Collective dynamics of small-world networks

Fredrik Jonasson, Fredrik Gustafsson,
Simon Huss, Simon Johansson Norman, Olle Torstensson

June 2019

Abstract

In this paper we investigate an extension of the Watts-Strogatz small-world network model, modifying it to rewire the edges within the network using a combination of random and preferential attachment (vertices with more connections to it are more likely to be chosen). For parameters g and p defining the ratio between the two types of rewiring we find that purely using preferential attachment gives the network a small-world type structure, while the combination of random and preferential attachment results in the network having properties of a regular network.

In recent years there has been an increasing interest in small-world networks due to their ability to describe various biological processes and social networks etc. A small-world network is a combination between a regular network and a random graph. Similarly to a regular network it shows large clustering, while the path-lengths within the network are small - characteristic to random graphs [1]. Extending the small-world network by changing it from choosing the long-range connections using random attachment to using a combination of random attachment and preferential attachment can be used to model more complex social networks [2]. However, in this paper we will consider a non-growing network with a "one-step" rewiring rule.

The model presented by Watts and Strogatz starts from a regular ring lattice, with n vertices connected to their k nearest neighbours. Each edge is then randomly rewired with probability p, while ensuring that all edges are unique, i.e no pair of vertices have multiple connections between each other. The rewiring starts with the edge leading to the closest neighbour of the vertex. Next, after cycling through all vertices clockwise, the edge connecting a vertex to its second closest neighbour is rewired. The procedure continues until each edge has had a chance to be rewired. Setting the parameter to $p = 0$ or $p = 1$ results in a regular or completely random network respectively [1].

The model previously described can be expanded by changing the rewiring rules. By introducing a new variable g, and a random variable on a uniform interval (0,1), r, the new rewiring rules will be the following. If $g \leq r \leq 0.5 + p$ then rewire with a preferential probability defined as $\frac{k_i}{\sum k}$, if $p + 0.5 \leq r \leq 1 - g$ then rewire it to a random vertex, else do not rewire the edge. This is once again done by starting with the edges connecting each vertex to its closest neighbour in the same way as the Watts and Strogatz network. We restrict the parameter g to $0 \leq g \leq 0.5$.

The structure of the networks resulting from the models discussed can be measured using their clustering coefficient, $C(p)$, and characteristic path length, $L(p)$ [1]. The characteristic path length is here defined as the mean value of the distance (i.e. the shortest path) between two vertices in the network. The (local) clustering coefficient for a vertex v can be calculated as the number of existing edges between the neighbours of v divided by the number of possible such edges. The (global) clustering coefficient $C(p)$ for the whole network is then the average of the local clustering coefficients. This makes $C(p)$ a measurement of the cliquishness of a neighbourhood in the network.

Figure 1 shows the characteristic path length and clustering coefficient of our extended model compared to the Watts-Strogatz model. With $g = 0.5$, resulting in all rewiring being made with preferential attachment, the small-world phenomenon is clearly observed and the computed $L(p)$ and $C(p)$ are practically the same as in the Watts and Strogatz approach. For smaller values of g there is no small-world behaviour since the characteristic path length remain large even for large p. For $g = 0.25$ there is eventually a drop in the clustering, when p becomes close to 1. A very small decrease of $L(p)$ can also be observed in this region.

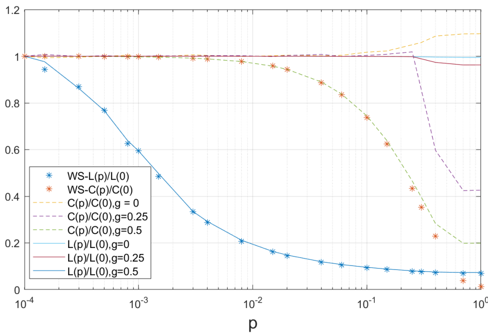

Figure 1: Characteristic path length, L, and clustering coefficient, C, for different values of p. Here WS denotes the small world network proposed by Watts and Strogatz and g is the parameter deciding upper bound for random rewiring and lower bound for preferential attachment in the combined network. The parameter p is the probability to rewire an edge in the small world network and what decides lower bound for random rewiring and upper bound for preferential attachment in the combined network. Each network have $n = 1000$ and $k = 10$, that is 1000 vertices each originally connecting to its 5 nearest neighbours on each side.

