3. Watts Strogatz Model

Random networks, in which nodes are connected randomly, are small worlds but not clustered. Watts-Strogatz (WS) proposed a network model that is both clustered and a small world, like real networks to some extent. They started off with the idea that people know their neighbors, so they connect each node with their *n* nearest neighbors.

This produces a clustered network, but it is a large world. They randomly rewired some of the links. WS discovered that rewiring only *a few* links is enough to decrease the average distance between nodes, i.e. make it a small world, while not affecting the clustering considerably.

Investigate this claim by WS. Start with nearest neighbor network (provided) of 100 nodes, where each node connects to its 6 nearest neighbors (similar to Fig. 1).

a. Calculate the average clustering coefficient and average path length of this network (C_{ave} and d_{ave})
b. Choose one link randomly and rewire one end to a different, randomly chosen, node
c. Calculate C_{ave} and d_{ave} again
d. Repeat parts b and c until all links have been rewired
e. Perform this simulation a sufficient number of times* to obtain representative averages of C_{ave} and d_{ave} for each number of rewired edges.
f. Plot these averages as a function of the number of rewired edges. Argue the claim made by WS.

* How can you tell if you've completed a sufficient number of simulations? This is a good question to consider.

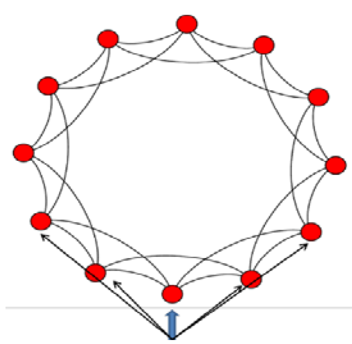


Figure 1. Each node connects with nearest neighbors. This produces a clustered but large-world network

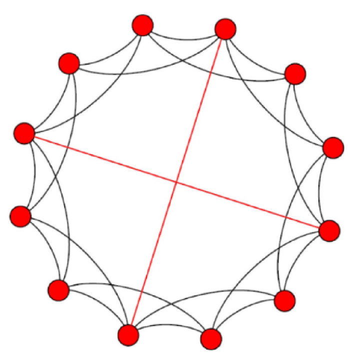


Figure 2. Network with rewired links (red). This is a clustered, small world network. (Note, red links in image were added not rewired

Introduction and Procedure

This project started with a regular graph with 100 nodes, each connected to 6 nearest neighbors, 3 on each side. A random node from a random edge was then selected to rewire to a random non-neighbor. This results in a slightly different graph. This continued till all edges were randomly selected and rewired. Eventually, a random graph was resulted. See depiction in Figure 1. Each time after an edge was rewired, two global graph metrics were calculated to evaluate the effect of edge rewiring: a) the average clustering coefficient, and b) the average shortest path length (distance) between all paired nodes. Clustering coefficient, C_{ave}, of a node is a measure of how connected the node's neighbors. To calculate the average shortest path length (distance), d_{ave}, of the graph, the graph must be connected, which means that each node must be connected to at least one other node. Average shortest path length of the graph is number of edges in the shortest path between two nodes, averaged over all pairs of nodes. This process was repeated for three set of 40 trials.

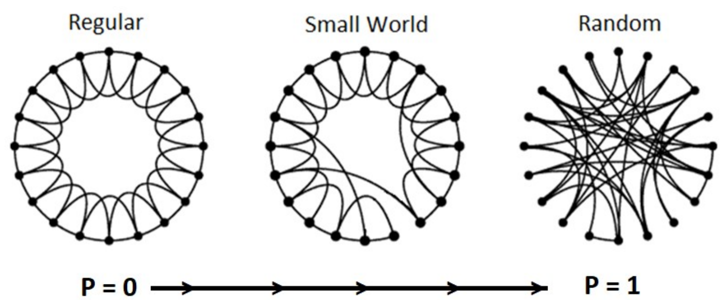


Figure 1. Regular Graph, Small World and a Random Graph

The Watts-Strogatz (WS) algorithm generates a regular graph consisting n nodes, each connected to k closest neighbors. The WS algorithm then progressively goes through one edge at a time rewiring it at random with p probability till all edges are gone through. Watts-Strogatz claims that rewiring only *a few* links is enough to decrease the average distance between nodes, i.e. make it a small world, while not affecting the clustering considerably.

