Social and Technological Networks Edinburg, 2015

Lecture 3. Graph Properties & Random Graphs, continued.

Rik Sarkar *Class notes*

In class, we covered the following ideas:

1. Expander graphs have small diameter.
2. Random graphs have small clustering coefficient in expectation
3. The degree distribution of any vertex in a random graph follows a normal distribution.
4. The web graph shows a power law distribution.

In all these discussions for random graphs, we are interested in random graphs close to the connectivity threshold, ie. average is degree a little above $\ln n$. We can consider $p = (c \ln n)/(n-1)$, for a constant c. As we know, below this, the graph has many isolated vertices, which is not so interesting. much above the threshold, the graph approaches a complete graph, which is also not so interesting. It is close to the threshold, that we see interesting properties.

1 Expansion

The proof of small diameter for expanders can be found in the notes linked from the web page. In the following exercise we show the expansion properties of ER graphs.

*** Exercise 1.1.** *Show that an ER random graph of n nodes has an expansion rate larger than some constant. [Hint: What can you say about the number of edges from any set S to nodes in $V \setminus S$?]*

2 Clustering coefficient

We showed that the expected clustering coefficient of ER graphs at $p = (c \ln n)/n$ is vanishingly small. (taking n instead of $n-1$ in the denominator simply makes calculations easier, does not change the essential; properties. You can try it with $n-1$ if you wish.) The analysis goes as follows.

There are $\binom{n}{3}$ possible triples in G. The probability that all 3 edges of a triple exist is p^3. Thus the expected number of closed triads is: $\binom{n}{3}p^3$. Simplifying:

$$\begin{aligned} E[\#\text{ closed triads}] &= \binom{n}{3}p^3 \\ &\leq n^3 p^3 \\ &\leq n^3 (c \ln n)^3 / n^3 \\ &\leq (c \ln n)^3 \end{aligned}$$

Next, observe that the total number of all triads (open and closed) in a connected graph is $\geq c'n$ for some constant c'.

Exercise 2.1. *Prove the statement above.*

(One of you pointed out after class that this is not $O(n)$. That is correct. This is not $O(n)$, it is $\Omega(n)$. Which is even better, since we are trying to show that the total number of triads is large.)

Thus,

$$CC = \frac{\#\text{ closed triads}}{\#\text{ all triads}} \leq \frac{(c\ln n)^3}{c'n}.$$

This ratio approaches arbitrarily close to zero as n grows sufficiently large. And since CC is smaller, it must be vanishingly small too.

Local tree structure. We can show that not only are cycles of length 3 unlikely, cycles of length k are also unlikely in random graphs.

Suppose we are interested in cycles of length k through a vertex v. Each step of this cycle is one out of $n-1$ nodes. So we can put an upper bound on the number of possible cycles as $\leq n^{k-1}$. Thus the probability that a cycle exists is less than $n^{k-1}p^k \leq n^{k-1}(\ln n/n)^k \leq (\ln^k n)/n$.

For any constant k, this is probability is small. Meaning that any constant sized neighborhood of a vertex v almost certainly looks like a tree if the graph is large enough.

Exercise 2.2. *For what sort of values of k are the neighborhoods likely to have cycles? That is, when does the above deduction fail, and when do cycles become likely?*

3 Degree distribution

The distribution of degrees, that is, the histogram of degrees $(1, 2, 3, \ldots)$ vs the number of nodes at each degree is another way to distinguish graphs.

In an ER graph, the degree of an individual vertex is drawn from a normal distribution or gaussian distribution. (This follows from the fact that the edges are independent, and by the central limit theorem, sum of independent random variables is a normal distribution.)

Note that this only says what to expect for one random variable. In class we plotted the distribution of degrees of all vertices in a graph, which are not independent. This actually produces Poisson distribution. There is a reason that that particular plot looks like a normal distribution too, but we will skip that for now.