Networks and Complexity

Solution 12-9

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

Ex 12.9: Price's model [3]

When a scientific paper is published it cites m other papers. Papers that have already been cited more often have higher visibility and hence attract new citations at a higher rate. We assume that a paper with k citations has the visibility $v = a + k$, where a represents an intrinsic visibility the paper has even without citations.

a) Let p_k be the in-degree distribution of the papers, i.e. p_k is the proportion of papers that have been cited k times. Consider a single citation from a new paper. Assume that the chance that a paper is cited is proportional to its visibility. Show that the probability v_k that a single new citation cites a paper that has previously had k citations is

$$v_k = p_k \frac{a+k}{a+z}$$

Solution

From the question we know that the probability is proportional to the visibility $a + k$ is should also be proportional to the proportion of target nodes with the respective degree. So we can use the Ansatz

$$V_k = Ap_k(a+k) \tag{1}$$

where A is a factor that we have to pick such that the distribution is normalized, i.e.

$$1 = \sum v_k = \sum Ap_k(a+k) = A(a+z) \tag{2}$$

where z is the mean in-degree. We can now read off

$$A = \frac{1}{a+z} \tag{3}$$

This means

$$v_k = p_k \frac{a+k}{a+z} \tag{4}$$

b) Show that in the long run the in-degree distribution p_k approaches a power-law for $k \gg 1$, Also determine the power-law exponent.

Solution

Following the same reasoning as for the BA-model in the lecture yields the ODE

$$\dot{c}_k = \frac{1}{N}\left(-c_k + \frac{mp_k(a+k)}{a+z}\right) \tag{5}$$

Here we need to be careful because the mean in-degree will only approach $z = m$ instead of $z = 2m$ for the undirected network. Therefore,

$$\dot{c}_k = \frac{1}{N}\left(-c_k + \frac{mp_k(a+k)}{a+m}\right) \tag{6}$$

In the steady state

$$c_k = \frac{mp_k(a+k)}{a+m}. \tag{7}$$

and hence

$$p_k = c_k \frac{a+m}{m(a+k)} \tag{8}$$

which together with $\partial_k c_k = -p_k$ yields

$$\partial_k c_k = -c_k \frac{a+m}{m(a+k)}. \tag{9}$$

We integrate

$$\begin{aligned} \int \frac{1}{c_k} \mathrm{d}c_k &= -\frac{a+2m}{m} \int \frac{1}{a+k} \mathrm{d}k & (10) \\ \ln(c_k) &= -\frac{a+m}{m} \ln(a+k) + C & (11) \\ c_k &= A(a+k)^{-1-a/m} & (12) \end{aligned}$$

where $A = \mathrm{e}^C$. Hence we can compute

$$p_k = -\partial_k c_k = B(a+k)^{-2-a/m} \tag{13}$$

where $B = A(1 + a/m)$ is another prefactor. In the limit $k \gg a$ we get the power law

$$p_k \sim k^{-2-a/m} \tag{14}$$

So, distribution of citations that we get from Price's model follows a power law with exponent $\gamma = -2 - a/m$.