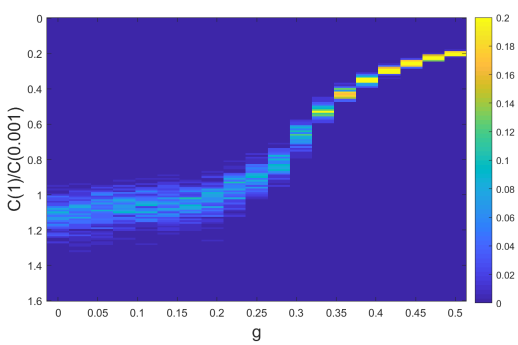

Figure 2: Phase transition over the g parameter and its effect on the clustering coefficient for the combined random and preferential attachment network.

To gain a better understanding of the small-world phenomenon, it is tempting to investigate *analytically* just how the quantities L and C depend on p in the two models. For the regular lattice (i.e. when $p = 0$), we can quite easily derive both the characteristic path length and the clustering coefficient.

Starting with L, we begin by noting that each vertex has $\frac{k}{2}$ neighbours at distance 1 in each direction. Thus it has k neighbours, k neighbours of neighbours, k neighbours of neighbours of neighbours etc. until we reach the opposite side of the ring where we find $(n-1 \mod k)$ vertices at distance $\left\lfloor \frac{n-1}{k} \right\rfloor + 1$ from the original vertex. For convenience, we assume that $(n-1 \mod k) = 0$ and get that

$$\begin{aligned}
L(0) &= \frac{1}{n}\left(k + 2k + \ldots + \left\lfloor \frac{n-1}{k} \right\rfloor k \right. \\
&\quad \left. + \left(\left\lfloor \frac{n-1}{k} \right\rfloor + 1\right)(n-1 \mod k)\right) \\
&= \frac{1}{n}\sum_{i=1}^{\frac{n-1}{k}} ik = \frac{1}{n}\left(\frac{n-1}{k} \cdot \frac{k + \frac{n-1}{k} \cdot k}{2}\right) \\
&= \frac{n-1}{n} \cdot \frac{n+k-1}{2k} \approx \frac{n}{2k},
\end{aligned}$$

for $n \gg k \gg 1$.

Continuing with C, we first note that there for each vertex are $\binom{k}{2} = \frac{k(k-1)}{2}$ possible edges between its neighbours. Then, assuming that $n > 2k$, we note that there are $\sum_{i=0}^{\frac{k}{2}-1} i$ edges between the neighbours in each direction and just as many edges going between neighbours in one direction and the other. Thus,

$$C(0) = \frac{3\sum_{i=0}^{\frac{k}{2}-1} i}{\left(\frac{k(k-1)}{2}\right)} = \frac{\left(\frac{3k(k-2)}{8}\right)}{\left(\frac{k(k-1)}{2}\right)} = \frac{3(k-2)}{4(k-1)} \approx \frac{3}{4},$$

for $k \gg 1$.

However, as Newman and Watts [3] point out, two problems with the Watts-Strogatz model (and by extension, our composite model) arise when $p > 0$ and we have to take rewiring into account. The first is the non-uniformity of the distribution of the rewired links. The second is the non-zero probability of disconnecting the graph in the rewiring process, in which case L is poorly defined. None of these facts pose a significant problem when studying the quantities using numerical methods, but complicate an analytical approach. To circumvent the issues when studying L, they propose a slightly altered version of the Watts-Strogatz model, in which each edge with probability p, instead of being rewired, gives rise to a *new* edge between *any* two vertices, chosen at random. Using this altered model, Newman et al. [4] manage to derive a mean-field approximation (for large n) up to relatively large values of p, but that fails when p gets close to 1.

$$L(p) \approx \frac{n}{\left(\frac{k}{2}\right)} \cdot f\left(n\frac{k}{2}p\right),$$

where

$$f(x) = \frac{1}{2\sqrt{x^2+2x}} \tanh^{-1}\left(\frac{x}{\sqrt{x^2+2x}}\right).$$

Thus, after some algebra and the use of the identity $\tanh^{-1} x = \frac{1}{2}\ln\frac{1+x}{1-x}$, we arrive at

$$L(p) \approx \frac{n}{k} \cdot \frac{\ln\left(\frac{\sqrt{(nkp)^2+4nkp}+nkp}{\sqrt{(nkp)^2+4nkp}-nkp}\right)}{\sqrt{(nkp)^2+4nkp}}.$$

Barrat and Weigt [5] solve the issue of analysing C by slightly redefining it from the *mean of the ratios* of "neighbourhood edges" to be the *ratio of the means*, i.e.

$$C(p) := \frac{\text{mean \#edges between nbrs. of a vtx.}}{\text{mean \#possible links between nbrs. of a vtx.}}.$$

As n grows larger, the probability that a closed triple remains that way approaches $(1-p)^3$, and they then elegantly arrive at the expression

$$C(p) \approx C(0) \cdot (1-p)^3 = \frac{3(k-2)}{4(k-1)}(1-p)^3,$$

which they go on to show has an error of order $\frac{1}{n}$ compared to the Watts-Strogatz definition.

