Lecture 3: Small world effects

Eduardo López

After presenting some of the basic ingredients of a general formulation of networks in terms of link indicators, and then using these to define many important properties of networks, in this lecture we begin to use some of these techniques in more concrete ways. While a number of results will still feel theoretical, we being to develop tools that are useful in applied settings.

1 The Milgram experiment

Let us start this lecture with the study of path length in social networks. It also happens to be the origin of the name "six degrees of separation".

The phrase "six degrees of separation" captures the notion that any pair of randomly chosen individuals in the population is separated by no more than a few intermediaries. The idea that, socially speaking, the world is small and increasingly getting smaller is not new, but it was not until the 20th century that a measurement of its *smallness* was attempted. Stanley Milgram [1] designed a study to test how many intermediaries it would take to connect people in Omaha, NE, and Boston, MA, and limiting the intermediaries to be on a first name basis with the senders.

Let us cite the abstract of Milgram's article with Travers [2]:

"Arbitrarily selected individuals (N=296) in Nebraska and Boston are asked to generate acquaintance chains to a target person in Massachusetts, employing the "small-world method" (Milgram, 1967). Sixty-four chains reached the target person. Within this group the number of intermediaries between starters and targets is 5.2. Boston starting chains reach the target person with less intermediaries than those starting in Nebraska; subpopulations among the Nebraska group do not differ among themselves. The funneling of chains through sociometric "stars" is noted, with 48 per cent of

chains passing through 3 persons before reaching the target. Applications of the method to studies of large scale social structure are discussed."
The fact that Travers and Milgram find that there are on average 5 individuals that act as intermediaries between starter and target leads to the conclusion that there are chains of 6 individuals that include starter and target persons. It turns out that statistically speaking there is a small inconsistency here because there are chains that never make it to their targets and this would mean that the average path length is indeed larger.

However, what this work achieved was to confirm the suspicion that was already present in the social sciences, that acquaintanceship leads to social *small worlds*. This is very important because things that travel in social networks would be able to travel quickly among nodes of the network: infectious diseases, critical information, etc. The phrase six degrees of separation has become the colloquial reference to this very important notion.

2 Building some theory behind this small-world effects

We start a more in-depth study of this phenomenon by focusing on a theoretical model developed by Watts and Strograts [3] that tests what are the main features of a network that exhibits the small world effect.

In those days, as Watts and Strogatz recognized, there were two well accepted limits that a network was known to have. On the one hand, a network could be fully random, such as the mathematical model due to the work by Erdős, Rényi and separately by Gilbert. These random networks were characterized by the idea that any node pair has an equal probability to be connect by a link. To be clear, Erdős, Rényi or Gilbert random networks are not one network, but rather entire families of networks that share very similar network features. These networks were known to have small average shortest path length between pairs or randomly chosen nodes. However, they lack more complicated local structure, such as that seen in real-world social networks; for instance, clustering in these kinds of networks approaches 0 as n grows.

The other limit of networks was embodied in actual empirical networks measured through data collection. Many examples of empirical networks were available at the time of Watts and Strogatz's investigations, as social network analysis was already a mature field. However, one of the limiting features of the networks was that they were universally small. It was tremendously hard in those days to collect data for social networks that

