The evolution of $G_{n,m}$

TOM BOHMAN

CARNEGIE MELLON UNIVERSITY

OUTLINE

1. **Background.** Definitions, history and a brief account of the evolution.

2. **Methods.** An example illustrating expected value and concentration.

3. **The giant component.** An explanation of the sudden emergence of the giant component.

OUTLINE

1. **Background.** Definitions, history and a brief account of the evolution.

2. **Methods.** An example illustrating expected value and concentration.

3. **The giant component.** An explanation of the sudden emergence of the giant component.

$G_{n,m}$: A graph chosen uniformly at random from the set of all graphs on vertex set $[n]$ with m edges.

So, if G is a graph on vertex set $[n]$ with m edges

$$\Pr(G_{n,m} = G) = \frac{1}{\binom{\binom{n}{2}}{m}}$$

e.g.

$$\Pr\left(G_{4,p} = H\right) = p^3(1-p)^3$$

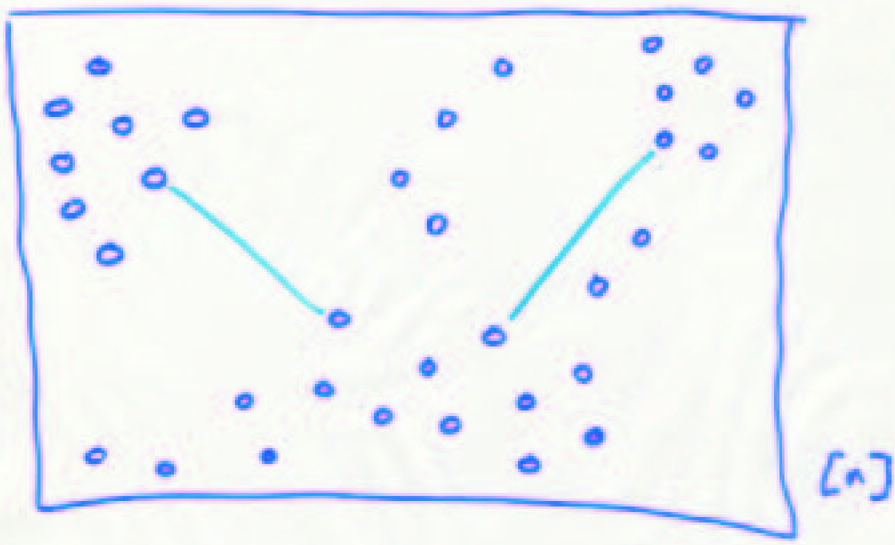
$$\Pr\left(G_{4,3} = H\right) = \frac{1}{\binom{6}{3}} = \frac{1}{20}$$

note: An event $\mathcal{E}_n$ is a collection of graphs on vertex set $[n]$. We are interested in

$$\lim_{n\to\infty} \Pr(\mathcal{E}_n)$$

where p, m are functions of n

$G_{n,m}$ can be viewed as a random graph process

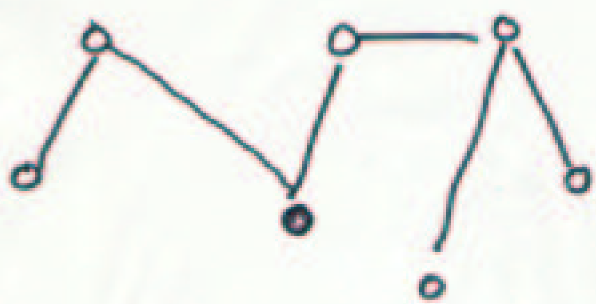


$E(G_{n,m}) = \{e_1, \ldots, e_m\}$ where e_i is chosen uniformly at random from $\binom{[n]}{2} \setminus \{e_1, \ldots, e_{i-1}\}$

Definition: The girth of a graph G is the size of the shortest cycle in G.

The chromatic number of G is the smallest number of colors in a proper coloring of G.

Note: large girth $\Rightarrow$ G is locally 2-colorable



Theorem (Erdös, 1959):

For all k, ℓ there exists a graph G with girth $> \ell$ and chromatic number $> k$.

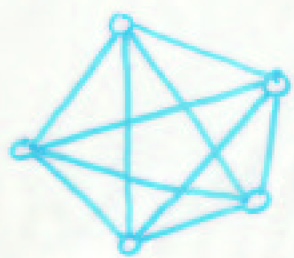
Pf: Consider $G_{n,p}$ for n sufficiently large and p carefully chosen.

$G_{n,p}$ can be altered to give a graph with the desired properties with nonzero probability.