Observation and Conclusion

My founding agrees with Watts-Strogatz's claim. After a few edges (2 to 8 edges) are rewired in each of the 120 trials, the graph changed from a regular graph to a small world graph. The graph has an average distance between all paired nodes, d_{ave}, less than 6 and the average clustering coefficient, C_{ave}, was not affected much.

Figure 2 below visualizes how rewiring of each random edge alters the graph slightly differently in the beginning. The randomized process resulted in a different random graph in each of the trials A through D. Some of the new wires reached nodes farther away, while some reached closer non-neighbors. A long-range rewired edge would significantly reduce the shortest path length (distance) between the nodes connected by the new edge. The shortest path ($d(i) = 1$) between the nodes connected by the i^{th} rewired edge went directly across the ring instead of the original series of hops. Such rewiring also reduced the distance between the ***nodes near the new edge***, since the neighboring nodes on each side could also utilize the edge to get to the neighboring nodes on other side of the ring. Hence the effect on the average distance, d_{ave}, between paired nodes of the graph amplifies even with one long-range rewired edge.

On the other hand, the clustering coefficient, C(i), of node i measures how much its friends are connected. To the node that was let go because an edge was rewired, it lost only one friend out of the 6 original neighbors. But this did not alter how the node's remaining friends are connected at all. On the other hand, to the node that acquired a new neighbor, its clustering coefficient was reduced since the new neighbor was not connected to any of the original 6 neighbors. So the rewiring only affected the C(i) of one of the originally connected nodes. Thus, the effect on the average clustering coefficient, C_{ave}, of the graph as a whole is not as significant as on the average shortest path length or distance, d_{ave}, of the graph as a whole. See Table 1 for a listing and changes of the metrics for these trials. Observed variance of metrics from trial to trial is further discussed at the end of this report.

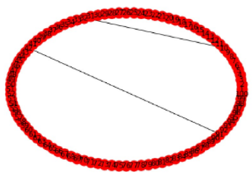
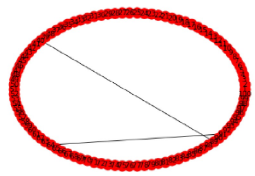
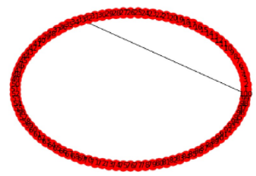
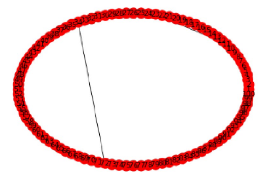
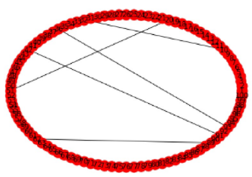
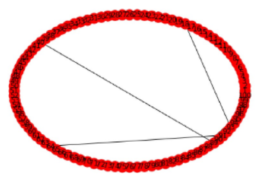
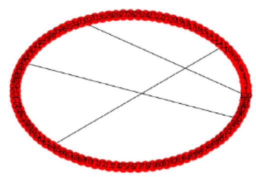
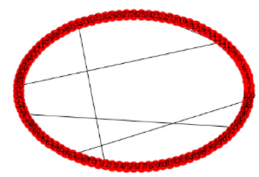
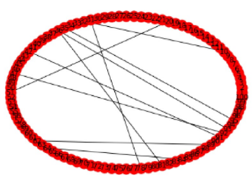
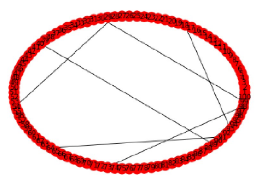
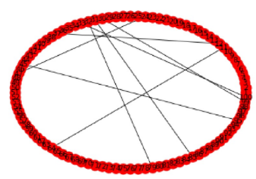
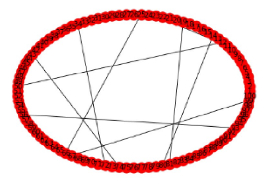
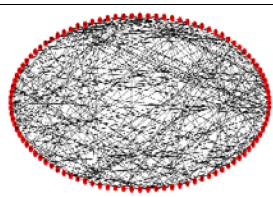
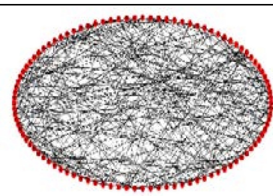
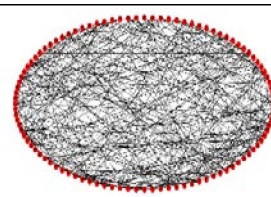
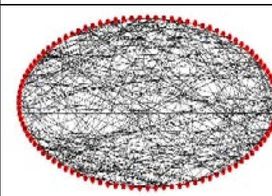
Trial	Trial A	Trial B	Trial C	Trial D
Two random edges rewired				
5 random edges rewired				
10 random edges rewired				
All 300 edges rewired				

