Diameter and distances

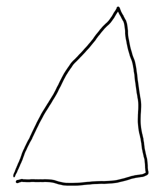
Friday, February 16, 2018 10:31 AM

Distances in $G(n, d/n)$:

— Balls grow like d^r so heuristic says distances $\sim \log_d n$.

$\mathbb{E}[Z_r \mid Z_{r-1}] \le d Z_{r-1}$

so $\mathbb{E} Z_r \le d^r$



$\Rightarrow$ w.h.p $d(u,v) \ge (1-\varepsilon) \log_d n$.

Theorem: If $u, v \in V$ of $G(n, d/n)$ $d > 1$ then Conditional on $u, v \in$ Giant,

$$\frac{d(u,v)}{\log_d n} \xrightarrow{P} 1$$

Also $\dfrac{\text{Diam}(\text{Giant})}{\log_d(n)} \in (1+\delta, C)$ w.h.p.

If Z_n is a B.P. with offspring distribution X, $\mathbb{E}X = \mu > 1$, $\sigma^2 = \text{Var} X < \infty$

$$\mu^{-n} Z_n \xrightarrow{a.s.} W$$

where $\mathbb{P}[W>0] = \mathbb{P}[Z_n \text{ survives}]$, $\mathbb{E}W = 1$.

<u>Proof</u>: Recursively calculate

$$\mathrm{Var}(Z_n | Z_{n-1}) = Z_{n-1} \cdot \sigma^2, \quad \mathbb{E}[Z_n | Z_{n-1}] = \mu Z_{n-1}$$

$$\mathrm{Var}(\mu^{-n} Z_n - \mu^{-(n-1)} Z_{n-1})$$

$$= \mathbb{E}(\mathrm{Var}(\mu^{-n} Z_n - \mu^{-(n-1)} Z_{n-1} | Z_{n-1})) + \underbrace{\mathrm{Var}(\mathbb{E}[\quad " \quad | Z_{n-1}])}_{=0}$$

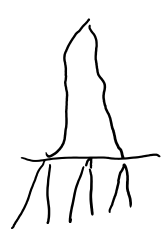
$$= \mathbb{E}(\mu^{-2n} Z_{n-1} \sigma^2) = \mu^{-(n+1)} \sigma^2$$

So $\sum \| \mu^{-n} Z_n - \mu^{-(n-1)} Z_n \|_2 < \infty$

$\Rightarrow \mu^{-n} Z_n \xrightarrow{a.s.} W$. $\mathbb{E}W = \lim \mathbb{E} \mu^{-n} Z_n = 1$.

So $\mathbb{P}[W>0] = \alpha > 0$.

Let $\tau_M = \inf\{n : Z_n \geq M\}$.



$$\mathbb{P}[W=0 | \tau_M < \infty] \leq (1-\alpha)^M$$

So $\mathbb{P}[W=0, W \text{ survives}] = 0$.

<u>Coupling</u> TV distance

$$d_{TV}(X, Y) = \min_A \mathbb{P}[X \in A] - \mathbb{P}[Y \in A]$$

$$= \frac{1}{2} \sum_{k} \left| \mathbb{P}[X=k] - \mathbb{P}[Y=k] \right|$$

Minimize over couplings,

$\exists (\tilde{X}, \tilde{Y})$ s.t. $\tilde{X} \overset{d}{=} X$, $\tilde{Y} \overset{d}{=} Y$,

$\mathbb{P}[\tilde{X} \neq \tilde{Y}] = d_{TV}(X, Y)$.

