BRANCHES IN BUCKET RECURSIVE TREES WITH VARIABLE CAPACITIES OF BUCKETS

Ramin KAZEMI

Bucket recursive trees with variable capacities of buckets (BRT-VCB) introduced by Kazemi (2012). In this paper, we study the random variable which counts the number of branches of size $a$ attached to the bucket containing label $j$ in a BRT-VCB of size $n$ (the number of subtrees of size $a$ rooted at the children of bucket containing label $j$).

Keywords: BRT-VCB, branches, limiting distribution, joint distribution.

MSC2010: 05C05, 60F05

1. Introduction

Trees are defined as connected graphs without cycles, and their properties are basics of graph theory. A tree on $n$ nodes labeled 1, 2, ..., $n$ is a recursive tree if the node labeled 1 is distinguished as the root, and for each $2 \leq k \leq n$, the labels of the nodes in the unique path from the root to the node labeled $k$ form an increasing sequence [6]. Bucket trees are a generalization of the ordinary trees where buckets (or nodes) can hold up to $b \geq 1$ labels. Mahmoud and Smythe [5] introduced bucket recursive trees as a generalization of ordinary recursive trees. In this model the capacity of buckets is fixed. They applied a probabilistic analysis for studying the height and depth of the largest label in these trees and Kuba and Panholzer [4] analyzed these trees as a special instance of bucket increasing trees which is a family of some combinatorial objects. Kazemi [2] introduced a new version of bucket recursive trees where the nodes are buckets with variable capacities labelled with integers 1, 2, ..., $n$. In fact, the capacity of buckets is a random variable in these models. He studied the depth quantity and the first Zagreb index in these models [3].

A bucket recursive tree with variable capacities of buckets (BRT-VCB) starts with the root labelled by 1 that has $r \geq 0$ descendants each of them making a subtree. The nodes in the subtrees have capacities $c < b$ or $c = b$. The nodes with capacities $c < b$ are connected together with 1 edge and the nodes with capacities $b$ have descendants $\geq 0$ again each of them making a subtree such that the labels within these nodes are arranged in increasing order. The tree is completed when the label $n$ is inserted in the tree. Figure 1 illustrates such a tree of size 19 with $b = 3$. For constructing a tree of size $n + 1$ (attracting label $n + 1$ to a tree of size $n$), if a leaf $v$ has the capacity $c < b$, then we add the label $n + 1$ to this node and make a node with capacity $c + 1$ or produce a node $n + 1$. But for a node with capacity $b$, we only produce a new node $n + 1$. The last nodes with $c \leq b$ labels at the end of subtrees are called leaves and other nodes are called non-leaves. The probability $p$, which gives the probability that label $n + 1$ is attracted by node $v$ in the model is $\frac{c(v)}{n - |\gamma|}$, where

$$\gamma = \{v \in T; c = c(v) = k < b, \text{ and } v \text{ is a non-leaf}\}.$$ 

This model can be considered as a generalization of random recursive trees [1].

---

1Assistant Professor, Department of Statistics, Imam Khomeini International University, Qazvin, Iran, e-mail: kazemi@ikiu.ac.ir
of non-negative numbers \((\alpha_k)_{k \geq 0}\) with \(\alpha_0 > 0\) and a sequence of non-negative numbers \(\beta_1, \beta_2, ..., \beta_{b-1}\) is used to define the weight \(w(T)\) of any ordered tree \(T\) by \(w(T) := \Pi_v w(v)\), where \(v\) ranges over all nodes of \(T\). The weight \(w(v)\) of a bucket \(v\) is given as follows:

\[
w(v) := \begin{cases} 
\alpha_d(v), & v \text{ is root or complete (} c(v) = b) \\
\beta_c(v), & v \text{ is incomplete (} c(v) < b). 
\end{cases}
\]

Let \(L(T)\) denotes the set of different increasing labelings of the tree \(T\) with distinct integers \(\{1, 2, ..., |T|\}\), where \(L(T) := |L(T)|\) denotes its cardinality. Then the family \(\mathcal{T}\) consists of all trees \(T\) together with their weights \(w(T)\) and the set of increasing labelings \(L(T)\). For a given degree-weight sequence \((\alpha_k)_{k \geq 0}\) with a degree-weight generating function \(\phi(t) := \sum_{k \geq 0} \alpha_k t^k\) and a bucket-weight sequence \(\beta_1, \beta_2, ..., \beta_{b-1}\), we define the exponential generating function

\[
T_{n,b}(z) := \sum_{n=1}^{\infty} T_{n,b} \frac{z^n}{n!},
\]

where \(T_{n,b} := \sum_{|T| = n} w(T) \cdot L(T)\) is the total weights. Kazemi [2] showed