Figure 2. Visualization of Changes to Graph for Trials A To D

*Table 1. Trials A -D: Average Clustering Coefficient (**C_{ave}**) and Average Distance (**d_{ave}**) of Graph*

	Trial A		**Trial B**		**Trial C**		**Trial D**	
# of Rewired Edges	**C_{ave}**	**d_{ave}**	**C_{ave}**	**d_{ave}**	**C_{ave}**	**d_{ave}**	**C_{ave}**	**d_{ave}**
1	0.6000	8.7576	0.6000	8.7576	0.6000	8.7576	0.6000	8.7576
2	0.5943	7.7311	0.5943	7.1277	0.5943	7.3511	0.5966	8.3061
3	0.5886	6.7968	0.5862	6.2887	0.5888	7.1313	0.5909	6.8657
4	0.5807	6.5655	0.5784	6.1394	0.5850	6.6323	**0.5829**	**5.9784**
5	**0.5750**	**5.8851**	**0.5703**	**5.4354**	0.5817	6.4966	0.5771	5.6291
6	0.5693	5.3747	0.5702	5.4202	**0.5736**	**5.5673**	0.5714	5.1598
7	0.5648	5.2697	0.5645	5.2630	0.5702	5.3392	0.5636	4.9970
8	0.5596	4.9036	0.5575	5.2343	0.5664	5.1826	0.5579	4.7337
9	0.5556	4.8699	0.5529	5.2226	0.5609	5.1341	0.5551	4.6921
10	0.5476	4.7026	0.5495	4.9913	0.5523	4.7899	0.5471	4.6125
11	0.5401	4.6558	0.5420	4.7085	0.5481	4.7297	0.5411	4.4412
12	0.5387	4.5438	0.5364	4.6089	0.5401	4.6592	0.5392	4.4119
13	0.5390	4.5315	0.5317	4.6044	0.5339	4.5400	0.5358	4.3463
14	0.5340	4.3885	0.5273	4.5465	0.5288	4.4671	0.5299	4.2200
15	0.5300	4.3754	0.5213	4.4859	0.5239	4.4475	0.5223	4.1877
16	0.5257	4.3087	0.5181	4.3246	0.5212	4.2214	0.5144	4.1295
17	0.5208	4.2954	0.5147	4.2560	0.5143	4.1784	0.5119	4.0855
18	0.5164	4.2002	0.5114	4.2356	0.5085	4.0747	0.5085	4.0232
19	0.5138	4.1582	0.5030	4.1461	0.5006	4.0339	0.5053	4.0063
20	0.5114	4.1143	0.5009	4.0305	0.4929	4.0119	0.5020	4.0026
21	0.5070	4.0768	0.5032	4.0368	0.4900	3.9149	0.5000	3.9780
22	0.5000	3.9622	0.4977	3.9943	0.4854	3.8822	0.4950	3.9053
23	0.4970	3.9349	0.4904	3.9622	0.4826	3.8275	0.4904	3.8721
24	0.4925	3.9079	0.4823	3.9267	0.4798	3.8040	0.4869	3.8566
25	0.4868	3.8469	0.4775	3.8265	0.4733	3.7503	0.4813	3.8137

Data Analysis and Results

The rewiring process was repeated for 40 trials and the mean C_{ave} and mean d_{ave} were plotted as a function of the number of rewired edges. Figure 3 presents the normalized $\overline{C}_{ave}$ and $\overline{d}_{ave}$ values on the y-axis against the number of rewired edges on the x-axis in a semi-logarithmic scale. The plot of the mean average distance $\overline{d}_{ave}$ drops considerably at the beginning of the process when only a few edges are rewired, while the plot of mean average clustering coefficient $\overline{C}_{ave}$ declines on more gradual slope. For $\overline{d}_{ave}$ to drop below 6, which is 0.68 of the original average shortest path length 8.76, thus making the graph a small world, it only took as few as 5 edges to be rewired. At that point, $\overline{C}_{ave}$ was still well above 0.9 of the original average clustering coefficient. See green dotted line in Figure 3 showing the average metrics when 5 edges were rewired.

