Cutting Resillient Networks

Xing Shi Cai Cecilia Holmgren  Fiona Skerman 

Uppsala University (Sweden)

Luc Devroye 

Mcgill University

~~Cutting Resillient Networks~~

k-cut on Paths and Some Trees

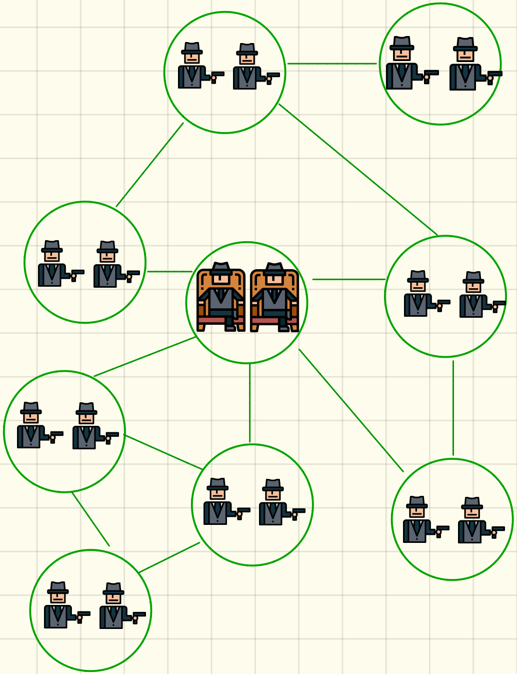
Xing Shi Cai Cecilia Holmgren  Fiona Skerman 

Uppsala University (Sweden)

Luc Devroye 

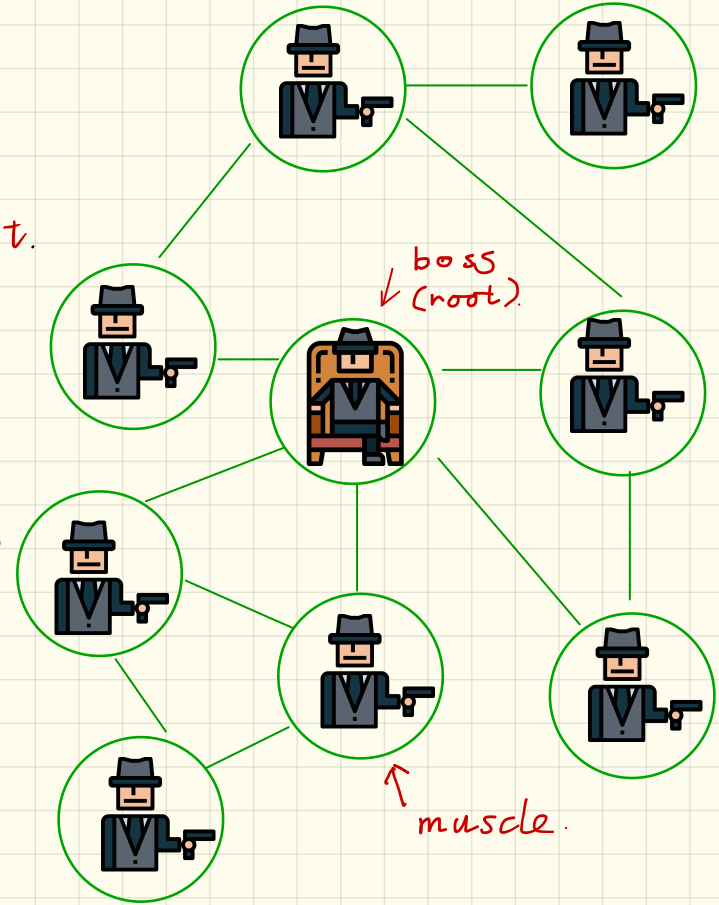
Mcgill University

The Model



Rooted Graph

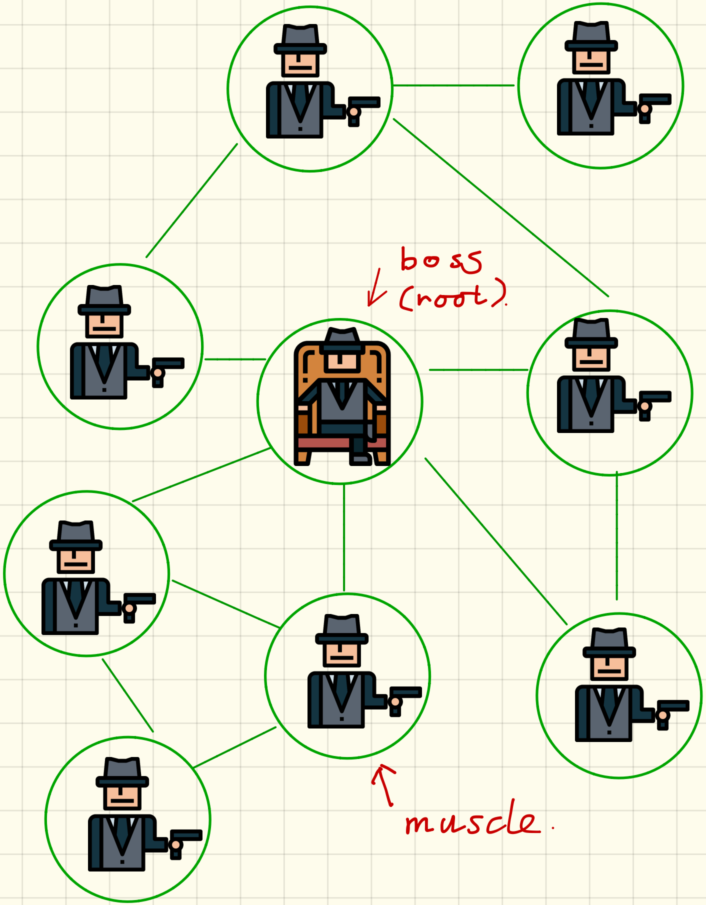
- A rooted graph has one node labelled as the root.
- Can be viewed as models for criminal networks, terrorist cells, or botnets (mallicious P2P networks).



