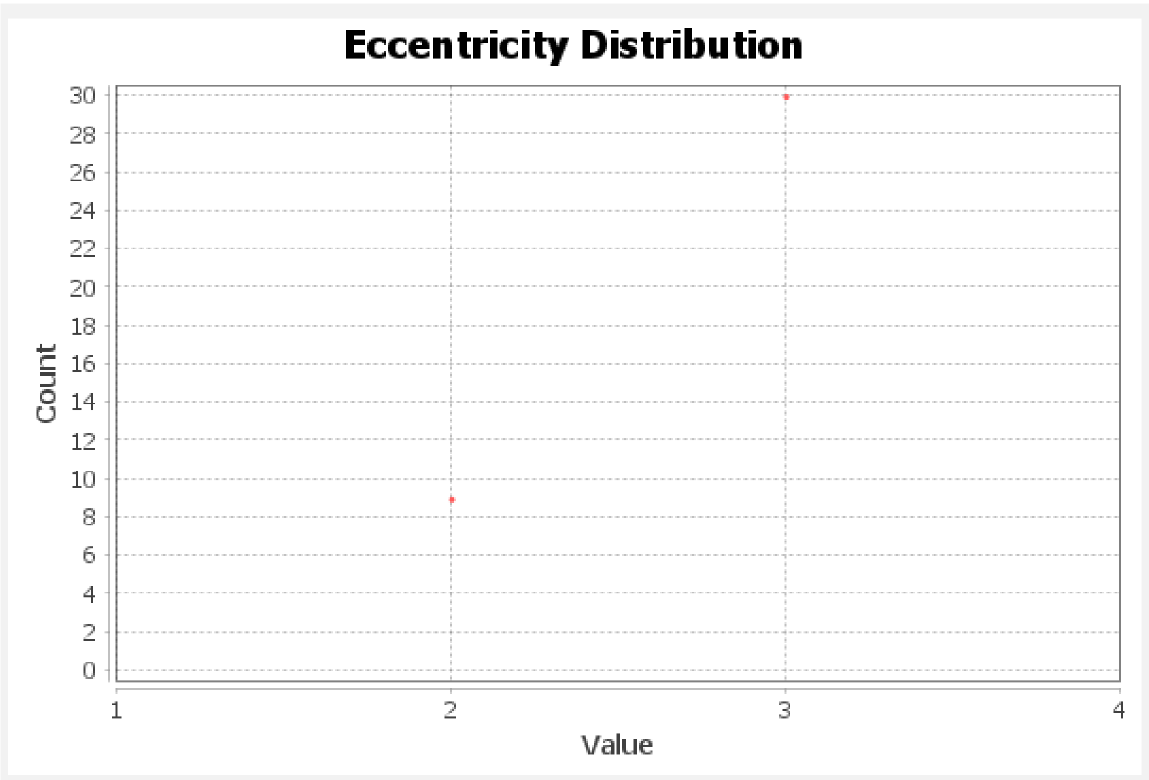

Figure: Represents Eccentricity distribution

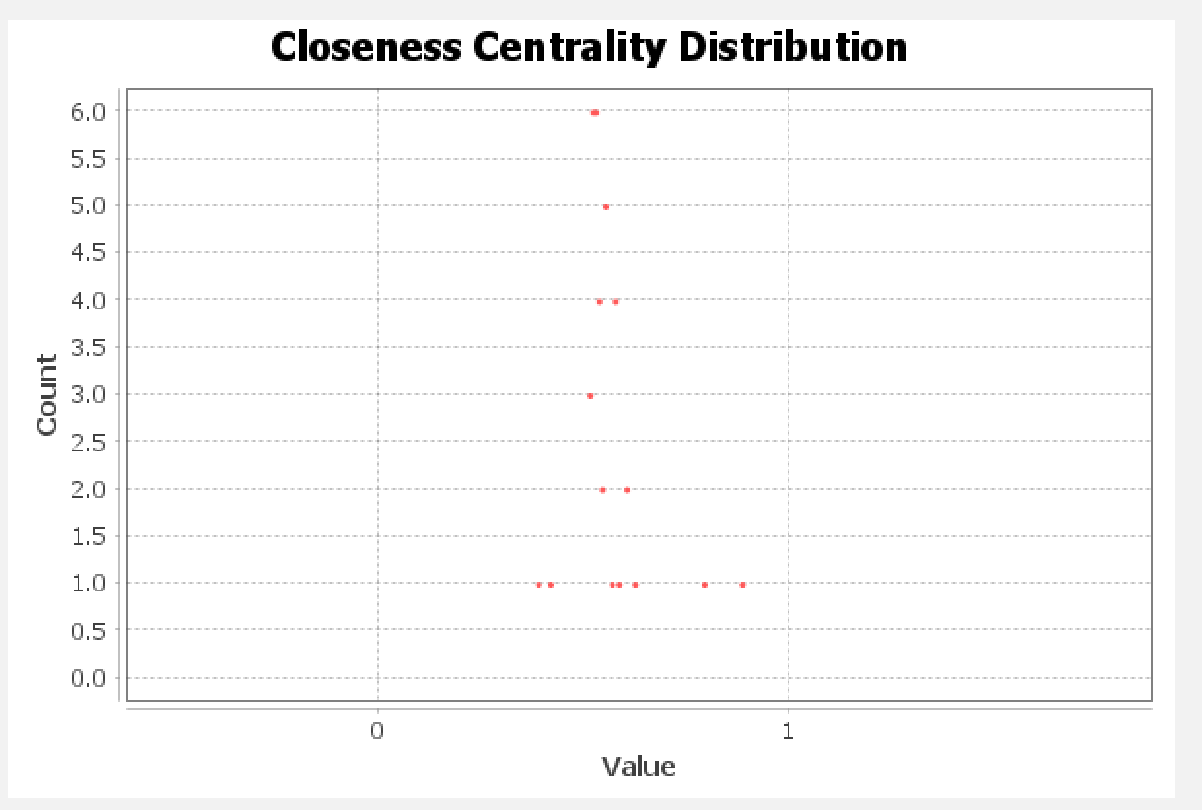

Figure: Represents closeness centrality distribution

References

1. American Psychological Association (2017). Ethical Principles of Psychologists and Code of Conduct. American Psychological Association. [online] Available at: https://www.apa.org/ethics/code.
2. Barabási, A., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.
3. Hong Zhang, 2011. Uncovering Road Network Structure Through Complex Network Analysis, the Hong Kong polytechnic university.
4. Michael T. Goodrich and Evrim Ozel. 2022. Modelling the Small-World Phenomenon with Road Networks, University of California, Irvine, USA.https://doi.org/10.1145/3557915.3560981
5. Milgram, S., 1967. The small world problem. Psychology Today 2, 60–67.
6. Milgram, S., 1977. The Individual in a Social World: Essays and Experiments. Addison Wesley, Reading, MA.
7. Newman M.,2018, Networks, Oxford University press , University of Michigan.

8. Sebastian Schnettler A structured overview of 50 years of small-world research Centre for Research on Inequalities and the Life Course, Department of Sociology, Yale University, New Haven, CT, United States
9. Watts, D.J., & Strogatz, S.H.,1998. Collective dynamics of ‘small-world’ networks, Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA.
10. Watts, D.J., 1999a. Small Worlds. The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton, NJ.