Colóquio Brasileiro de Matemática

IMPA, Rio de Janeiro, July 2005

Counting

CONNECTED

Graphs

using ERDŐS MAGIC

Joel Spencer

Courant Institute

Joint with

Remco van der Hofstad (EURRANDOM)

Complexity = E-V+1

$C(k,l)$ = Number of

CONNECTED Labelled Graphs

k Vertices

Complexity l

$C(k,0) = k^{k-2}$ Cayley

$C(k,l) \sim c_l k^{3l/2} k^{k-2}$ Wright

$l > (\frac{1}{2}+\epsilon)k \ln k$ Erdős-Rényi

General $k, l \to \infty$ Bender, Canfield, McKay (1990)

via 2-core: Pittel, Wormald (2005)

Areas

- Discrete Mathematics
- Theoretical Computer Science

  Breadth First Search Algorithms
- Probability
- Percolation

Tilted Balls in Bins

$k-1$ balls, k bins, $p \in (0,1]$

Truncated Geometric

Ball j in Bin T_j

$$\Pr[T_j = i] = \frac{p(1-p)^{i-1}}{1-(1-p)^k}$$

Z_i balls in bin i

$Y_0 = 1$, $Y_i = Y_{i-1} + Z_i - 1$ (so $Y_k = 0$)

TREE: $Y_t > 0$, $0 \le t < k$ (=Parking Function)

$$M := \sum_{i=0}^{k} (Y_i - 1) = \binom{k}{2} - \sum_{j=1}^{k} T_j$$

$G(k,p)$ Random Graph

Vertices $0,1,\ldots,k-1$

Adjacency Prob p

THM: Prob $G(n,p)$ Connected $= A_1A_2$

THM: Prob $G(k,p)$ Connected & Complexity l

$= C(k,l)p^{k+l-1}(1-p)^{k(k-1)/2-k-l+1} = A_1A_2A_3$

with

$A_1 = (1-(1-p)^k)^{k-1}$

$A_2 = \text{Pr}[\text{TREE}]$

$A_3 = \text{Pr}[BIN[M,p] = l|\text{TREE}]$

Strategy: A_2, A_3 determine $C(k,l)$

Breadth First Search

1	2	3	4	5
N	N	Y	Y	N
N	N	-	-	N
Y	N	-	-	N
-	Y	-	-	Y
-	-	-	-	-
-	-	-	-	-

$T_3 = T_4 = 1$, $T_1 = 3$, $T_2 = T_5 = 4$

A_1: All T_j defined

$\vec{Z} = (2, 0, 1, 2, 0, 0)$

Walk $\vec{Y} = (1, 2, 1, 1, 2, 1, 0)$

TREE: BFS doesn't terminate early

Tree Edges $03, 04, 41, 12, 15$

$M = 2$ Unexposed $34, 25$

Setting the Tilt p

$\mu := E[M]$, $\sigma^2 := Var[M]$

$$p\mu = l$$

Three Regimes

`Small` $l = o(k)$, $k^{-3/2} \ll p \ll k^{-1}$

`Large` $l = \Theta(k)$, $p = \Theta(k^{-1})$

`Very Large` $l \gg k$, $p \gg k^{-1}$

($l > ck \ln k$; $p > c' \frac{\ln k}{k}$ Erdős-Rényi)

($l \neq 0$, fixed); $p = \Theta(k^{-3/2})$; complicated)

