UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2020:54

History and properties of random recursive trees

Victor Wikström

Examensarbete i matematik, 15 hp
Handledare: Colin Desmarais
Examinator: Veronica Crispin Quinonez
Augusti 2020

Department of Mathematics
Uppsala University

History and properties of random recursive trees

Victor Wikström

August 27, 2020

Abstract

The purpose of this thesis is to highlight the history and properties of random recursive trees. Recursive trees were first studied in 1967 by Tapia and Myers where they introduced it as *node-weighted trees*. Recursive trees have since been used to study pyramid schemes, amongst others. A random recursive tree is a recursive tree chosen uniformly at random amongst all recursive trees of order n. Using Pólya urns we can study the degree distribution and with the help of harmonic numbers we can find the expected insertion depth. These kind of trees also exhibit the small-world phenomenon.

Contents

1 Introduction

1.1 Introduction to recursive trees

In this thesis, we discuss the history and properties of random recursive trees. We begin by defining a graph. A *graph* is a collection of vertices (also known as nodes) connected by edges. A *tree* is a connected graph with no cycles, and a rooted tree is a tree with one vertex labelled the root. A *recursive* tree is a tree with vertices labelled from 1 to n ,where 1 is the root of the tree and $n \in \mathbb{N}$ is the number of vertices in the tree. The labels on all paths away from the root are strictly increasing. The number of vertices in a recursive tree is its order and a recursive tree of order n is denoted by t_n. A *random* recursive tree is a recursive tree which is chosen uniformly at random amongst all recursive trees of order n (see section 2.3).

The neighbour of a vertex on the path towards the root is called its *parent* and all the other neighbouring vertices are called *children*. The vertices with no children are called *leaves* and the rest of the vertices are called *internal vertices*. The number of children a vertex v has is the *outdegree* of said vertex and is denoted by $outdeg(v)$. The *degree* of a vertex is the number of neighbours of v, so a leaf would have degree 1.

Recursive trees were first studied in 1967 by Tapia and Myers [17], where they called it concave node-weighted trees. In their paper, the root is labelled n and the vertices are labelled in decreasing order.

Recursive trees have been used, amongst others, when studying pyramid & chain letter schemes [9], which was initially developed for attorneys faced with prosecuting promoters of such activity. It was also used to study the spread of contamination from a single source within some organism [15].

1.2 Network science

Network science, which dawned in the beginning of the 21st century, is the study of networks, like the power grid, trade networks, social networks, neural networks etc. Random graphs are used to model real world networks. The properties of random graphs are studied and compared to those of real world networks. See [2] for an overview. Two properties often studied are *distances* and *degree distributions*.

Definition 1.1. ***Distance.*** *The distance between two vertices is the number of edges in the shortest path between them.*

Definition 1.2. ***Degree distribution.*** *The degree distribution P(K) is the probability distribution defined as the proportion of vertices with degree K.*

1.2.1 Small-world phenomenom

Abbas Mehrabian [14] wrote about using random recursive trees to prove the small-world phenomenon in several other random graph models. The

small-world phenomenon, also known as "Six degrees of separation", goes back almost a hundred years and was first discussed by Frigyes Karinthy [12] in a short story called *Chain-links*. He started it as a game where he and his friends tried to connect any two persons through five or less individuals. The premise behind this "game" is that the amount of people you can contact grows exponentially with every individual you go through. Karinthy played this game plenty and although he couldn't prove the phenomenon, he also could not find any two persons who couldn't be connected through no more than five links. Travers and Milgram [18] did a study on this where they asked 296 arbitrary individuals in Nebraska and Boston to generate acquaintance chains to target a specific person in Massachusetts. From the study, 64 chains succeeded in reaching the target person and the mean number of links were 5.2.

In recent studies, small-world networks are usually defined to be networks where the distance is on the order of $\ln n$ as the number of vertices n in the network goes to infinity (see [2]). A more precise definition is given in [19], definition 1.7.
Let $(G_n)_{n=1}^{\infty}$ be a sequence of graphs, and let D_n be the typical distance between any two vertices in G_n, i.e.,

$$D_n = \text{dist}(u, v)$$

where u and v are two vertices chosen uniformly at random from G_n, and $\text{dist}(u, v)$ is the distance between these two vertices. Then we have the following definition:

Definition 1.3. *The graph sequence $(G_n)_{n=1}^{\infty}$ is* small world *if there exists a finite number K such that*

$$\lim_{n\to\infty} \mathbb{P}(D_n \leq K \ln n) = 1.$$

1.3 Outline

First we study the combinatorial properties of recursive trees, including the number of recursive trees and its relation to permutations. We define random recursive trees. Next, we look at degree distribution and finally we study distances, including depth and height in random recursive trees.