Destroying a Network

We do not know where is the boss. So we

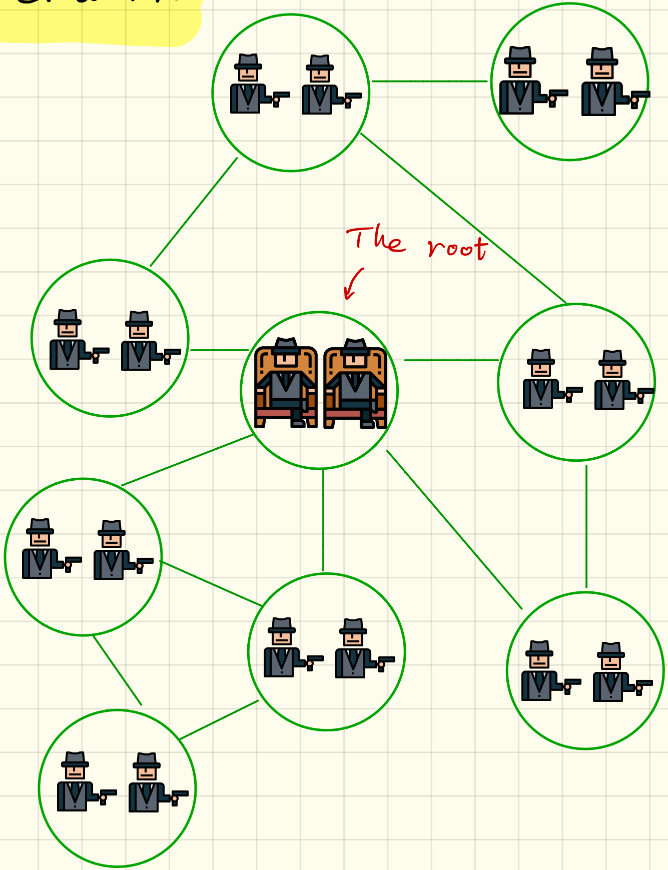
1. Choose a node unit at random. Remove it.
2. If the graph becomes disconnected, keep only the component containing the boss.
3. Repeat until the root is removed.



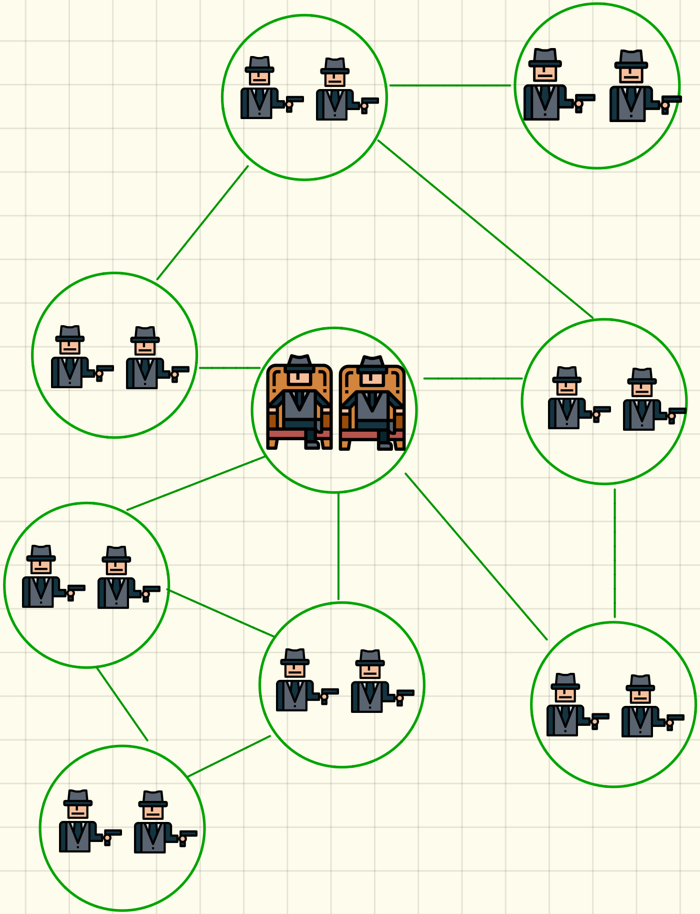
Destroying a Resillient Network

Assume each node has $k \in \mathbb{N}$ backups. We

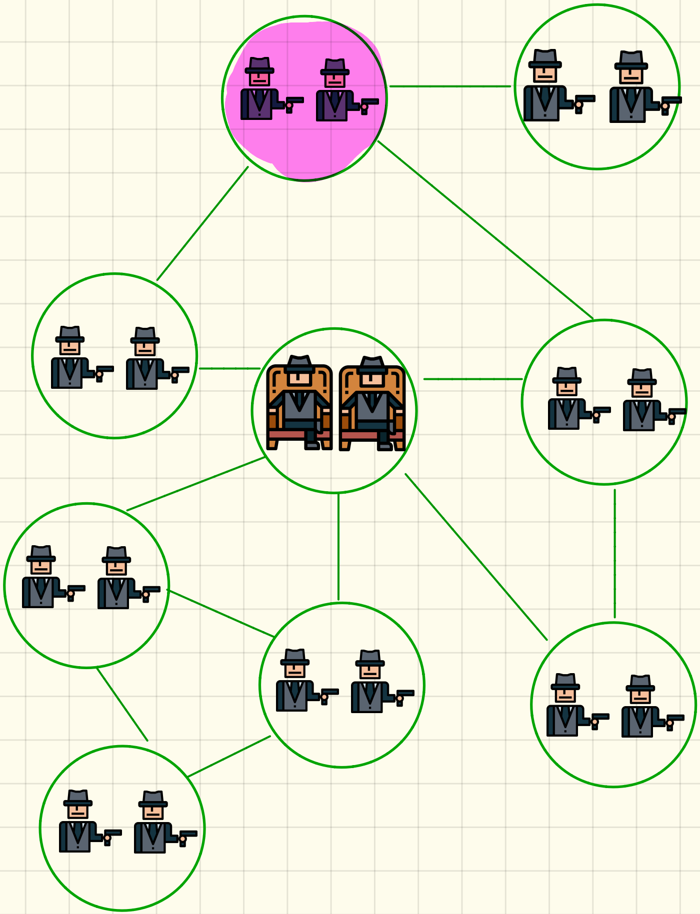
1. Choose a node unif. at random. Remove one of its back up (cut it once)
2. Remove a node if all k backups are gone.
3. Keep only the component containing the root.
4. Repeat untill the root is gone.