`Small`: $k^{-3/2} \ll p \ll k^{-1}$

$\epsilon = \frac{1}{2}pk$

Left Z_i Poisson $1+\epsilon$

Galton-Watson $\Pr[\text{ESC}] \sim 2\epsilon$

Right $Z_i^* = Z_{k-i}; Y_i^* = Y_{k-i}$

$Y_0^* = 0$, $Y_i^* = Y_{i-1}^* + 1 - Z_i^*$

Z_i^* Poisson $1-\epsilon$

$\Pr[\text{ESC}^*] \sim \epsilon$

*** Scaling for ESC, ESC^* is $\epsilon^{-2} \ll k$

$\Pr[\text{TREE}] \sim \Pr[\text{ESC}] \Pr[ESC^*] \sim 2\epsilon^2$

Note: $p \ll k^{-3/2} \Rightarrow \Pr[\text{TREE}] \sim k^{-1}$

$p = \Theta(k^{-3/2}) \Rightarrow \Pr[\text{TREE}]$ complicated

`Large`: $p \sim \frac{c}{k}$

Left Z_i Poisson $\frac{c}{1-e^{-c}}$

Galton-Watson Pr[ESC] $\sim 1 - e^{-c}$

Right $Z_i^* = Z_{k-i}; Y_i^* = Y_{k-i}$

$Y_0^* = 0$, $Y_i^* = Y_{i-1}^* + 1 - Z_i^*$

Z_i^* Poisson $\frac{ce^{-c}}{1-e^{-c}}$

Pr[ESC*] $\sim 1 - \frac{ce^{-c}}{1-e^{-c}}$

Chernoff: $Y_i > 0$ in middle

Pr[TREE] $\sim$ Pr[ESC] Pr[ESC^*] $\rightarrow 1-(c+1)e^{-c}$

`Very Large` $p \gg k^{-1}$

Chernoff: $Y_i > 0$ all i

Pr[TREE] $\rightarrow 1$

Gaussian M

$M := \binom{k}{2} - \sum_{j=1}^{k} T_j$

Esseen: $\Pr[M < \mu + u\sigma] \rightarrow \Pr[N(0,1) < u]$

*** Still holds *conditional on* $TREE$

Hardest when p barely $\gg k^{-3/2}$

Easy when p `Large`

Trivial when p `Very Large`

From CLT to Local Stats

$E[W] = \mu$, $\mathrm{Var}[W] = \sigma^2$

$V = BIN[W, p]$, $l = \mu p = E[V]$

Formally: Infinite sequence of W_k, p_k, σ_k, l_k

Assume ** $p^2\sigma^2 = O(p\mu)$ **

$(\sigma^+)^2 := p^2\sigma^2 + p\mu$

Assume $\sigma^{-1}(W - \mu) \to N(0, 1)$

THM: V Local CLT, Mean l, Var $(\sigma^+)^2$

$$\Pr[V = l] \to \frac{1}{\sqrt{2\pi}\sigma^+}$$

Apply to $M|$TREE

$$A_3 = \Pr[BIN[M,p] = l | TREE]$$

`Small`: $p\mu = l = p^2\sigma^2$, $A_3 \sim (4\pi l)^{-1/2}$

`Very Large`: $p\mu = l \gg p^2\sigma^2$, $A_3 \sim (2\pi l)^{-1/2}$

`Large`: $p \sim \frac{c}{k}$. $p^2\sigma^2 = \Theta(l)$, $A_3 \sim g(c) l^{-1/2}$

The Giant/Dominant Component

$G(n,p)$

$p = \frac{c}{n}$, $c > 1$ Erdős-Rényi Giant

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

Supercritical: Dominant Component

THM: Probability $C(v)$ has k vertices, complexity l is $\sim A_1^* A_2 A_3$ with

$$A_1^* = \Pr[BIN[n-1, 1-(1-p)^k] = k-1]$$

Corollary: Local Stats for k, l of Giant/Dominant Component. Correlated Gaussians

Generating Random Connected Graph

with Nitin Arora

Time $\Theta(K+L)$ (!!) for $L=\Omega(\ln K)$

p with $p\mu = L$

Tilted Balls into Boxes

$L=\Omega(K)$ get BFS Tree with prob. $\Omega(1)$

$L=o(K)$ use Fast Abort.

Add precisely L of M unexposed with prob.

$$\frac{\Pr[\mathrm{BIN}[M,p]=L]}{\max_m \Pr[\mathrm{BIN}[m,p]=L]}$$

Any new possibility that existence acquires,

even the least likely,

transforms everything about existence.

– Milan Kundera