CS 4540/CS 8803: Algorithmic Theory of Intelligence Fall 2023

Lecture 8: Random Graphs Continued

Instructor: Mirabel Reid *Lecture date: 10/11*

Scribed by: Mirabel Reid

Recall: definition of monotone graph property, threshold in random graph.

Theorem 1. $p^* = \frac{\log n + c}{n}$ *is the threshold for the number of isolated vertices in* $G(n, p)$. *That is,*

$$\lim_{n\to\infty} Pr(G_{n,p} \text{ has an isolated vertex}) = \begin{cases} 1 & \text{if } c = c(n) \to -\infty \\ 0 & \text{if } c = c(n) \to \infty \end{cases}$$

Proof. Let X be the number of isolated vertices in the graph. Let $I_v = 1$ if v is an isolated vertex, and 0 otherwise. Then, we can write $X = \sum_{v\in V} I_v$.

A vertex v is isolated if none of its edges to other vertices are present. Since each edge is independently present with probability p, we have $Pr(I_v) = (1-p)^{n-1}$

$$\mathbb{E}X = \sum_{v\in V} \mathbb{E}I_v = n(1-p)^{n-1}$$

First, we will look at the case when $c \to \infty$. Substituting $p = \frac{\log n + c}{n}$,

$$\begin{aligned}\mathbb{E}X &= n\left(1 - \frac{\log n + c}{n}\right)^{n-1} \\ &\le n \exp\left(-\frac{\log n + c}{n}(n-1)\right) \\ &= n \exp\left(-(\log n + c) + \frac{\log n + c}{n}\right) \\ &< e^{-c+1}\end{aligned}$$

As $c \to \infty$, this goes to zero.

By Markov's Inequality (and as you will prove in the homework),

$$Pr(X > 0) \le \mathbb{E}X < e^{-c+1}$$

This proves the top inequality.

For $c \to -\infty$, we can use the inequality $1 - x \ge e^{x/(1-x)}$ for $x < 1$

$$\begin{aligned}\mathbb{E}X = n\left(1 - \frac{\log n + c}{n}\right)^{n-1} &\ge n \exp\left(-\frac{\log n + c}{n - \log n - c}(n-1)\right) \\ &= n \exp\left(-(\log n + c) - \frac{(\log n + c)(\log n + c - 1)}{n - \log n + c}\right) \\ &> e^{-c-1}\end{aligned}$$

We have $\mathrm{E}[X] \to \infty$ as $n \to \infty$; can we say that $X > 0$ with high probability? Not necessarily!

Now using the second moment method: notice we can use variance to bound $Pr(X = 0)$.

$$\begin{aligned}\mathbb{E}X^2 &= \mathbb{E}(\sum_{v\in V} I_v)^2 = \sum_v \mathrm{E}[I]_v^2 + \sum_{u\ne v\in V} \mathrm{E}[I]_u I_v \\ &= \sum_v Pr(I_v = 1) + \sum_{u\ne v\in V} Pr(I_u = 1 \text{ and } I_v = 1) \\ &= n(1-p)^{n-1} + n(n-1)(1-p)^{2n-3} = \mathrm{E}[X] + (\mathrm{E}[X])^2(1 + o(1))\end{aligned}$$

Here, we will use Chebyshev's inequality to bound the probability that $X = 0$. If $X = 0$, then $|X - \mathrm{E}[X]| \geq \mathrm{E}[X]$;

$$Pr(X = 0) \leq Pr(|X - \mathrm{E}[X]| \geq \mathrm{E}[X]) \leq \frac{VarX}{(\mathrm{E}[X])^2}$$

Substituting $Var(X) = \mathrm{E}[(]X^2) - (\mathrm{E}[X])^2 = \mathrm{E}[X] + o(1)(\mathrm{E}[X])^2$:

$$Pr(X = 0) \leq \frac{\mathrm{E}[X] + o(1)(\mathrm{E}[X])^2}{(\mathrm{E}[X])^2} = \frac{1}{\mathrm{E}[X]} + o(1) < e^c + o(1)$$

As $n \to \infty$, $c \to -\infty$, so this goes to 0. Therefore, $Pr(X > 0)$ goes to 1. □

1 Alternative Graph Models

Degree concentration in $G(n,p)$: let d_v =degree of v in $G(n,p)$.

$$\mathrm{E}[d]_v = (n-1)p$$

$$Pr(d_v \geq (n-1)p + t\sqrt{(n-1)p}) \lesssim \exp(-t^2/2)$$

Motivation: Modelling real world networks, e.g. social media sites. Degree distribution decays polynomially rather than exponentially.

Preferential attachment model: define a sequence of n graphs, $G_1, G_2, \ldots, G_n$. Each G_t has t vertices and tm edges (loops and multi-edges allowed).

Algorithm 1 Preferential Attachment Model, AKA Barabási–Albert Model

```
Parameter: Integer m > 0
Construct G_1: 1 vertex, wi
for t = 2, 3, ... do
    Add a vertex v_t to the graph
    Sample m vertices u_1, ..., u_m. For any u ∈ G_t

        Pr(u_i = u) = deg(u, G_t) / 2mt

end for
```