2 Combinatorial properties of recursive trees

2.1 Number of recursive trees

Proposition 2.1. *The number of recursive tress of order n is given by $(n-1)!$*

Proof. Let r_n be the number of recursive trees of order n.
Base case: $r_1 = 0! = 1$

Induction step: For every recursive tree of order $k-1$, there are $k-1$ options to be the parent of the k-th vertex, so

$$r_k = (k-1)r_{k-1} = (k-1)(k-2)! = (k-1)!$$

□

2.2 Relation to permutations

There is a bijection between recursive trees and permutations, see for example [7]. Permutations can be written in many ways but in this thesis we will only consider the cycle decomposition. Consider the following permutation π of the numbers 1 through 8.

$$\pi = \begin{pmatrix} 1\,2\,3\,4\,5\,6\,7\,8 \\ 7\,3\,2\,6\,1\,4\,8\,5 \end{pmatrix}$$

which has the cycle decomposition

$$\pi = (1,7,8,5)(2,3)(4,6)$$

These permutations can easily be modified by either adding or removing objects in the permutation. If we remove n we get a new permutation $\tilde{\pi}$. If n belonged to a cycle of length 1 we just remove that cycle. If n is included in a cycle of length greater than 1 then there exist i and j such that $\pi(i) = n$ and $\pi(n) = j$. By deleting n we set that $\tilde{\pi}(i) = j$. For the example above we get

$$\tilde{\pi} = \begin{pmatrix} 1\,2\,3\,4\,5\,6\,7 \\ 7\,3\,2\,6\,1\,4\,5 \end{pmatrix}$$

and the cycle decomposition becomes

$$\tilde{\pi} = (1,7,5)(2,3)(4,6)$$

If we instead add n to $\tilde{\pi}$ we can either include n by creating a new cycle containing only n, or we can include it in an already existing cycle. By adding n to a cycle of size k we have k possible ways to add it which means that there are a total of $n-1$ possible ways to include object n into the already existing cycles in $\tilde{\pi}$.

Recursive trees grow in a similar manner as permutations in cycle notation. To show that we first need to change the labels on the trees by starting with 0 instead of 1. We begin by adding a vertex labeled 1 to the root and at the same time creating the permutation (1). The next steps are to continue adding vertices, either to the root or to another vertex. If the k-th vertex is added to the root then we create a new cycle (k). If it is added to a non-root vertex i then we consider the cycle that i belongs to and place k in between i and the vertex that i was mapped to. This process continues until all vertices have been added (see figure 1 for a few examples).