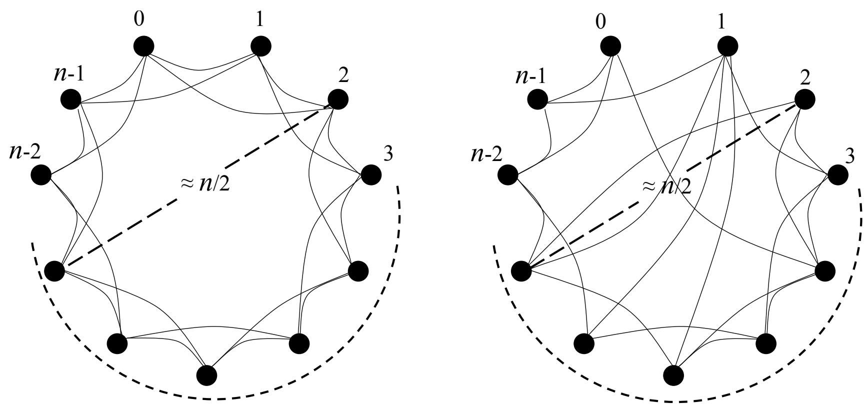

Figure 1: Illustration of the structure of the Watts-Strogatz model network with $k = 4$ before any rewiring has taken place (left), and then after some random rewiring has taken place (right). Locally, after the random rewiring, there are still triangles and thus clustering. However, those node pairs that would require long shortest paths to be reached because of needing to travel through half of the network now become short due to the introduction of "short-cuts" from the random rewiring.

could exceed anything beyond the hundreds of individuals. Therefore, Milgram's experiment stood as the main example of a large scale study of a social system in which there were simultaneously many nodes and an experimental test of the network properties of the underlying network. Note however that the network in which Milgram's messages were travelling was still unknown; only those paths that were exploited could be seen.

To build a bridge between random networks that could be of any size, but ultimately random, and empirical networks with real-world structure that stayed limited to small sizes (were automatically path length would remain low), Watts and Strogatz developed a new model that was able to preserve local complex structure in the form of large clustering coefficient and at the same time, transition seamlessly between between a large and a small-world. This network has come to be known as a *the small-world network* colloquially, although this is an unfortunate term because many networks display this property. In these notes, we do not use this term to avoid abuse of the term and instead choose to simply call the model the Watts-Strogatz or WS network.

To begin to understand the network and the model, let us study Fig. 1

for a moment. The figure follows closely Fig. 1 of Ref. [3]. The characteristics of the network on the *left* side of the illustration are that, with the nodes labelled sequentially $0, 1, \ldots, n-1$, each node is connected to its most immediate $k/2$ neighbors either side of the node according to this labelling. The parameter k indicates the degree of the nodes (all nodes have the same degree k), and it is conditioned to be even and $\leq n-1$. These labels act as coordinates with the caveat that the network is ring-like so, for instance, nodes 0 and $n-1$ are only one link apart, 0 and $n-2$ are two links apart, nodes 1 and $n-1$ are also two links apart, etc. Symbolically, nodes are connected according to the rule

$$a_{ij} = \begin{cases} 1, & i_\uparrow - i_\downarrow \leq k/2 \textbf{ or } n + i_\downarrow - i_\uparrow \leq k/2 \\ 0, & \text{otherwise} \end{cases}, \quad i \neq j, \quad k \text{ even}, \tag{1}$$

where

$$i_\downarrow = \min(i,j), \quad i_\uparrow = \max(i,j) \tag{2}$$

To reiterate, two nodes are connected if they are within $k/2$ links of each other. Note that $k \geq 2$, and when $k = 2$ exactly, WS reduces to a ring of n nodes. The following `python` code can generate Watts Strogatz networks in their ordered state

```
>>> def WS(n,k):
...     R=nx.Graph()
...     for i in range(n-1):
...         for c in range(1,k/2+1):
...             j=i+c
...             if j>n-1:
...                 j=j-n
...             R.add_edge(i,j)
...      return(R)
```

The network in its ordered state $r = 0$ (a measure of disorder explained below) has a local clustering that can be calculated with some care. Since the main purpose of this lecture is to explain the origin of the small-world effect we not dwell on the details of this calculation. For reference, I've placed the calculation of the local clustering in the appendix section. The main result is that

$$c_i = \frac{3}{4}\left(\frac{k-2}{k-1}\right) + 2\frac{\binom{\frac{3}{2}k-n+2}{2}}{k(k-1)}, \tag{3}$$

which has two behaviors, one for $k < 2n/3$ and another for $n-1 \geq k \geq 2n/3$. In the first behavior, only the first term of the equation above contributes,

