Protection Number in Trees

Zbigniew Gołębiewski

Wrocław University of Science and Technology
Faculty of Fundamental Problems of Technology
Department of Computer Science

Protection number

- *The protection number of a tree* is the length of the shortest path from the root to a leaf.
- *The protection number of a vertex* v in tree T is the protection number of a maximal subtree of T having v as a root.

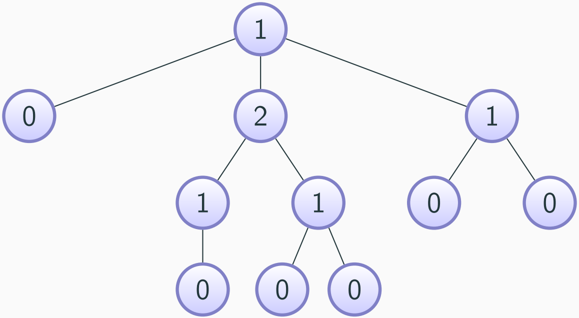


Figure 1: Tree with vertices holding their protection numbers.

Protection number of a random simply generated tree

Simply generated trees

If t_n denotes the sum of the weights of all trees with n vertices, then the generating function $T(z) = \sum_{n \geq 0} t_n z^n$ satisfies

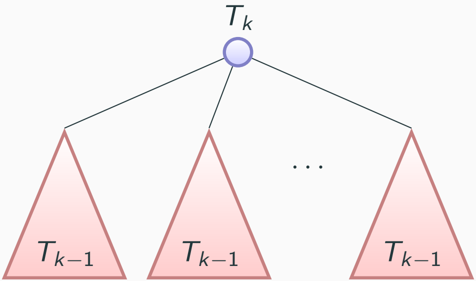
$$T(z) = z\phi(T(z)),$$

where the power series $\phi(t) = \sum_{j \geq 0} \phi_j t^j$ has only non-negative coefficients, $\phi_0 > 0$, and there is a $j \geq 2$ such that $\phi_j > 0$. Moreover, it is required that the equation $\tau\phi'(\tau) = \phi(\tau)$ has a unique positive solution.

Protection number of a random simply generated tree

- $T_k(z)$ denote the OGF of the class of simply generated trees that have protection number **at least** k:

$$\begin{cases} T_k(z) = z\left(\phi\left(T_{k-1}(z)\right) - \phi_0\right), \\ T_0(z) = T(z). \end{cases}$$



Protection number of a random simply generated tree

Lemma

All generating functions $T_k(z)$ have the same dominant singularity as $T(z)$, and it is a square root singularity.

Protection number of a random simply generated tree

Lemma

All generating functions $T_k(z)$ have the same dominant singularity as $T(z)$, and it is a square root singularity.

$$\mathbb{P}(X_n \geq k) = \frac{[z^n]T_k(z)}{[z^n]T(z)}$$

$$\mathbb{E}X_n = \sum_{k \geq 1} \mathbb{P}(X_n \geq k) \quad \text{and} \quad \mathbb{E}(X_n^2) = \sum_{k \geq 1} (2k-1)\mathbb{P}(X_n \geq k)$$

Protection number of a random simply generated tree

Theorem (Gittenberger, G., Larcher, Sulkowska, 2019)

Let X_n be the protection number of a random simply generated tree of size n. Then

$$\lim_{n\to\infty} \mathbb{E}X_n = \sum_{k\geq 1} \rho^{k-1} \prod_{i=1}^{k-1} \phi'(T_i(\rho)),$$

and

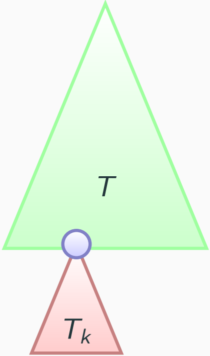
$$\lim_{n\to\infty} \mathbb{V}X_n = \sum_{k\geq 1} (2k-1)\rho^{k-1} \prod_{i=1}^{k-1} \phi'(T_i(\rho)) - \left(\lim_{n\to\infty} \mathbb{E}X_n\right)^2.$$

with ρ denoting the dominant singularity of the generating function $T(z) = z\phi(T(z))$ of the class of simply generated trees.

