Networks and Complexity

Solution 6-10

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

Ex 6.10: Diameter of heterogeneous networks [4]

Revisit our estimate of the network diameter (average node distance) from the previous chapter. In the derivation we assumed that our initial node has degree z. Further we assumed that if follow the links form this node we reach nodes that have in average z links in addition to the one we are arriving on. This means that we assumed $q = z$. Let us now relax this assumption. Follow the reasoning of Chap. 5 to derive a formula for the diameter of a locally treelike network with a given number of nodes N, mean degree z, and mean excess degree q, but clustering coefficient $c = 0$. Then, use your formula to estimate the diameter of a network with 10000 nodes, of which half have degree 2 and half have degree 8.

<u>*Solution*</u>

Following the same reasoning as in the previous chapter we first estimate n_d the expected number of nodes at distance d from a starting node for different values of d.

$$n_0 = 1 \tag{1}$$
$$n_1 = z \tag{2}$$
$$n_2 = zq \tag{3}$$
$$n_3 = zq^2 \tag{4}$$

We start with one person at distance 0, who has z neighbors. These neighbors have q additional neighbors, who in turn have q additional neighbors. Continuing this logic we find

$$n_d = zq^{d-1} \tag{5}$$

for all $d > 1$. To find the diameter D we ask at what distance we reach practically everybody. Which leads to the condition

$$N = zq^{D-1} \tag{6}$$

We can solve this equation for D by dividing by z and applying a logarithm:

$$N/z = q^{D-1} \tag{7}$$
$$\log_q(N/z) = D - 1 \tag{8}$$
$$D = 1 + \log_q(N/z) \tag{9}$$

If we want we can use the logarithms rules to rewrite this as

$$D = 1 + \log_q(N) - \log_q(z) = 1 + \frac{\ln(N) - \ln(z)}{\ln q}. \tag{10}$$

The example network from the question has $N = 10,000$ and

$$p_k + 0.5\delta_{k,2} + 0.5\delta_{k,8}. \tag{11}$$

We compute

$$z = \sum kp_k = 0.5 \cdot 2 + 0.5 \cdot 8 = 5 \tag{12}$$

and construct the excess degree distribution

$$p_k = (k+1)p_{k+1}/z = (\delta_{k,1} + 4\delta_{k,7})/5 = 0.2\delta_{k,1} + 0.8\delta_{k,7} \tag{13}$$

and the mean excess degree

$$q = \sum kq_k = 0.2 + 7 \cdot 0.8 = 5.8. \tag{14}$$

Substituting into our diameter formula we find

$$D = 1 + \log_{5.8}(10000/5) = 1 + \ln(2,000)/\ln(5.8) \approx 5.32, \tag{15}$$

which answers the second part of the question.

As a little bonus let's compare this to our previous network, which did not take the heterogeneity of the degree distribution into account. In the previous chapter we already derived

$$D = \log_z(N) \tag{16}$$

substituting yields

$$D = \log_5(10000) \approx 5.77 \tag{17}$$

So by taking the heterogeneity into account, we discovered that the network is actually a little bit smaller than we would have expected if we neglected the heterogeneity. In this example the difference is small because the heterogeneity is mild ($q \approx z$). By contrast, very strongly heterogeneous networks can have very small diameters, we will see this again in a later lecture.