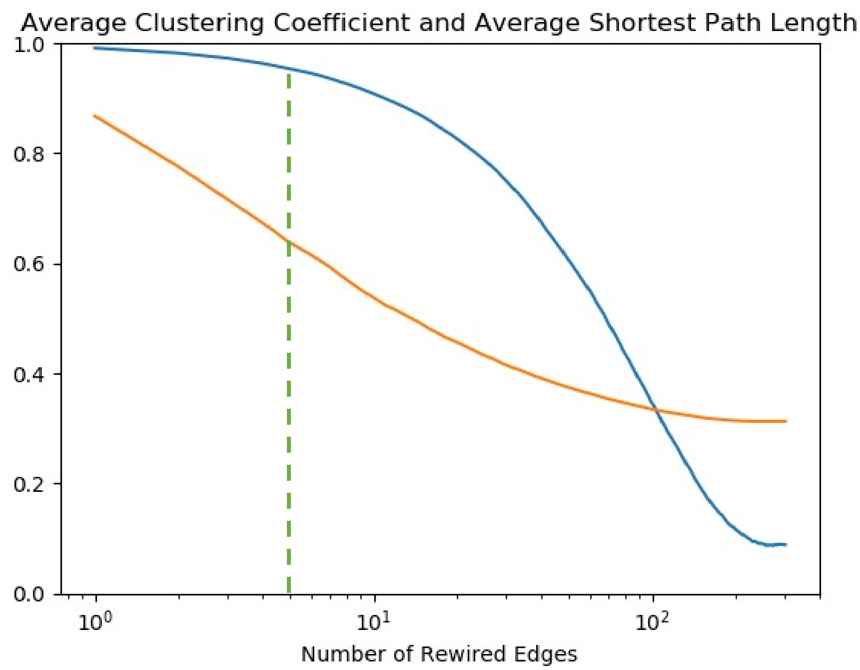


Figure 3. Mean Average Shortest Path Length $\overline{d}_{ave}$ and Mean Average Clustering Coefficient C_{ave} of Graph as a Function of the Number of Rewired Edges

<u>Variance from Trial to Trial and Conclusion</u>

Because of random chance, a rewired edge may be short-range, long-range or somewhere between. So, the effect of rewiring to the average distance, d_{ave}, varies from trials to trials each time an edge is rewired, as shown in Table 1 above for Trials A through D. So it varied from 4 to 6 edge rewiring to change the graph to a small world.

Table 2 below shows the average results over 10, 20, 40, 80 and 120 trials. It shows that it took 20 trials and up for the variance of the mean of the average distance, d_{ave}, to stay below 0.2. And it shows consistently that 5 edge rewiring was enough to change the original regular graph to a small-world with the mean average distance $\overline{d}_{ave}$ below 6. So, I conclude that 20 trials are sufficient with variance below 0.2 for purpose of this exercise. Additionally, Table 3 shows that the results over 20, 40, 80 and 120 trials consistently reduced the mean average distance $\overline{d}_{ave}$ of the graph by more than 0.68 of the original d_{ave} to below 6 and the mean average clustering coefficient $\overline{C}_{ave}$ well above 0.95 of the original C_{ave}, of the original graph.

