Fourth Colloquium on Mathematics and Computer Science DMTCS proc. **AG**, 2006, 423–426

Recursive random trees with product-form random weights

Konstantin Borovkov[1†] and Vladimir Vatutin[2‡]

[1]*Department of Mathematics and Statistics, The University of Melbourne, Parkville 3010, Australia*
[2]*Department of Discrete Mathematics, Steklov Mathematical Institute, Gubkin St. 8, 119991 Moscow, Russia*

We consider growing random recursive trees in random environment, in which at each step a new vertex is attached according to a probability distribution that assigns the tree vertices masses proportional to their random weights. The main aim of the paper is to study the asymptotic behavior of the mean numbers of outgoing vertices as the number of steps tends to infinity, under the assumption that the random weights have a product form with independent identically distributed factors.

Keywords: random recursive trees, random environment, Sptizer's condition, distance to the root, outdegrees

Introduction

We consider random recursive trees which are constructed incrementally as follows. Starting with a single vertex $v(0)$ with weight $w(0) = 1$ and label 0, at the first step a new vertex $v(1)$ is added to the tree as a child of the initial vertex. It is labelled by 1 and assigned a random weight $w(1) > 0$. At step $j > 1$, given all the weights $w(0), w(1), \dots, w(j-1)$, first a node $v(j^*)$ is chosen at random from the existing nodes $v(0), v(1), \dots, v(j-1)$ with probabilities proportional to the nodes' weights, and then a new vertex $v(j)$ is added to the tree as a child of $v(j^*)$. The new vertex has label j and a random weight $w(j) > 0$. As at the initial step (where we put $1^* = 0$), the edge is directed from $v(j^*)$ to its child vertex $v(j)$.

If $w(j) = 1$ for all j, then we get the standard recursive tree. If $w(j) = a^j$, where $a \neq 1$ is positive, we get the recursive tree considered in Borovkov and Motyer (3).

In this paper we consider the case when $w(j) = a_1 \cdots a_j$, where a_k are i.i.d. random variables. Our aim is to study the asymptotic behavior of the mean value of the outdegree $N_n(j)$ of the vertex $v(j)$, $j \leq n$, (i.e. the number of edges coming out of $v(j)$ in the tree having n nonrooted vertices) as $n \to \infty$.

Denote by $\mathcal{T}_n$, $n = 0, 1, 2, \dots$, the set of all rooted recursive trees with n nonrooted vertices. That is, $\mathcal{T}_n$ consists of the rooted trees whose root is labelled by 0 and whose nonrooted vertices are labelled by numbers $1, 2, \dots, n$ in such a way that for any $j \leq n$ the shortest path leading from the vertex with label j to the root traverses only the vertices with labels $k \leq j$. For a $t_n \in \mathcal{T}_n$, let $t_n(j) \in \mathcal{T}_{n+1}$ be the recursive tree which is obtained from t_n by adding a vertex labelled by $n+1$ as a child of the vertex with the label $j \in \{0, 1, \dots, n\}$.

Now we can describe our construction of random recursive trees. Let θ_j, $j = 1, 2, \dots, n$, be i.i.d. r.v.'s. First, we run a random walk

$$S_0 = 0, \qquad S_j = \theta_1 + \cdots + \theta_j, \quad j \geq 1. \tag{1}$$

Second, given S_j, $j = 0, 1, \dots, n$, we construct a (conditional) Markov chain $T_0, T_1, \dots, T_n$ with $T_k \in \mathcal{T}_k$, $k = 0, 1, \dots, n$, by assigning the weight $w(j) := e^{-S_j}$ to the vertex labelled by $j \geq 0$ (so that $w(j) = a_1 \cdots a_j$, $j \geq 1$, with $a_j := e^{-\theta_j}$ being i.i.d. r.v.'s), setting, for $r = 0, 1, \dots, n$,

$$W_r := \sum_{q=0}^{r} w(q) = \sum_{q=0}^{r} e^{-S_q}, \qquad p_r(j) := e^{-S_j} W_r^{-1} = e^{-S_j} \Big(\sum_{q=0}^{r} e^{-S_q} \Big)^{-1}, \quad j = 0, 1, \dots, r,$$

†Supported by the ARC Centre of Excellence for Mathematics and Statistics of Complex Systems.
‡Supported by RFBR grant 05-01-00035 and by the program "Contemporary Problems of Theoretical Mathematics" of the Russian Academy of Sciences.