Protection number of a random vertex in a simply gen. tree

$S_k(z)$ denote the OGF of the number of k-protected vertices summed over all trees of a given size:

$$S_k(z) = \underbrace{z^{-1}}_{\text{remove pointed leaf}} \cdot \underbrace{T_k(z)}_{\text{attach a tree with protection num. at least k}} \cdot \underbrace{\partial_u T(z,u)|_{u=1}}_{\text{pointing at a leaf}}$$



Protection number of a random vertex in a simply gen. tree

Theorem (Gittenberger, G., Larcher, Sulkowska, 2019)

Let Y_n be the protection number of a randomly chosen vertex in a random simply generated tree of size n. Then,

$$\lim_{n\to\infty} \mathbb{E}\,(Y_n) = \frac{\phi_0}{T(\rho)} \sum_{k\geq 1} T_k(\rho),$$

and

$$\lim_{n\to\infty} \mathbb{V}\,(Y_n) = \frac{\phi_0}{T(\rho)} \sum_{k\geq 1} (2k-1)\,T_k(\rho) - \left(\lim_{n\to\infty} \mathbb{E}\,(Y_n)\right)^2.$$

Models comparison

- X_n - protection number of a random tree of size n,
- Y_n - protection number of a randomly chosen vertex in a random tree of size n.

	$\lim_{n\to\infty} \mathbb{E}(X_n)$	$\lim_{n\to\infty} \mathbb{V}(X_n)$	$\lim_{n\to\infty} \mathbb{E}(Y_n)$	$\lim_{n\to\infty} \mathbb{V}(Y_n)$
Plane trees	1.62297	0.71569	0.72764	0.81689
Motzkin trees	2.54637	1.67934	1.30760	1.73061
Cayley trees	2.28619	1.59847	1.18652	1.63220
Incomplete bin. trees	3.53647	3.76388	1.99181	3.63825
Complete binary trees	1.56298	0.37298	1.26568	0.22659

Protection number of a random Pólya tree

Pólya trees – unlabelled non-plane trees

$$\mathcal{T} \simeq \mathcal{Z} \times MSet\,(\mathcal{T})$$

$T(z)$ denote the OGF of the class of Pólya trees:

$$T(z) = ze^{T(z)} \exp\left(\sum_{i\geq 2} \frac{T(z^i)}{i}\right),$$

$$T(z) \underset{z\to\rho}{\sim} 1 - b\sqrt{1-\frac{z}{\rho}} + \frac{b^2}{3}\left(1-\frac{z}{\rho}\right) + d\left(1-\frac{z}{\rho}\right)^{3/2} + \cdots,$$

$\rho \approx 0.33832$ and $b \approx 1.55949$.

Protection number of a random Pólya tree

$T_k(z)$ denote the OGF of the class of Pólya trees that have protection number **at least** k:

$$\begin{cases} T_k(z) = z e^{T_{k-1}(z)} \exp\left(\sum_{i \geq 2} \frac{T_{k-1}(z^i)}{i} \right) - z, \\ T_0(z) = T(z). \end{cases}$$

Lemma

All the generating functions $T_k(z)$ have their (unique) dominant singularity at ρ, and the singularity is a square root singularity.

Protection number of a random Pólya tree

$$\mathbb{E}(X_n) = \sum_{k\geq 1}\prod_{i=1}^{k}\mathbb{P}(X_n \geq k | X_n \geq k-1),$$

where

$$\mathbb{P}(X_n \geq k | X_n \geq k-1) = \frac{[z^n]T_k(z)}{[z^n]T_{k-1}(z)}.$$

Lemma

$$[z^n]T_{k-1}(z) = \frac{\gamma_k \rho^{-n} n^{-\frac{3}{2}}}{\Gamma(-1/2)}\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right),$$

$$[z^n]T_k(z) = \frac{(T_k(\rho)+\rho)\gamma_k \rho^{-n} n^{-\frac{3}{2}}}{\Gamma(-1/2)}\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right),$$

as $n \to \infty$, with a constant $\gamma_k > 0$.