\[
T_{n,b} = \frac{(n-1)!(b)!^{n(1-\sum_{i=1}^{r} |P_{n_i}|)}}{b}, \quad n \geq 1,
\]

\[
T_{n,b}(z) = -\frac{1}{b} \log \left(1 - b!^{1-\sum_{i=1}^{r} |P_{n_i}|} z\right),
\]

\[
\phi(T_{n,b}(z)) = \frac{(b-1)!}{1 - b!^{1-\sum_{i=1}^{r} |P_{n_i}|} z}, \quad (1)
\]

where \(|P_{n_i}|\) denotes the size of the set of all trees of size \(n_i\)(\(i = 1, 2, ..., r\)). In the equation (1), if \(i\)-th subtree starts with a bucket with capacity \(c = 1\), then we set \(|P_{n_i}| = 0\).

The motivation for studying the bucket recursive trees with variable capacities of buckets is multifold. For example, if \(n\) atoms in a branching molecular structure (such as dendrimer) are stochastically labelled with integers \(1, 2, ..., n\), then atoms in different functional groups can be considered as the labels of different buckets of a bucket recursive tree.

2. Number of Branches of Size \(a\)

Let \(S_{n,j,a}\) counts the number of branches of size \(a\) attached to the bucket containing label \(j\) in our model of size \(n\) (or in other words, the number of subtrees of size \(a\) rooted at the children of node containing label \(j\), in a random grown tree of size \(n\)). We use a
combinatorial approach to find the differential equation corresponds to the random variable $S_{n,j,a}$ and give a closed formula for the probability distribution and the factorial moments of $S_{n,j,a}$. Furthermore limiting distribution result of $S_{n,j,a}$ is given, not only for $j$ fixed, but a full characterization dependent on the growth $j = j(n)$ compared to $n$ is presented. Moreover, the joint distribution of $S_{n,1,1}, S_{n,1,2}, \ldots, S_{n,1,n-j}$ is computed for all grown trees.

Let $S_{n,1,a}$ be the number of branches of size $a$ attached to the root node (label 1) in our model of size $n$. For encoding the behavior of this quantity we introduce the bivariate generating function

$$M(z,v) = \sum_{n \geq 1} \sum_{m \geq 0} P(S_{n,1,a} = m) T_{n,b} \frac{z^n}{n!} v^m.$$ 

By definition of the model one gets the following explicit result for the probabilities $P(S_{n,1,a} = m)$:

$$P(S_{n,1,a} = m) = \sum_{r \geq m} \alpha_r \left( \frac{n}{m} \right) \sum_{n_1 + \cdots + n_r = n-1} T_{n_1,b}^* \cdots T_{n_r,b}^* \left( \frac{n-1}{n_1, \ldots, n_r} \right),$$

where for $1 \leq i \leq m$, $n_i = a$ and for $m+1 \leq j \leq r$, $n_j \neq a$. Also $T_{n,r,b}^*$ is the total weights of the $i$th subtree. Since $T_{n_1,b}^* \cdots T_{n_r,b}^* = b^r \Sigma_{j=1}^{\left|P_{n,j}\right|} T_{n_1,b} \cdots T_{n_r,b}$, (c.f [2]) by multiplying with $T_{n,b} z^{n-1} v^m / (n-1)!$ and summing up over $n \geq 1, m \geq 0$, the equation (2) yields to an explicit formula for $\frac{\partial}{\partial z} M(z,v)$, which is given below:

$$\frac{\partial}{\partial z} M(z,v) = b^r \Sigma_{j=1}^{\left|P_{n,j}\right|} \left\{ \varphi(T_{n,b}(z) + T_{a,b} z^a(v-1)) \right\} \exp \left\{ \frac{b T_{a,b} z^a(v-1)}{a!} \right\}.$$

For describing the behavior of arbitrary label $j > 1$ we introduce the trivariate generating function

$$N(z,u,v) = \sum_{k \geq 0} \sum_{j \geq 1} \sum_{m \geq 0} P(S_{k+j,j,a} = m) T_{k+j,b} \frac{z^{j-1}}{(j-1)!} \frac{u^k}{k!} v^m.$$ 

For our model of size $n$ with root-degree $r$ and subtrees with sizes $n_1, \ldots, n_r$, enumerated from left to right, where the bucket containing label $j$ lies in the leftmost subtree and is the $i$-th bucket in this subtree, we can reduce the computation of the probabilities $P(S_{n,i,a} = m)$ to the probabilities $P(S_{n,i+a} = m)$, when the parameter does only depend on the subtree of bucket containing label $j$. We get as factor the total weight of the $r$ subtrees and the root node $\alpha_r b^r \Sigma_{j=1}^{\left|P_{n,j}\right|} T_{n_1,b} \cdots T_{n_r,b}$, divided by the total weight $T_{n,b}$ of trees of size $n$ and multiplied by the number of order preserving relabellings of the $r$ subtrees, which are given here by

$$\left( \frac{j-2}{i-1} \right) \left( \frac{n-j}{n_1-i} \right) \left( \frac{n-1-n_1}{n_2, n_3, \ldots, n_r} \right).$$