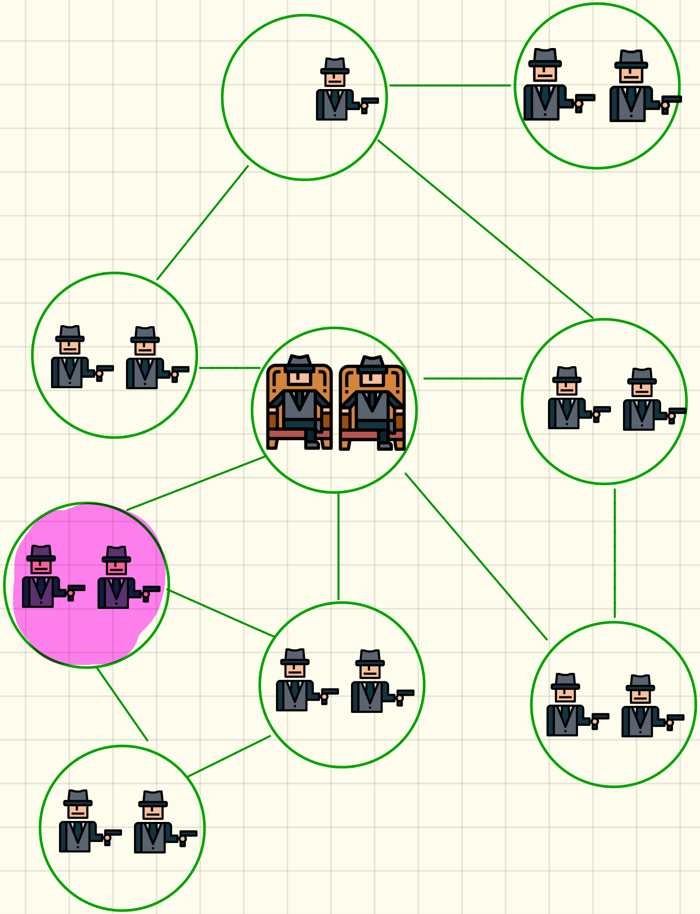
Example



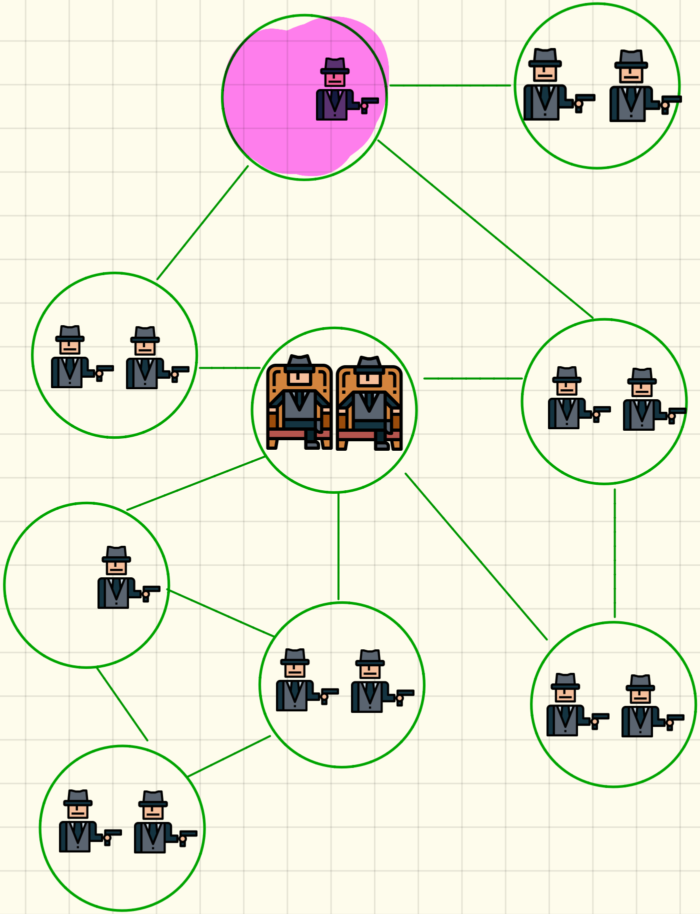
Example



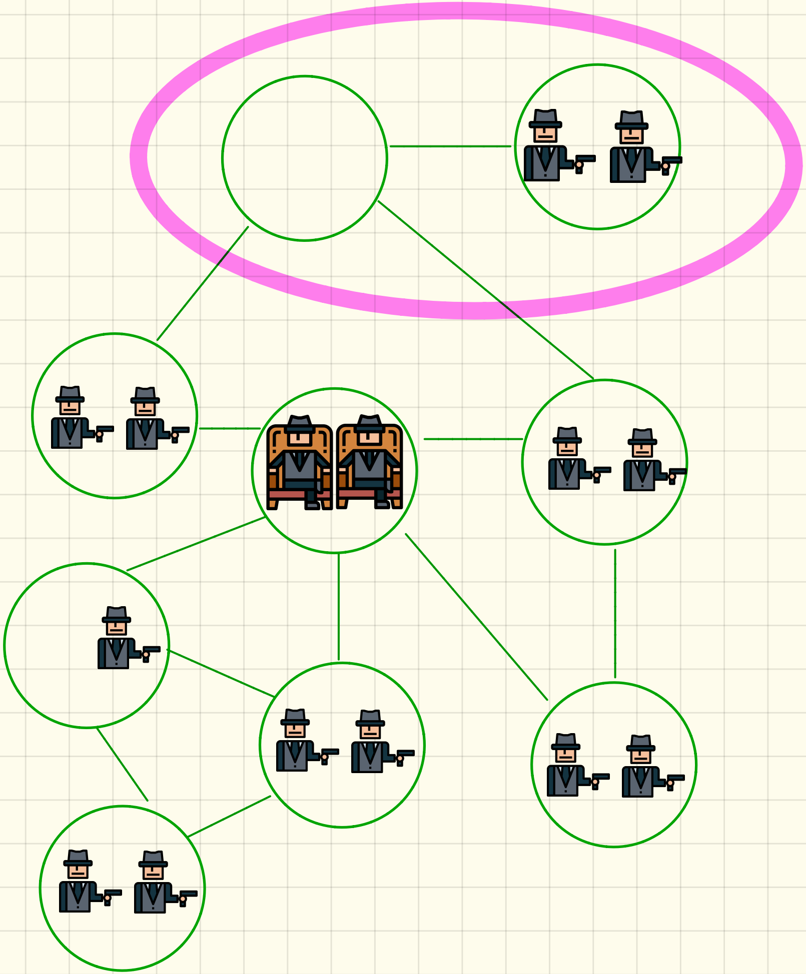
Example



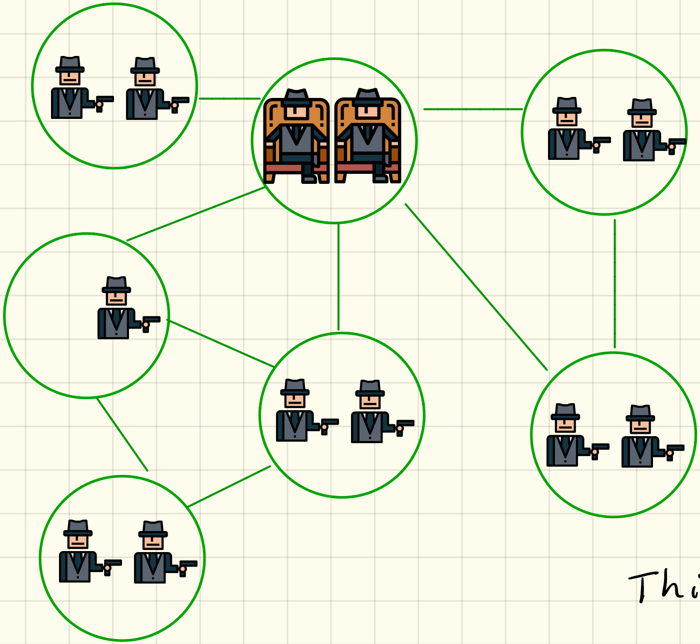
Example



Example

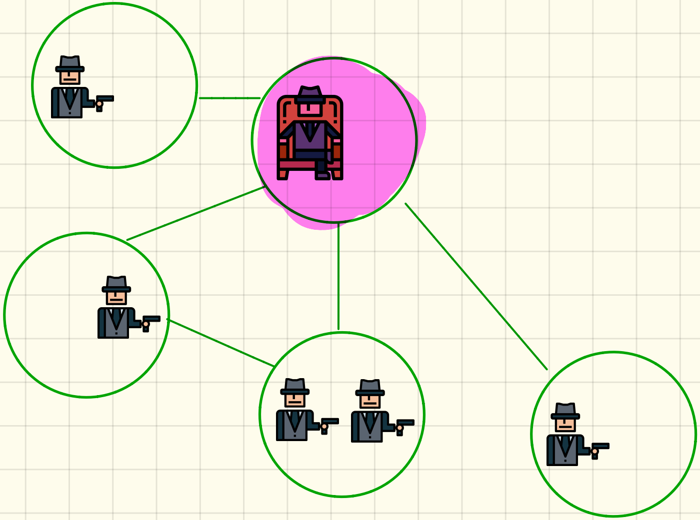


Example



This continues ...

Example

