Chernoff Bounds and Diameter Bounds

Thursday, October 18, 2007
9:20 AM

In $G(n,p)$, we
expect $\binom{n}{2}p$ edges,

so for $p=\frac{1}{n}$, expect $\frac{n}{2}$ edges --
so expected degree of each vertex is 1

Let's see that it is unlikely to deviate

To bound chance it differs,
apply Chernoff - Hoeffding bounds:

let $X_1,\ldots,X_k$ be ind rand variables
s.t.

$$\Pr[X_i=1]=p_i$$

$$\Pr[X_i=0]=1-p_i.$$

let $X=\sum_i X_i$

let $\mu=\sum_i p_i = E[X]$

Then $\forall\, \delta<1$

$$\Pr[X<(1-\delta)\mu]<e^{-\frac{\delta^2\mu}{2}}$$

$$\Pr[X>(1+\delta)\mu]<e^{-\frac{\delta^2\mu}{3}}$$

<u>Ex.</u> create a variable for each
potential edge

X = total # edges
$p=\frac{1}{n-1}$, $\mu=\frac{n(n-1)}{2}\cdot p=\frac{n}{2}$

So, $\Pr[\#\text{edges} < \frac{n}{3}] = \Pr[X < (1-\delta)\mu]$, $\delta = \frac{1}{3}$

$$\le e^{\frac{-\mu\delta^2}{2}} = e^{\frac{-n}{2\cdot 2\cdot 3^2}} = e^{\frac{-n}{36}}$$

for n big, is exceedingly small.

So, actual # edges tightly concentrated around $p\binom{n}{2}$.

We will prove crude bounds on diameter of $G(n,p)$ for $p = 6\ln n/n$

First, let's prove a lower bound on the diameter.
Let d_{max} be the maximum degree.

Claim 1: $\Pr[d_{max} > 11\ln n] \to 0$ as $n \to \infty$.

pf. for any particular node, $E[d] = 6\ln n = \mu$
So applying Chernoff bound with $\delta = \frac{5}{6}$, we have
$\Pr[d > 11\ln n]$
$$= \Pr[d > (1+\delta)E[d]] \le e^{\frac{-\delta^2\mu}{3}} \le e^{-\frac{25\cdot 6\ln n}{36\cdot 3}} \le n^{-25/18}$$

So, $\Pr[d_{max} > 11\ln n] \le n \cdot n^{-25/18} \le n^{-7/18}$

Let $B_r(x) = \{v : d(v,x) \le r\}$, $B_0(x) = \{x\}$

Now, $|B_r(x)| \le d_{max} |B_{r-1}(x)| \le d_{max}^r$

So, if $d_{max}^r < \frac{n}{2}$, then most nodes have distance greater than r to x.

Similarly, if $d_{max}^r < n$ then $diam(G) \ge r$,

happens if $r \le \frac{\lg n}{\lg d_{max}} = \frac{\lg n}{\lg(11 \ln n)} = \frac{\ln n}{\ln(11) + \ln\ln n}$

We will try to prove a comparable upper bound on the diameter of $\frac{\ln n}{\ln \ln n} + 1$.

Our approach will have two steps.

Let $S_r(x) = B_r(x) - B_{r-1}(x)$

1. We will show that for $l = \frac{\ln n}{2 \ln\ln n}$

   with high probability $|S_l(x)| \ge \sqrt{n}$.

2. Then, for each pair x, y

   with high probability there is an edge between $S_l(x)$ and $S_l(y)$ (or they over-lap)

Let's do part 2 first:

Given that $|S_l(x)| \ge \sqrt{n}$ and $|S_l(y)| \ge \sqrt{n}$,

and that they don't overlap,

There are at least n pairs (u,v) $u \in S_l(x)$, $v \in S_l(y)$

There are at least n pairs (u,v) $u \in S_\ell(x)$, $v \in S_\ell(y)$

The probability that none appears in the graph is at most

$$(1-p)^n = \left(1 - \frac{6\ln n}{n}\right)^n \le \left(e^{-\frac{6\ln n}{n}}\right)^n = e^{-6\ln n} = n^{-6}.$$

So, the probability there exists a pair (x,y) with

$$\mathrm{dist}(x,y) > 2\ell+1 \le \binom{n}{2} n^{-6} \le n^{-4}$$

Let's return to part 1.

