Expected Number of Structures in Random Graphs and Randomized Graph Algorithms

Bhumrapee Soonjun & Nawat Ngerncham

March 25, 2024

Overview

With high probability (w.h.p.)

Definition (With high probability)

We say that event A happens *with high probability (w.h.p.)* if, for some constant $c > 0$,

$$\mathbf{Pr}\left[A\right] \geq 1 - \frac{1}{n^c} = 1 - O\left(\frac{1}{n^c}\right)$$

Definition (Almost surely)

We say that an event A happens *almost surely (a.s.)* if

$$\mathbf{Pr}\left[A\right] \to 1 \text{ as } n \to \infty$$

In this case, n is often the number of observations we make or the *size of the problem*. Note that we may use them interchangeably.

Simple Concentration Bounds

Theorem (Markov's Inequality)

Let X be a non negative random variable and $\lambda > 0$, then

$$\mathbf{Pr}\left[X \geq \lambda\right] \leq \frac{\mathbb{E}[X]}{\lambda}$$

Theorem (Chebyshev's Inequality)

Let X random variable with finite $\mathbb{E}[X]$ and $Var(X)$, then for any $t > 0$

$$\mathbf{Pr}\left[|X - \mathbb{E}[X]| \geq t\right] \leq \frac{Var(X)}{t^2}$$

Chernoff-Hoeffding Bounds

Theorem (Chernoff-Hoeffding Bounds)

Let $X_1, X_2, \cdots, X_n$ be an independently and identically distributed random variables taking values in $\{0, 1\}$, $\mathbf{X} = \sum_i X_i$, and $\mu = \mathbf{E}[\mathbf{X}]$. Then, for some $0 < \delta < 1$, the following inequalities hold,

1. $\mathbf{Pr}[\mathbf{X} \geq (1+\delta)\mu] \leq \exp\left(-\frac{\delta^2\mu}{2+\delta}\right)$
2. $\mathbf{Pr}[\mathbf{X} \leq (1-\delta)\mu] \leq \exp\left(-\frac{\delta^2\mu}{2}\right)$
3. $\mathbf{Pr}[|\mathbf{X} - \mu| \geq \delta\mu] \leq 2\exp\left(-\frac{\delta^2\mu}{3}\right)$

The above inequalities are called Chernoff's bounds. Note that this and the general form can be derived by applying Markov's inequality on the moment-generating function of the random variable. Moreover, there are multiple forms of these but we will work mostly with the above three.

Diameter on fixed p random graphs is 2 w.h.p.

Theorem

Let $G(n,p)$ be a random graph of order n and each edge appears with fixed probability $p \in (0,1)$. Then, $\operatorname{diam}(G) \leq 2$ *w.h.p.*

Lemma

Let $G(n,p)$ be a random graph of order n and each edge appears with fixed probability $p \in (0,1)$. Then, every pair of distinct vertices u, v have shared neighborhood w.h.p.

Diameter on fixed p random graphs is 2 w.h.p. — Proof

Proof of Lemma.

Define a random variable $X_{u,v}$ to be 1 if $N(u) \cap N(v) = \emptyset$ and 0 otherwise and define X to be the sum of all $X_{u,v}$'s. Note that the probability that two vertices share no common neighbor is

$$\mathbf{Pr}\left[N(u) \cap N(v) = \emptyset\right] = (1 - p^2)^{n-2}$$

where p^2 is the probability that u and v pick the same neighbor where there are $n-2$ of such possibilities. By Markov's Inequality, we have

$$\mathbf{Pr}\left[X \geq 1\right] \leq \mathbf{E}\left[X\right] = \binom{n}{2}(1 - p^2)^{n-2} \to 0 \text{ as } n \to \infty \qquad \text{[SageMath verified]}$$

Since $\mathbf{Pr}\left[X < 1\right] = 1 - \mathbf{Pr}\left[X \geq 1\right]$, the claim of the lemma satisfies w.h.p. □

Diameter on fixed p random graphs is 2 w.h.p. — Proof

Proof of Theorem.

Note that the probability that $\text{diam}(G) \leq 2$ is the same as the probability that every pair of distinct vertices u, v are adjacent or share a common neighbor. Hence, the probability that every distinct pair of vertices share a common neighbor is a lower bound for $\mathbf{Pr}[\text{diam}(G) \leq 2]$. Let X be the number of pairs of distinct vertices that do not share neighbors. We can conclude that

$$\mathbf{Pr}[\text{diam}(G) \leq 2] \geq \mathbf{Pr}[X < 1] \to 1 \text{ as } n \to \infty$$

Therefore, the diameter of random graph $G(n, p)$ with fixed p is at most 2 w.h.p. $\square$

Approximating Number of Triangles

Algorithm APPROXTRICOUNT(Graph G, Integer T)

$\alpha \leftarrow 0$
for i=1, 2, ..., T **do**
  Pick an edge (u, v) u.a.r. from $E(G)$
  Pick a vertex w u.a.r. from $V(G) \setminus \{u, v\}$
  if (u, w) and $(v, w) \in E(G)$ **then**
    Increment α
$\hat{\tau} \leftarrow \frac{\alpha}{T} \cdot \frac{m(n-2)}{3}$
return $\hat{\tau}$

ApproxTriCount Analysis

Theorem (Expected output)

Let τ be the true number of triangles in G. ApproxTriCount returns a random variable $\hat{\tau}$ with expected value

$$\mathbf{E}\left[\hat{\tau}\right] = \tau$$

ApproxTriCount Analysis

Proof.