Protection number of a random Pólya tree

$$\mathbb{P}(X_n \geq k | X_n \geq k-1) \sim T_k(\rho) + \rho.$$

Theorem (Gittenberger, G., Larcher, Sulkowska, 2019)

Let X_n be the protection number of a random Pólya tree of size n. Then

$$\lim_{n\to\infty} \mathbb{E}\left(X_n\right) = \sum_{k\geq 1} \prod_{i=1}^{k} \left(T_k(\rho) + \rho\right) \approx 2.15488,$$

and $\lim_{n\to\infty} \mathbb{V}X_n \approx 1.36999$.

Protection number of a random vertex in a Pólya tree

- Marking a leaf and replacing it by a tree with protection number k does not work here!
- But the number of k-protected nodes in a tree is an additive tree parameter!

$$S_k(z) = \frac{T(z)\sum_{j\geq 2} S_k(z^j) + T_k(z)}{1 - T(z)},$$

and

$$\mathbb{P}(Y_n \geq k) \sim \frac{2}{b^2}\left(\sum_{i\geq 2} S_k(\rho^i) + T_k(\rho)\right).$$

Protection number of a random vertex in a Pólya tree

Theorem (Gittenberger, G., Larcher, Sulkowska, 2019)

Let Y_n be the protection number of a random vertex in a random Pólya tree of size n. Then

$$\lim_{n\to\infty} \mathbb{E}\left(Y_n\right) = \sum_{k\geq 1} \frac{2}{b^2}\left(\sum_{i\geq 2} S_k(\rho^i) + T_k(\rho)\right) \approx 0.99532$$

and $\lim_{n\to\infty} \mathbb{V}\left(Y_n\right) \approx 1.38187$.

Protection number of a random plane oriented recursive tree

Plane oriented recursive trees

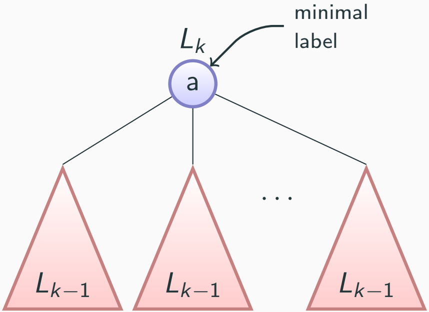
$$\mathcal{L} \simeq \mathcal{Z}^{\Box} \star Seq\,(\mathcal{L})$$

$L(z)$ denote the EGF of the plane oriented recursive trees (PORTs):

$$L(z) = \int_0^z \frac{dt}{1 - L(z)},$$
$$L(z) = 1 - \sqrt{1 - 2z}.$$

Protection number of a random PORT

$L_k(z)$ denote the EGF of the class of PORTs that have protection number **at least** k:



$$\begin{cases} L_k(z) = \int_0^z \frac{dt}{1 - L_{k-1}(z)} - z, \\ L_0(z) = 1 - \sqrt{1-2z}. \end{cases}$$

$$L_1(z) = 1 - \sqrt{1-2z} - z,$$

$$L_2(z) = \frac{1}{2}\Big(-\ln 4 + \left(\sqrt{2}+2\right)\ln\left(\sqrt{2}+1-\sqrt{1-2z}\right) - \left(\sqrt{2}-2\right)\ln\left(\sqrt{2}-1+\sqrt{1-2z}\right)\Big) - z.$$

Protection number of a random PORT

Lemma

All the generating functions $L_k(z)$ have their (unique) dominant singularity at $\frac{1}{2}$.

Lemma

For any $k \geq 1$ and $z \in \mathbb{C}$ we have $|z| \leq \frac{1}{2} \implies |L_k(z)| < 1$.