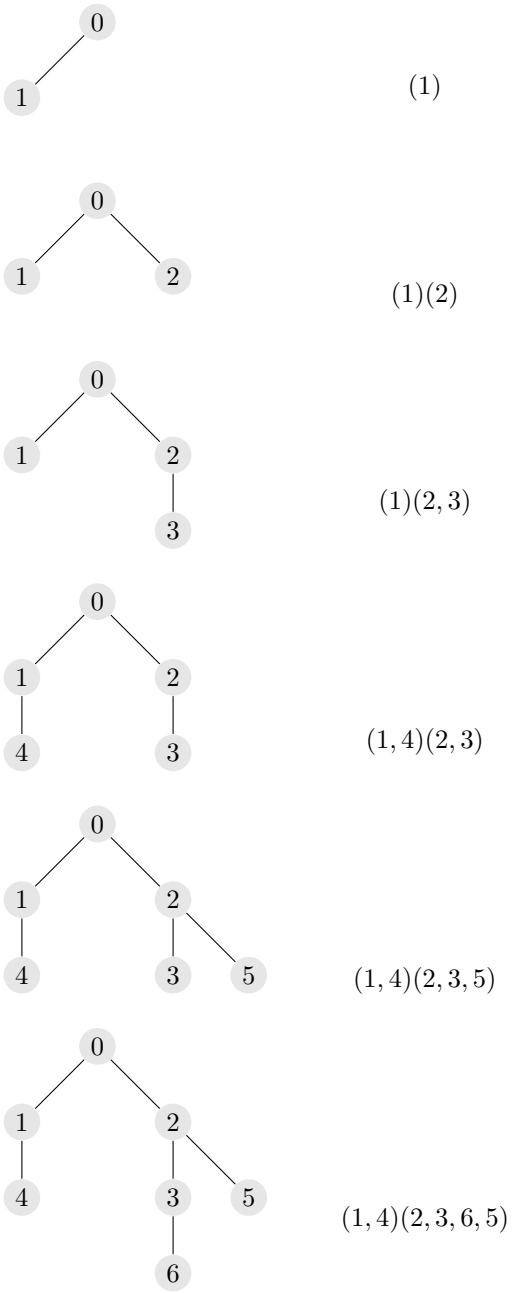

Figure 1: Recursive trees with root 0 and their corresponding cycle decompositions

2.3 Random recursive trees

A random recursive tree T_n is a recursive tree chosen uniformly at random amongst all recursive trees of order n, so for a particular tree t_n, $\mathbb{P}(T_n = t_n) = 1/(n-1)!$. We can also obtain T_n by starting with the root and adding vertices 2,3, ..., n one at a time such that every vertex k chooses a parent uniformly at random among the $k-1$ vertices in the tree. In this way we see that we get the same distribution,

$$\mathbb{P}(T_n = t_n) = \frac{1}{1} \cdot \frac{1}{2} \cdot \frac{1}{3} \cdot \ldots \cdot \frac{1}{n-1} = \frac{1}{(n-1)!}$$

3 Degree distributions

3.1 Number of leaves in a recursive tree

We start by looking at leaves. The number of trees of order n with k leaves can be expressed by Eulerian numbers, defined by the recursion

$$\left\langle {n \atop 0} \right\rangle = \left\langle {n \atop n-1} \right\rangle = 1,$$

$$\left\langle {n \atop k} \right\rangle = (n-k)\left\langle {n-1 \atop k-1} \right\rangle + (k+1)\left\langle {n-1 \atop k} \right\rangle \tag{1}$$

Proposition 3.1. *The number of trees of size n with k leaves is given by*

$$p(n,k) = \left\langle {n-1 \atop k-1} \right\rangle \tag{2}$$

Proof. Let $p(n,k)$ be the number of trees with n vertices and k leaves. We can see immediately that

$$p(1,1) = 1, \quad p(2,1) = 1, \quad p(3,1) = 1, \quad p(3,2) = 1.$$

We can also see that there is only one tree with n vertices and 1 leaf (the path with n vertices) and there is also only one tree with n vertices and $n-1$ leaves (the star where vertices $2, 3, \ldots, n$ are children of the root). In other words,

$$p(n,1) = 1 \quad \textit{and} \quad p(n,n-1) = 1 \tag{3}$$

For all other k, with $1 < k < n-1$, $n \geq 4$ we have that trees with n vertices and k leaves can be built from trees with $n-1$ vertices. There are two ways we can get k leaves when we add a vertex to a tree t_{n-1}. The first scenario is when the new vertex is added to an internal vertex and a new leaf is created. For this to happen there has to be $k-1$ leaves in t_{n-1} and there are $n-k$ internal vertices in t_{n-1}.
The second way is when the new vertex gets attached to an already existing leaf, and the number of leaves is unchanged. There are k leaves in t_{n-1}. We therefore get that the total number of trees with n vertices and k leaves is given by:

$$p(n,k) = (n-k) \cdot p(n-1,k-1) + k \cdot p(n-1,k) \tag{4}$$

This recursion is very similar to the Eulerian numbers (*see eq. 1*). Since

$$p(n, n-1) = \left\langle {n-1 \atop n-2} \right\rangle = 1$$
$$p(n, 1) = \left\langle {n-1 \atop 0} \right\rangle = 1$$

and $p(n,k)$ follows the same recursion as $\left\langle {n-1 \atop k-1} \right\rangle$, we see that the numbers are the same. □

The probability that the tree L_n has k leaves is then

$$P(L_n = k) = \frac{p(n,k)}{(n-1)!} = \frac{1}{(n-1)!} \left\langle {n-1 \atop k-1} \right\rangle$$

since there are $(n-1)!$ trees of order n.