- If (X_i), (Y_i) independent

$$d_{TV}\left(\sum X_i, \sum Y_i\right) \le \sum d_{TV}(X_i, Y_i)$$

- For $Pois(p)$, $Ber(p)$,

$$d_{TV}(Ber(p), Pois(p))$$
$$= \frac{1}{2}\left(\left|e^{-p} - (1-p)\right| + \left|pe^{-p} - p\right| + \mathbb{P}[Pois(p) \ge 2]\right)$$
$$= O(p^2)$$

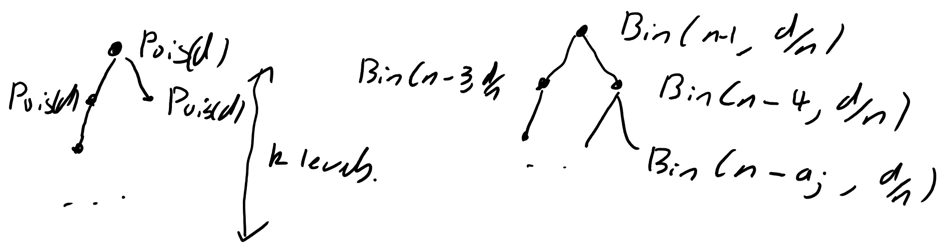
$$d_{TV}(Pois(np), Bin(n,p)) \le O(np^2).$$

$$d_{TV}(Bin(n,p), Bin(n-k,p)) \le kp.$$

Typical distance in $G(n, d/n)$ $d > 1$.

- Couple with BP $Pois(d)$ offspring dist.

$Pois(d)$



$\ldots \quad \downarrow$

$$\mathbb{P}[\text{Coupling fails}] \leq \sum_j O\left(a_j \cdot \frac{d}{n}\right) + O\left(n \frac{d^2}{n^2}\right)$$

$$\leq O\left(\left(\sum_{i=1}^{k} z_i\right)^2\right)$$

If $k \leq \left(\frac{1}{2} - \varepsilon\right) \log_d n$,

$Z_k \approx n^{\frac{1}{2} - \varepsilon}$ and couple the two processes.

Growth rate

$$Z_r = \underbrace{\mathrm{Bin}\left(Z_{r-1} \cdot (n - \sum z_i), \frac{d}{n}\right)}_{Y} - \text{cycles}$$

$$\mathbb{E}\, Y = d Z_{r-1} - \frac{d}{n} O\left(\sum z_i^2\right)$$

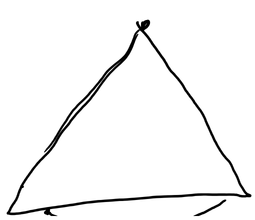
$\leq d^{2r}/n$ in expectation

<u>Chernoff Bound</u>

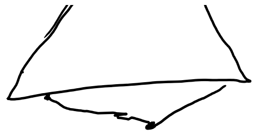
$$\mathbb{P}\left[\,|Y - \mathbb{E}\, Y| \geq x\,\right] \leq \exp\left(-\frac{x^2}{2 Z_{r-1} \cdot n \cdot \frac{d}{n} \cdot \left(1 - \frac{d}{n}\right)}\right)$$

Choose $x = \sqrt{3 d Z_r \log n}$

LHS $\leq o\left(\frac{1}{n}\right)$, take union bound.



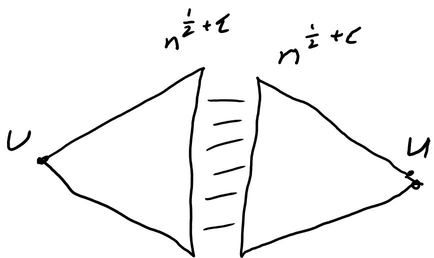
First cycle



First cycle forms when boundary $\approx n^{1/2}$.

$$\#\text{ cycles} \sim \mathrm{Bin}\left(\binom{Z_k}{2}, \left(\frac{d}{n}\right)^2\right) \ll Z_k.$$

$$Z_{k+1} \approx d Z_k - Z_k^2/n.$$



$k = \left(\frac{1}{2}+\varepsilon\right)\log_d n$

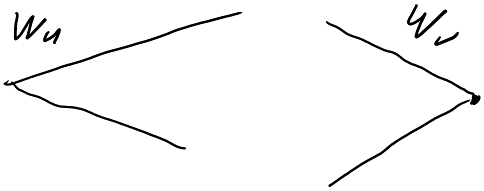
$$\mathbb{P}[\text{fail to connect}]$$
$$\leq \mathbb{P}\left[\mathrm{Bin}\left(n^{1+2\varepsilon}, \frac{d}{n}\right) = 0\right]$$
$$\leq e^{-n^{2\varepsilon}}$$