Protection number of a random PORT

Theorem (G., Klimczak, 2019)

For any $k \geq 1$ we have

$$L_k(z) = b_k(z) + c_k(1-2z)^{k-\frac{1}{2}} + O\left((1-2z)^k\right) \quad \text{as} \quad z \to \frac{1}{2}$$

where $b_k(z)$ is a polynomial of degree at most $k-1$ and c_k is some constant.

$$n![z^n]L_k(z) = n! \cdot 2^n \left(\frac{d_k}{n^{k+\frac{1}{2}}} + O\left(\frac{1}{n^{k+1}}\right)\right),$$

where d_k is a constant dependent of c_k and $d_k \leq \frac{2^{k-2}}{\sqrt{\pi}}$.

Protection number of a random PORT

$$\mathbb{P}(X_n \geq k) = 2d_k\sqrt{\pi} \cdot n^{1-k} + O\left(n^{\frac{1}{2}-k}\right) \qquad \text{as } n \to \infty$$

Theorem (G., Klimczak, 2019)

Let X_n be the protection number of a random plane oriented recursive tree of size n. Then

$$\mathbb{E}\left(X_n\right) = 1 + \frac{2}{n} + O\left(n^{-\frac{3}{2}}\right),$$
$$\mathbb{V}\left(X_n\right) = \frac{2}{n} - \frac{4}{n^2} + O\left(n^{-\frac{5}{2}}\right).$$

Thank you!

Protection number of a random non-plane recursive tree

$$\mathcal{U}_{\geq k} = \left(\mathcal{Z}^{\square} \star \mathrm{Set}_{\geq 1}(\mathcal{U}_{\geq k-1})\right)$$

The symbolic equation leads to the following equation system

$$\begin{cases} U_k(z) = \int_0^z e^{U_{k-1}(t)} dt - z \\ U_0(z) = \ln \frac{1}{1-z} \end{cases}$$

We can compute $U_1(z)$ and $U_2(z)$:

$$U_1(z) = \ln \frac{1}{1-z} - z, \qquad U_2(z) = \frac{\mathtt{Ei}(1) - \mathtt{Ei}(1-z)}{e} - z,$$

where $\mathtt{Ei}(x) = -\int_{-x}^{\infty} \frac{e^{-t}}{t} dt$ for $x \in \mathbb{R}^+$.

Protection number of a random non-plane recursive tree

There is the bijection between non-plane recursive trees of size n and permutations of $n-1$ elements. The bijection transfers some shape characteristics:

- root degree $\leftrightarrow$ number of cycles
- root subtrees sizes $\leftrightarrow$ cycle sizes

Using this information we observe that a tree $T \in \mathcal{U}_{n,2}$ corresponds to a permutation $\sigma \in S_{n-1}$ whose all cycles are of length at least 2. Such a permutation is called *derangement*.

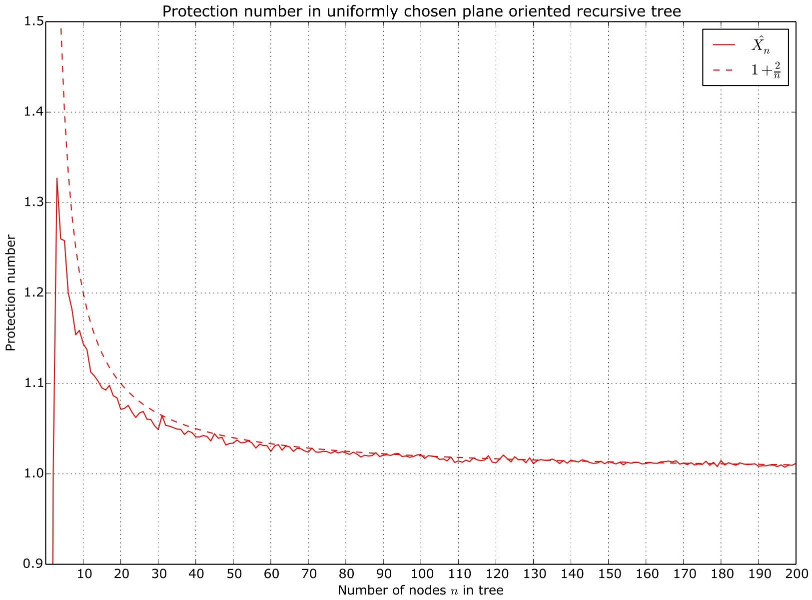
Protection number of a random non-plane recursive tree