Since the Eulerian numbers are symmetric, i.e,

$$\left\langle {n \atop k} \right\rangle = \left\langle {n \atop n-k-1} \right\rangle$$

we get that $p(n,k) = p(n, n-k)$ from which we conclude that

$$\mathbb{E}[L_n] = \frac{n}{2}.$$

We now know the number of vertices of degree 1 (leaves). We also want to study the distribution of vertices with degrees > 1. To do that we first look at Pólya urns.

3.2 Pólya urns

Pólya urns were first discussed by George Pólya and Florian Eggenberger back in early 1900s [8]. They considered an urn containing 1 white ball and 1 black ball. When a ball is drawn, replace the ball with an additional ball of the same colour. This is a model for contagious diseases.

Now consider a random recursive tree as introduced in section 1.1. We introduce a new type of Pólya urn. Let's say that the white balls represent the leaves and the black balls represent internal vertices. If a white ball is drawn from the urn, it is placed back with one additional black ball, i.e a leaf is chosen and that leaf turns into a vertex of outdegree one and a new leaf gets attached to that vertex. If a black ball gets drawn, it is placed back with one additional white ball since when an internal vertex is chosen, it remains an internal vertex while a new leaf is added.

In figure 2, at step 6 we had three black balls and three white balls in our

Figure 2: Example of a black ball drawn in simple Pólya urn represented by a recursive tree

urn. The ball was randomly chosen at step 7 and we end up with an urn containing four white balls and three black balls. If we want to look at the distributions of vertices of higher degrees than leaves we have to add more colours to represent the different outdegrees of the vertices in the tree.

We consider a class of generalized Pólya urns (also known as generalized Pólya-Eggenberger urn or generalized Friedman urn) as follows:

- There are m types (or colours) of balls, labelled $1, 2, \ldots, m$.
- For each type $j = 1, \ldots, m$, we associate a (column) vector $\boldsymbol{A}_j = (A_{1,j}, A_{2,j}, \ldots, A_{m,j})^T$ such that $A_{i,j}$ is a non-negative integer for $i \neq j$, and $A_{j,j}$ is an integer greater than or equal to -1.
- Let $\boldsymbol{X}_n = (X_{n,1}, \ldots, X_{n,m})$, where $X_{n,j}$ is the number of balls of type j in the urn at time n.
- The urn starts with a given vector $\boldsymbol{X}_0$.
- At each step n, a ball is chosen uniformly at random amongst all the balls in the urn. So the probability of choosing a ball of type j is

$$\frac{X_{n,j}}{\sum_{i=1}^{m} X_{n,i}}.$$

- If a ball of type j is chosen at step n, then it is replaced with $A_{i,j}$ balls of type i for each i. In other words, $\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \boldsymbol{A}_j^T$.
- The *replacement matrix* of the urn is defined to be the matrix

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,m} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,m} \end{bmatrix}$$

Svante Janson [11] showed that under certain conditions,

$$\frac{\boldsymbol{X}_n - n\boldsymbol{\mu}}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \Sigma)$$

for some vector $\boldsymbol{\mu}$ and some matrix Σ. The notation $\mathcal{N}(0, \Sigma)$ describes a multivariate normal distribution. We can explicitly calculate $\boldsymbol{\mu}$ and Σ, although Σ is quite complicated to calculate. As for $\boldsymbol{\mu}$, this is an eigenvector of A, and we will see later what eigenvector explicitly.

For the convergence stated above to hold, one of the conditions needed is that the urn has to be *irreducible*.

Definition 3.2. *An urn process is* irreducible *if starting with any ball of type j, it is possible that the urn eventually has a ball of type i.*

Suppose that A has a simple (algebraic multiplicity 1) real eigenvalue λ_1 such that $\lambda_1 > Re(\lambda)$ for all other eigenvalues λ, where $Re(\lambda)$ is the real part of a complex number.

Finally, let $\boldsymbol{v} = (v_1, \ldots, v_m)$ be the eigenvector of A associated with the eigenvalue λ_1 such that $v_1 + \cdots + v_m = 1$. Since λ_1 is simple, $\boldsymbol{v}$ is unique since the eigenspace associated with λ_1 has dimension 1.