Let us first identify the probability of the algorithm picking triangle t at iteration i. Consider a triangle t. WLOG, the edge (u, v) that we picked u.a.r. is part of t w.p. $3/m$. Then, the probability that we pick a specific w such that $\{u, v, w\}$ form t is $\frac{1}{(n-2)}$. Hence, the probability that we pick a triangle at iteration i is

$$\frac{3}{m} \cdot \frac{1}{n-2} = \frac{3}{m(n-2)}$$

Since there are τ triangles in G, the probability that we pick *any* triangle at iteration i is

$$\mathbf{Pr}\left[\text{Algorithm picks any triangle}\right] = \frac{3\tau}{m(n-2)}$$

ApproxTriCount Analysis

Proof.

Hence, it follows that after T iterations, the expected value of α is

$$\mathbf{E}\left[\alpha\right] = \frac{3\tau \cdot T}{m(n-2)}$$

Finally, we have that

$$\begin{aligned}
\mathbf{E}\left[\hat{\tau}\right] &= \mathbf{E}\left[\frac{\alpha}{T} \cdot \frac{m(n-2)}{3}\right] \\
&= \mathbf{E}\left[\alpha\right] \cdot \frac{m(n-2)}{3T} \\
&= \frac{3\tau T}{m(n-2)} \cdot \frac{m(n-2)}{3T} = \tau
\end{aligned}$$

Therefore, the expected output of ApproxTriCount is the true number of triangles. □

Picking a good T

Theorem

Suppose that τ is the true number of triangles in G and let $\epsilon > 0$ and $0 < \delta \leq 1/2$. Running $\textsc{ApproxTriCount}(G, T)$ *with* $T \geq \frac{3\tau}{\epsilon^2} \ln(2/\delta)$ *returns $\hat{\tau}$ satisfying $|\hat{\tau} - \tau| < \epsilon$ w.p.*

$$\mathbf{Pr}\left[|\hat{\tau} - \tau| < \epsilon\right] \leq 1 - \delta$$

Picking a good T

Proof.

First, note that $\mathbf{Pr}\left[|\hat{\tau} - \tau| < \epsilon\right] \geq 1 - \delta$ is equivalent to $\mathbf{Pr}\left[T|\hat{\tau} - \tau| \geq T\epsilon\right] \leq \delta$. To fit this problem into something C-H can handle, we let $\epsilon = \delta\tau$. Then, it follows that

$$\begin{aligned}
\mathbf{Pr}\left[T|\hat{\tau} - \tau| \geq T\epsilon\right] &= \mathbf{Pr}\left[T|\hat{\tau} - \tau| \geq \delta(T\tau)\right] \\
&\leq 2\exp\left(-\frac{\delta^2 T\tau}{3}\right) \\
&= 2\exp\left(-\frac{\epsilon^2 T}{3\tau}\right)
\end{aligned}$$

Picking a good T

Proof.

Since $T \geq \frac{3\tau}{\epsilon^2}\ln(2/\delta)$, we have the following.

$$\begin{aligned} 2\exp\left(-\frac{\epsilon^2 T}{3\tau}\right) &= \frac{2}{\exp\left(\frac{\epsilon^2 T}{3\tau}\right)} \\ &\leq \frac{2}{\exp\left(\frac{\epsilon^2}{3\tau}\cdot\frac{3\tau}{\epsilon^2}\ln(2/\delta)\right)} \\ &= \frac{2}{e^{\ln(2/\delta)}} \\ &= \delta \end{aligned}$$

Therefore, the claim of the theorem holds. □

Streaming Setting

Not all data is static—what if the data is given to you one by one?

Reservoir Sampling

When it comes to sampling from a stream, one of the most popular algorithm used is reservoir sampling [5].

Algorithm $\textsc{Sample}$(Data stream X) from [5]

$s \leftarrow X[1]$
for $i = 2, 3, \ldots, n$ **do**
  if $\textsc{Coin}(1/i)$ lands heads **then**
    $s \leftarrow X[i]$
return s

Note: $\textsc{Coin}(p)$ flips a biased coin that lands heads w.p. p.
Main property of RS: At the end, we maintain a true random sample of X.

Streaming Algorithm

Algorithm $\textsc{SampleTriangleStream}$(Stream of edges E) from [3]

$i \leftarrow 1$
for edge $e = (u, v)$ from stream **do**
  if $\textsc{Coin}(1/i)$ lands heads **then**
    $a \leftarrow u$; $b \leftarrow v$
    Pick a vertex w u.a.r. from $V \setminus \{a, b\}$
    $x \leftarrow F$; $y \leftarrow F$
  if $e = (a, w)$ **then** $x \leftarrow T$
  if $e = (b, w)$ **then** $y \leftarrow T$
  Increment i
if $x \wedge y$ **then return** 1 **else return** 0

Note: We assume that we already know the vertices beforehand.

SampleTriangleStream Analysis

Lemma (Output value)

Suppose that SampleTriangle *outputs a value* β. *The expected output is*

$$\mathbf{E}[\beta] = \frac{|T_3|}{|T_1| + 2|T_2| + 3|T_3|}$$

where T_i *refers to vertex triplets with* i *edges in its induced subgraph.*

SampleTriangleStream Analysis

Proof.

First, we will show that there are in total

$$m(n-2) = |T_1| + 2|T_2| + 3|T_3|$$

choices for the algorithm to sample an edge and a vertex. Since we pick an edge then another vertex from what remains, the LHS is obvious.