$\square$

Theorem (Erdös, Rényi 1960)

$$\lim_{n\to\infty} \Pr\left(G_{n,m} \text{ contains a copy of } K_k\right) = \begin{cases} 0 & \text{if } m = o\left(n^{2-\frac{2}{k-1}}\right) \\ 1 & \text{if } m = \omega\left(n^{2-\frac{2}{k-1}}\right) \end{cases}$$



K_5

$f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$

$f(n) = \omega(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = +\infty$

Theorem (Erdös, Rényi 1959)

(i) If $m = \frac{n \log n}{2} + f(n)$ where $\frac{f(n)}{n} \to +\infty$ then

$$\lim_{n \to \infty} \Pr(G_{n,m} \text{ is connected}) = 1$$

$$\lim_{n \to \infty} \Pr(G_{n,m} \text{ has a perfect matching}) = 1$$

(ii) If $m = \frac{n \log n}{2} + f(n)$ where $\frac{f(n)}{n} \to -\infty$ then

$$\lim_{n \to \infty} \Pr(G_{n,m} \text{ is connected}) = 0$$

$$\lim_{n \to \infty} \Pr(G_{n,m} \text{ has a perfect matching}) = 0$$

Theorem (Erdös, Rényi 1960)

(i) If $m = cn + o(n)$ and $c < 1/2$ then a.a.s. the largest component of $G_{n,m}$ has $O(\log n)$ vertices.

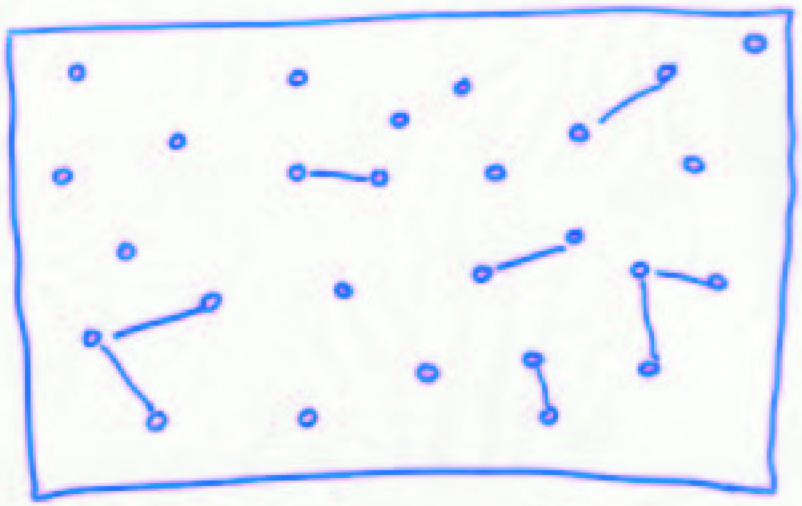
(ii) If $m = cn + o(n)$ and $c > 1/2$ then a.a.s. the largest component of $G_{n,m}$ has $\Omega(n)$ vertices and the second largest component has $O(\log n)$ vertices.

$f(n) = o(n)$ if $\frac{f(n)}{n} \to 0$

$f(n) = O(\log n)$ if $f(n) < C \log n$ for some constant C.

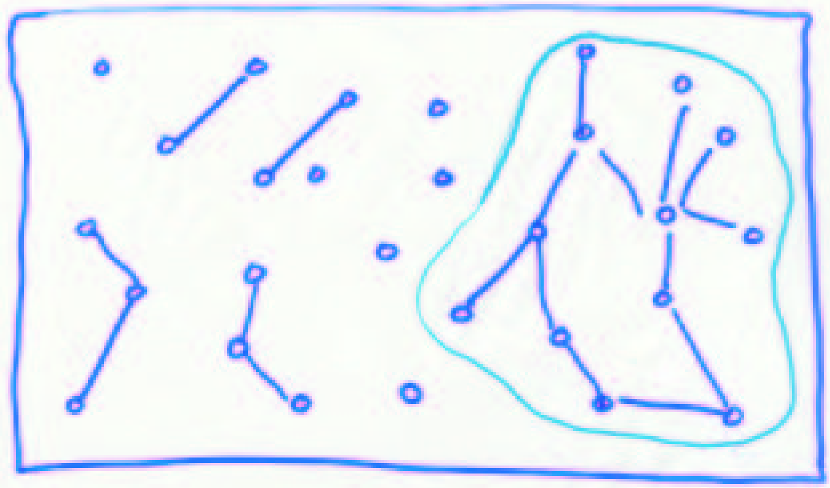
$f(n) = \Omega(n)$ if $f(n) > C \log n$ for some constant C.

$m < n/2$: all components are small



- $m = o(\sqrt{n})$. Isolated vertices and isolated edges.
- $m = \omega(\sqrt{n})$, $m = o(n^{2/3})$. Isolated vertices, isolated edges and paths with 2 edges.
- $m = cn$, $c < 1/2$. Every component has $O(\log n)$ vertices. No K_4's.