We then have the two following theorems; a law of large numbers, and a central limit theorem:

Theorem 3.3 ([1, Section V.9.3], [11, Theorem 3.21])**.** *Suppose an urn process $(\boldsymbol{X}_n)_{n=1}^{\infty} = (X_0, X_1, X_2, \ldots)$ as described above is irreducible. Then*

$$\frac{\boldsymbol{X}_n}{n} \xrightarrow{a.s.} \lambda_1 \boldsymbol{v}_1$$

as $n \to \infty$.

Theorem 3.4 ([11, Theorem 3.22], see also [10, Theorem 3.1])**.** *Suppose an urn process $(\boldsymbol{X}_n)_{n=1}^{\infty} = (X_0, X_1, X_2, \ldots)$ as described above is irreducible with replacement matrix A, and suppose further that $\lambda_1 > 2Re(\lambda)$ for all other eigenvalues λ of A. Then*

$$\frac{\boldsymbol{X}_n - n\lambda_1 \boldsymbol{v}}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(\boldsymbol{0}, \Sigma)$$

as $n \to \infty$.

3.3 Degree distributions of vertices with degree >1

We can describe the vertices in a random recursive tree as balls in an urn. Let d be a positive integer: this will correspond to the largest degree we wish to look at. Suppose we grow a sequence of random recursive trees $T_1, T_2, T_3, \ldots, T_n, \ldots$, and for each $i = 1, \ldots, d$ we let $N_{n,i}$ be the number of vertices of degree i in the tree T_n. Consider an urn process $(\boldsymbol{X}_n)_{n=1}^{\infty}$ with $d+1$ types of balls, where for each $\boldsymbol{X}_n = (X_{n,1}, \ldots, X_{n,d}, X_{n,d+1})$, the number $X_{n,i}$ is the number of vertices of degree i in T_n for $i = 1, \ldots, d$ (so $X_{n,i} = N_{n,i}$), and $X_{n,d+1}$ is the number of vertices whose degree is greater than d.

We see that the urn process has the following $(d+1)\times(d+1)$ replacement matrix:

$$A = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 & 1 & 1 \\ 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

Consider when a vertex of degree k gets chosen as the parent of a new vertex. That vertex becomes a vertex with degree $k+1$ when the leaf gets attached. In the urn scenario, this means that the ball representing a vertex of degree k gets removed and a ball representing a vertex of degree $k+1$ is placed inside the urn along with a ball representing a leaf. When an internal vertex is chosen, a new ball representing a leaf is added, while the numbers of balls representing leaves is unchanged when a leaf is chosen. The last column in A only contains a 1 at $A_{1,d+1}$ and zeros everywhere else since we classify all vertices with a degree higher than d as the same colour. So that means when a vertex of degree $k \geq d$ is chosen, the ball representing it gets placed back in the urn along with an additional ball representing a leaf.

This matrix is irreducible because as the urn grows, the number of different degrees gets higher and higher and therefore we will probably eventually get a ball of the desired degree (see definition 2.1).

Now to calculate the eigenvalues of A. The following argument is a simplified version of the proof of Lemma 3.2 in [4]. The characteristic polynomial of the matrix A is given by

$$det(A-\lambda I_{n\times n}) = \begin{vmatrix} -\lambda & 1 & 1 & \cdots & 1 & 1 & 1 \\ 1 & -1-\lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1-\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1-\lambda & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -1-\lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -\lambda \end{vmatrix}$$

The determinant of a matrix is unchanged by adding a multiple of a row to another row or by subtracting a multiple of a column to another column. Therefore, consider subtracting column $d+1$ of $A-\lambda I$ from all

other columns. Then

$$det(A-\lambda I_{n\times n}) = \begin{vmatrix} -\lambda-1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & -1-\lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1-\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1-\lambda & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -1-\lambda & 0 \\ \lambda & \lambda & \lambda & \cdots & \lambda & 1+\lambda & -\lambda \end{vmatrix}$$

In the matrix above, add row i to row $d+1$ for all $i = 1, \ldots, d$. Then

$$det(A-\lambda I_{n\times n}) = \begin{vmatrix} -\lambda-1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & -1-\lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1-\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1-\lambda & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -1-\lambda & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & -\lambda+1 \end{vmatrix}$$