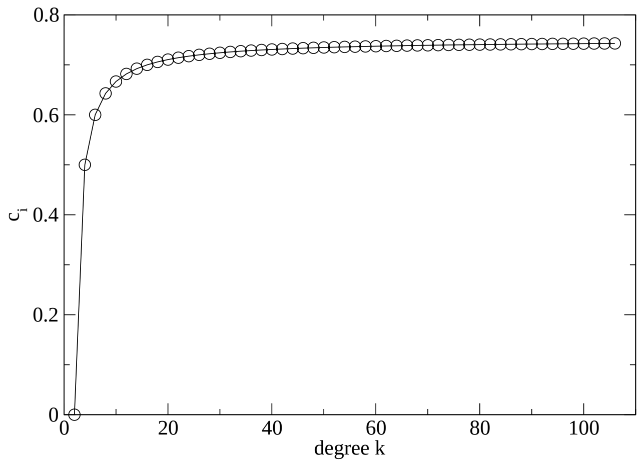

Figure 2: Local clustering coefficient for any node in the ordered WS network with $n = 1000$ and $2 \leq k \leq 102$. Clustering quickly becomes large, a feature that is not attainable with a typical random network with equivalent values of n and k.

and has a limit for large k (but below $2n/3$) equal to $3/4$. When the extra term comes into play, c_i starts to grow again until reaching 1 when k becomes $n - 1$. This calculation would also be done numerically, but having the equations makes it unnecessary.

Equation 3 shows that indeed the WS network is endowed with a large local clustering coefficient when $k \geq 4$ (see Fig. 2). For instance, any WS with $n > 6$ and $k = 4$, $c_i = 1/2$ for each and every node, and the value continues to grow with k. When $k = n - 1$, $c_i = 1$, as one would expect since all nodes are connected to all other nodes.

Another important feature to track is the average shortest paths in the network. If we look at the left network in Fig. 1 (no disorder), we can see that in order to navigate this network we can take steps of length $k/2$. Since it is mainly a ring, two nodes diametrically opposed to each other in this ring are roughly $n/2$ "nodes apart" (in terms of the labels of the nodes) in either direction of travel (one has to mind whether n is even or odd, but for large n, $n/2$ is a very good approximation). In addition, since we can hop along a single link from node i to node $i + k/2$, we find that the longest shortest path of the network without disorder is approximately $(n/2)/(k/2) = n/k$. This is confirmed in Fig. 3 for a WS with $n = 1000$, $k = 4$, and $r = 0$, where the last point on the horizontal axis is 250, exactly as predicted by

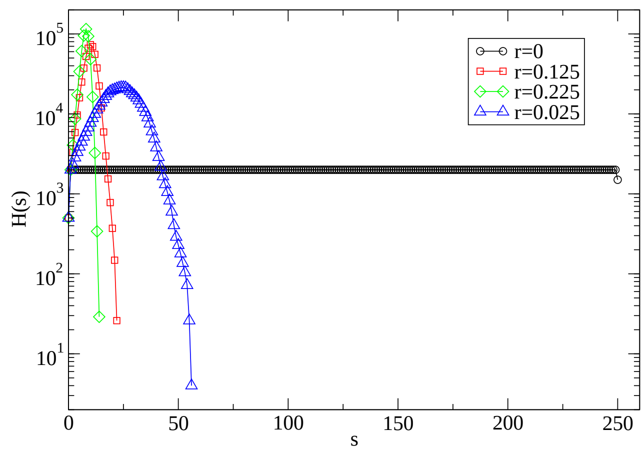

Figure 3: Histogram of shortest path lengths for increasing r in a WS network with $n = 1000$ and $k = 4$.

$n/k = 1000/4$.