$\underline{m > n/2}$: $\exists$ a unique "giant component."



- $m = cn$, $c > 1/2$. There is one connected component that has $\Omega(n)$ vertices. All other components have $O(\log n)$ vertices.
- $m = cn \log n$, $c < 1/2$. Isolated vertices remain.
- $m = cn \log n$, $c > 1/2$. $G_{m,n}$ is connected and has a perfect matching.

An example

Claim: Let $p = n^{-\alpha}$, $4/3 < \alpha < 3/2$.

A.a.s. $G_{n,p}$ has at least one 2-path but no component with 4 or more vertices.

2-path 3-path 3-star.

this appears.

these do not appear

Pf: Let

X = # of 2-paths in $G_{n,p}$

Y = # of 3-paths

Z = # of 3-stars

complete graph on n vertices

For each 2-path A in K_n define

$$X_A = \begin{cases} 1 & \text{if } A \text{ is in } G_{n,p} \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$X = \sum_A X_A$$

there are $\frac{n(n-1)(n-2)}{2}$ terms in this sum

$$E[X] = E\left[\sum_A X_A\right]$$

$$= \sum_A E[X_A]$$

$$= \sum_A \Pr(X_A = 1)$$

$$= \sum_A p^2$$

$$= \frac{n(n-1)(n-2)}{2} p^2$$

Simillarly,

$$E[Y] = \frac{n(n-1)(n-2)(n-3)}{2} p^3$$

$$E[Z] = n \binom{n-1}{3} p^3$$

Now,

$$E[Y] \leq n^4 p^3 = n^{\overbrace{4-3\alpha}} \to 0$$

$\alpha > 4/3$

$\Rightarrow 4 - 3\alpha < 0$

$$E[Z] \leq n^4 p^3 = n^{4-3\alpha} \longrightarrow 0$$

<u>Markov's inequality</u>: If X takes only non-negative values and $\lambda > 0$ then

$$\Pr[X \geq \lambda] \leq \frac{E[X]}{\lambda}$$

So, $\Pr[Y \geq 1] \leq E[Y] \to 0$

$\Pr[Z \geq 1] \leq E[Z] \to 0$

$$E[X] = \frac{n(n-1)(n-2)}{2} p^2$$

$$\geq \frac{n^3}{8} p^2$$

assuming $n \geq 4$

$$= \frac{n^{3-2\alpha}}{8}$$

$$\longrightarrow +\infty$$

$$\boxed{\begin{array}{l} \alpha < 3/2 \\ \Rightarrow 3-2\alpha > 0 \end{array}}$$

Does this imply

$\Pr(\exists \text{ a 2-path in } G_{n,p})$

$\longrightarrow 1$

?

Recall: $\mathrm{Var}(X) = E\left[(X - E[X])^2\right]$

Claim: $\Pr(X=0) \leq \dfrac{\mathrm{Var}[X]}{E^2[X]}$

Proof:

$$\begin{aligned}
\Pr(X=0) &\leq \Pr\left((X - E[X])^2 \geq E^2[X]\right) \\
&\leq \frac{E\left[(X-E[X])^2\right]}{E^2[X]} \\
&= \frac{\mathrm{Var}[X]}{E^2[X]}
\end{aligned}$$

$\square$

To Show: $\mathrm{Var}[X] = o\left(E^2[X]\right)$

Since $X = \sum_{A} X_A$

$$
\begin{aligned}
\mathrm{Var}[X] &= E\left[(X - E[X])^2\right] \\
&= E[X^2] - E^2[X] \\
&= E\left[\left(\sum_A X_A\right)^2\right] - \left(\sum_A E[X_A]\right)^2 \\
&= \sum_{A,B} E[X_A X_B] - E[X_A]\,E[X_B] \\
&= \sum_{A \neq B} \underbrace{E[X_A X_B] - E[X_A]\,E[X_B]}_{\mathrm{Cov}(X_A, X_B)} \\
&\quad + \underbrace{\sum_A \mathrm{Var}[X_A]}
\end{aligned}
$$

This term is bounded above by $E[X]$ and

$E[X] = o(E^2[X])$

If A and B share an edge then

$$\mathrm{Cov}(X_A, X_B) = p^3 - p^4$$

If A and B do not share an edge then

$$\mathrm{Cov}(X_A, X_B) = 0$$

$$\sum_{A \neq B} \mathrm{Cov}(X_A, X_B) = \sum_{A} \sum_{\substack{B: \\ A \cap B \neq \emptyset}} p^3 - p^4$$

$$\leq \sum_{A} 2n p^3$$

$$\leq n^4 p^3$$

Since $a > 4/3$

$$\longrightarrow 0$$