We will show that, with high probability,

for $r \le \frac{\ln n}{2\ln\ln n}$, $\quad |S_r(x)| \ge (\ln n)\,|S_{r-1}(x)|$,

which implies $\quad |S_\ell(x)| \ge (\ln n)^\ell = (\ln n)^{\frac{\ln n}{2\ln\ln n}} = n^{\frac{1}{2}}$

To do this, we consider the probability a node in $V - B_{r-1}(x)$ is in $S_r(x)$. The probability it is not in $S_r(x)$ is the probability it has no edge from $S_{r-1}(x)$

$$= (1-p)^{|S_{r-1}(x)|} \le 1 - p\,|S_{r-1}(x)| + \binom{|S_{r-1}(x)|}{2} p^2 \qquad (*)$$

To estimate this, first calculate

Assuming $d_{max} \le 11\ln n$, $\quad |B_{r-1}(x)| \le (11\ln n)^\ell = n^{\frac{1}{2}} \cdot n^{\frac{\ln(11)}{2\ln\ln n}}$

$$\le n^{1/3},$$

for n sufficiently large

And, for n sufficiently large, we get

$$(*) \le 1 - \left(\tfrac{11}{12}\right) p\,|S_{r-1}(x)|$$

So, the prob a node in $V-B_{r-1}(x)$ is in $S_r(x)$ is ~~at~~ least

$$\frac{11}{12}\, p\,|S_{r-1}(x)|.$$

For n sufficiently large, $|B_{r-1}(x)| \le \frac{n}{11}$, so $|V-B_{r-1}(x)| \ge \frac{10}{11}n$,

and so the expected number of nodes in $S_r(x)$ is at least $\frac{10}{12}\cdot p\cdot n\cdot |S_{r-1}(x)| = (5\ln n)\,|S_{r-1}(x)|$

So, applying a Chernoff bound with $\mu = (5\ln n)|S_{r-1}(x)|$, $\delta = \frac{4}{5}$, we find

$$\Pr\left[\,|S_r(x)| < (\ln n)\,|S_{r-1}(x)|\,\right] \le e^{-\frac{\delta^2\mu}{2}} \le \left(e^{-\frac{4^2\cdot 5\ln n}{5^2\cdot 2}}\right)^{|S_{r-1}(x)|},$$

$$\le n^{-8/5}$$

To finish, we now ~~s~~ observe that the probability there is an x and $r \le l$ for which $|S_r(x)| < (\ln n)\,|S_{r-1}(x)|$ is at most $n\cdot l\cdot n^{-8/5} < n^{-\frac{1}{2}}$ for n suff. large

So, with prob ~~at~~ least $1-(n^{-7/18}+n^{-4}+n^{-\frac{1}{2}})$, for n suff large, the diameter is ~~at~~ most $\frac{\ln n}{\ln\ln n}+1$

Let me clean up a little bit of that logic. Let A be the event

$d_{max} \ge 11\ln n.$

We showed $\Pr[A] \le n^{-7/18}$.

let $C_{x,r}$ be the event $|S_r(x)| < (\ln n)|S_{r-1}(x)|$

and $D_{x,r-1}$ be the event $|B_{r-1}(x)| > (11 \ln n)^{r-1}$

We proved $\Pr[C_{x,r} \mid \overline{D_{x,r-1}}] < n^{-8/5}$.

And, $\overline{D_{x,r-1}} \Rightarrow A$.

We want an upper bound on

$\Pr[\underset{x,r}{\mathrm{OR}}\ C_{x,r}]$, so write

$$\Pr[\underset{x,r}{\mathrm{OR}}\ C_{x,r}] = \Pr[A \text{ and } \underset{x,r}{\mathrm{OR}}\ C_{x,r}] +$$

$$\Pr[\bar{A} \text{ and } \underset{x,r}{\mathrm{Or}}\ C_{x,r}]$$

$$\leq n^{-7/18} + \Pr[\underset{x,r}{\mathrm{Or}}\ (\bar{A} \text{ and } C_{x,r})]$$

now, $\Pr[\bar{A} \text{ and } C_{x,r}]$

as $\bar{A} \Rightarrow D_{x,r-1}$

$$\leq \Pr[D_{x,r-1} \text{ and } C_{x,r}]$$

$$= \Pr[C_{x,r} \mid D_{x,r-1}] \cdot \Pr[D_{x,r-1}]$$

$$\leq \Pr[C_{x,r} \mid D_{x,r-1}]$$

$\leq n^{-8/5}$

So,

$$\Pr\left[\operatorname*{OR}_{x,r} C_{x,r}\right] \leq$$

$$n^{-7/18} + l \cdot n \cdot n^{-8/5}$$

$$\leq n^{-7/18} + n^{-1/2}$$ for n sufficiently large.