From here, by expanding along the bottom row, we see that

$$det(A-\lambda I_{n\times n}) = (1-\lambda) \begin{vmatrix} -\lambda-1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -1-\lambda & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1-\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1-\lambda & 0 \\ 0 & 0 & 0 & \cdots & 1 & -1-\lambda \end{vmatrix}$$

What remains is a lower triangular matrix, the determinant of which is the product of the diagonal entries. Therefore,

$$det(A - \lambda I_{n\times n}) = (1-\lambda)(-\lambda-1)^d$$

The roots of the characteristic polynomial are the eigenvalues of A. So we see that the eigenvalues of A are 1 with multiplicity 1 and -1 with multiplicity d.
Next, the eigenvector $\boldsymbol{v} = (v_1, \ldots, v_{d+1})$ associated with λ_1 such that $v_1 + \cdots + v_{d+1} = 1$ is given by

$$\boldsymbol{v} = \left(\frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2^i}, \ldots, \frac{1}{2^d}, \frac{1}{2^d}\right).$$

To see this, look at

$$A\boldsymbol{v} = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 & 1 & 1 \\ 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{8} \\ \vdots \\ \frac{1}{2^{d-1}} \\ \frac{1}{2^d} \\ \frac{1}{2^d} \end{bmatrix}$$

The first row gives

$$\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots+\frac{1}{2^d}+\frac{1}{2^d}$$

And since

$$\sum_{i=1}^{n}\frac{1}{2^i}=1-\frac{1}{2^n},$$

we get that

$$\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots+\frac{1}{2^d}=1-\frac{1}{2^d}-\frac{1}{2}.$$

So the first row gives

$$1-\frac{1}{2^d}-\frac{1}{2}+\frac{1}{2^d}=\frac{1}{2}.$$

For rows $2 \leq i \leq d$ we get

$$\frac{1}{2^{i-1}}-\frac{1}{2^i}=\frac{1}{2^i},$$

since $\frac{1}{2^i}$ is always half of $\frac{1}{2^{i-1}}$. The last row, $d+1$, only gives $\frac{1}{2^d}$. The new vector then becomes

$$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{8} \\ \vdots \\ \frac{1}{2^{d-1}} \\ \frac{1}{2^d} \\ \frac{1}{2^d} \end{bmatrix}$$

We can therefore use Theorem 3.3 to see that

$$\frac{\boldsymbol{X}_n}{n} \xrightarrow{a.s.} \left(\frac{1}{2},\frac{1}{4},\ldots,\frac{1}{2^i},\ldots,\frac{1}{2^d},\frac{1}{2^d}\right),$$

and use Theorem 3.4 to see that

$$\frac{\boldsymbol{X}_n-n\boldsymbol{v}}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(\boldsymbol{0},\Sigma),$$

both results were proved by Janson [10]. In particular, by letting $d \to \infty$, we get a law of large numbers and a central limit theorem for the number of vertices of degree i in random recursive trees;

$$\frac{N_{n,i}}{n} \xrightarrow{a.s.} \frac{1}{2^i},$$

and

$$\frac{N_{n,i}-\frac{n}{2^i}}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,\sigma_i)$$

for some variance σ_i [10].

4 Distances

For a vertex v, let $d(v)$ be the *depth* or distance from the root to vertex v. First, a few random variables need to be defined:

- I_n : the *insertion depth of vertex n*. This is the distance from the n-th added vertex v_n to the root of the tree (so $d(v)$). Since the first vertex is the root, $I_1 = 0$, and since the second vertex is a child of the root, then $I_2 = 1$.
- TP_n : the *total path length* of a random recursive tree T_n with n vertices. This value is the sum of the depths over all vertices v in T_n.
- H_n : the *height* of a random recursive tree with n vertices. This is the maximum value over all depths of the vertices in the tree T_n, i.e the longest distance away from the root.