So $\mathbb{P}[d(u,v) \geq (1+\varepsilon)\log_d n \mid u,v \in \text{Giant}] \to 0.$

Hence

$$\mathbb{P}[d(u,v) \leq (1-\varepsilon)\log_d n] \to 0.$$

<u>Fluctuations:</u> $\quad d(u,v) - \log_d n$?



$$\frac{\partial B_r(u)}{d^r} \sim W_u$$

Choose l_u, l_v so that

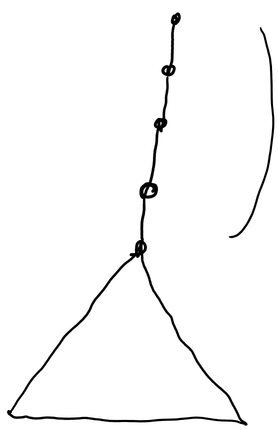
$W_u \, d^{\ell_u} = n^{1/2}$

- Then a constant probability that $B_{\ell_u}(u)$ and $B_{\ell_v}(v)$ intersect each level

---

Points of large distance:

We can find points whose local nbhd is a path of length $\delta \log_d n$



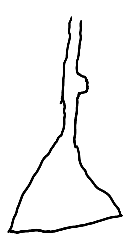
$\delta \log n$

Prob: $(de^{-d})^{\delta \log_d n} \approx n^{-\delta(d - \log d)/\log d}$

Typical distance $\sim (1+\delta)\log_d n$.

Diameter:

Explore neighbourhood one level at a time.



$A_k \geq \delta k$ for $k \geq C \log n$

with prob $\geq 1 - o(n)$.

Let level r be the first level with $\sum_{i=1}^{r} Z_r \geq B \log n$ (so $r \leq B \log n$)

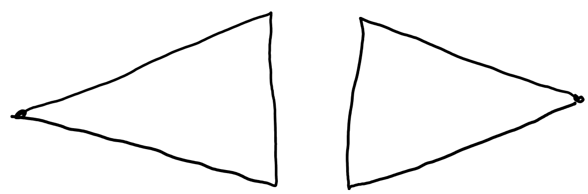
$$Z_{r+1} \sim \mathrm{Bin}\left(\left(n - \sum_{i=0}^{r} Z_i\right), 1 - \left(1 - \frac{d}{n}\right)^{Z_r}\right).$$

If $\sum_{i=1}^{r} Z_i \in (B\log n, n^{3/4})$ then $\approx Z_r \frac{d}{n}$.

$\mathbb{P}[Z_{r+1} \geq (d-\varepsilon')Z_r] \geq 1 - n^{-3}$.

- by Azuma - Hoeffding

$\Rightarrow$ $Z_{(B+\frac{1}{2}+\varepsilon)\log n} \geq n^{\frac{1}{2}+\varepsilon''}$ for all vertices in the giant component w.h.p.



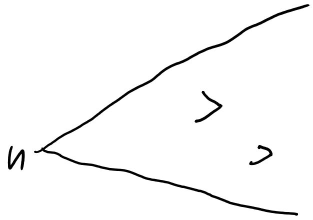
connections $\sim \mathrm{Bin}(n^{1+2\varepsilon''}, \frac{d}{n})$

So w.h.p. $\mathrm{Diam}(\text{Giant}) \leq 2(B+\frac{1}{2}+\varepsilon)\log n + 1$

Random d-regular:

Theorem

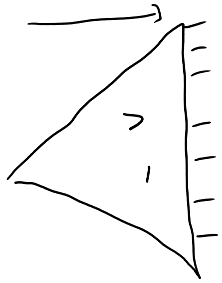
$$\frac{d(u,v)}{\log_{d-1} n}, \quad \frac{\mathrm{Diam}(G)}{\log_{d-1} n} \xrightarrow{p} 1$$



growth rate $(d-1)^r$

cycles $o(1)$ for

$$r \leq \left(\tfrac{1}{2} - \varepsilon\right) \log_{d-1} n.$$



$d(d-1)^r$ half edges to match.