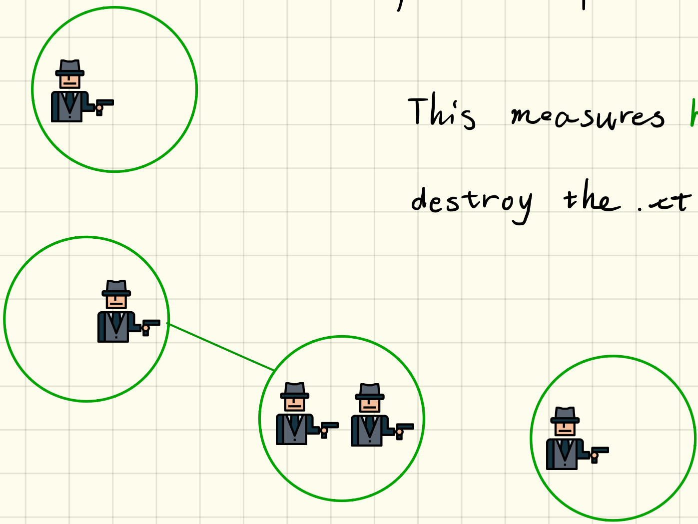


untill the root is gone.

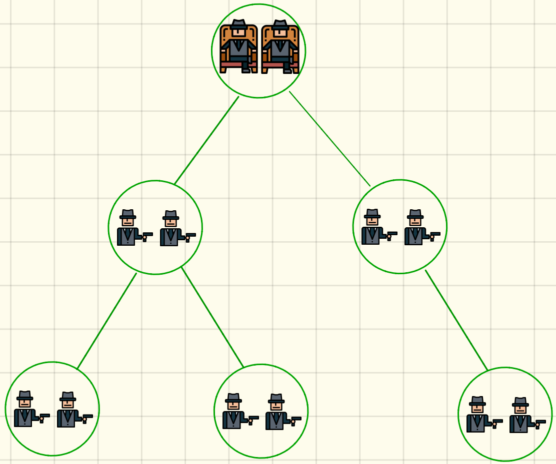
Example

We mostly care $K(G_n)$ — the number of cuts needed to for the process to end.

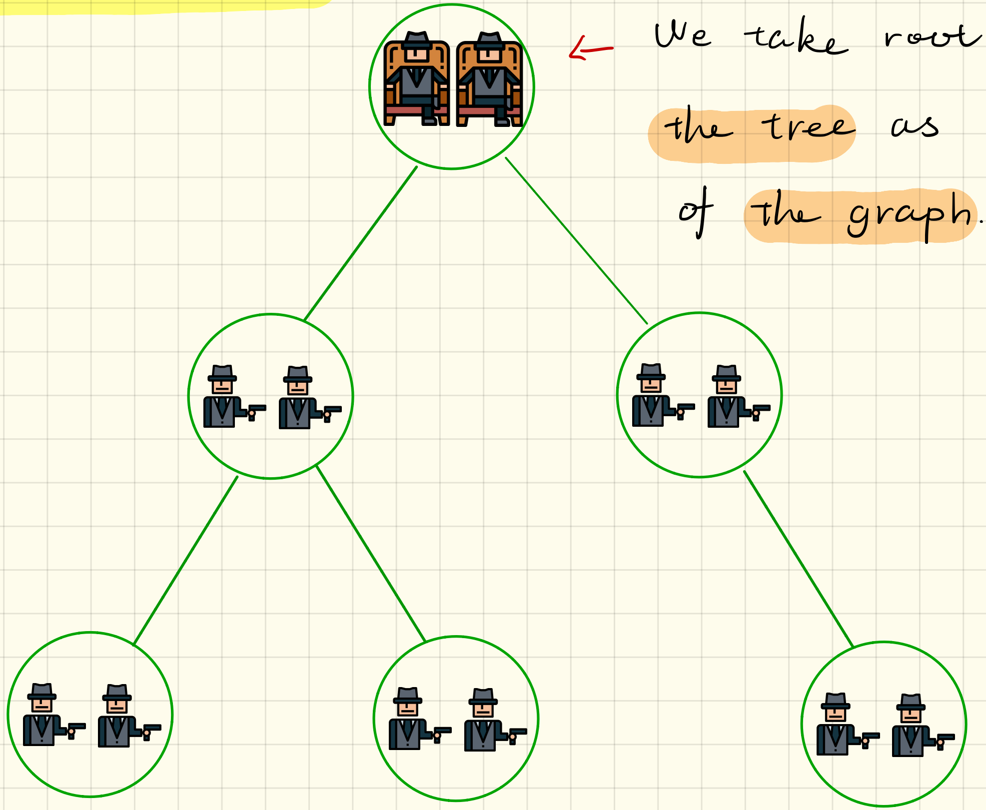
This measures how hard it is to destroy the network.