Let's begin with I_n, the insertion depth. Let $h_n = \sum_{i=1}^{n} 1/i$ be the n-th harmonic number. In the following proof, we'll use the following fact:

$$\sum_{j=1}^{n} h_j = (n+1)h_n - n. \tag{5}$$

We can prove this by induction. The base case for $n = 1$ is

$$\sum_{j=1}^{1} h_j = h_1 = 1 = (1+1)h_1 - 1 = 2(1) - 1 = 1$$

so (5) holds for the base case. For the inductive step, suppose

$$\sum_{j=1}^{k} h_j = (k+1)h_k - k$$

for some positive integer k. Then

$$\begin{aligned}
\sum_{j=1}^{k+1} h_j &= \sum_{j=1}^{k} h_j + h_{k+1} \\
&= (k+1)h_k - k + h_{k+1} \\
&= (k+1)\left(h_{k+1} - \frac{1}{k+1}\right) - k + h_{k+1} \\
&= (k+1)h_{k+1} - 1 - k + h_{k+1} \\
&= (k+2)h_{k+1} - (k+1)
\end{aligned}$$

which concludes the inductive step.

Theorem 4.1 (See [5])**.** *Let I_n be the insertion depth of the n-th vertex. Then $\mathbb{E}[I_1] = 0$ and for all $n \geq 2$, $\mathbb{E}[I_n] = h_{n-1}$.*

Proof. We prove this by using *strong induction*. For the base cases, $\mathbb{E}[I_1] = 0$ and $\mathbb{E}[I_2] = 1$ since $I_1 = 0$ and $I_2 = 1$ (see discussion above). Now suppose that for $k \geq 2$, we have that $\mathbb{E}[I_j] = h_{j-1}$ for all $2 \leq j \leq k$. When we add the $k+1$-th vertex v_{k+1}, we first choose a vertex v_i uniformly at random (with probability $1/k$). If v_i is chosen, then the depth of v_{k+1} will be $I_i + 1$. Therefore, given the tree T_k,

$$\mathbb{E}[I_{k+1}|T_k] = \sum_{j=1}^{k} \frac{1}{k}(I_j + 1).$$

By the law of total expectation ($\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$),

$$\mathbb{E}[I_{k+1}] = \mathbb{E}[\mathbb{E}[I_{k+1}|T_k]] = \mathbb{E}\left[\sum_{j=0}^{k} \frac{1}{k}(I_j + 1)\right].$$

Then by linearity of expectation, the induction hypothesis, and eq 5,

$$\begin{aligned}
\mathbb{E}[I_{k+1}] &= \mathbb{E}\left[\sum_{j=1}^{k} \frac{1}{k}(I_j + 1)\right] \\
&= \sum_{j=1}^{k} \frac{1}{k}\mathbb{E}[(1 + I_j)] \\
&= \frac{1}{k}\left((1 + \mathbb{E}[I_1]) + \sum_{j=2}^{k}(1 + \mathbb{E}[I_j])\right) \\
&= \frac{1}{k}\left((1 + 0) + \sum_{j=2}^{k}(1 + h_{j-1})\right) \\
&= \frac{1}{k}\left(k + \sum_{j=1}^{k-1} h_j\right) \\
&= \frac{1}{k}(k + kh_{k-1} - (k-1)) \\
&= h_{k-1} + \frac{1}{k} \\
&= h_k.
\end{aligned}$$

This completes the induction hypothesis, and therefore the proof of the theorem □

One fact we know about the harmonic numbers h_n is that they closely resemble the area underneath the curve $1/x$ between 1 and n, which is $\ln n$. As a consequence, we have the following convergence:

$$\frac{\mathbb{E}[I_n]}{\ln n} \to 1$$

as $n \to \infty$.

For the insertion depth, we have from Devroye [5] and Mahmoud [13] that

$$\frac{I_n - \ln n}{\sqrt{\ln n}} \xrightarrow{d} \mathcal{N}(0,1).$$

For total path length, by using linearity of expectation and eq. 5, we have that

$$\mathbb{E}[TP_n] = \sum_{i=1}^{n} \mathbb{E}[I_i] = \sum_{i=1}^{n-1} h_i = nh_{n-1} - (n-1) = nh_n - \frac{n}{n} - n + 1 = nh_n - n.$$

Therefore,

$$\frac{\mathbb{E}[TP_n]}{n \ln n} \to 1.$$

Mahmoud [13] showed the following:

$$\frac{TP_n - n \ln n}{n} \xrightarrow{d} W,$$

for some non-degenerate random variable W. Dobrow and Fill [6] showed that W does not have a normal distribution.

As for the height, Pittel [16] showed that

$$\frac{H_n}{\ln n} \xrightarrow{a.s.} e.$$

This means that the *height* is not much further from the expected *depth* of any vertex as $n \to \infty$.