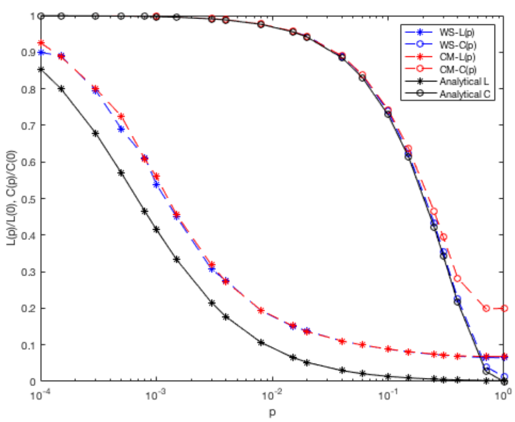

Figure 3: $L(p)$ and $C(p)$ as measured on the Watts-Strogatz(WS) network and our composite model(CM) network with $g = 0.5$, compared to their analytical approximations. Here, we use $n = 1000$ and $k = 10$. This value of n is really too small to produce a very good analytical approximation of $L(p)$ on any portion of the interval. For larger n we would see a closer approximation for a large portion of the interval, which would then start to diverge as p grows larger [4].

Others have embraced the approach by Newman and Watts, constructing models where random edges are added rather than rewired. Jespersen and Blumen developed a generalized model introducing a distance dependency, where instead of the additional edges being uniformly randomly distributed the probability $p(l)$ of two vertices being connected follows a power law. More precisely, new links are added according to $p(l) = \frac{a}{l^\alpha}$, where l is the distance between the vertices and a is a normalization constant. In the limit $\alpha \to 0$ the small-world network of Newman et al. is obtained. In their work they studied the behaviour of a random walker on the network and was able to identify regions at which the properties of the network was much similar to either a small-world network or a regular network. The transition between these regions appeared to be around $\alpha \approx 2$ [6]. Sen and Chakrabarti came to the same result, also studying networks in which long-range connections are added according to a power law, while further finding that non-small-world behaviour may appear in the small-world region if there are restrictions on the total number of long range connections in the network. Thus the phase boundary α being a function, $\alpha = \alpha(q)$, and $\alpha(q) \to 2$ for large q, where q is the mean number of long-range connections [7].

So far we have considered the properties, $L(p)$ and $C(p)$ of two different network models. Next we will simulate the spread of a disease on each of the networks while varying the probability of rewiring, p. The disease model is the following. In each time step each infected person interacts with all neighbours who become infected with probability s. After each infected has interacted with their neighbours each infected can recover with probability q. An individual who has recovered can not get the disease as long as they stay recovered. Finally the recovered individuals becomes healthy with probability h.

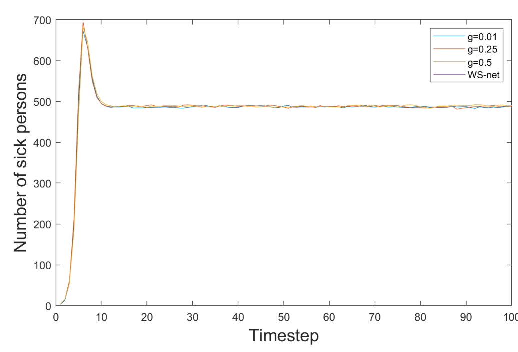

Figure 4: Number of sick persons at each unit of time for four network structures - the WS small world network and our extended model with combined random and preferential attachment with $g = 0.01$, $g = 0.25$ and $g = 0.5$. When simulating the spread the disease it starts at the vertex with index 1 and spreads to the neighbours of an infected person with probability s. An infected person recovers with probability r and can not get infected as long as it belongs to the recovered. A recovered individual becomes healthy with probability h.

