Tianjin

June 2007

The Erdős-Rényi Phase Transition

Joel Spencer

TP! trivial being! I have received your letter, you should have written already a week ago.

The spirit of Cantor was with me for some length of time during the last few days, the results of our encounters are the following . . .

letter, Paul Erdős to Paul Turán
November 11, 1936

Paul Erdős and Alfred Rényi

On the Evolution of Random Graphs

Magyar Tud. Akad. Mat. Kutató Int. Közl

volume 8, 17-61, 1960

$\Gamma_{n,N(n)}$: n vertices, random $N(n)$ edges

[...] the largest component of $\Gamma_{n,N(n)}$ is of order $\log n$ for $\frac{N(n)}{n} \sim c < \frac{1}{2}$, of order $n^{2/3}$ for $\frac{N(n)}{n} \sim \frac{1}{2}$ and of order n for $\frac{N(n)}{n} \sim c > \frac{1}{2}$. This double "jump" when c passes the value $\frac{1}{2}$ is one of the most striking facts concerning random graphs.

The (Traditional) "Double Jump"

$G(n, p)$, $p = \frac{c}{n}$ (or $\sim \frac{c}{2}n$ edges)

(Average Degree c, "natural" model)

- $c < 1$

  Biggest Component $O(\ln n)$

  $|C_1| \sim |C_2| \sim \ldots$

  All Components simple (= tree/unicyclic)

- $c = 1$

  Biggest Component $\Theta(n^{2/3})$

  $|C_1|n^{-2/3}$ nontrivial distribution

  $|C_2|/|C_1|$ nontrivial distribution

  Complexity of C_1 nontrivial distribution

- $c > 1$

  Giant Component $|C_1| \sim yn$, $y = y(c) > 0$

  All other $|C_i| = O(\ln n)$ and simple

The Five Phases

Subcritical: $p = \frac{c}{n}$ and $c < 1$

Barely subcritical: $p \sim \frac{1}{n}$ and $p = \frac{1}{n} - \lambda(n) n^{-4/3}$ with $\lambda(n) \to \infty$

The Critical Window

$$p = \frac{1}{n} + \lambda n^{-4/3}$$

λ arbitrary real, but constant.

Barely supercritical: $p \sim \frac{1}{n}$ and $p = \frac{1}{n} + \lambda(n) n^{-4/3}$ with $\lambda(n) \to \infty$

Supercritical: $p = \frac{c}{n}$ and $c > 1$

• Barely Subcritical

$p \sim \frac{1}{n}$ and $p = \frac{1}{n} - \lambda(n)n^{-4/3}$ with $\lambda(n) \to \infty$

All components simple.

Top k components about same size

$|C_1| = o(n^{2/3})$

• Barely Supercritical

$p \sim \frac{1}{n}$ and $p = \frac{1}{n} + \lambda(n)n^{-4/3}$ with $\lambda(n) \to \infty$

Dominant Component

$|C_1| \gg n^{2/3}$, High Complexity

All other $|C| \ll n^{2/3}$, Simple

Duality: Remove Dominant Component and get Subcritical Picture.

Math Physics Bond Percolation

Z^d. Bond "open" with probability p

There exists a critical probability p_c

• Subcritical, $p < p_c$.

All C finite, $E[|C(\vec{0})|]$ finite

Pr$[|C(\vec{0})| \geq u]$ exponential tail

• Supercritical, $p > p_c$.

Unique Infinite Component

$E[|C(\vec{0})|]$ infinite

Pr$[|C(\vec{0})| \geq u|$finite$]$ exponential tail

• Critical, $p = p_c$.

All C finite, $E[|C(\vec{0})|]$ infinite, heavy tail

Key topic: $p = p_c \pm \epsilon$ as $\epsilon \to 0$.

Random 3-SAT

n Boolean $x_1, \ldots, x_n$

$L = \{x_1, \overline{x_1}, \ldots, x_n, \overline{x_n}\}$ literals

Random Clauses $C_i = y_{i1} \vee y_{i2} \vee y_{i3}$, $y_{ij} \in L$

$f(m) :=$ Pr[$C_1 \wedge \cdots \wedge C_m$satisfiable]

Conjecture: There exists critical c_0

- Subcritical, $c < c_0$, $f(cn) \sim 1$
- Supercritical, $c > c_0$, $f(cn) \sim 0$

Friedgut: Criticality, but possibly nonuniform

Critical Window **???**: $m_0(n)$ with $f(m_0) = \frac{1}{2}$.

Is there scaling $m = m_0 + \lambda n^\alpha$ to "see" $f(m)$ go ~ 1 to ~ 0.