To determine the average shortest distance in the ordered case, all we need to do is determine it for a single node to all other nodes of the network. This is because all nodes are in the same situation, and thus they too would have the same distances. notice that there are direct links between any pair of nodes i and j with labels that differ from one another by Δ in the range $1 \leq \Delta \leq k/2$ (minding the cycling of labels if we need to pass through node 0 in either direction). If Δ is larger than $k/2$ but less than or equal to $2 \times k/2 = k$, two hops are needed, one to a appropriate node inside the nearest $k/2$, and then a single hop from that node to the node outside the nearest $k/2$. This continues for any node along the ring until the nodes labelled diametrically opposed to i, that is $i \pm n/2$ (again minding the cyclic labelling). We can see this uniformity in travel distances on display in Fig. 3 for $r = 0$. Thus, the shortest distances are uniformly distributed between 1 and n/k. This makes the average distance from any node to any other node satisfy

$$\langle s \rangle = k \frac{\sum_{s=1}^{n/k} s}{n} = \frac{k \binom{n/k+1}{2}}{n} = \frac{1}{2}\left(\frac{n}{k} + 1\right) \approx \frac{n}{2k}, \tag{4}$$

where the k in front of the expression takes into account the fact that each block of nodes in a group of $k/2$ of them encompass k nodes altogether (taking into account opposite directions of movement from i. This $\langle s \rangle$ for the ordered WS network is what is known as a *large-world network*, a network

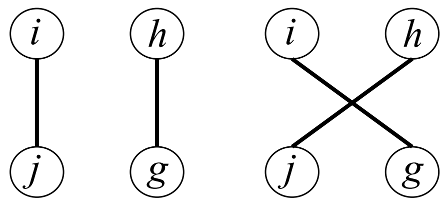

Figure 4: Link swapping procedure used by Watts and Strogratz

in which the distance between nodes is proportional to the number of nodes. We find such large world networks in many contexts. A very simple day-to-day example is a network of roads where each intersection is a node (allowing us to change paths in various directions) and road segments are links. In road networks, the number of intersections that a typical traveller needs to cross going from a random origin to a random destination is a polynomial function of n, the number of intersections. Typically, the function is proportional to $n^{1/2}$, slower in growth than the result in Eq. 4 so WS in its ordered state is a particularly bad network to commute.

To introduce disorder, Watts and Strogatz allowed links to be rewired randomly with a rewiring rate r. Given that this network has $m = nk/2$ links, r means that, on average, $rnk/2$ links are different than those in the fully ordered WS network (assuming n is large and k is small by comparison). The way that the rewiring takes place is by taking two links, say ij and hg, and swapping the ends of them so that in the rewired network, the links are now ig and hj (see Fig. 4. For the rewiring to be indeed random, the two original node pairs ij and hg can be anywhere in the network. This means that as the random rewiring takes place, nodes that originally were a number of hops away from each other can now be much closer.

What did Watts and Strogatz find? The ultimate effect of even a small amount of random rewiring is that WS networks, while preserving clustering, can also be made to display very short paths. And because these path lengths are highly influenced by any available short-cuts, it means that the small-world phenomenon is really a consequence of the presence of a *sufficient* amount of disorder/randomness in the network. Just to appreciate how quickly disorder can shorten the paths, see Fig. 3. Figure 5 shows the behaviour of $\langle s \rangle$ as a function of n, and compares it to a logarithmic function

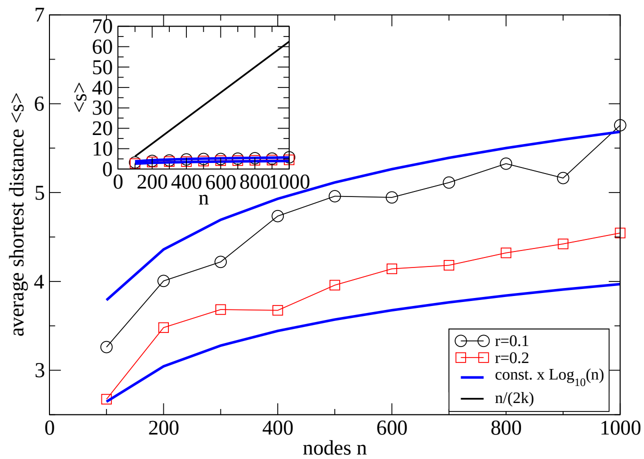

Figure 5: $\langle s \rangle$ for simulated WS networks with n between 100 and 1000 for $r = 0.1, 0.2$. Also, two curves with functional form $c \log n$ with values of $c \approx 1$. In the inset, the main plot is repeated but another line is added, which corresponds to $n/(2k)$, the average distance calculated in Eq. 4. Clearly, the logarithmic function is very slow compared to the linear, large-world distance.