For the RHS, let $t = \{u, v, w\}$ be a triple of vertices.

- Let $t \in T_1$. There is only one way of picking t. WLOG, we do this by picking edge (u, v) and vertex w.
- Let $t \in T_2$. There are 2 ways of picking t. WLOG, we could pick the edge (u, v) along with vertex w or pick the edge (u, w) along with vertex v.
- Let $t \in T_3$. There are 3 ways of picking t. Indeed, we can pick any of the three edges along with the other vertex.

SampleTriangleStream Analysis

Proof.

Notice that SampleTriangleStream returns 1 if it picks (a, b) and w such that $\{a, b, w\} \in T_3$ and (a, b) is the first edge of the triplet that shows up in the stream. Hence,

$$\mathbf{E}[\beta] = \frac{|T_3|}{|T_1| + 2|T_2| + 3|T_3|}$$

$\square$

SampleTriangleStream Analysis

Theorem

Let s be the number of parallel instances of SampleTriangleStream *where instance i returns β_i. Similarly, let*

$$\hat{T}_3 = \left(\frac{1}{s} \sum_{i=1}^{s} \beta_i \right) \cdot m(n-2)$$

Then,

$$\mathbf{Pr}\left[\left| \hat{T}_3 - |T_3| \right| < \epsilon \right] \geq 1 - \delta$$

for any s satisfying

$$s \geq \frac{3|T_3|}{\epsilon^2} \cdot \ln(2/\delta)$$

SampleTriangleStream Analysis

Proof.

The proof is the same as when we were picking a good T for static G. $\square$

SampleTriangleStream Analysis

It is easy to see that we only need $O(s)$ space to run this algorithm where each instance requires $O(1)$ to keep track of a, b, w.

Note also that the running time on each instance is $O\left(1 + \frac{\log m}{m}\right)$ in amortized expectation.

This is because *expensive operations* are when Coin lands heads which happens $O(\log m)$ times in expectation [3].

Degree Theorem

Theorem (Degree Theorem [4])

Let G be a random sparse graph. That is, $p = o(1)$. Now, let $\Delta(\mathcal{G}(n,p))$ and $\delta(\mathcal{G}(n,p))$ denotes the maximum and minimum degree of $\mathcal{G}(n,p)$, respectively, then

1. *if $np = \omega \log n$ where $\omega \to \infty$, then w.h.p.*

$$\delta(\mathcal{G}(n,p)) \approx \Delta(\mathcal{G}(n,p)) \approx np$$

2. *Let $p = c/n$ for some constant $c > 0$, then w.h.p*

$$\Delta(\mathcal{G}(n,p)) \approx \frac{\log n}{\log \log n}$$

Degree Theorem Proof - Warm Up

Degree Theorem - (i).

Let X_i be an indicator random variable counting if edge i^th connected to u is presented or not, so.

$$X_i = \begin{cases} 1, \text{if } i \text{ is presented} \\ 0, \text{otherwise} \end{cases}$$

and,

$$X = \sum_{i=1}^{n-1} X_i$$

Clearly, X is binomial, so $\mathbb{E}[X] = (n-1)p = O(np) \approx np$.

Degree Theorem Proof - Warm Up

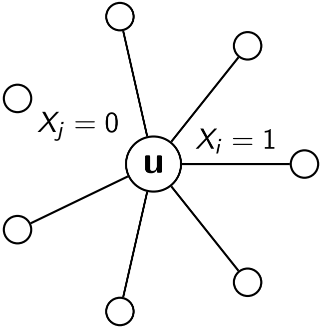

Figure: Degree Visualisation

Degree Theorem Proof - Warm Up

Degree Theorem - (i).

If $np = \omega \log n$ then

$$\mathbf{Pr}[\exists v : |\deg(v) - np| \geq \delta np] \leq 2\exp\left\{-\frac{\omega^{-2/3} np}{3}\right\} = 2\exp\left\{-\frac{\omega^{1/3}\log n}{3}\right\} = \frac{2}{n^{\omega^{1/3}/3}}$$

Thus,

$$\mathbf{Pr}[\forall v : |\deg(v) - np| < \delta np] \geq 1 - O\left(n^{-\omega^{1/3}/3}\right)$$

which occurs with high probability. Hence,

$$\delta(\mathcal{G}(n,p)) \approx \Delta(\mathcal{G}(n,p)) \approx np$$

$\square$

Degree Theorem Proof

Degree Theorem - (ii) Case 1 / Case 2.

We will prove the upper and lower bounds on d. Let $d^{-} = \lceil \frac{\log n}{\log\log n - 2\log\log\log n} \rceil$, then,

$$\mathbf{Pr}[\exists v : \deg(v) \geq d] \leq n \binom{n-1}{d} \left(\frac{c}{n}\right)^{d}$$

recall from many classes that $\binom{n-1}{d} \leq \left(\frac{en}{d}\right)^{d}$, so

$$\begin{aligned}
&\leq n\left(\frac{ec}{d}\right)^{d} = \frac{e^{\log n} \cdot e^{d} \cdot e^{d\log c}}{e^{d\log d}} \\
&= \exp\left\{\log n - d\log d + d + d\log c\right\} = \exp\left\{\log n - d\log d + O(d)\right\} \\
&\leq \exp\left\{\log n - \log n \log\log n + O(\lceil \frac{\log n}{\log\log n - 2\log\log\log n} \rceil)\right\} \\
&\leq \exp\left\{-\log n\log\log n + O(\log n)\right\} = \exp\left\{\log n\left(O(1) - \log\log n\right)\right\} \\
&\leq \exp\{-\log n\} = 1/n
\end{aligned}$$