Evolution of n-Cube

Ajtai, Komlos, Szemeredi

Bollobas, Luczak, Kohayakawa

Borgs, Chayes, Slade, JS, van der Hofstad

$p = c/n$

$c < 1$ subcritical

$c > 1$ giant $\Omega(2^n)$ component

Critical $p_0 \sim n^{-1}$

At $p_0(1-\epsilon)$ all "small"

At $p_0(1+\epsilon)$. For $\epsilon = \Omega(n^{-100})$ and more:

Giant $2\epsilon n$. Second *open*

Critical Window (dominant emerges): *open*

Poisson Birth Process

Root node "Eve"

Parameter c

Each node has $Po(c)$ children

(Poisson: $\Pr[Po(c) = k] = e^{-c}c^k/k!$)

$Z_t \sim Po(c)$, iid

t-th node has Z_t children

Queue Size Y_t. $Y_0 = 1$ (Eve)

$Y_t = Y_{t-1} + Z_t - 1$ (Has children and dies)

Fictional Continuation: Y_t defined though process stops when some $Y_s = 0$.

Size $T = T_c^{po}$ is *minimal* t with $Y_t = 0$.

$T = \infty$: All $Y_t > 0$.

$T = T_c$ is total size

Binomial Birth Process

Parameters m, p

$Z_t \sim B[m,p]$, iid

$T = T^{bin}_{m,p}$ total size.

For m large, p small, mp moderate:

Binomial is very close to Poisson $c = mp$.

Binomial Birth Process very close to Poisson Birth Process

Graph Birth Process

Parameters n, p

Generate $C(v)$ in $G(n, p)$. `BFS`

Queue: $Y_0 = 1$, $Y_t = Y_{t-1} + Z_t - 1$

Points Born: $Z_t \sim B[N_{t-1}, p]$

Dead Points (popped): t

Live Points (in Queue): Y_t

Neutral Points(in Reservoir): N_t

$t + Y_t + N_t = n$

$N_0 = n-1$, $N_t = N_{t-1} - Z_t$, $N_t \sim B[n-1, (1-p)^t]$

$T = T^{gr}_{n,p}$: minimal t with $Y_t = 0$

$T = t$ implies $N_t = n - t$

Poisson Birth Trichotomy

- $c < 1$

  T finite

- $c = 1$

  T finite

  $E[T]$ infinite (heavy tail)

- $c > 1$

  $\Pr[T = \infty] = y = y(c) > 0$

Poisson Birth Exact

$$\Pr[T_c = u] = \frac{e^{-uc}(uc)^{u-1}}{u!}$$

$$\Pr[T_1 = u] = \frac{e^{-u}u^{u-1}}{u!} = \Theta(u^{-3/2})$$

For $c > 1$, $\Pr[T = \infty] = y = y(c) > 0$ where

$$1 - y = e^{-cy}$$

For $c < 1$, $\alpha := ce^{1-c} < 1$

$\Pr[T_c > u] = O(\alpha^u)$ Exponential Tail

Poisson Birth Near Criticality

$c = 1 + \epsilon$, $T = T_c^{po}$

$\Pr[T = \infty] \sim 2\epsilon$

$\Pr[T = u] \sim (2\pi)^{-1/2} u^{-3/2} (ce^{1-c})^k$

$\ln[ce^{1-c}] \sim -\epsilon^2/2$

- u small: $u = o(\epsilon^{-2})$

$\Pr[T_c = u] \sim \Pr[T_1 = u] = \Theta(u^{-3/2})$

Scaling: $u = A\epsilon^{-2}$

$\Pr[\infty > T_{1+\epsilon} > A\epsilon^{-2}] = \epsilon e^{-(1+o(1))A/2}$

$\Pr[T_{1-\epsilon} > A\epsilon^{-2}] = \epsilon e^{-(1+o(1))A/2}$

Poisson Birth $\sim$ Graph Birth

$Z_1 \sim B[n-1,p]$ roughly $Po(c)$, $c = pn$.

Ecological Limitation: $Z_t \sim B[N_{t-1},p]$.

Process succeeds, N_{t-1} gets smaller

Fewer new vertices

Death is inevitable

Upper: $\Pr[T^{gr}_{n,p} \geq u] \leq \Pr[T^{bin}_{n-1,p} \geq u]$

Proof: Replenish reservoir

Lower: $\Pr[T^{gr}_{n,p} \geq u] \geq \Pr[T^{bin}_{n-u,p} \geq u]$

Proof: Hold reservoir to $n-u$.

Why $n^{-4/3}$ for Critical Window

$p = (1+\epsilon)/n$, $\epsilon > 0$, $\epsilon = o(1)$.

$\Pr[T^{po}_{1+\epsilon} = \infty] \sim 2\epsilon$.