Theorem (G., Klimczak, 2019)

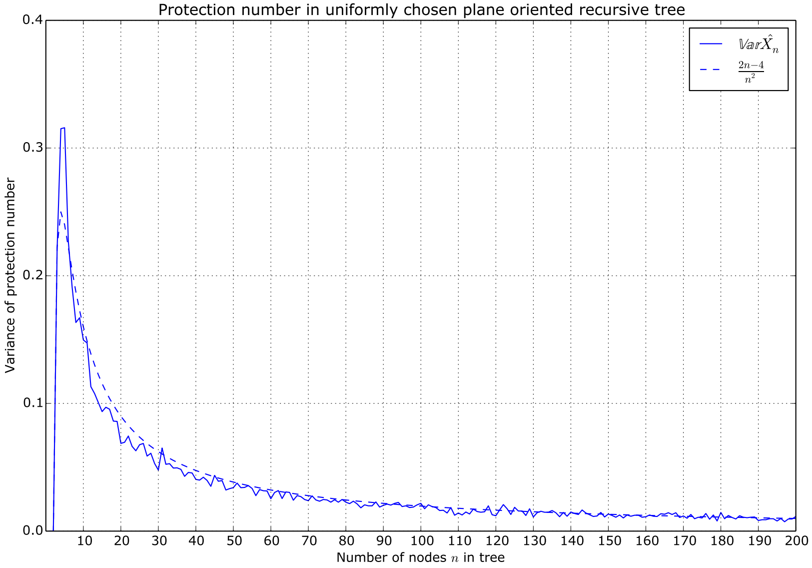
Let X_n be the protection number of a random non-plane oriented recursive tree of size n. Then

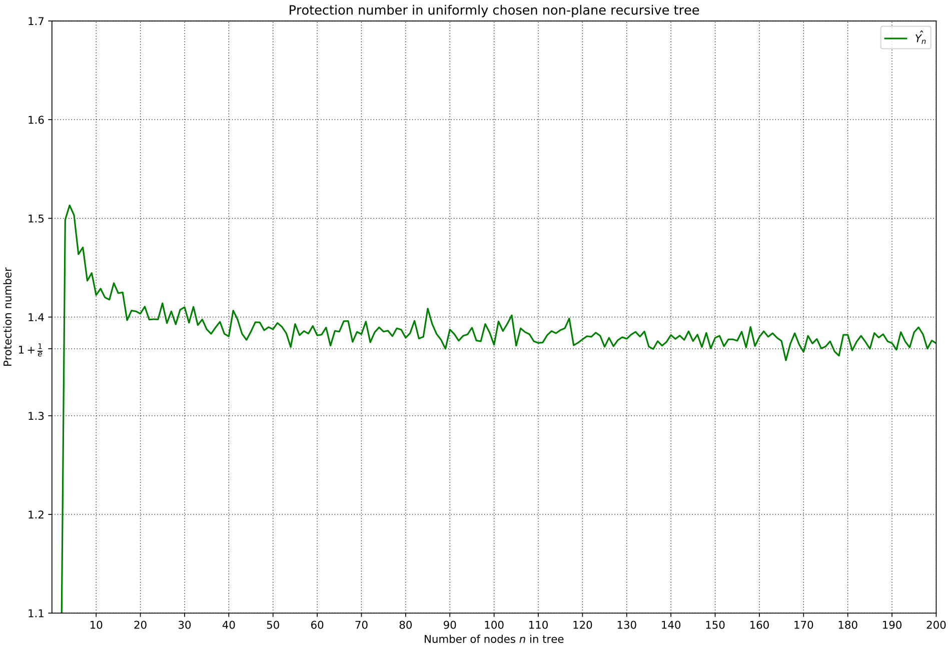
$$\mathbb{E}(X_n) = 1 + \frac{1}{e} + O\left(\frac{1}{n^{1-\frac{1}{e}}}\right),$$

