May 11th, 2004
Kevin Costello

Math 262 Lectures Notes

The Rise of the Giant Component in $G(w)$

1 Statement of Main Result

Last time we proved that if a random graph $G(w)$ satisfied

$$d > \frac{4}{e(1-\epsilon^2)}$$

then almost surely every connected component S satisfies either $Vol(S) > \epsilon\, Vol(G)$ or $|S| < c\log(n)$ (where $|S|$ denotes the number of vertices in S). In other words, the components partition themselves into "Giant" components (those which contain a positive fraction of the volume of the graph) and smaller components. Today, we aim to prove a theorem which holds for **any** degree d which is larger than 1. Namely, we wish to show:

Theorem 1 *Let G(w) be a random graph whose expected degree sequence satisfies $d > 1 + \delta$, where $\delta > 0$. Then there is almost surely a unique giant component*

The proof of this theorem is rather complicated, so will be spread out over more than one lecture (alternatively, the proof is available in [1]). Today we focus on the existence portion of the theorem under the additional assumption that there is an M such that

$$w_i < M \,\forall\, i$$

2 A Concentration Inequality

We will make use of the following concentration inequality:

Lemma 1 *Let $\{X_1, X_2, ..., X_n\}$ be a sequence of independent random variables satisfying Pr($X_i = 1$)=p_i, Pr($X_i = 0$)=$1 - p_i$, and let $X = \sum_{i=1}^n a_i x_i$, where $a_i > 0$. Let $\nu = \sum_{i=1}^n a_i^2 p_i$. Then*

$$Pr(X < E(X) - t) \leq e^{\frac{-t^2}{2\nu}}$$

This is a weighted version of the standard concentration inequality for Bernoulli variables, and reduces down to the standard inequality if all the a_i are equal to 1. It is proven as Lemma 3.1 in the notes from April 6, where a bound on $Pr(X > E(X) + t)$ is also given.

3 Constructing Large Components Via Branching Processes

Ideally what we would like to do is take a vertex v of large expected degree and build up a large component of our graph by taking the vertices adjacent to v, the vertices adjacent to vertices adjacent to v, the vertices adjacent to those, and so on. We can describe this more formally in terms of a "branching process" using a depth-first search tree.

Fix a vertex v in our graph, which we also call S_0. Define S_{k+1} to be the set of all vertices not already in $\cup_{i=0}^{k} S_i$ which are adjacent to at least one vertex in S_k. By construction the S_k are pairwise disjoint and all lie in the component of our graph containing v

Let X_k denote the sum of the weights of vertices in S_k (so $X_k = Vol(S_k)$). We continue this process until either (for some ϵ to be chosen later) $\sum_{k=0}^{k_0} X_k > \epsilon Vol(G)$ (in which case we have our desired giant component), or some S_k is empty (in which case we've exhausted the whole component containing v, so must pick a new vertex as our starting point).

As per the usual, we will let ρ denote $\frac{1}{Vol(G)}$, so the probability vertices i and j are connected by an edge is equal to $w_i w_j \rho$. We call the vertex j **unexposed** if $j \notin S_1 \cup S_2 \cup ... \cup S_k$.

$$
\begin{aligned}
E(X_{k+1}) &= \sum_{j unexposed} w_j Pr(j \in S_{k+1}) \\
&= \sum_{j unexposed} w_j (1 - \prod_{i \in X_k} (1 - w_i w_j \rho)) \\
&\approx \sum_{j unexposed} w_j (x_k w_j \rho) \ \text{(Provided n is large enough so that ρ is small)} \\
&= X_k (\sum_{j \in V(G)} w_j^2 - \sum_{j exposed} w_j^2) \rho \\
&\approx x_k \widetilde{d}
\end{aligned}
$$

The last approximation holds since we are only continuing our branching process so long as the exposed vertices constitute only a small portion of the volume of the graph, so we can safely ignore the contribution from the exposed vertices.

$\widetilde{d} \geq d > 1 + \delta$, so **on average** the size of our component is growing exponentially.

4 Converting Fast Expected Growth into an Almost Sure Bound

What we would like to do is transform this into some sort of bound which with a given probability will hold for the entire process. To do this, we use the concentration inequality from above, with $p_i = x_k w_i \rho$ and $a_i = w_i$ to get that:

$$Pr(X_{k+1} < \widetilde{d} X_k - \lambda) \leq e^{\frac{-\lambda^2}{2\nu}}$$

Here,

$$\nu = \sum a_i^2 p_i = (\sum w_j^3) X_k \rho \leq M(\sum w_j^2) X_k \rho = M\widetilde{d} X_k$$

Choosing $\lambda = X_k(\widetilde{d} - 1)/2$, we get

$$Pr(X_{k+1} < X_k \frac{\widetilde{d} + 1}{2}) < e^{-\frac{(\widetilde{d}-1)^2}{8M\widetilde{d}} X_k} \leq (1-c)^{X_k}$$

where c is a positive constant chosen suitably close to 0. We can now redefine our process slightly as follows:

Starting from our vertex v, we continue branching out an additional level at each step until either $X_k > \epsilon Vol(G)$ (in which case we have a giant component) or for some k we have $X_{k+1} < X_k(\widetilde{d} + 1)/2$ (in which case our attempt failed).

At any given step the probability that we fail at that step is bounded above by $(1-c)^{X_k}$. Assuming that X_1 is sufficiently large, success at each stage guarantees that $X_{k+1} > X_k + 1$. For sufficiently large X_1 we can now bound the total probability of failure (given S_1) by

$$\sum_{j=X_1}^{\infty} (1-c)^j$$

Since this sum is a convergent geometric series, we can choose a t_0 such that if $X_1 > t_0$, the total probability of failure is bounded by an absolute constant c_0 which is smaller than 1. We can guarantee this condition is satisfied by picking the starting point in our branching process to be a vertex whose degree is at least t_0.

In fact (as will be proved in the lecture notes from May 13), there is a constant $B > 0$ such that almost surely at least $B|V(G)|$ vertices have degree at least t_0 (here again we need the upper bound on the weight of a vertex).

We can therefore construct the giant component as follows:

Pick a vertex v_1 of degree at least t_0. If the component of G containing v_1 is giant, we are done. If this is not the case, remove the component containing v_1 from G and take a v_2 in our new G with degree at least t_0. Repeat the process, constructing $v_3, v_4, ...$ until either a giant component is found or there are no more vertices of the required degree.

Since components which are not giant have volume $o(Vol(G))$, the number of vertices we can successively choose via this algorithm approaches infinity as the size of the graph does. Since at each stage we have probability at least $(1 - c_0)$ of success, the total probability of success increases to 1 as the size of the graph approaches infinity.

References

[1] Chung, F. and Lu, L. "Connected Components of Random Graphs With Given Expected Degree Sequences", 2004.