Cutting Rooted Trees



Cutting Rooted Trees



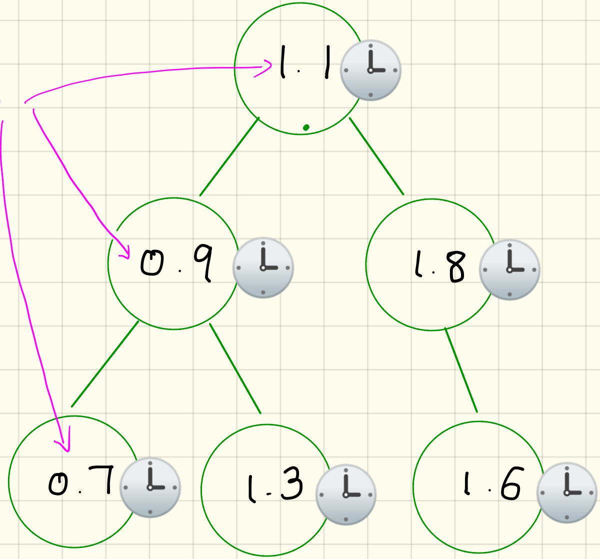
We take root of the tree as root of the graph.

The case $k=1$

- Let $\mathbb{T}_n$ be a rooted tree with n vertices.
- Let $K(\mathbb{T}_n)$ be the num. of cuts needed to destroy $\mathbb{T}_n$.
- $K(\mathbb{T}_n)$ has been studied for $k=1$ in
  - Cayley trees Meir and Moon (1970)
  - Complete Binary trees Janson (2004).
  - Conditional Galton-Watson trees Janson (2006).
    Addario-Berry, Broutin, Holmgren (2014).
  - Binary Search Trees and Split Trees Holmgren (2011, 2012).
  - RRT Meir and Moon (1974) Drmota et al. (2009).

An equivalent model $k=1$

- We give each node v a time stamp $T_v \sim \text{Exp}(1)$ iid



- We cut a node v at time T_v if v is still in the tree.

- Each time we are still cutting a uniform random node.

Idea comes from Svante (2004).

Records

v is in tree at time T_v

$\Longleftrightarrow$

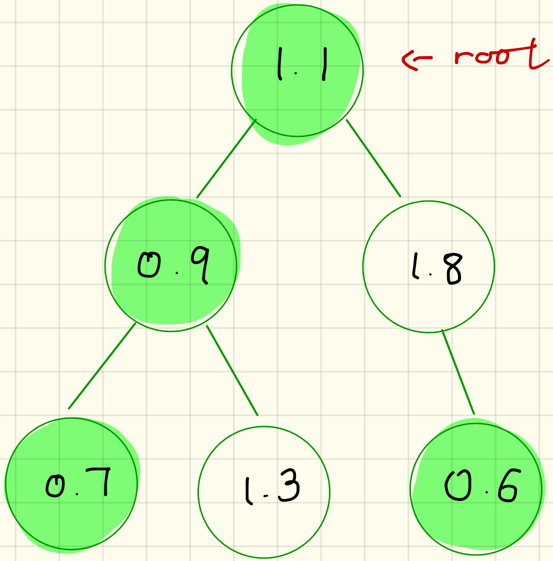
No ancestor of v died before T_v

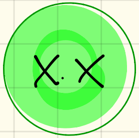
$\Longleftrightarrow$

$$\boxed{T_v < \min_{u:\, u \prec v} T_u.}$$

Janson call T_v or (simply v)

a record.





$$\boxed{\#\text{ of records} = \#\text{ of cuts}}$$

Generalize to $k \geq 1$.

- Each node v get timestamps

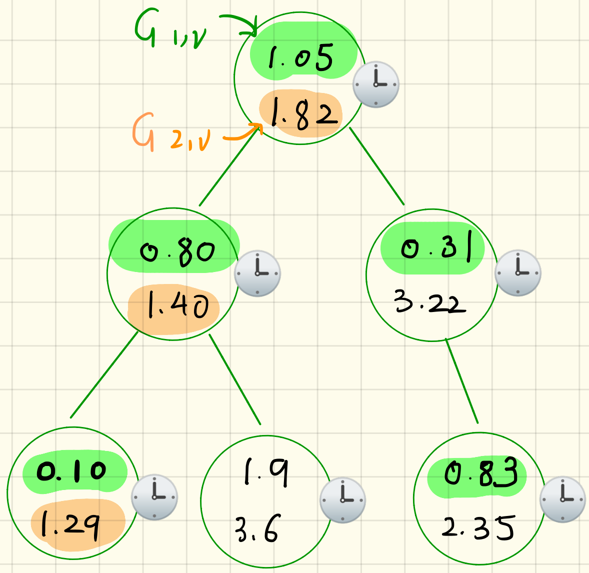
$$T_{1v}, T_{2v}, \ldots \sim \mathrm{Exp}(1) \text{ iid}.$$

- Let $G_{rv} = \sum_{i=1}^{r} T_{iv} \sim \mathrm{Gam}(k, 1)$
- Cut v at time G_{rv} if v is still in tree $\longleftrightarrow$

Time u dies.

