

NETWORKS AND COMPLEXITY

Solution 7-7

*This is an example solution from the forthcoming book Networks and Complexity.
Find more exercises at <https://github.com/NC-Book/NCB>*

Ex 7.7: The most common cycle [3]

We know our random graphs are locally treelike, but we also know that if $z > 1$ then there must be some cycles. To explore this, consider an ER graph with $N = 101$ nodes and mean degree $z = 2$. Find out what the most common length l for a cycle is.

Solution

So we expect that there few very short cycles because the network is locally tree-like. We also expect that there are few very long cycles because the network isn't actually that large. So there must be some kind of optimum in the middle. For a first rry we can use the approximate formula

$$n_{\text{motif}} = \frac{1}{s} z^k N n - k \quad (1)$$

We have already figured out above that for a cycle of length l (an l -cycle) the symmetry number is $s = 2l$. Moreover $n = k = l$ is true in every cycle. So the equation simplifies to

$$n_{l\text{-cycle}} = \frac{1}{2l} z^l \quad (2)$$

This is a weird result. If we increase l we find ever more numerous cycles. But this makes no sense because there can't be any cycles that have a length greater than 101. OK, here is an idea. The error of the approximate formula grows for large motifs. Likely this makes our estimates for long cycles inaccurate. So let's use the statistically exact formula instead.

$$n_{\text{motif}} = \frac{n!}{s} p^k B(N, n) \quad (3)$$

we know $N = 101$ which allows us to $p = z/(N - 1) = 2 \cdot 10^{-2}$. We can now for example compute

$$n_{3\text{-cycle}} = \frac{3!}{6} (2 \cdot 10^{-2})^3 \frac{99 \cdot 100 \cdot 101}{1 \cdot 2 \cdot 3} = 8 \cdot 10^{-6} \frac{999900}{6} = \frac{8}{6} 0.9999 \quad (4)$$

so that's just 1 or two three-cycles. However, maybe we have learned something from this calculation. Notice that the $n!$ enters from the denominator of the Binomial Formula will always cancel with the $n!$ that is in the front of the motif equation. we can thus use

$$n_{l\text{-cycle}} = \frac{1}{2l} p^l 101 \cdot 100 \cdots 101 - l \quad (5)$$

Now consider what happens when we increase l we multiply one more factor of p the denominator of the fraction in the beginning increases by 2 and we multiply a factor of $101 - l$ in the end. Does this mean there are more or less $l + 1$ -cycles than l -cycles? Let's consider the ratio

$$\frac{n_{l+1\text{-cycle}}}{n_{l\text{-cycle}}} = \frac{\frac{1}{2l+2} p^{l+1} 101 \cdot 100 \cdots 101 - l \cdot 101 - l - 1}{\frac{1}{2l} p^l 101 \cdot 100 \cdots 101 - l} = \frac{l+1}{l} p (100 - l) \quad (6)$$

The first factor hardly matters because $(l + 1)/l \approx 1$ for all but the first few l . The factor p is 0.02 (so a division by 50). However for the first few l the last factor constitutes a multiplication of almost 100. for example for $l = 3$ we have

$$\frac{n_{4\text{-cycle}}}{n_{3\text{-cycle}}} = \frac{4}{3} \frac{97}{50} \approx 2.6 \quad (7)$$

so we expect approximately 2.6 times as many 4-cycles as there are 3-cycles. We can now also see that the ratio will stay greater than one until the last factor drops below 50. So let's try

$$\frac{n_{51\text{-cycle}}}{n_{50\text{-cycle}}} = \frac{51}{50} \frac{50}{50} = 1.02 \quad (8)$$

so there are just a tiny bit more cycles of length 51 than of length 50. Let's try the next step

$$\frac{n_{52\text{-cycle}}}{n_{51\text{-cycle}}} = \frac{52}{51} \frac{49}{50} \approx 0.999 \quad (9)$$

So there are fewer cycles of length 52 than 51. So the cycle of length 51 is actually the most common cycle in the network. So the most common length of cycle is one that visits half the nodes if the network. The number of these cycles that we expect is very high, because there are so many ways in which we can select 51 nodes from 101 nodes.

OK we are really done with this exercise, but if we wanted to we could write

$$\frac{n_{l+1\text{-cycle}}}{n_{l\text{-cycle}}} \approx \frac{z}{N} (N - l) \quad (10)$$

where we used $p \approx z/N$ and $(l + 1)/l \approx 1$. We can now solve for the point where this ratio becomes 1

$$1 = z(1 - l/N) \quad (11)$$

$$1/z = 1 - l/N \quad (12)$$

$$l/N = (z - 1)/z \quad (13)$$

so

$$l = \frac{z - 1}{z} N \quad (14)$$

So in networks with mean degree greater than one the length of the typical cycle scales with the number of nodes and in networks with significant mean degree $z \geq 2$ the typical cycles visit the majority of nodes.