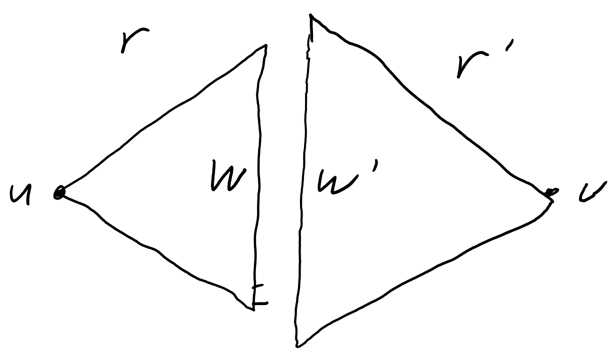
- let I_j be indicator that the j-th is matched to a new vertex so

$$\mathbb{P}[I_j] \leq \frac{d(d-1)^{r+1}}{(d-o(1))n}$$

so $\sum I_j \preceq \mathrm{Bin}\left(d(d-1)^{r+1}, \frac{d(d-1)^{r+1}}{dn}\right)$

$$\mathbb{E} \sum I_j \leq O\left((d-1)^{2r}/n\right)$$

- <u>Claim</u>: At most one cycles in any ball of radius $\frac{1}{5}\log_{d-1} n$.



$r = \frac{1}{2}\log_{d-1} n$

$r' = \frac{1}{2}\log_{d-1} n + l$

w.h.p $\quad |W| \approx (d-o(1))(d-1)^{r-1}$

$\quad |W'| \approx (d-o(1))(d-1)^{r'-1}$

- match edges from W to W' one by one,

I_j indicator that j-th half edge of W matched to W'.

$$\mathbb{P}[I_j] \approx \frac{d(d-1)^{r'}}{(d-o(1))n}$$

$$\mathrm{Pois}\left(\frac{(1-\varepsilon)(d-1)^{r'}}{n}\right) \lesssim I_j \lesssim \mathrm{Pois}\left(\frac{(1+\varepsilon)(d-1)^{r'}}{n}\right)$$

So # paths length $r+r'+1$

$$= \sum I_j \sim \mathrm{Pois}\left(\frac{d(d-1)^{r+r'+1}}{n}\right)$$

$$= \mathrm{Pois}\left(d(d-1)^{\ell}\right)$$

$\Rightarrow$ Typical distance $\approx \log_{d-1} n$

Diameter $\leq \log_{d-1} n + C \log\log n$.

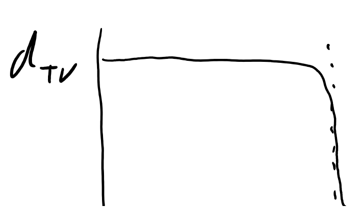
Random Walk on Random d-regular

- X_t^x is a SRW, pick a random neighbour in each step.
- π stationary distribution is uniform

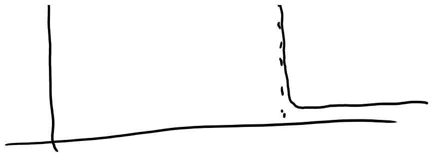
$$t_{mix}(s) = \min\{t : \max_x d_{TV}(X_t^x, \pi) \leq s\}$$

For all $0 < s < 1$,

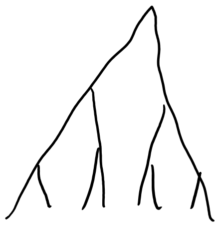
$$\frac{t_{mix}(s)}{\frac{d}{d-2}\log_{d-1}(n)} \xrightarrow{P} 1$$



Cutoff



On a tree speed of S.R.W. is



$$\frac{d-2}{d}$$

So time $\frac{d}{d-2}$ to reach $\mathrm{Diam}(G)$.

- Depth after t steps is $Z_t = \frac{d}{d-2} \cdot t + \sqrt{t} \cdot N(0, \sigma^2)$

- A non-backtracking RW (NBRW) has $Y_{t+1} \neq Y_{t-1}$.