Due to symmetry arguments we obtain a factor $r$, if the bucket containing label $j$ is the $i$-th bucket in the second, third, ..., $r$-th subtree. Summing up over all choices for the rank $i$ of bucket containing label $j$ in its subtree, the subtree sizes $n_1, \ldots, n_r$, and the degree $r$ of the root node gives for $n \geq j \geq 2$ the following recurrence:

$$P(S_{nj,a} = m) = \sum_{r \geq 1} r \alpha_r \sum_{n_1 + \cdots + n_r = n-1} T_{n_1,b}^* \cdots T_{n_r,b}^* \frac{T_{n,b}}{T_{n_1,b} \cdots T_{n_r,b}} \times \min_{i=1}^{\min(n_1,j-1)} P(S_{n_1,i,a} = m) \left( \frac{j-2}{i-1} \right) \left( \frac{n-j}{n_1-i} \right) \left( \frac{n-1-n_1}{n_2, n_3, \ldots, n_r} \right).$$

(4)
With the same method of [2],

\[ \frac{\partial}{\partial z} N(z, u, v) = b! n^{-\sum_{i=1}^{\left\lfloor \frac{n}{a} \right\rfloor} |\mathcal{P}_{n,i}|} \varphi(T_{n,b}(z + u))N(z, u, v) \]

with the initial condition

\[ N(0, u, v) = \sum_{k \geq 0} \sum_{m \geq 0} \mathbb{P}(S_{k+1,1,a} = m)T_{k+1,b} \frac{u^k}{k!} v^m = \frac{\partial}{\partial u} M(u, v). \]

Thus

\[ N(z, u, v) = b! n^{-\sum_{i=1}^{r} |\mathcal{P}_{n,i}|} \varphi(T_{n,b}(z + u)) \frac{\partial}{\partial u} M(u, v) \frac{\varphi(T_{n,b}(u))}{\varphi(T_{n,b}(z + u))}. \]

3. The Main Results

In the following theorem we show that the marginal probabilities \( \mathbb{P}(S_{n,j,a} = m) \) are independent of \( b \).

**Theorem 3.1.** The probability that there are \( m \) branches of size \( a \) attached to bucket containing label \( j \) is given as follows:

\[ \mathbb{P}(S_{n,j,a} = m) = \frac{1}{a^m a! \left( \frac{n-1}{j-1} \right)} \sum_{\ell=0}^{|n-\ell-a|} \left( -1 \right)^\ell \left( n - a(m + \ell) \right). \]

**Proof.** Let \( [z^n]f(z) \) denote the operation of extracting the coefficient of \( z^n \) in the formal power series \( f(z) = \sum f_n z^n \). Thus

\[ \mathbb{P}(S_{n,j,a} = m) = \frac{(j-1)! (n-j)! [z^{j-1} u^{n-j} v^m] N(z, u, v)}{a^m a! \left( \frac{n-1}{j-1} \right)} \frac{\frac{\partial}{\partial u} M(u, v)}{1 - b! \sum_{i=1}^{r} |\mathcal{P}_{n,i}| u z} \]

\[ = b! n^{-\sum_{i=1}^{r} |\mathcal{P}_{n,i}|} \frac{b^{n-1} (j-1)(1 - \sum_{i=1}^{r} |\mathcal{P}_{n,i}|)}{b! a^m \left( \frac{n-1}{j-1} \right)} \frac{\varphi(T_{n,b}(u))}{(1 - b! \sum_{i=1}^{r} |\mathcal{P}_{n,i}| u z)} \frac{e^{b\sum_{i=1}^{r} |\mathcal{P}_{n,i}|} u^a e^{-b_T \sum_{i=1}^{r} |\mathcal{P}_{n,i}| u z}}{1 - b! \sum_{i=1}^{r} |\mathcal{P}_{n,i}| u z} \]

\[ = \frac{1}{a^m a! \left( \frac{n-1}{j-1} \right)} \sum_{\ell=0}^{|n-\ell-a|} \left( -1 \right)^\ell \left( n - a(m + \ell) \right). \]

since \([z^n]f(qz) = q^n[z^n]f(z)\).