of the form $c \times \log n$. Symbolically,

$$\langle s \rangle \propto \log n \quad \text{[disordered WS network]}, \tag{5}$$

where $\propto$ is the symbol for proportionality. This last result means that the network is indeed small-world.

A thorough analysis of how the randomness changes path lengths requires some mathematics that are, once again, not central to the lecture. The basic notion behind the calculation is the following: if rewiring occurs randomly, then for any given node in the network, a shortcut to a place that is, on average, $n/(2k)$ links away is likely to occur in $rn/2k$ random rewirings. The mathematical origin of the logarithmic behavior is that on average the trip across any link in the network makes accessible a constant number μ of new nodes to be reached. This constant rate of new nodes per link crossed means that after ℓ crossings, the number of accessible nodes becomes

$$n_\ell \propto \mu^\ell. \tag{6}$$

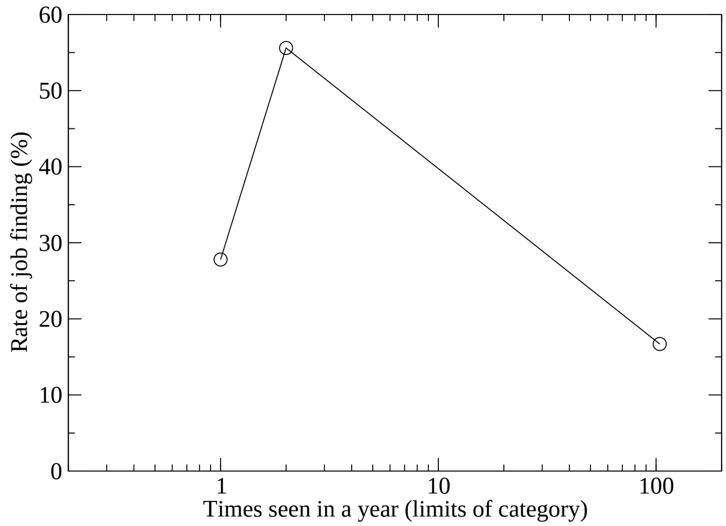

Figure 6: Rate of job finding through social contacts. Granovetter categorized the frequency of contact as rarely (meaning less than once a year), occasionally (once a year or more but less than twice a week), and often (twice a week or more). For often we use 104, for occasionally we use 2, and for rarely we use 1.

When ℓ is large enough, n_ℓ becomes of the order of n, and at that point

$$n \propto \mu^s \quad \Rightarrow \quad s \propto \log n. \tag{7}$$

How should we understand this phenomenon? The high clustering in empirical social networks means that the balance of preference of how to use links is biased towards using said links to close triangles. This means that there are not as many links available to form connections in the network between generally unrelated nodes. However, the WS model shows that very few links are needed to make the shortcuts that create a small-world effect.

3 Weak Ties

The work by Watts and Strogatz connects strong with the ideas of Mark Granovetter [4] who in the early 1970s discussed for the first time the importance of week ties in social networks. This article essentially changed the way that social networks were viewed.

As explained in the article, important social notions require the study of weak ties, even if many social phenomena takes place only among individuals that have strong ties among them. For instance, while it is true that most of our time is spent among our close friends and family, aspects of our life such as information diffusion are highly affected by these weak times. Finding a job, for instance, is most effectively done through contacts we rarely see who would have information about positions that our close contacts would not have (see Fig. 6 for a representation of his results on this respect). Or more concretely, the information that our close contacts have is very similar to ours. It is only when we reach out to farther contacts that we are likely to find new information.