Degree Theorem Proof

Degree Theorem - (ii) Case 1 / Case 2.

$$\mathbf{Pr}[\forall v : \deg(v) < d^-] \geq 1 - \frac{1}{n}$$

with high probability. Now on to case 2, let $d^+ = \lceil \frac{\log n}{\log\log n + 2\log\log\log n} \rceil$ and X_d be the number of vertices of degree d, and $X_{d,i}$ be respective Bernoulli r.v in $\mathcal{G}(n,p)$, then

$$\mathbb{E}[X_d] = \sum_{i \in I} \mathbf{Pr}[X_{d,i} = 1] = \sum_{i \in I} \left(\frac{c}{n}\right)^d \left(1 - \frac{c}{n}\right)^{n-d-1} = n\binom{n-1}{d}\left(\frac{c}{n}\right)^d \left(1 - \frac{c}{n}\right)^{n-d-1}$$

$$= \exp\left\{\log n - \frac{\log n}{\log\log n}(\log\log n - \log\log\log n + o(1)) + O(d)\right\} \to \infty$$

From here, the text also went to infinity and beyond and skipped directly to the results. So whatever comes after is from a fragment of my imagination adapting from another special case from another Theorem.

Degree Theorem Proof

Degree Theorem - (ii) Case 2.

Let's first do some acrobatics. Notice that,

$$\binom{n-1}{d} = \frac{(n-1)!}{d!(n-d)!} = \frac{(n-1)(n-2)(n-3)...(n-d)}{d!} = \frac{\prod_{i=1}^{d-1}(n-i)}{d!}$$

Consider the first few expansions of the numerator, we get

$$\begin{aligned}
\prod_{i=1}^{d}(n-i) &= n^d - \left(\sum_{i=1}^{d} i\right) \cdot n^{d-1} + (\cdot)O\left(n^{d-2}\right) + \dots \\
&= n^d + O(d^2)n^{d-1} + O(n^{d-2}) + \dots \\
&= n^d \left[1 - O(d^2)O(1/n) + O(d^3)O(1/n^2) + \dots\right] \\
&= n^d \left[1 + O(d^2)O(1/n)\right] = n^d \left[1 + O\left(\frac{d^2}{n}\right)\right]
\end{aligned}$$

Degree Theorem Proof

Degree Theorem - (ii) Case 2.

$$\begin{aligned}
\mathbb{E}[X_d] &= n\binom{n-1}{d}\left(\frac{c}{n}\right)^d\left(1-\frac{c}{n}\right)^{n-d-1} \\
&= n\frac{n^d}{d!}\left(1+O\left(\frac{d^2}{n}\right)\right)\left(\frac{c}{n}\right)^d \exp\left\{-(n-1-d)\left(\frac{c}{n}+O\left(\frac{1}{n^2}\right)\right)\right\} \\
&= n\frac{c^d}{d!}\left(1+O\left(\frac{d^2}{n}\right)\right)\exp\left\{-c+O\left(\frac{1}{n}\right)+\frac{c}{n}+\frac{1}{n^2}-\frac{cd}{n}+O\left(\frac{d}{n^2}\right)\right\} \\
&= n\frac{c^d}{d!}\left(1+O\left(\frac{d^2}{n}\right)\right)\exp\left\{O\left(\frac{1}{n}\right)+(1-d)\frac{c}{n}+O\left(\frac{d}{n^2}\right)\right\} \\
&= n\frac{c^d e^{-d}}{d!}\left(1+O\left(\frac{1}{n}\right)\right) \\
&= O(n)
\end{aligned}$$

Degree Theorem Proof

Degree Theorem - (ii) Case 2.

We now compute the following quantity,

$$\begin{aligned}\mathbf{Pr}[\deg(u) = \deg(v) = d] &= \frac{c}{n}\left(\binom{n-2}{d-1}\left(\frac{c}{n}\right)^{d-1}\left(1-\frac{c}{n}\right)^{n-1-d}\right)^2 \\ &+ \left(1-\frac{c}{n}\right)\left(\binom{n-2}{d}\left(\frac{c}{n}\right)^{d}\left(1-\frac{c}{n}\right)^{n-2-d}\right)^2\end{aligned}$$

First line takes care of where there is an edge between u and v, and the second is there when there isn't. So

$$= \mathbf{Pr}[\deg(u) = d]\mathbf{Pr}[\deg(v) = d]\left(1 + O\left(\frac{1}{n}\right)\right)$$

Degree Theorem Proof

Degree Theorem - (ii) Case 2.

Recall that $Var(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, so

$$\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left[X_i X_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{Pr}[X_i = d, X_j = d]$$

Now,

$$\mathbb{E}[X]^2 = \left(\sum_{i=1}^{n}\mathbf{Pr}[X_i = d]\right)\left(\sum_{j=1}^{n}\mathbf{Pr}[X_j = d]\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{Pr}[X_i = 1]\mathbf{Pr}[X_j = d]$$

So,

$$\mathrm{Var}(X_d) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left[\mathbf{Pr}[X_i = d, X_j = d] - \mathbf{Pr}[X_i = 1]\mathbf{Pr}[X_j = d]\right] = \sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}(X_i, X_j)$$

Degree Theorem Proof

Degree Theorem - (ii) Case 2.