$$G_{rv} < \min_{u:\, u \prec v} G_{ku}$$

- Call such G_{rv} or (simply v) an r-record

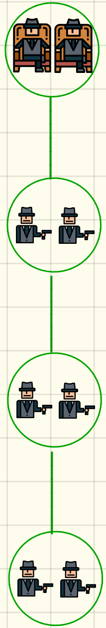


X.X ← 1-record

X.X ← 2-record

cuts → $$\boxed{K(\mathbb{T}_n) = \sum_{r=1}^{k} K_r(\mathbb{T}_n)}$$ ← Number of r-records.

k-cut on a path



The simplest graph – path

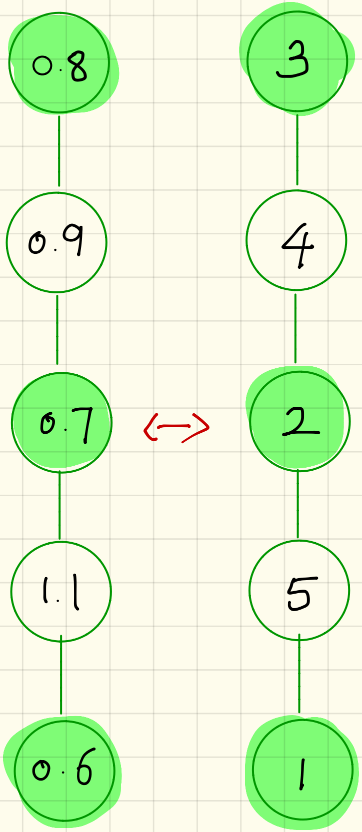
Let $\mathbb{P}_n$ be a path a n nodes.

For all graphs $\mathbb{G}_n$ of n nodes,

$$\boxed{K(\mathbb{P}_n) \precsim K(\mathbb{G}_n)}$$

i.e., a path is the easiest to cut.

Quiz: Which graph is the hardest to cut?



For $k=1$, $K(\mathbb{P}_n) \sim$ # of records

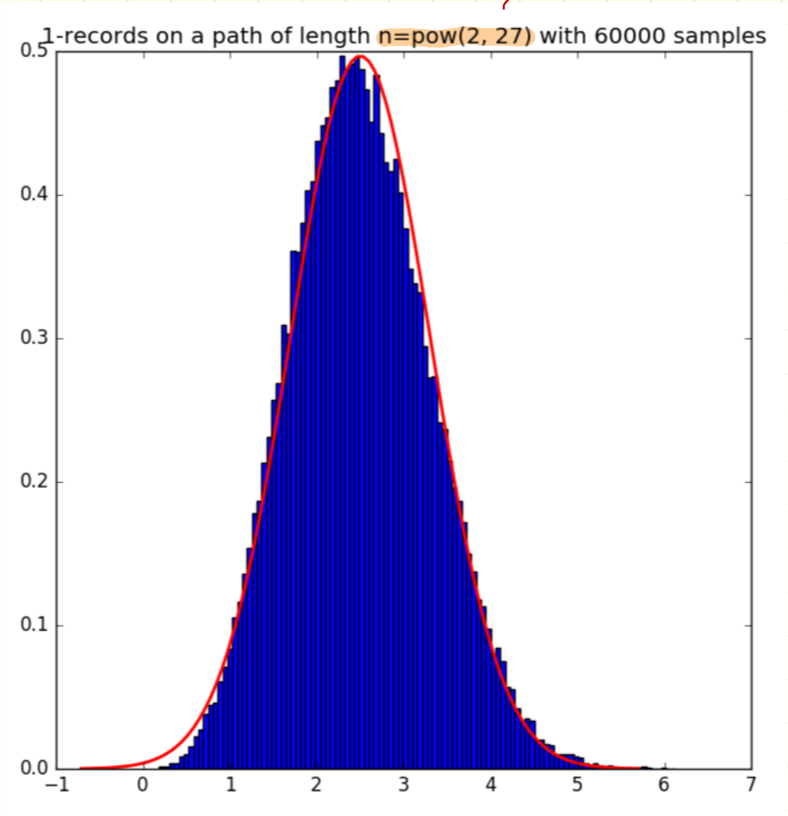
in unif. rand. permutation.

$$\frac{K(\mathbb{P}_n) - \log(n)}{\sqrt{\log(n)}} \xrightarrow{d} N(0,1) \quad (\text{normal})$$

$$\mathbb{E}(K(\mathbb{P}_n)) = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} = H_n \sim \log n$$

Simulation for $k=2$.

$2^{27} \simeq 10^8$



Looks like a normal.

Tried to prove.

But failed!

More simulations. $k=2$

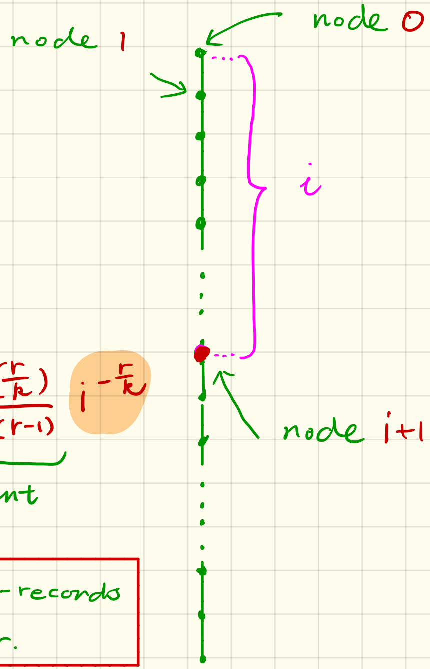
```
==========================================================
Analysis 30000 samples for path length 536870912 (2^29)
==========================================================

Mean 58141.156467
Mean divided by sqrt(n) 2.509278

Variance 358057278.161121
Variance divided by n 0.666934

The 3 moment divided by variance^(3/2) is 0.238525
The 4 moment divided by variance^(4/2) is 2.919119
The 5 moment divided by variance^(5/2) is 2.326212
The 6 moment divided by variance^(6/2) is 14.591784
```

$\simeq 10^9$

- This cannot be a normal distribution.
- The expectation is order $\sqrt{n}$
- The variance is order n.

Can we find the constant?

The moment – Expectation.



- Let $I_{r,i+1}$ be the indicator that $i+1$ is a r-record.