Table 2. Mean Average Clustering Coefficient ($\overline{C}_{ave}$) and Mean Average Shortest Path Length ($\overline{d}_{ave}$) of Graph

	From 10 trials			From 20 trials			From 40 trials			From 80 trials			From 120 trials		
# of Rewired	Mean C_{ave}	Mean d_{ave}	Variance d_{ave}	Mean C_{ave}	Mean d_{ave}	Variance d_{ave}	Mean C_{ave}	Mean d_{ave}	Variance d_{ave}	Mean C_{ave}	Mean d_{ave}	Variance d_{ave}	Mean C_{ave}	Mean d_{ave}	Variance d_{ave}
0	0.6000	8.7576	0.0000	0.6000	8.7576	0.0000	0.6000	8.7576	0.0000	0.6000	8.7576	0.0000	0.6000	8.7576	0.0000
1	0.5946	7.9639	0.2295	0.5940	7.8303	0.2511	0.5943	7.7950	0.2487	0.5945	7.7770	0.2436	0.5945	7.7603	0.2404
2	0.5881	7.2795	0.1811	0.5880	7.2729	0.2694	0.5887	7.1059	0.2753	0.5890	6.9991	0.3505	0.5889	7.0211	0.3411
3	0.5828	6.5783	0.3498	0.5827	6.5773	0.3292	0.5828	6.4656	0.2675	0.5834	6.4088	0.2863	0.5833	6.4466	0.2927
4	0.5759	6.1321	0.2240	0.5767	6.0955	0.2152	0.5771	6.0481	0.2021	0.5779	6.0322	0.2086	0.5778	6.0588	0.2002
5	**0.5708**	**5.7009**	0.2223	**0.5721**	**5.6843**	0.1458	**0.5722**	**5.7314**	0.1644	**0.5727**	**5.7075**	0.1595	**0.5729**	**5.7227**	0.1636
6	0.5648	5.5190	0.2626	0.5666	5.4471	0.1826	0.5670	5.4688	0.1411	0.5674	5.4218	0.1086	0.5676	5.4299	0.1011
7	0.5600	5.3847	0.2588	0.5613	5.2347	0.1639	0.5617	5.2206	0.1082	0.5624	5.1891	0.0905	0.5625	5.2117	0.0896
8	0.5571	5.1796	0.1005	0.5568	5.0561	0.0813	0.5567	5.0299	0.0753	0.5574	5.0155	0.0699	0.5573	5.0292	0.0666
9	0.5536	4.9458	0.0735	0.5519	4.8562	0.0575	0.5518	4.8747	0.0600	0.5522	4.8529	0.0623	0.5523	4.8735	0.0604
10	0.5483	4.8314	0.0688	0.5467	4.7374	0.0546	0.5464	4.7490	0.0493	0.5469	4.7213	0.0510	0.5471	4.7403	0.0506
11	0.5433	4.7149	0.0707	0.5417	4.6515	0.0508	0.5410	4.6417	0.0496	0.5420	4.6084	0.0434	0.5422	4.6270	0.0459
12	0.5378	4.5926	0.0552	0.5364	4.5469	0.0393	0.5360	4.5337	0.0370	0.5368	4.5006	0.0335	0.5372	4.5128	0.0324
13	0.5327	4.4627	0.0438	0.5316	4.4316	0.0282	0.5309	4.4189	0.0296	0.5313	4.4079	0.0263	0.5321	4.4286	0.0299
14	0.5274	4.3240	0.0205	0.5259	4.3198	0.0147	0.5255	4.3308	0.0202	0.5261	4.3240	0.0215	0.5269	4.3398	0.0244
15	0.5225	4.2232	0.0164	0.5210	4.2312	0.0113	0.5203	4.2436	0.0156	0.5211	4.2438	0.0180	0.5220	4.2663	0.0221
16	0.5172	4.1662	0.0120	0.5159	4.1696	0.0098	0.5153	4.1803	0.0143	0.5163	4.1777	0.0156	0.5172	4.2001	0.0202
17	0.5124	4.1182	0.0148	0.5104	4.1167	0.0135	0.5094	4.1272	0.0146	0.5111	4.1220	0.0145	0.5120	4.1409	0.0181
18	0.5072	4.0756	0.0156	0.5050	4.0688	0.0144	0.5043	4.0767	0.0157	0.5062	4.0704	0.0147	0.5072	4.0868	0.0183
19	0.5019	4.0397	0.0127	0.4995	4.0249	0.0136	0.4994	4.0283	0.0131	0.5011	4.0243	0.0135	0.5019	4.0356	0.0142
20	0.4978	3.9888	0.0175	0.4954	3.9734	0.0164	0.4951	3.9868	0.0145	0.4966	3.9811	0.0150	0.4972	3.9897	0.0143
21	0.4930	3.9249	0.0142	0.4904	3.9165	0.0134	0.4898	3.9299	0.0119	0.4913	3.9257	0.0112	0.4923	3.9384	0.0110
22	0.4883	3.8912	0.0117	0.4862	3.8886	0.0124	0.4853	3.8902	0.0117	0.4865	3.8837	0.0102	0.4874	3.8973	0.0105
23	0.4829	3.8465	0.0138	0.4805	3.8486	0.0126	0.4803	3.8595	0.0115	0.4818	3.8501	0.0099	0.4828	3.8615	0.0097
24	0.4780	3.8027	0.0105	0.4760	3.8040	0.0086	0.4753	3.8166	0.0095	0.4766	3.8114	0.0091	0.4777	3.8249	0.0089
25	0.4732	3.7541	0.0091	0.4711	3.7658	0.0073	0.4706	3.7815	0.0084	0.4717	3.7775	0.0082	0.4729	3.7919	0.0081