Substituting in earlier results yields,

$$\begin{aligned}
\mathrm{Var}(X_d) &= \sum_{i=1}^{n}\sum_{j=1}^{n}\Bigg[\mathbf{Pr}[\deg(u)=d]\mathbf{Pr}[\deg(v)=d]\left(1+O\left(\frac{1}{n}\right)\right) \\
&\qquad\qquad - \mathbf{Pr}[\deg(u)=d]\mathbf{Pr}[\deg(v)=d]\Bigg] \\
&= \sum_{i\neq j=1}^{n}\left[\mathbf{Pr}[\deg(u)=d]\mathbf{Pr}[\deg(v)=d]O\left(\frac{1}{n}\right)\right] + \sum_{i=j}^{n}\mathrm{Cov}(X_i,X_i) \\
&= \sum_{i\neq j=1}^{n}\left[\mathbf{Pr}[\deg(u)=d]\mathbf{Pr}[\deg(v)=d]O\left(\frac{1}{n}\right)\right] + \sum_{i=j}^{n}\mathrm{Var}(X_i)
\end{aligned}$$

Degree Theorem Proof

Degree Theorem - (ii) Case 2.

Variance of indicator (Bernoulli) is just pq, so

$$\mathrm{Var}(X_d) \le \sum_{i \ne j = 1}^{n} O\left(\frac{1}{n}\right) + \sum_{i=j}^{n} \mathbf{Pr}[X_i = d](1 - \mathbf{Pr}[X_i = d])$$

Since $0 < p, q, < 1$ and $\mathbb{1} = p$, $\mathrm{Var}(\mathbb{1}) < \mathbb{E}[\mathbb{1}]$, hence

$$\begin{aligned} &\le \sum_{i \ne j = 1}^{n} O\left(\frac{1}{n}\right) + \mathbb{E}[X] \\ &= \sum_{i=1}^{n} \sum_{j \ne i}^{n} O\left(\frac{1}{n}\right) + \mathbb{E}[X] \\ &= O(n) + O(n) \\ &= O(n) \end{aligned}$$

Degree Theorem Proof

Degree Theorem - (ii) Case 2.

Using Chebyshev's,

$$\mathbf{Pr}[X_d = 0] \leq \mathbf{Pr}[|X_d - \mathbb{E}[X_d]| \geq \mathbb{E}[X_d]] \leq \frac{\mathrm{Var}(X_d)}{(\mathbb{E}[X])^2} \leq \frac{O(n)}{(O(n))^2} \leq O\left(\frac{1}{n}\right)$$

Hence,

$$\mathbf{Pr}[X_d > 0] \geq 1 - O\left(\frac{1}{n}\right)$$

Which occurs with high probability. From these two, we get that $\forall v \in G, \deg(v) < d^-$, and $\exists u \in G : \deg(u) > d^+$, so

$$\frac{\log n}{\log\log n + \log\log\log n} \leq \deg(u) \leq \frac{\log n}{\log\log n - 2\log\log\log n}$$

$$\Omega\left(\frac{\log n}{\log\log n}\right) \leq \deg(u) \leq O\left(\frac{\log n}{\log\log n}\right)$$

$\square$

Global Mincut History

1. Recall that a cut means a set of edge E where G/E separate G into two disjoint set A, B
2. Note that **global mincut** is not the same as $s - t$ mincut aka. max-flow min-cut
3. **Global mincut** problem that asks how many edges you have to remove to disconnect a graph
4. Computing this is hard. The search space is large, exponentially large, however, there is a polynomial time algorithm from Stoer-Wagner computing this in $O(|V||E| + |V|^2 \log |V|)$ time.
5. Today, let's look at an approximation that runs in $O(|V|^2 \text{polylog}|V|)$ time and returns the minimum cut w.h.p.

Global Mincut Algorithm - Karger

1. The following subroutine will be useful. The subroutine will keep contracting the graph randomly until there are t vertices left.
2. It picks an edge described by two vertices on the edge uniformly at random and calls a contract on it.
3. Contract subroutine is different from what have seen in class. Contract will replace the two vertices, u, v, with a new vertex w and take all the adjacent edges of u, v with it. Note that it doesn't merge edges, instead, it keeps all edges so this can be a multigraph.

Algorithm CONTRACTION(Graph G, Integer t) Karger's

while $|V(G)| > t$ **do**
  $(u, v) \leftarrow$ sample edge e u.a.r from G
  $G \leftarrow$ CONTRACT(u, v)
return G

Global Mincut Algorithm - Karger

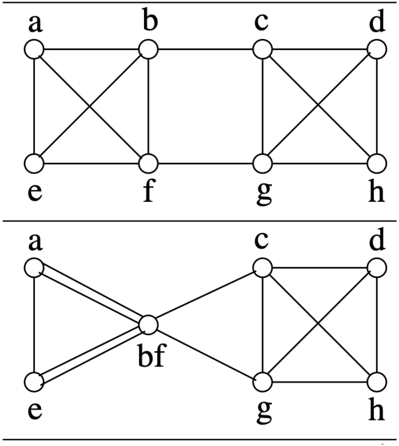

Figure: Contraction example

Global Mincut Algorithm

1. To find mincut, we call contraction(2). This produces a graph with two vertices at the end, therefore we easily get a cut which is the leftover edges between them.

