October 5, 2007
*Prepared by Adam Guetz*

## 1. Random graphs II: connectivity and diameter

In this lecture we will compute the threshold value for another basic property of random graphs: the value at which a graph is almost surely connected. We also show that almost surely the diameter of the graph is logarithmic when it is connected.

**Theorem 2.1** *Let $c > 0$ be a constant and $p = n^{-1}(\log n + c + o(1))$. Then for $G \in G(n,p)$, the probability that $G$ is connected tends to $e^{-e^{-c}}$ as $n \to \infty$.*

**Proof:** For $2 \leq r \leq n/2$, let $f(r)$ be the probability that $G$ has a connected component of size $r$.

$$\begin{aligned} f(r) &\leq \binom{n}{r} r^{r-2} p^{r-1} (1-p)^{r(n-r)} \\ &\leq \left(\frac{en}{r}\right)^r r^{r-2} p^{r-1} (1-p)^{rn/2} \\ &\leq e^r r^{-2} n^r \left(\frac{2\log n}{n}\right)^{r-1} \left(1 - \frac{\log n}{n}\right)^{rn/2} \\ &\leq \frac{(2e)^r}{2} r^{-2} n (\log n)^{r-1} n^{-r/2+1} \end{aligned}$$

We have to handle the case of $r = 2$ separately:

$$\begin{aligned} f(2) &\leq cn^2 \frac{\log n}{n} \left(1 - \frac{\log n}{n}\right)^{2(n-2)} \\ &\leq c \log n \frac{1}{n^2} \left(1 - \frac{\log n}{n}\right)^{-4} \end{aligned}$$

Therefore,

$$\sum_{r=2}^{n/2} f(r) = o(1)$$

By this calculation, we see that as $n$ goes to infinity the limit of

$$\Pr(G \text{ is connected}) = \Pr(G \text{ does not have any isolated vertices}).$$

In other words, with high probability the graph is a giant connected component with a few isolated vertices. The probability that a vertex is isolated is

$$(1-p)^{n-1} = \left(1 - \frac{\log n + c}{n}\right)^{n-1}$$
$$\leq (1-p)^{-1} e^{-\log n - c} \leq \frac{1}{n} e^{-c}$$

The expected number of ordered $k$-tuples of isolated vertices is

$$\frac{n!}{(n-k)!}\left(1 - \frac{\log n + c}{n}\right)^{n+(n-1)+\cdots+(n-k+1)} \rightarrow e^{-ck}.$$

By using the method of moments the distribution of isolated vertices converges to a Poisson distribution and the probability that no vertex is isolated is then $e^{-e^{-c}}$ ■

**Corollary 2.1** *Let $p = \frac{a \log n}{n}$, with a constant. If $a > 1$, then w.h.p. $G$ is connected. If $a < 1$, then $G$ is not connected with at least a constant probability.*

We now establish the threshold value for logarithmic diameter.

**Theorem 2.2** *If $p = \frac{c \log n}{n}$ for $c > 1$ then with high probability the diameter of $G \in G(n,p)$ is $O(\frac{\log n}{\log np})$.*

**Proof:** Fix a vertex $v$. Define

$$B_k(v) = \{u | \text{distance}_G(u,v) = k\}.$$

Remember the proof of the theorem for the existence of connectivity threshold. For small $k$, we expect $B_k(v)$ to be very close to $(np)^k$. By repeated application of Chernoff bounds one can show that for all $k$ such that $(np)^{k-1} < n^{-2/3}$,

$$\Pr(|B_k(v)| > (n^- p)^{k-1}) \geq 1 - o(1/n),$$

where $n^- = n - n^{2/3}$. Set $k = 1 + \frac{1+\epsilon}{2} \frac{\log n}{\log np}$.

$$\Pr(|B_k(v)| > n^{\frac{1+\epsilon}{2}}) = 1 - o(1/n).$$

Fix another vertex $u$. By a birthday paradox argument, the probability that $B_k(v)$ and $B_k(u)$ do not intersect

$$\Pr(B_k(v) \cap B_k(u) = \emptyset) \leq \left(1 - \frac{n^{\frac{1+\epsilon}{2}}}{n}\right)^{n^{\frac{1+\epsilon}{2}}}$$
$$\leq e^{-n^\epsilon}$$

■

The above argument is building on the fact that for small $p < n^{-\epsilon}$, the immediate neighborhood of every vertex in $G(n,p)$ is a tree. In other words, there are very few small cycles in these graphs. This does not mean that these graphs are not very well-connected. These graphs have a large number of very large cycles.

Consider the following process in the graph: remove all vertices of degree less than $k$ until there are no such vertices left. The remaining set is the $k$-core of the graph. The following statement shows that $G(n,p)$ graphs have a large $k$-core way before the connectivity threshold:

**Theorem 2.3** *For every $k \geq 2$, there are constants $\lambda_k$ and $c_k$ such that w.h.p. the giant connected component of $G(n,p)$ for $p = \lambda_k/n$ has a $k$-core with $c_k n$ vertices.*