- X_t same law as NBRW Y_{Z_t}.

$$\mathbb{P}_x[Y_\ell = y] = \frac{\#\text{ path } x \to y \text{ length } \ell}{d(d-1)^{\ell-1}}$$

$$\approx \frac{\mathrm{Pois}(d(d-1)^{\ell-1}/n)}{d(d-1)^{\ell-1}}$$

for $\log_{d-1} n < \ell < (2 - o(1)) \log_{d-1} n$.

When $\ell \geq \log_{d-1} n + \log\log n$,

$\mathrm{Pois}(d(d-1)^{\ell-1}/n) \approx d(d-1)^{\ell-1}/n$ w.h.p.

so $\mathbb{P}_x[Y_\ell = y] = \frac{1+o(1)}{n}$ for most x, y

$\Rightarrow \quad d_{TV}(X_t^x, \pi) \to 0$

for $t \geq \frac{d}{d-2} \log_{d-1}(n) + (\log_{d-1} n)^{\frac{1}{2}+\varepsilon}$.

- What about $G(n, d/n)$?

  Paths of length $\delta \log n$

  mean $t_{mix} \geq c \log^2 n$.

  This is the right order, no cutoff.

Heavy Tailed / Scale free graphs

Degree distribution

$$\mathbb{P}[d_v = k] \sim k^{-\alpha-1} \quad \text{for } 1 < \alpha < 2.$$

So $\mathbb{E} d_v = \mu < \infty$, $\operatorname{Var} d_v = \infty$.

$$\max_v d_v \sim n^{1/\alpha} \gg \sqrt{n}.$$

In configuration model

$$\mathbb{P}[\text{Self loop}] \geq \mathbb{P}[\text{Self loop at max degree vertex}]$$

$$\geq 1 - \prod_{i=1}^{n^{1/\alpha}} \frac{n^{1/\alpha} - i}{dn - 2i}$$

$$\geq 1 - \exp(-c n^{\frac{2}{\alpha} - 1})$$

- So w.h.p. configuration model is not simple.

- <u>Alternative model</u>:

$$\mathbb{P}[(u,v) \in E \mid d_u, d_v] = \frac{d_u d_v / n}{}$$

$$\mathbb{P}[(u,v)\in E \mid d_u, d_v] = \frac{d_u d_v/n}{1+d_u d_v/n}$$

Size biased

$\mathbb{P}[d_v^* = k] \sim k^{-\alpha}/n$ infinite mean.

– Can't calculate $\mathbb{E} Z_r$.

$$\boxed{\begin{aligned}
&\text{If } d_i^* \text{ index } 0<\beta<1, \text{ then}\\
&\mathbb{P}[\max_{1\le i\le n} d_i^* \geq t] = \mathbb{P}[\,\mathrm{Bin}(n, c\,t^{-\beta}) \geq 1]\\
&\quad \text{So} \quad \max d_i^* \approx n^{\frac{1}{\beta}}.\\
&\text{But also} \quad \sum_{i=1}^n d_i^* \approx n^{1/\beta}\\
&\quad \text{since} \quad \mathbb{E}\sum_{i=1}^n d_i^* I(d_i^* \le n^{1/\beta})\\
&\qquad = n\,\mathbb{E}\, d_i^* I(d_i^* \le n^{1/\beta})\\
&\qquad = n \sum_{k=1}^{n^{1/\beta}} k^{-\beta-1}\cdot k\\
&\qquad \asymp n^{1/\beta}.
\end{aligned}}$$

If level r has Z_r vertices

maximum size $\sim (Z_r)^{\frac{1}{\alpha-1}}$.

So $c(Z_r)^{1/\alpha-1} \le Z_{r+1} \le C'(Z_r)^{1/\alpha-1}$

Roughly

$$Z_r \sim \exp\left(c \left(\frac{\cdot}{\alpha-1}\right) \right)$$

so $Z_r \geq n^{1/2}$ when $r \approx \log\log n$.

Once $Z_r \geq n^{1/2}$, in the next level it w.h.p connects to the vertex of maximal degree.