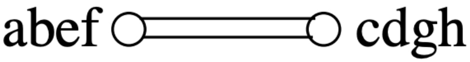

Figure: contraction(2) result. [1]

Edge Survivability Probability Bound

Lemma

Let contraction(t) be defined as above, then after t iterations, the probability of min-cut edge surviving is,

$$\mathbf{Pr}[\mathit{mincut\ edges\ survive}] \geq \frac{t(t-1)}{n(n-1)} = \Omega\left(\frac{t^2}{n^2}\right)$$

Lemma (Edge Survivability Probability Bound)

The probability of min-cut edges being selected for contraction after i^{th} round is,

$$\mathbf{Pr}[\mathit{min\text{-}cut\ edge\ selected}] \leq \frac{2}{n-i}$$

Edge Survival Probability After t Iterations

Proof.

We first calculate the probability of a mincut edge being selected. Let k be the minimum edge cut set size, and i^{th} be the iteration, then the probability that we will select any e_j is

$$\mathbf{Pr}[Any e_j \text{ selected}] = k/|E|$$

Consider the number of edges in the graph. We know that,

$$|E| \geq (n-i)k/2$$

Otherwise, we can produce a smaller min edge cut set. Combining this yields,

$$\mathbf{Pr}[e_j \text{ selected}] = \frac{k}{|E|} \leq \frac{k}{k(n-i)/2} = \frac{2}{n-i}$$

Therefore, the probability that no e_j is selected is

$$\mathbf{Pr}[\text{no } e_j \text{ selected}] \geq 1 - \frac{2}{n-i}$$

Karger's Analysis

Proof.

Let the event that after t iterations the mincut edges survive by E_s, and E_i be an event that in the iteration i^{th}, the edges survive. Then, by the above inequality

$$\begin{aligned}\mathbf{Pr}[E_s] &= \prod_{i=0}^{n-t-1} \mathbf{Pr}[E_i] \geq \prod_{i=0}^{n-t-1} \left(1 - \frac{2}{n-i}\right) = \prod_{i=0}^{n-t-1} \left(\frac{n-i-2}{n-i}\right) \\ &= \frac{n-2}{2} \cdot \frac{n-3}{n-1} \cdot \frac{n-4}{n-2} \cdot \ldots \cdot \frac{t+1}{t+3} \cdot \frac{t}{t+2} \cdot \frac{t-1}{t+1} \\ &= \frac{t(t-1)}{n(n-1)} = \Omega\left(\frac{t^2}{n^2}\right)\end{aligned}$$

□

If $t = 2$, the probability is $\Omega(1/n^2)$, therefore we should run this n^2 times to get the mincut. Contraction runs in $O(n^2)$ so total running time is $O(n^4)$.

Fast Mincut [2]

Algorithm $\textsc{FastMincut}$(Graph G) Karger-Stein's

if $|V(G)| \leq 6$ **then**
    $C \leftarrow$ brute force mincut
    return $|C|$
else
    $t = \lceil 1 + n\sqrt{2} \rceil$
    $H_1, H_2 \leftarrow \textsc{contraction}(G,\ t), \textsc{contraction}(G,\ t)$
    $C_1, C_2 \leftarrow \textsc{FastMincut}(H_1), \textsc{FastMincut}(H_2)$
    return $\min\{C_1, C_2\}$

Fast Mincut Running Time

Theorem (Fast Mincut Running Time)

(i) Let fast-mincut, contraction be defined as above, and G be any connected graph, then the total running time of fast-mincut on G is

$$T(n) = O(n^2 \log n)$$

(ii) Moreover, to produce a minimum cut of G, we runs $c \log n$ copies leading to

$$T(n) = O(n^2 \log^2 n)$$

with high probability.

Theorem (Minimum Cut Probability)

Running $c \log n$ copies of fast-mincut produces a mincut in at least one of the copies with high probability.

Fast Mincut Running Time - Proof

Fast Mincut Running Time.

To prove (i), we consider the following recurrence of fast-mincut.

$$\begin{aligned} T(n) &= 2T(\lceil 1 + n/\sqrt{2} \rceil) + O(n^2) \\ &\approx 2T(n/\sqrt{2}) + O(n^2) \\ &= \sum_{i=1}^{2\log n} 2^i \cdot \frac{n^2}{2^i} = \sum_{i=1}^{2\log n} n^2 \\ &= O(n^2 \log n) \end{aligned}$$

To prove (ii), we use theorem 16, therefore we run $c \log n$ copies, so it follows that

$$T(n) = O(n^2 \log^2 n)$$

And it produces a minimum cut with high probability.

$\square$

Fast Mincut Analysis Intuition

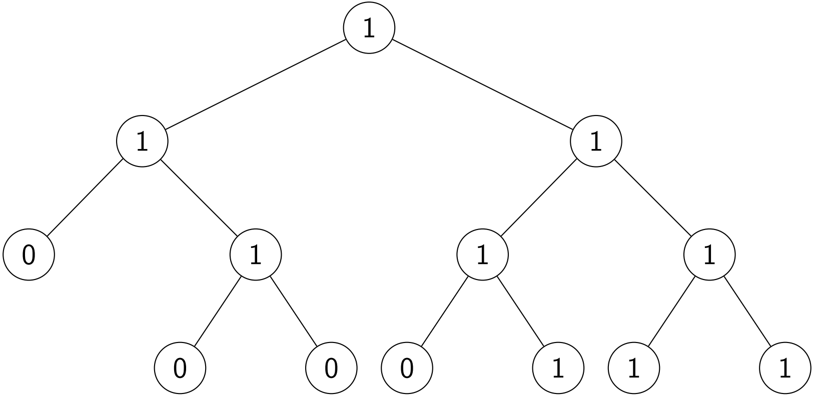

Figure: Binary Tree representing the process. 1 represents the min-cut still being in the subproblem at recursion depth l, and 0 otherwise.

