Social and Technological Networks Edinburg, 2015

Lecture 2. Graph Properties & Random Graphs

Rik Sarkar *Class notes*

The plan is to cover some basic properties of graphs and see how they turn out for random graphs. We looked at:

1. Models of random graphs. Erdos-Renyi Model, configuration model.
2. Probability of isolated vertices in ER model.
3. Expansion: How fast sets can grow as we take neighbors at different distances.

Probability of isolated vertices. This turned out to have a threshold phenomenon at $p = (\ln n)/(n-1)$. Where isolated vertices are very likely at smaller values: $p = (1-\varepsilon)(\ln n)/(n-1)$ and unlikely at larger values: $p = (1+\varepsilon)(\ln n)/(n-1)$ for $\varepsilon > 0$. We proved the second case in class, showing that obove the threshold, isolated vertices are unlikely:

Theorem 0.1. *The probability that an ER graph of n vertices with $p = (1+\varepsilon)(\ln n)/(n-1)$ and $\varepsilon > 0$, has one or more isolated vertices is less than $n^{-\varepsilon}$.*

There was a small bug in the proof on the board in using "=" instead of "$\leq$". You can find the correct version in proof of Exercise 3 in solutions to sample problems (See material for Lecture 1).

The other theorem gives a lower bound for the probability of an individual vertex being isolated below the threshold:

Theorem 0.2. *The probability that in an ER graph of n vertices with $p = (1-\varepsilon)(\ln n)/(n-1)$ and $\varepsilon > 0$, a particular vertex v is isolated is $\geq (2n)^{-(1-\varepsilon)}$.*

We are omitting the proof of this for now (you can try proving yourself!). A corollary is that under the assumptions of the theorem, there will be $\geq n^{\varepsilon}/2$ isolated vertices.

In all these cases, we are are relying on the fact that a polynomial like n^{α}, for constant $\alpha > 0$ grows fast for large n. This is also the source of the notion of *high probability*:

Definition 0.3. *We say that a property X is true with high probability if $\Pr[\bar{X}] \leq \frac{1}{\mathrm{poly}(n)}$. That is, the probability of the complement of X being true is small. We can also rewrite as $\Pr[X] \geq \left(1 - \frac{1}{\mathrm{poly}(n)}\right)$.*

In such definitions, n is usually the size of the input in consideration. For example, number of vertices.

Just as $\text{poly}(n)$ can be assumed to be large, $\ln n$ can be assumed to be small. As a consequence $\text{poly}(\log n)$ can be considered much smaller than n or $\text{poly}(n)$ for large n. In many cases, it is reasonable to assume $\log n$ is about as small as a constant.

Exercise 0.1. *Show that* $\ln n = \Theta(\lg n)$*, and* $\lg n = \Theta \log(n)$*.*

$\ln n$, $\lg n$ and $\log n$ are the usual notations for $\log$ to base e, 2 and 10 respectively. The exercise above is to show that log functions to different constant bases differ only by constant factors.

Exercise 0.2. *Set up the ipython notebook on a system of your choice with networkx. Try it out.*

Exercise 0.3. *Write code to create plots showing the threshold phenomenon for existence of isolated vertices.*

* **Exercise 0.4.** *In class, we showed that above the threshold, isolated vertices are unlikely. However, this does not say that the graph overall is connected. It is possible that the graph itself stays in to two different connected components with no edge bridging the two.*

In this exercise, show that this is also unlikely. That is, above the threshold, with high probability, the graph has only a single connected component.