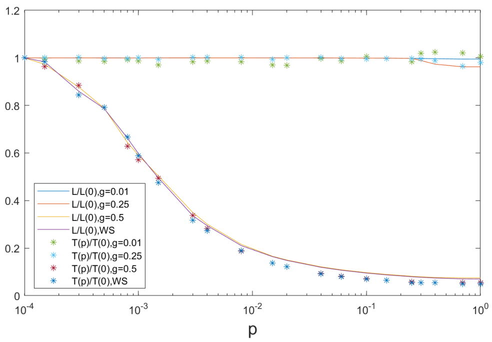

Figure 5: Time until every vertex has become infected when $s = 1$ and $q = h = 0$ compared with the path length for the Watt Strogatz network and the combined network structure with different values of g.

In Figure 1 we saw that as the value of g in the extended model approaches 0.5 the network structure becomes closer to the original small-world network by Watts and Strogatz. Thus replacing all random rewiring to preferential seem to have little effect. However, combining random and preferential rewiring shows a different result. It seems that the small-world phenomenon of small characteristic path length combined with large clustering no longer occurs. Instead, for small values of g, our extended model seem to be closer to a regular network with clustering and path lengths both large. This type of transition between a regular and small-world region was also observed in networks in which long-range edges are added according to a power law, as shown by Jespersen and Blumen. Although we have not pinpointed the phase boundary in our model it seems that g must be close to its upper bound, 0.5, for any small-world behaviour to be noticeable. Further Figure 2 shows that the variance of the networks clustering coefficient decreases as g increases. The structure of the network does not appear to affect the total number of diseased individuals which reaches a steady state in all types of networks, as Figure 4 shows. Of more interest is the time it takes until everyone becomes infected which follows the same curve as the characteristic path lengths of the networks, seen in Figure 5. A disease spreads much more rapidly in small-world networks and in the Watts-Strogatz model it becomes obvious how little random rewiring is needed in a regular network for it to resemble a small-world network. Our extended model however shows less sensitivity to the rewiring and will for most cases behave as a regular network. This might be due to that duplicate edges are forbidden in the model. We believe that it would be interesting to compare our results to models without this restriction. It would also be interesting to see the result of new edges being *added* by preferential attachment instead of rewired.

References

[1] Duncan J. Watts and Steven H. Strogatz. "Collective dynamics of 'small-world' networks". In: *Nature* 393.6684 (1998), pp. 440–442. ISSN: 1476-4687. DOI: 10.1038/30918. URL: https://doi.org/10.1038/30918.

[2] Riitta Toivonen et al. "A model for social networks". In: *Physica A: Statistical Mechanics and its Applications* 371.2 (2006), pp. 851–860. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2006.03.050. URL: http://www.sciencedirect.com/science/article/pii/S0378437106003931.

[3] M. E. J. Newman and D. J. Watts. "Scaling and percolation in the small-world network model". In: *Phys. Rev. E* 60.6 (Dec. 1999), pp. 7332–7342. DOI: 10.1103/PhysRevE.60.7332. URL: https://link.aps.org/doi/10.1103/PhysRevE.60.7332.

[4] M. E. J. Newman, C. Moore, and D. J. Watts. "Mean-Field Solution of the Small-World Network Model". In: *Phys. Rev. Lett.* 84.14 (Apr. 2000), pp. 3201–3204. DOI: 10.1103/PhysRevLett.84.3201. URL: https://link.aps.org/doi/10.1103/PhysRevLett.84.3201.

[5] A. Barrat and M. Weigt. "On the properties of small-world network models". In: *The European Physical Journal B - Condensed Matter and Complex Systems* 13.3 (Feb. 2000), pp. 547–560. ISSN: 1434-6036. DOI: 10.1007/s100510050067. URL: https://doi.org/10.1007/s100510050067.

[6] S. Jespersen and A. Blumen. "Small-world networks: Links with long-tailed distributions". In: *Phys. Rev. E* 62.5 (Nov. 2000), pp. 6270–6274. DOI: 10.1103/PhysRevE.62.6270. URL: https://link.aps.org/doi/10.1103/PhysRevE.62.6270.

[7] P. Sen and B. K. Chakrabarti. "Small-world phenomena and the statistics of linear polymers". In: *J. Phys. A: Math. Gen.* 34 (Sept. 2001), pp. 7749–7455. DOI: 10.1088/0305-4470/34/38/303. URL: https://iopscience-iop-org.ezproxy.its.uu.se/article/doi.