Minimum Cut Probability Proof

1. To prove theorem 16 (minimum cut probability) we need more advanced tools
2. We here present two proofs one based on **Stochastic Processes** and sheer power and another we will use Chernoff's bound on offspring distribution
3. The main idea is to analyze this as a **Galton-Watson Branching Processes** with approximately binomial offspring distribution
4. The setup is as follows, for each recursive call of fast-mincut, it produces H_1 and H_2. If the minimum cut is in H_i we say Y_i produces an offspring. So each Y_i can produce $0, 1$, or 2 offspring(s). If $Y_i = 0$, then both H_1 and H_2 do not contain minimum cut, if $Y_i = 1$ then either one of them contains the cut, and if $Y_i = 2$ both of them contain a cut.
5. The quantity of interest is then the **survival probability** after $2 \log n$ generations. Or simply, $1 - \rho_n$, where ρ_n is the **probability of extinction** on generation n.

Stochastic Processes - Branching Processes

Definition (Branching Processes)

A **Galton-Watson Branching Processes** is a **Discrete Time Homogenous Markov Chain** that satisfies the following properties. Let Z_t be a random variable counting the number of individuals in generation t^{th} and Y_i^t be individual i^{th} at time t.

1. A population starts with one individual at time $t = 0$ so $Z_0 = 1$.
2. Every individual lives on a unit of time and produces Y offspring.
3. Y can tales values $0, 1, 2, \ldots$ with probability $\mathbf{Pr}[Y = k]$.
4. All individuals reproduce independently.

The collection of $Z_n = \{Z_0, Z_1, Z_2, \ldots\}, n \in \mathbb{N}$ is called a branching process.

Stochastic Processes - Main Theorems

Theorem (Distribution at time n)

Let $G(s) = \mathbb{E}[s^Y] = \sum_{y=0}^{\infty} \mathbf{Pr}[y = 1]s^y$ be the **probability generating function** *of each individual Y. Let $Z_0 = 1$, and Z_n be the population size at time n. We now define $G_(s)$ to be the probability generating function of Z_n, then*

$$G_n(s) = G\left(G\left(G\left(\ldots G\left(s\right)\ldots\right)\right)\right)$$

is the composition of G n times.

Theorem (Expected Population Size At Time n)

Let $\{Z_0, Z_1, Z_2, \ldots\}$ be a branching process with $Z_0 = 1$. Let Y denote the family size distribution, and suppose that $\mathbb{E}[Y] = \mu$. Then,

$$\mathbb{E}[Z_n] = \mu^n$$

Stochastic Processes - Main Theorems

Theorem (Extinction Probability)

Let the branching process be defined as above, $G(s)$ be the generating function of Y, G_n be the generating function of Z_n, and ρ_n be the extinction probability at generation n. Then,

$$\rho_n = G(\rho_{n-1}) = G_n(0)$$

Moreover,

$$\rho^* \iff s = G(s)$$

where $\rho^* = \lim_{n\to\infty} \rho_n$.

Minimum Cut Probability Proof

Minimum Cut Probability - Stochastic Processes.

Let E be the event where mincut edges survive. Then, using t, for finite n, we can show that the probability that edge in min-cut will survive contraction(t) is,

$$\mathbf{Pr}[E] \geq \frac{t(t-1)}{n(n-1)} = \frac{(\lceil 1 + n/\sqrt{2} \rceil)(\lceil 1 + n/\sqrt{2} \rceil - 1)}{n(n-1)} = \frac{n+\sqrt{2}}{2n-2} > \frac{1}{2}$$

Since we call this recursively and independent of each other, we can view this as a **branching process** where any node can have zero, one or two children depending on if the mincut survived in zero, one or both children. Let Y_i denotes the number of children that individual Y_i has, then

$$\mu = \mathbb{E}[Y_i] = np > 2 \cdot \frac{1}{2} = 1$$

Fast Mincut Probabilistic Analysis - Stochastic Processes

Minimum Cut Probability - Stochastic Processes.

Recall from **Stochastic Processes** the probability that a branching process will be extinct by generation n is $\rho_n = G(\rho_{n-1}) = G_n(0)$ where $G(\cdot)$ is probability generating function of Y, and $G_n(\cdot)$ is n^{th} composition of G. Then, taking $p = 1/2$, the lower bound,

$$G(s) = \sum_{i=0}^{\infty} \mathbf{Pr}[Y = i]s^i = \frac{1}{4} + \frac{1}{2}s + \frac{1}{4}s^2 = \frac{1}{4}(1+s)^2$$

Then, for ρ_n

$$\rho_n = \frac{1}{4}(1 + \rho_{n-1})^2$$

Now we just need to solve this recurrence.

Recurrence - Oh no...

Input

$$\left\{T(n) = \frac{1}{4}(1 + T(n-1))^2, T(0) = 0\right\}$$

Alternate form

$$\left\{(T(n-1)+1)^2 = 4\,T(n), T(0) = 0\right\}$$

Enlarge Data Customize A P

Expanded form

$$\left\{T(n) = \frac{1}{4}T(n-1)^2 + \frac{1}{2}T(n-1) + \frac{1}{4}, T(0) = 0\right\}$$

Figure: Wolfram cannot solve this. GG

Recurrence - But Wait.

