Social and Technological Networks Edinburgh, 2018

Notes 1. Random Graphs.

Rik Sarkar *Notes*

Clustering coefficient. In class we touched upon the fact that in an ER graph the CC tends to zero. let's prove that here.

Q 1. *Show that a connected graph has $\Omega(n)$ triads (counting both open and closed).*

In class we saw that a triangle or closed triad is is three vertices a, b, c with all edges ab, bc, ca between them. The number of possible triangles or triads is clearly $\binom{n}{3}$, which is $\Theta(n^3)$. The probability that a particular triangle exists is p^3.

Q 2. *What is the expected number of triangles or closed triads in the graph?*

Q 3. *Clustering coefficient is the ratio of number of closed triads to number of all triads. Show that for ER graphs with $p = \frac{\ln n}{n}$ (where n is an unknown variable) the clustering coefficient cannot be bounded from below buy a constant. (That is, there is no constant number such that CC is always greater than that.)*

Emergence of giant component. ER graphs show the emergence of a giant component at $p = 1/n$, where one component is relatively large, while other components are mostly small. (in fact, at this point, we know that many vertices are isolated. Why?)

This is another threshold phenomenon. At $p = (1 - \varepsilon)/n$, w.h.p, no giant component exists, and existing components are of size $O(\log n)$. At $p = (1 + \varepsilon)/n$, the size of giant component is about εn.

Giant component is a phenomenon, not only in random graphs, but also in many real graphs. In many real situation, the graph is not fully connected, but does contain one component containing a large fraction of the nodes. Thus, this is the portion used in network analysis.

Q 4. *Write code in your iPython notebook to create the threshold plot for isolated vertices or giant components. How does the look of the plot change with number of vertices in the graph?*