Discrete Mathematics and Theoretical Computer Science DMTCS vol. **13:3**, 2011, 21–38

The Variance of the Profile in Digital Search Trees

Ramin Kazemi[†] Mohammad Q. Vahidi-Asl[‡]

Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Evin, Tehran, Iran

received 31st January 2011, *accepted 24th August 2011*.

What today we call digital search tree (DST) is Coffman and Eve's sequence tree introduced in 1970. A digital search tree is a binary tree whose ordering of nodes is based on the values of bits in the binary representation of a node's key. In fact, a digital search tree is a digital tree in which strings (keys, words) are stored directly in internal nodes. The profile of a digital search tree is a parameter that counts the number of nodes at the same distance from the root. In this paper we concentrate on external profile, i.e., the number of external nodes at level k when n strings are sorted. By assuming that the n input strings are independent and follow a (binary) memoryless source the asymptotic behaviour of the average profile was determined by Drmota and Szpankowski (2011). The purpose of the present paper is to extend their analysis and to provide a precise analysis of variance of the profile. The main (technical) difference is that we have to deal with an inhomogeneous part in a proper functional-differential equations satisfied by the second moment and Poisson variance. However, we show that the variance is asymptotically of the same order as the expected value which implies concentration. These results are derived by methods of analytic combinatorics such as generating functions, Mellin transform, Poissonization, the saddle point method and singularity analysis.

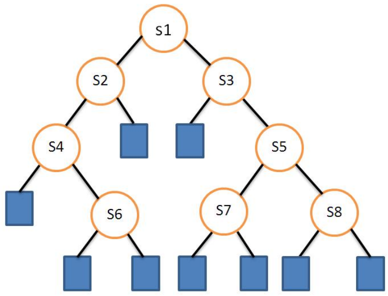
Keywords: Digital search trees, tree profiles, Poisson variance, generating functions, saddle point method, singularity analysis, Mellin transform, Poissonization.

1 Introduction

In 1970, Coffman and Eve published a paper that introduced several modification to hashing schemes using trees [1]. The paper included three type of trees and some analysis of their basic properties. These trees are known today by different names. Coffman and Eve's sequence trees are what today we call digital search trees. The method is well suited for storing data when the bit composition (or letter composition) of the data is available [16]. In fact, digital search trees and tries are two classes of so-called $\log n$–trees and their construction is based on digital keys and not on the order structure of the keys as in the case of binary search trees [2]. They are important in many computer science applications like data compression, pattern matching and hashing. For example, the popular Lempel-Ziv compression scheme is strongly related to digital search trees.

Digital search trees (especially, digital search trees and tries) are fundamental data structures on words [19, 20, 21, 22, 23]. Jacquet and Szpankowski [7] considered the asymptotic behavior of the Lempel-Ziv

[†]Email: `rst.kazemi@gmail.com`
[‡]Email: `m-vahidi@sbu.ac.ir`



$$\begin{aligned} B_{8,0} &= 0 \\ B_{8,1} &= 0 \\ B_{8,2} &= 2 \\ B_{8,3} &= 1 \\ B_{8,4} &= 6 \end{aligned}$$

Fig. 1: A digital search tree built on eight strings $s_1, ..., s_8$ (i.e., $s_1 = 0...$, $s_2 = 1...$, $s_3 = 01...$, $s_4 = 11...$, etc.) with internal (ovals) and external (squares) nodes, and its profiles.

parsing scheme and digital search trees by assuming a random model based on independent memoryless input strings. They also analyzed digital tries with Markovian dependency [6]. There is a more general random model of dynamical sources but in this case the technicalities are usually more involved [13, 23]. Furthermore, Louchard and Szpankowski [15] studied the average profile and limiting distribution for the phrase size in the Lempel-Ziv parsing algorithm, Jacquet and Régnier [9] discussed the limiting distribution in a trie partitioning process, and Knessl and Szpankowski [10] analyzed the asymptotic behavior of the height in a digital search tree and the longest phrase of the Lempel-Ziv scheme. They also considered the average profile of symmetric digital search trees [11] (which means that all letters in the memoryless source appear with the same frequency). Louchard [14] provided an exact and asymptotic distribution in digital search trees.

Recently Park et al. [17, 18] studied the external and internal profiles of tries (with memoryless input strings) and provided a very detailed picture on the asymptotic behaviour of the provile (expected value, variance, limiting distribution). For digital search trees (DST) Drmota and Szpankowski [3] proved asymptotic results on the average profile which are similar to those for tries. The purpose of this paper is to extend their analysis to asymptotics for the variance of the profile.

Before we state our results we recall the construction of a digital search tree (DST) which stores (binary) input strings in its internal nodes. The root contains the first string, and the next string occupies the right or the left child of the root depending on whether its first symbol of the next string is 0 or 1. At each level of the tree a different bit of the key is checked; if the bit is 0, the search continues down the left subtree, if it is 1, the search continues down the right subtree. The remaining strings are stored in available nodes which are directly attached to nodes already existing in the tree, as shown in Figure 1[12, 16, 21].