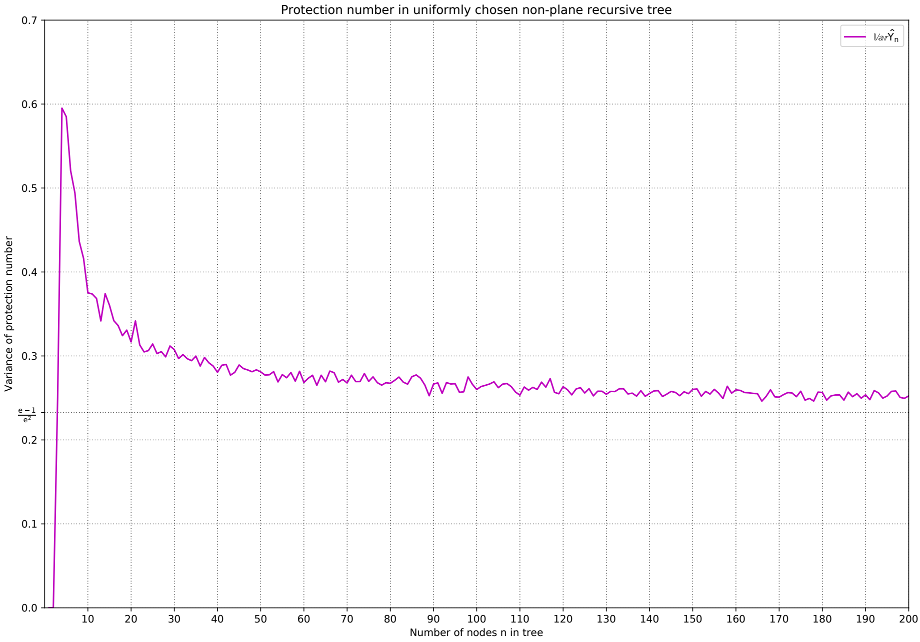
$$\mathbb{V}(X_n) = \frac{1}{e} - \frac{1}{e^2} + O\left(\frac{1}{n^{1-\frac{1}{e}}}\right).$$



Protection number in uniformly chosen plane oriented recursive tree
$\hat{X}_n$
$1+\frac{2}{n}$
Protection number
Number of nodes n in tree
1.5
1.4
1.3
1.2
1.1
1.0
0.9
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200


Protection number in uniformly chosen plane oriented recursive tree
Variance of protection number
Number of nodes n in tree
$\mathbb{V}ar\hat{X}_n$
$\frac{2n-4}{n^2}$
0.0
0.1
0.2
0.3
0.4
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200


Protection number in uniformly chosen non-plane recursive tree
Protection number
Number of nodes n in tree
$\hat{Y_n}$
1.7
1.6
1.5
1.4
$1+\frac{1}{e}$
1.3
1.2
1.1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200



Protection number in uniformly chosen non-plane recursive tree
Variance of protection number
Number of nodes n in tree
0.0
0.1
0.2
$\frac{e-1}{e^2}$
0.3
0.4
0.5
0.6
0.7
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
$\mathbb{V}ar\hat{Y}_n$

This presentation was based on the following papers:

1. *Protection Numbers in Simply Generated Trees and Pólya Trees*, Bernhard Gittenberger, Zbigniew Gołębiewski, Isabella Larcher, Małgorzata Sulkowska, submitted.
2. *Protection Number of Recursive Trees*, Zbigniew Gołębiewski, Mateusz Klimczak, ANALCO19, 2019.