Table 3. Normalized Average Clustering Coefficient, Cave(i)/Cave(0),and Average Shortest Path Length, dave(i)/dave(0), of Graph

	From 20 trials		From 40 trials		From 80 trials		From 120 trials	
# of Rewired Edges	$C_{ave}(i)/C_{ave}(0)$	$d_{ave}(i)/d_{ave}(0)$	$C_{ave}(i)/C_{ave}(0)$	$d_{ave}(i)/d_{ave}(0)$	$C_{ave}(i)/C_{ave}(0)$	$d_{ave}(i)/d_{ave}(0)$	$C_{ave}(i)/C_{ave}(0)$	$d_{ave}(i)/d_{ave}(0)$
1	0.9901	0.8941	0.9904	0.8901	0.9908	0.8880	0.9909	0.8861
2	0.9800	0.8305	0.9812	0.8114	0.9816	0.7992	0.9815	0.8017
3	0.9711	0.7510	0.9713	0.7383	0.9723	0.7318	0.9722	0.7361
4	0.9612	0.6960	0.9618	0.6906	0.9631	0.6888	0.9630	0.6918
5	0.9534	0.6491	0.9537	0.6544	0.9546	0.6517	0.9549	0.6535
6	0.9444	0.6220	0.9449	0.6245	0.9457	0.6191	0.9459	0.6200
7	0.9354	0.5977	0.9362	0.5961	0.9374	0.5925	0.9375	0.5951
8	0.9280	0.5773	0.9279	0.5744	0.9289	0.5727	0.9289	0.5743
9	0.9198	0.5545	0.9197	0.5566	0.9204	0.5541	0.9204	0.5565
10	0.9111	0.5409	0.9106	0.5423	0.9115	0.5391	0.9118	0.5413
11	0.9029	0.5311	0.9017	0.5300	0.9033	0.5262	0.9036	0.5283
12	0.8940	0.5192	0.8934	0.5177	0.8946	0.5139	0.8953	0.5153
13	0.8861	0.5060	0.8848	0.5046	0.8855	0.5033	0.8868	0.5057
14	0.8765	0.4933	0.8758	0.4945	0.8768	0.4937	0.8781	0.4955
15	0.8684	0.4831	0.8672	0.4846	0.8686	0.4846	0.8700	0.4872
16	0.8599	0.4761	0.8588	0.4773	0.8605	0.4770	0.8620	0.4796
17	0.8507	0.4701	0.8490	0.4713	0.8518	0.4707	0.8534	0.4728
18	0.8417	0.4646	0.8406	0.4655	0.8437	0.4648	0.8454	0.4667
19	0.8326	0.4596	0.8323	0.4600	0.8351	0.4595	0.8365	0.4608
20	0.8256	0.4537	0.8251	0.4552	0.8276	0.4546	0.8287	0.4556
21	0.8174	0.4472	0.8163	0.4487	0.8189	0.4483	0.8205	0.4497
22	0.8103	0.4440	0.8088	0.4442	0.8108	0.4435	0.8123	0.4450
23	0.8009	0.4395	0.8006	0.4407	0.8029	0.4396	0.8046	0.4409
24	0.7933	0.4344	0.7922	0.4358	0.7943	0.4352	0.7961	0.4367
25	0.7852	0.4300	0.7843	0.4318	0.7861	0.4313	0.7881	0.4330