The $\sim 2\epsilon n$ points "going to infinity" merge to form dominant component.

T^{po} finite is $O(\epsilon^{-2})$, corresponds to component sizes $O(\epsilon^{-2})$.

Finite/Infinite Poisson Dichotomy becomes Small/Dominant Graph Dichotomy *if* $\epsilon^{-2} \ll 2n\epsilon$, or $\epsilon \gg n^{-1/3}$.

The Barely Subcritical Region

$p = (1-\epsilon)/n$, $\epsilon = \lambda n^{-1/3}$,

$Pr[|C(v)| \geq u] \leq \mathrm{Pr}[T_{1-\epsilon} \geq u]$

$u = K\epsilon^{-2} \ln n \Rightarrow \mathrm{Pr} = o(n^{-1})$

No Such component.

More delicately:

Parametrize $u = K\epsilon^{-2} \ln \lambda = Kn^{2/3}\lambda^{-2} \ln \lambda$

K big: $Pr[|C(v)| \geq u] = O(\epsilon\lambda^{-10})$

Expected $n\epsilon\lambda^{-10} = n^{2/3}\lambda^{-9}$ vertices in components of size $\geq Kn^{2/3}\lambda^{-2} \ln \lambda$

No such component!

Barely Supercritical

$p = (1+\epsilon)/n$, $\epsilon = \lambda n^{-1/3}$, $\lambda \to +\infty$

Trichotomy on Component Size

Small: $|C| < K\epsilon^{-2} \ln n$ [can be impoved!]

Large: $(1-\delta)2\epsilon n < |C| < (1+\delta)2\epsilon n$

Awkward: All else

No Middle Ground

No Awkward Components

Suffices: $\Pr[C(v) \text{ awkward}] = o(n^{-1})$

No Middle Ground

$Y_t = n - t - N_t = B[n-1, 1-(1-p)^t] - (t-1)$

At start $E[Y_t] \sim \epsilon t$ [Negligible EcoLim]

When $t \gg \epsilon^{-2} \ln n$, $E[Y_t] \gg \text{Var}[Y_t]^{1/2} \sim t^{1/2}$,

$\Pr[Y_t = 0] = o(n^{-10})$

Later $E[Y_t] = (n-1)[1-(1-p)^t] - (t-1) \sim \epsilon t - \frac{t^2}{2n}$

For $t \sim 2\epsilon n$, $E[Y_t] \sim 0$, dominant component.

$|C(v)| = t$ implies $Y_t = 0$.

For $t \sim y\epsilon n$, $y \neq 2$:

$\Pr[|C(v)| = t] \leq \Pr[Y_t = 0] = o(n^{-10})$.

Escape Probability

$S := K\epsilon^{-2}\ln n$, $\alpha := \Pr[|C(v)| \geq S]$

$\Pr[|C(v)| \geq S] \leq \Pr[T^{bin}_{n-1,p} \geq S]$

$np = 1+\epsilon$, $S \gg \epsilon^{-2}$ so $\sim 2\epsilon$

$\Pr[T^{bin}_{n-S,p} \geq S] \leq \Pr[|C(v)| \geq S]$

(Here $\epsilon \gg n^{-1/3}\ln^{1/3} n$ but with care . . .)

- As $Sp = o(\epsilon)$ EcoLim negligible!

$p(n-S) = 1+\epsilon+o(\epsilon)$ so $\Pr \sim 2\epsilon$

Sandwich: Escape Prob $\sim 2\epsilon$

Almost Done

Not Small implies Large $\sim 2\epsilon n$

Expected $2\epsilon n$ in Large components

BUT

Can we have two

of size $2\epsilon n$

half the time?

Sprinkling

Add sprinkle of $n^{-4/3}$, $p \leftarrow p^{+}$

If $G(n,p)$ had two Large they would merge

That would give $\geq 4\epsilon n$ in $G(n,p^{+})$

But $p^{+} = (1+\epsilon+o(\epsilon))/n$ has nothing $\geq 4\epsilon n$

Conclusion:

- $G(n,p)$ has precisely one Large component
- It has size $\sim 2\epsilon n$
- As no middle ground:

All other component sizes $\leq K\epsilon^{-2}\ln n$.

So Large Component is Dominant Component

Computer Experiment (Try It!)

$n = 500000$ vertices. Start: Empty

Add random edges

Parametrize $e/\binom{n}{2} = (1 + \lambda n^{-1/3})/n$

Merge-Find for Component Size/Complexity

$-4 \leq \lambda \leq +4$, $|C_i| = c_i n^{2/3}$

See biggest merge into dominant

It is six in the morning.

The house is asleep.

Nice music is playing.

I prove and conjecture.

– Paul Erdős, in letter to Vera Sós