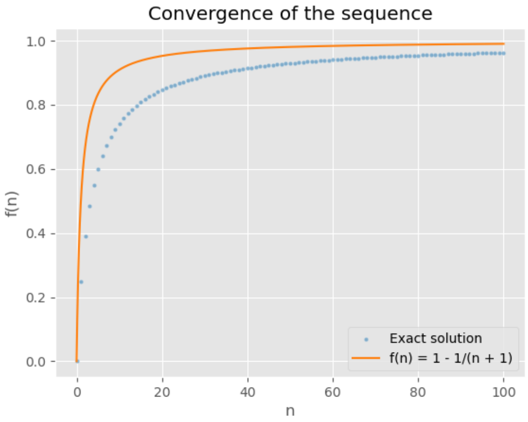

Figure: Asymptotic Recurrence Solution

Fast Mincut Probabilistic Analysis - Stochastic Processes

Minimum Cut Probability - Stochastic Processes.

With divine intervention (linear regression on sampled points), we claim $\rho_n = O(1 - 1/n)$, then we perform inductive verification. Let $\rho_0 = 0$ and $c = 1$, then

$$
\begin{aligned}
\rho_n = \frac{1}{4}\left(1 + \rho_{n-1}\right)^2 &\leq \frac{1}{4}\left(1 + c - \frac{c}{n-1}\right)^2 = \frac{(n - 1 + nc - 2c)^2}{4\,(n-1)^2} \leq \frac{(2n-3)^2}{4\,(n-1)^2} \\
&= \left(\frac{2n-3}{2n-2}\right)^2 = \left(1 - \frac{1}{2n-2}\right)^2 = 1^2 - 2 \cdot \frac{1}{2n-2} + \left(\frac{1}{2n-2}\right)^2 \\
&= 1 - \frac{1}{n-1} + \frac{1}{(2n-2)^2} \\
&\leq 1 - \frac{1}{n}
\end{aligned}
$$

For $n \geq 2$. So, the recurrence has asymptotic solution of $O(1 - 1/n)$

Fast Mincut Probabilistic Analysis - Stochastic Processes

Minimum Cut Probability - Stochastic Processes.

Since we only consider up to $2\log n$ layers, it follows that

$$\rho_n = 1 - 1/n = 1 - 1/\log n \leq \exp\{-1/\log n\}$$

Then, if we repeat this $c\log^2 n$ times, we get that

$$\rho_n = \prod_{i=0}^{c\log^2 n} \exp\{-1/\log n\} = \exp\left\{-c\frac{\log^2 n}{\log n}\right\} = \exp\left\{\log n^{-c}\right\} = \frac{1}{n^c}$$

So the probability of survival after $c\log^2 n$ runs is,

$$\mathbf{Pr}[\text{mincut edge survive}] \geq 1 - \frac{1}{n^c}$$

which occurs with high probability.

$\square$

Fast Mincut Probabilistic Analysis - Chernoff

Minimum Cut Probability - Chernoff.

Here, we also show Chernoff-based proof. Consider the number of individuals on generation n. Let,

$$X_i = \begin{cases} 1, \text{if the individual is present,} \\ 0, \text{otherwise, i.e parent doesn't reproduce} \end{cases}$$

and,

$$X = \sum_i X_i$$

Consider the quantity $\mathbb{E}[X]$ which is the expected number of individual in generation n. We know that this is the same as

$$\mathbb{E}[X] = \mathbb{E}[Z_{n-1}] = \mu^{n-1} = (np)^{n-1} > (1/2 \cdot 2)^{n-1} = 1$$

Fast Mincut Probabilistic Analysis - Chernoff

Minimum Cut Probability - Chernoff.

Since each $X_i \sim I(p)$ and $i.i.d$, for $\delta = 1/2$, using Chernoff's we obtain,

$$\mathbf{Pr}[X \leq (1-\delta)\mathbb{E}[X]] = \mathbf{Pr}[X \leq (1-\delta)\mathbb{E}[Z_n]] \leq \exp\left\{-\frac{\delta^2}{2}\mu^{n-1}\right\} \leq \exp\left\{-\frac{1}{8}\right\}$$

If we run the algorithm $c \log n$ times, the probability of failure is at most

$$\mathbf{Pr}[\text{fail}] \leq (\exp\{-1/8\})^{c \log n} = \exp\left\{\log n^{\frac{-c}{8}}\right\} = \frac{1}{n^{c/8}}$$

So, the probability of at least one success after $c \log n$ is

$$\mathbf{Pr}[\text{survive}] \geq \mathbf{Pr}[X \leq (1-\delta)\mathbb{E}[X]] = \mathbf{Pr}[X > 1/2] = \mathbf{Pr}[X \geq 1/2]$$

$$= 1 - \frac{1}{n^{c/8}} = 1 - \frac{1}{n^c} = 1 - O\left(\frac{1}{n}\right)$$

which occurs with high probability.

Thank you!

$+$ Q&A time

Reference

- Sanjeev Arora, Winter 2010.

- Surender Baswana and Sandeep Sen.

- Luciana S. Buriol, Gereon Frahling, Stefano Leonardi, Alberto Marchetti-Spaccamela, and Christian Sohler.
  Counting triangles in data streams.
  Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, Jun 2006.

- Alan Frieze and Michał Karoński.
  Introduction to Random Graphs.
  Cambridge University Press, 2015.

- Jeffrey S. Vitter.
  Random sampling with a reservoir.
  ACM Trans. Math. Softw., 11(1):37–57, mar 1985.