- Then

$$\mathbb{E}(I_{r,i+1}) = \int_0^\infty \underbrace{\frac{e^{-x}\,x^{r-1}}{r!}}_{\text{Density of } G_{r,i+1}} \overbrace{\mathbb{P}(\mathrm{Gam}(k) > x)^{i}}^{\text{Every node above } i+1 \text{ dies after } x} \approx \underbrace{\frac{(k!)^{\frac{r}{k}}}{k}\,\frac{\Gamma(\frac{r}{k})}{\Gamma(r-1)}}_{\text{constant}}\, i^{-\frac{r}{k}}$$

- Summing this up

$$\mathbb{E}(K_r(\mathbb{P}_n)) = \sum_{i=0}^{n} \mathbb{E}(I_{r,i}) \approx \eta_{k,r} \cdot n^{1-r/k}$$

$\boxed{\text{Only 1-records matter.}}$

- For $k=2$,

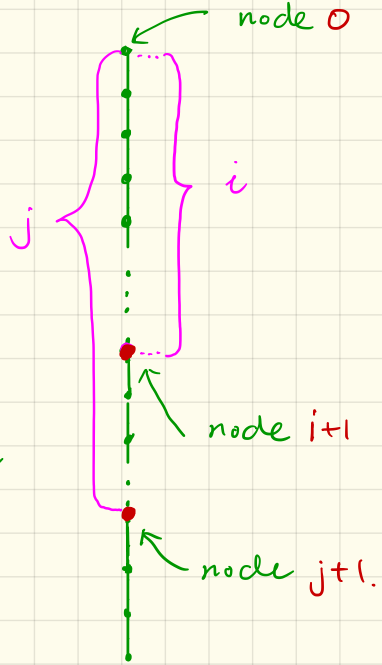
2.50$\sqrt{n}$ by simulation

$$\mathbb{E}(K_1(\mathbb{P}_n)) \sim \sqrt{2\pi n} \approx 2.5066\sqrt{n}$$

The moment – Variance

- We only care about 1-records.
- Similar to expectation

$$\frac{\mathbb{E}\left(K_1(\mathbb{P}_n)^2\right)}{n^{2-\frac{2}{k}}} \sim \underbrace{\frac{\Gamma(\frac{2}{k})\,(k!)}{k-1} + \frac{\pi \cot(\frac{\pi}{k})\,\Gamma(\frac{2}{k})\,(k!)^{\frac{2}{2}}}{2(k-2)(k-1)}}_{\text{Rather complicated constant}}$$



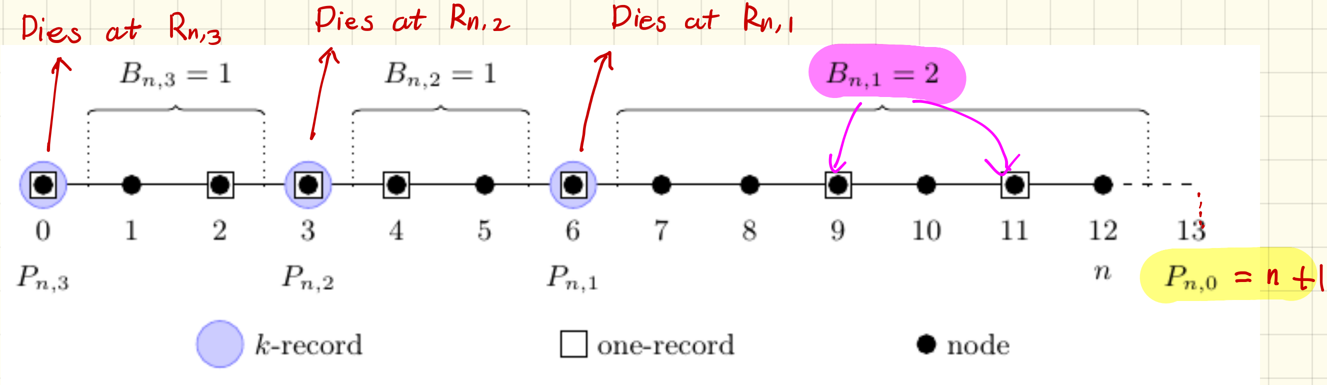
- When $k=2$

0.66 n by simulation

$$\mathrm{Var}(K_1(\mathbb{P}_n)) \sim \left(\frac{\pi^2}{2} + 2 - 2\pi\right)n \approx 0.651\,n.$$

- Higher moments seem to be harder!

The limit distribution



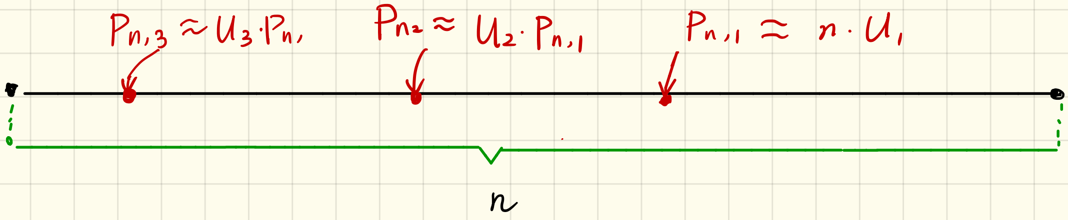
- Let $P_{n,1}, P_{n,2}, \cdots$ be the position of k-records. (where the path breaks)
- Let $R_{n,1}, R_{n,2}, \cdots$ be the time they die
- Conditioning on $P_{n,1}, P_{n,2}, \cdots, R_{n,1}, R_{n,2},$

$$\underset{\substack{\uparrow \\ \text{One-records between } P_{n,j-1},\, P_{n,j}}}{B_{n,j}} \approx \mathrm{Bin}\Big(\underbrace{P_{n,j-1} - P_{n,j}}_{\#\text{ of nodes}},\ \underbrace{\mathbb{P}(\mathrm{Exp}(1) < R_{n,j})}_{\text{Prob of being a 1-record.}}\Big).$$

Time segment breaks off.

The limit distribution

- Let $U_1, U_2, U_3 \cdots$ be iid $\mathrm{Unif}[0,1]$.