Over time, there has been a move to also characterize weak ties as being *long ties* because it is recognized that they connect parts of a social system that are distant from one another in various ways (homophily, cohesion, etc.). The designation of long ties can found in the literature with some regularity.

Ultimately, it is important to understand that social systems posses both strong (short) and weak (long) ties, and both are needed to maintain the structure of large social networks. Each kind of tie has influence over some aspects of social interaction.

Appendix

To calculate the clustering of each local node in the WS network before disorder sets in, we need to consider how many triangles involve each node. This is a straightforward calculation that, nevertheless, requires a bit of care.

Each node i is connected to k other nodes, which means we can immediately calculate the number of v-shapes involving i as $\binom{k}{2}$. To calculate $t_i = t$ (all nodes have the same number of triangles t connected to them), we must determine which of the k neighbors of i are connected to each other.

There is one certain contribution to the triangle count and a second possible contribution, depending on the relation between n and k. To see why this would occur, notice that from a given node i, one of the two neighbors with a label that differs most with i is node $h = i - k/2$, and other node is $j = i + k/2$. Now, for a critical value k_c of k that is large enough, h and j can also be such that $j + k_c/2 = h$ (remembering to cycle the labels around once we cross node 0). At this point, the ordered WS network starts to show triangles formed at the far end of i. From the perspective of the entire network, note that the critical value k_c is such that it splits the network into three sections, and therefore, we anticipate that $3 \times k_c/2 = n$, leading to $k_c = 2n/3$. We come back to this later.

Let us now embark in calculating these contributions to the total triangle count. Let us start by assuming that $n \gg k$ and thus there is no need to worry about the situation just explained. In this case, all we need to worry about are the possible triangles forming close to node i.

The near contribution is equal to

$$t_{\text{near}} = \frac{3}{8}k(k-2). \tag{8}$$

On the other hand, the far contribution, if any, is equal to

$$t_{\text{far}} = \binom{\frac{3}{2}k - n + 2}{2}, \tag{9}$$

where, conveniently, the definition of binomial coefficients takes care of the possibility that this far contribution is not present when $-n + 3k/2 < 0$[1].

These equations then allow us to write a total value for the local clustering coefficient, given by

$$\begin{aligned} c_i = \frac{t}{\binom{k_i}{2}} &= 2\frac{t_{\text{near}} + t_{\text{far}}}{k(k-1)} = \frac{2t_{\text{near}}}{k(k-1)} + \frac{2t_{\text{far}}}{k(k-1)} \\ &= \frac{3}{4}\frac{k(k-2)}{k(k-1)} + 2\frac{\binom{\frac{3}{2}k-n+2}{2}}{k(k-1)} = \frac{3}{4}\left(\frac{k-2}{k-1}\right) + 2\frac{\binom{\frac{3}{2}k-n+2}{2}}{k(k-1)}. \end{aligned} \tag{10}$$

While $3k/2 < n$, in the limit of large k, $c_i \to 3/4$. On the other hand, when $k \geq 2n/3$ (k is 66.67% of n), clustering starts to increase once again as now k is so large that the farthest nodes from i begin to connect again as their distance along the part of the network diametrically opposed to i is now of order $k/2$. The effect is that new triangles appear. Mathematically, c_i increases, starting from $3/4$, as a function to t_{far}.

[1] To be precise, by definition a binomial coefficient

$$\binom{a}{b} = 0, \quad a < b$$

which automatically deals with the possibility of no contributions from t_{far} from Eq. 9.

References

[1] Milgram S (1967) The Small World Problem. Psychology Today 22, 61-67

[2] Travers J, Milgram S (1969) An experimental study of the small world problem. Sociometry 32, 425-443.

[3] Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393, 440–442.

[4] Granovetter M (1973) The Strength of Weak Ties. American Journal of Sociology 78, 1360-1380.