**Theorem 3.2.** Let \( m^2 = m(m-1) \cdots (m-s+1) \). The factorial moments of the random variable \( S_{n,j,a} \) in our model of size \( n \) are given as follows:

\[ \mathbb{E}(S_{n,j,a}^2) = \sum_{m \geq 0} m^2 \mathbb{P}(S_{n,j,a} = m) = \frac{1}{a^s} \left( \frac{n-1}{j-1} \right). \]

**Proof.** Let \( D_x \) be the differential operator with respect to \( x \), and \( E_x \) be the evaluation operator at \( x = 1 \). Thus

\[ E_x D_v^s \frac{\partial}{\partial z} M(z, v) = \frac{(b-1)!}{1 - b! \sum_{i=1}^{r} |\mathcal{P}_{n,i}| z^a} \exp \left\{ \frac{b_T \sum_{i=1}^{r} |\mathcal{P}_{n,i}| z^a (v - 1)}{a!} \right\} \]

\[ = \frac{(b-1)!}{1 - b! \sum_{i=1}^{r} |\mathcal{P}_{n,i}| z^a} \left( \frac{b_T \sum_{i=1}^{r} |\mathcal{P}_{n,i}| z^a}{a!} \right)^s. \]
A) For $z^j {z^n}^j$ The joint distribution of $N(z, u, v)$ We have the factor $D_z(n, z, u, v)$ Thus $\rightarrow \infty$ Proof.

B) For $\alpha^n_j$ Proof.

C) For bucket recursive trees introduced by Mahmoud and Smythe are independent of $j$ for investigating the effect of bucketing on random recursive trees. All results obtained for these models $n, b$ since $n \rightarrow \infty$ follows from the Poisson distributed with parameter $\lambda > 0$.

**Theorem 3.3.** A) For $n \rightarrow \infty, j = o(n)$ and a fixed: $S_{n,j,a} \rightarrow S_a \sim \text{Poi}\left(\frac{1}{a}\right)$.

B) For $n \rightarrow \infty, j = \rho n$ with $0 < \rho < 1$ and a fixed: $S_{n,j,a} \rightarrow S_{\rho,a} \sim \text{Poi}\left(\frac{1-\rho}{a}\right)$.

C) For $n \rightarrow \infty, n - j = o(n)$ and a fixed: $S_{n,j,a} \rightarrow S_a \sim \text{P}(S_a = 0) = 1$.

**Proof.** The proof is quite similar to increasing trees [1].

**Theorem 3.4.** The joint distribution of $S_{n,1,1}, S_{n,1,2}, ..., S_{n,1,n-1}$ is given as follows:

$$
\mathbb{P}(S_{1,1} = m_1, ..., S_{n,1,n-1} = m_{n-1}) = \frac{\alpha^{\sum_{i=1}^{n-1} m_i} \left(\sum_{i=1}^{n-1} m_i\right)!}{b^{\sum_{i=1}^{n-1} \left|s_i\right|}} \prod_{i=1}^{n-1} \frac{b^{m_{i,i}(1-\sum_{i=1}^{n-1} \left|s_i\right|)}}{(b)_{m_{i,i}!}},
$$

for all sequences of non-negative integers satisfying $\sum_{i=1}^{n-1} m_i = n - 1$.

**Proof.** We have the factor \(\frac{\alpha^{\sum_{i=1}^{n-1} m_i} \left(\sum_{i=1}^{n-1} m_i\right)!}{b^{\sum_{i=1}^{n-1} \left|s_i\right|}}\) to the choices for the labels, the factor $\alpha^{\sum_{i=1}^{n-1} m_i}$ corresponds to the root degree and the factor \(\left(\sum_{i=1}^{n-1} m_i\right)\) to the different positions of the subtrees. Thus

$$
T_{n,b} \mathbb{P}(S_{1,1} = m_1, ..., S_{n,1,n-1} = m_{n-1}) = \alpha^{\sum_{i=1}^{n-1} m_i} \left(\sum_{i=1}^{n-1} m_i\right)! \prod_{i=1}^{n-1} \frac{b^{m_{i,i}(1-\sum_{i=1}^{n-1} \left|s_i\right|)}}{(b)_{m_{i,i}!}},
$$

Since the total weights of BRT-VCB with $n$ vertices is [2]:

$$
T_{n,b} = b^{-1}(n-1)! (b!)^{n-1} \left(1-\sum_{i=1}^{n-1} \left|s_i\right|\right),
$$

proof is completed.

**4. Conclusion**

In this paper we studied the branches of size $a$ attached to the bucket containing label $j$ for investigating the effect of bucketing on random recursive trees. All results obtained for bucket recursive trees introduced by Mahmoud and Smythe are independent of $b$ since for these models $T_{n,b} = (n-1)!$ [4]. For bucket recursive trees with variable capacities of...
buckets, although $T_{n,b} = n^{-1}(n-1)!((b)^n(1-\sum_{i=1}^{\mu} |P_{n_i}|))$, but only the joint distribution of $S_{n,1,1}$, $S_{n,1,2}$, ..., $S_{n,1,n-1}$ is dependent on $b$.

REFERENCES