- Let $E_1, E_2, E_3, \cdots$ be iid $\mathrm{Exp}(1)$. We can also approximate

$$\underbrace{R_{n,1} \approx n^{-\frac{1}{k}} (k!\, E_1)^{\frac{1}{k}}}_{\text{when } P_{n,1} \text{ dies}}, \quad \underbrace{R_{n,2} \approx (P_{n,1})^{-\frac{1}{k}} (k!\, E_2 U_1 + k!\, E_2)^{\frac{1}{2}}}_{\text{when } P_{n,2} \text{ dies.}}, \cdots$$

$$\boxed{\frac{K(\mathbb{P}_n)}{n^{1-\frac{1}{k}}} \xrightarrow{d} \sum_{p=1}^{\infty} B_p := \mathcal{B}_k}$$

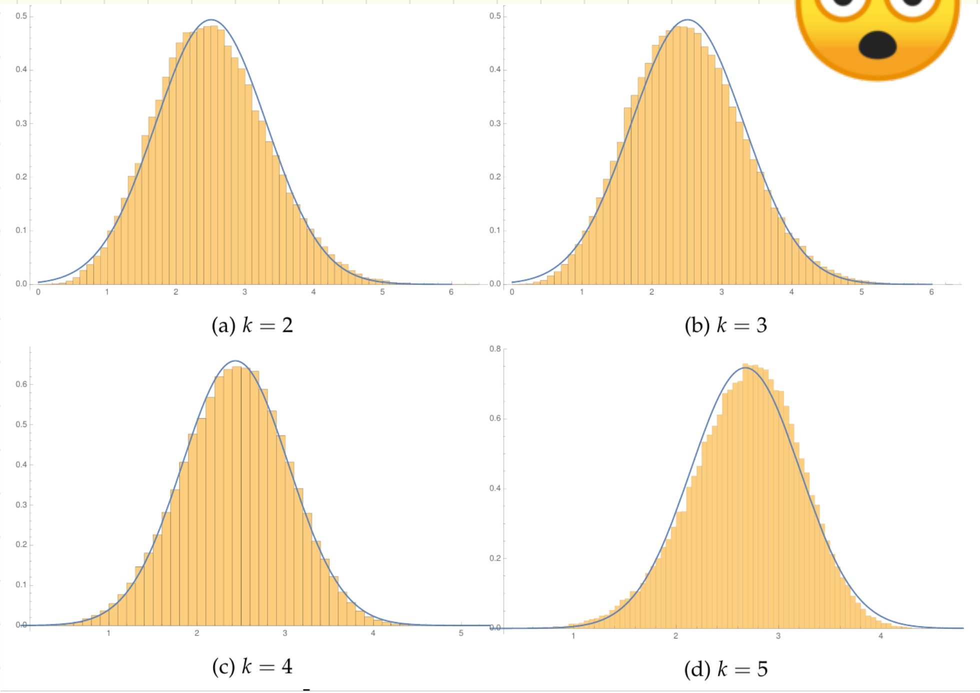
B_{nj} rescaled

All moments also converge.

complicated function of $U_1, U_2, \cdots, E_1, E_2 \cdots$

The simulation

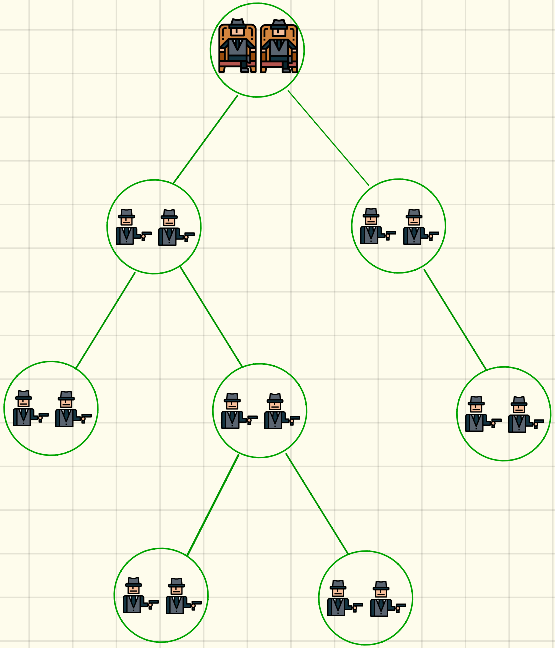
- We do not have the density function of B_k.
- But simulation suggests that it's close to a normal distribution.
- Not everything that looks normal is normal!



(a) $k = 2$ (b) $k = 3$

(c) $k = 4$ (d) $k = 5$

Complete Binary Trees, Conditional Galton-Watson Trees, and many more



The current landscape for $k \geq 2$.

- Cai, Devroye, Holmgren, Skerman 2019. EJP
- Cutting resilient networks — complete binary trees. Cai, Holmgren. 2018, arxiv
- The k-cut model in deterministic and random trees. Berzunza,  Cai, Holmgren. (would have been on arxiv if not for the boat trip).
- A note on the asymptotic expansion the Lerch's Transcendent. Cai, J.L. Lopez. 2019, Integral Transforms and Special functions.

Graph	LLL	Moments	CLT
Paths	😀	😀	😀
Complete graphs	😀	😀	😀
Complete binary trees	😀	😀	😀
Conditional GW trees	😀	😀	😀
Preferential attachment	😀	😀	
Random recursive trees	😀	😀	
Split trees	😀	😀	

The challenge

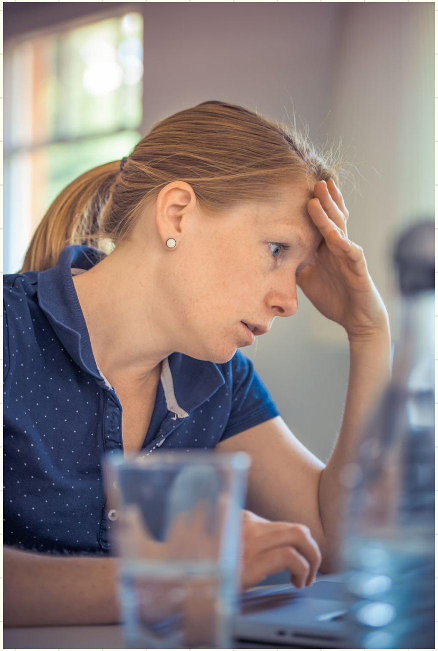
- Can even 1-cut be studied in any random graph?
- Cannot use record any more.
- Possible candidate: $G_{n,p}$ with $p = n^{-1} + n^{-4/3}$
- The giant is almost a GW tree, plus $O_p(1)$ edges. We choose root in the giant unif. rand.
- Simulation suggest, on average it takes $(k=1)$

$$\boxed{1.42 \cdot \sqrt{\text{giant size}}}$$

to cut the giant.

should be 1.25 for GW trees.

- Can we prove this?



Thanks for listening!