4.1 Small-world phenomenon in random recursive trees

Recall the definition of small world graph sequence, *definition 1.3*. Let's consider our graph sequence $(T_n)_{n=1}^{\infty}$ as a sequence of random recursive trees where T_n is grown from T_{n-1} as usual. The distance between two vertices u, v is at most the sum of the distances from u to the root and v to the root, i.e., the sum of the depth of u and the depth of v. The depth of these vertices is at most the height H_n of the tree, so $\text{dist}(u,v) \leq 2H_n$. Therefore, using the convergence in probability for height given above,

$$\begin{aligned}
\mathbb{P}(D_n \leq K \ln n) &\geq \mathbb{P}(2H_n \leq K \ln n) \\
&= \mathbb{P}(2H_n / \ln n \leq K) \\
&= \mathbb{P}(2H_n / \ln n - 2e \leq K - 2e) \\
&\geq \mathbb{P}(|2H_n / \ln n - 2e| \leq K - 2e)
\end{aligned}$$

The last line tends to 1 as $n \to \infty$ for any $K > 2e$ since $H_n / \ln n \xrightarrow{a.s.} e$ (and so $H_n / \ln n \xrightarrow{p} e$ as well). Therefore, from definition 1.3, random recursive trees exhibit the small-world phenomenon. This property of random recursive trees (and generalizations for random recursive trees) is exploited to prove that the small-world phenomenon holds in other types of random networks (see for example [3] and [14]).

References

[1] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Vol. 196. Springer-Vlg, 1972.

[2] Albert-László Barabási et al. *Network science*. Cambridge university press, 2016.

[3] Béla Bollobás and Oliver Riordan. "The diameter of a scale-free random graph". In: *Combinatorica* 24.1 (2004), pp. 5–34.

[4] Colin Desmarais and Cecilia Holmgren. *Normal limit laws for vertex degrees in randomly grown hooking networks and bipolar networks*. 2019. arXiv: `1910.13881 [math.PR]`.

[5] Luc Devroye. "Applications of the theory of records in the study of random trees". In: *Acta Informatica* 26.1-2 (1988), p. 123.

[6] Robert P Dobrow and James Allen Fill. "Total path length for random recursive trees". In: *Combinatorics, Probability and Computing* 8.4 (1999), pp. 317–333.

[7] Michael Drmota. *Random trees: an interplay between combinatorics and probability*. Springer Science & Business Media, 2009.

[8] Florian Eggenberger and George Pólya. "Über die statistik verketteter vorgänge". In: *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 3.4 (1923), pp. 279–289.

[9] Joseph L Gastwirth and PK Bhattacharya. "Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable". In: *Operations Research* 32.3 (1984), pp. 527–536.

[10] Svante Janson. "Asymptotic degree distribution in random recursive trees". In: *Random Structures & Algorithms* 26.1-2 (2005), pp. 69–83.

[11] Svante Janson. "Functional limit theorems for multitype branching processes and generalized Pólya urns". In: *Stochastic Processes and their Applications* 110.2 (2004), pp. 177–245.

[12] Frigyes Karinthy. "Chain-links". In: *Everything is different* (1929).

[13] Hosam M Mahmoud. "Limiting distributions for path lengths in recursive trees". In: *Probability in the Engineering and Informational Sciences* 5.1 (1991), pp. 53–59.

[14] Abbas Mehrabian. "Justifying the small-world phenomenon via random recursive trees". In: *Random Structures & Algorithms* 50.2 (2017), pp. 201–224.

[15] A Meir and JW Moon. "Cutting down recursive trees". In: *Mathematical Biosciences* 21.3-4 (1974), pp. 173–181.

[16] Boris Pittel. "Note on the heights of random recursive trees and random m-ary search trees". In: *Random Structures & Algorithms* 5.2 (1994), pp. 337–347.

[17] M Tapia and B Myers. "Generation of concave node-weighted trees". In: *IEEE Transactions on Circuit Theory* 14.2 (1967), pp. 229–230.

[18] Jeffrey Travers and Stanley Milgram. "An experimental study of the small world problem". In: *Social Networks*. Elsevier, 1977, pp. 179–197.

[19] Remco Van Der Hofstad. *Random graphs and complex networks*. Vol. 1. Cambridge university press, 2016.