and letting, for any $t_r \in \mathcal{T}_r$ and $j = 0, 1, \ldots, r$,

$$\mathbf{P}_w\left(T_{r+1} = t_r(j) \middle| T_r = t_r\right) \equiv \mathbf{P}\left(T_{r+1} = t_r(j) \middle| T_r = t_r;\, w(1), \ldots, w(r)\right) : p_r(j).$$

Recalling that $v(k^*)$ is the parent vertex for $v(k)$ and denoting by $I\{\mathcal{A}\}$ the indicator of the event $\mathcal{A}$, we can represent the outdegree $N_n(j)$ of the vertex $v(j)$, $j \leq n$, as

$$N_n(j) = \sum_{k=j+1}^{n} I\left\{v(k^*) = v(j)\right\}.$$

Therefore,

$$\begin{aligned} \mathbf{E}_w N_n(j) &:= \mathbf{E}\left[N_n(j) \middle| w(1), \ldots, w(n-1)\right] \\ &= \sum_{k=j+1}^{n} \mathbf{E}_w I\left\{v\left(k^*\right) = v(j)\right\} \sum_{k=j+1}^{n} p_{k-1}(j) = e^{-S_j} \sum_{k=j}^{n-1} W_k^{-1} \end{aligned}$$

and

$$\mathbf{E} N_n(j) = \sum_{k=j}^{n-1} \mathbf{E} e^{-S_j} W_k^{-1}.$$

In this note we formulate two theorems describing the asymptotic behavior (as $n \to \infty$) of the expectations $\mathbf{E}N_n(j)$ and the distributions of the random variables $\mathbf{E}_w N_n(j)$ in different ranges of the parameter j values when the random walk (1) satisfies the following *Spitzer* condition: there exists a $\rho \in (0,1)$ such that

$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{P}(S_k > 0) \to \rho \qquad \text{as} \quad n \to \infty. \tag{2}$$

Observe that if $\mathbf{E}\theta_1 = 0$ and the distribution of θ_1 belongs to the domain of attraction of a stable law of order $\alpha \in (1,2]$ then (see e.g. Bingham *et al.* (2)) there exists $\beta \in [-1,1]$ such that

$$\rho = \frac{1}{2} + \frac{1}{\pi\alpha} \arctan\Big(-\beta \tan \frac{\pi\alpha}{2}\Big).$$

Statement of results

Let

$$\gamma_0 := 0, \qquad \gamma_{j+1} := \min\{n > \gamma_j : S_n < S_{\gamma_j}\}, \quad j \geq 0,$$

and

$$\Gamma_0 := 0, \qquad \Gamma_{j+1} := \min\{n > \Gamma_j : S_n > S_{\Gamma_j}\}, \quad j \geq 0,$$

be the strict descending and ascending ladder epoches of $\{S_n\}_{n \geq 0}$, resp. Introduce the two renewal functions

$$\begin{aligned} U(x) &:= 1 + \sum_{j=1}^{\infty} \mathbf{P}(S_{\Gamma_j} < x), \quad x > 0; \quad U(0) = 1, \quad U(x) = 0, \quad x < 0, \\ V(x) &:= \sum_{j=0}^{\infty} \mathbf{P}(S_{\gamma_j} \geq -x), \quad x > 0; \quad V(0) = 1, \quad V(x) = 0, \quad x < 0. \end{aligned}$$

By means of $V(x)$ and $U(x)$ and the constant $\phi := \sum_{j=1}^{\infty} j^{-1}\mathbf{P}(S_j = 0)$ one can specify two sequences of probability measures $\{\mathbf{P}_n^-\}_{n \geq 1}$ and $\{\mathbf{P}_n^+\}_{n \geq 1}$ (defined on the σ-algebras $\Sigma_n = \sigma(S_1, \ldots, S_n)$, resp.), by setting for each $\mathcal{A} \in \Sigma_n$

$$\mathbf{P}_n^-(\mathcal{A}) := e^{\phi} \int_{\mathcal{A}} U(-S_n)\, I\{\tilde{M}_n < 0\}\, d\mathbf{P}, \qquad \mathbf{P}_n^+(\mathcal{A}) := \int_{\mathcal{A}} V(S_n) I\{L_n \geq 0\}\, d\mathbf{P}.$$

It is known (see e.g. Vatutin and Dyakonova (4)) that the measures are consistent, and therefore there exist measures $\mathbf{P}^-$ and $\mathbf{P}^+$ on the σ-algebra $\sigma(S_1, S_2, \ldots)$ such that their restrictions $\mathbf{P}^{\pm}|_{\Sigma_n}$ to Σ_n coincide with $\mathbf{P}_n^{\pm}, n = 1, 2, \ldots$.

One can show (see Lemma 2.7 in Afanasyev *et al.* (1)) that under the condition (2)

$$\eta_1 := \sum_{k=1}^{\infty} e^{S_k} < \infty \quad \mathbf{P}^{-}\text{- a.s.}, \qquad \eta_2 := \sum_{k=0}^{\infty} e^{-S_k} < \infty \quad \mathbf{P}^{+}\text{- a.s.} \tag{3}$$

Set

$$L_n := \min_{0\le k\le n} S_k, \quad \tilde{M}_n := \max_{1\le k\le n} S_k.$$

It is known (see e.g. Bingham *et al.* (2)) that under the condition (2) there exist functions $l_1(n)$ and $l_2(n)$ slowly varying at infinity, $l_1(n)l_2(n) \sim \pi^{-1}\sin\pi\rho$, $n\to\infty$, such that

$$\mathbf{P}\left(L_n \ge 0\right) \sim n^{\rho-1} l_1(n), \qquad \mathbf{P}(\tilde{M}_n < 0) \sim n^{-\rho} l_2(n) \qquad \text{as} \quad n\to\infty. \tag{4}$$

Theorem 1 *Let condition* (2) *hold. Then there exist positive sequences* $\{C_j\}_{j\ge0}$ *and* $\{D_j\}_{j\ge0}$ *such that for any fixed* $j\ge0$

$$\lim_{n\to\infty}\frac{\mathbf{E}N_n(j)}{n\mathbf{P}\left(L_n\ge0\right)} = C_j \qquad \textit{and} \qquad \lim_{n\to\infty}\frac{\mathbf{E}N_n(n-j)}{\mathbf{P}(\tilde{M}_n<0)} = D_j.$$

In addition,

$$\lim_{j,n-j\to\infty}\frac{\mathbf{E}N_n(j)}{(n-j)\mathbf{P}(\tilde{M}_j<0)\mathbf{P}\left(L_{n-j}\ge0\right)} = \frac{1}{\rho}.$$

Remark 2 *In view of* (4)*, the assertions of Theorem 1 can be rewritten as*

$$\mathbf{E}N_n(j) \sim C_j n^{\rho} l_1(n), \quad \mathbf{E}N_n(n-j) \sim D_j n^{-\rho} l_2(n) \quad \textit{as} \quad n\to\infty,$$
$$\mathbf{E}N_n(j) \sim \rho^{-1} j^{-\rho} l_2(j)\,(n-j)^{\rho}\, l_1(n-j) \quad \textit{as} \quad j, n-j\to\infty.$$

Our description of the asymptotic behavior of the distribution of the r.v. $\mathbf{E}_w N_n(j)$ will only be limited to the case when j is located either to the right of the random epoch $\tau(n)$ or in its small left vicinity.

Let $\tau(n) := \min\{k\ge0 : S_k \le S_l,\ l\in[0,n]\}$ be the left-most point at which the minimal value of the random walk $\{S_k\}_{0\le k\le n}$ is attained.

Theorem 3 *Let condition* (2) *hold and* $j = j(n)$ *be an arbitrary (possibly random) sequence such that* $(\tau(n)-j)_+ = o(n)$ *in probability as* $n\to\infty$*. Then*

$$\frac{e^{S_j - S_{\tau(n)}}}{n-j}\mathbf{E}_w N_n(j) \Longrightarrow \frac{1}{\eta_1^- + \eta_2^+},$$

where η_1^- *and* η_2^+ *are independent copies of the r.v.'s* η_1 *and* η_2 *from* (3)*, resp.*

References

[1] Afanasyev, V.I., Geiger, J., Kersting, G. and Vatutin, V.A. *Criticality for branching processes in random environment.* Ann. Probab., 33 (2005), 2, 645–673.

[2] Bingham, N.H., Goldie, C.M. and Teugels, J.L. *Regular variation.* Cambridge University Press, Cambridge (1987).

[3] Borovkov, K.A. and Motyer, A. *On the asymptotic benavior of a simple growing point process model.* Statistics and Probability Letters, 72 (2005), 265–275.

[4] Vatutin, V.A. and Dyakonova, E.E. *Galton-Watson branching processes in random environment. II: Finite-dimensional distributions.* Theory Probab. Appl., 49 (2004), 275–308.