Let $B_{n,k}$ be the number of external nodes at level k when n strings are sorted. We study the external profile built over n binary strings generated by a memoryless source, that is, we assume each string is a binary i.i.d. sequence with p being the probability of a "1" $(0 < p < 1)$; we also use $q := 1 - p$ and assume that $p < q$. We also mention that symmetric DST's, i.e. $p = q = \frac{1}{2}$, are not covered by our analysis because for $p < q$, $g(s, w)$ is an analytic function for $|w| < 1/T(s)$ where $T(s) = p^{-s} + q^{-s}$ and the saddle point analysis fails for $p = q$ [3]. In this case we expect a completely different behaviour and also the use of different methods, see also the discussion of the symmetric case in [3].

In order to state our main result we need the following notations. For a real number α with $(\log \frac{1}{p})^{-1} <$

$\alpha < (\log \frac{1}{q})^{-1}$, let $\rho = \rho(\alpha)$ be defined by the equation

$$\alpha = \frac{p^{-\rho} + q^{-\rho}}{p^{-\rho} \log \frac{1}{p} + q^{-\rho} \log \frac{1}{q}}.$$

Explicitly, $\rho(\alpha)$ is given by

$$\rho = \rho(\alpha) = \frac{1}{\log(p/q)} \log \left(\frac{1 - \alpha \log(1/p)}{\alpha \log(1/q) - 1} \right).$$

Furthermore set

$$\beta(\rho) = \frac{p^{-\rho} q^{-\rho} \log(p/q)^2}{(p^{-\rho} + q^{-\rho})^2}.$$

Theorem 1 *Let $\mathbb{V}ar(B_{n,k})$ denote the variance of the profile in unbalanced digital search trees with underlying probabilities $0 < p < q = 1 - p$. Let k and n be positive integers such that $k/\log n$ satisfies $(\log \frac{1}{p})^{-1} + \varepsilon \le k/\log n \le (\log \frac{1}{q})^{-1} - \varepsilon$. Then uniformly*

$$\mathbb{V}ar(B_{n,k}) = L\left(\rho_{n,k}, \log_{p/q} p^k n\right) \frac{(p^{-\rho_{n,k}} + q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}} \left(1 + \mathcal{O}\left(\frac{1}{\sqrt{k}}\right)\right), \tag{1}$$

where $\rho_{n,k} = \rho(k/\log n)$ and $L(\rho, x)$ is a non-zero periodic function with period 1 in x.

Remark 1 *The function $L(\rho, x)$ can be represented as $L(\rho, x) = \sum_{j \in Z} f(\rho + it_j)\Gamma(\rho + it_j)e^{-2j\pi ix}$, where $f(s)$ has the form*

$$f(s) = g(s-1, 1/T(s))D(s, 1/T(s))$$

and the functions $g(s, w)$ and $D(s, w)$ are described in (17) and in Lemma 3, respectively. However, it is not really explicit.

This theorem shows that the variance $\mathbb{V}ar(B_{n,k})$ is of the same order of magnitude as the expected value $\mathbb{E}(B_{n,k})$. In particular it follows that $B_{n,k}$ is concentrated around its expected value. The function L has small amplitude and their oscillations are consequences of an infinite number of saddle-points appearing in the integrand of the associated Mellin transform. The only difference between two asymptotic formulas is in function L for two different function f. The same phenomenon has been observed for tries [17, 18] where it was also shown that $B_{n,k}$ satisfies a central limit theorem. It would be natural to expect a corresponding behaviour for DST's, however, the methods of [17, 18] are not applicable in the present situation.

In order to analyze the variance we use the so-called Poisson variance $V_k(x)$ as a (hopefully) good approximation of the variance of the profile $\mathbb{V}ar(B_{n,k})$ [18] if we set $x = n$. It is defined by

$$V_k(x) := \widetilde{\Delta}_k''(x, 1) + \widetilde{\Delta}_k'(x, 1) - \left(\widetilde{\Delta}_k'(x, 1)\right)^2,$$

where

$$\widetilde{\Delta}_k(x, u) = \sum_{n=0}^{\infty} \mathbb{E}(u^{B_{n,k}}) e^{-x} \frac{x^n}{n!}$$

and $\widetilde{\Delta}'_k(x,1)$ and $\widetilde{\Delta}''_k(x,1)$ denote the first and the second derivative of $\widetilde{\Delta}_k(x,u)$ with respect to u at $u=1$, respectively.

It turs out that the Poisson variance satisfies a recurrence (10) that is quite similar to that of the "Poisson expectation" $\Delta_k^{(1)}(x) = \sum_{n=0}^{\infty} \mathbb{E}(B_{n,k}) e^{-x} \frac{x^n}{n!}$, see (5), in particular the inhomogeneous part in (10) is relatively small. This suggests suggests that the asymptotics of the variance should be of the same order of magnitude as for the expected value. In order to make this heuristics rigorous we have to overcome several technical difficulties. First we have to obtain an explicit solution for $F_k^{(3)}(s)$ of (11) which is the most difficult part (Section 2). Then we have to find proper asymptotics for $F_k^{(3)}(s)$ and to invert then Mellin transform of $V_k(x)$. This leads us to an infinite number of saddle points (cf. also [3] and [18]). The final step will be to show that the Poisson variance of $V_k(n)$ is asymptotically equal to $\mathbb{V}ar(B_{n,k})$. The reader is referred to [3] and [18] for a detailed discussion of the above mentioned tools that belong to analytic combinatorics.

2 Generating Functions

In this section we recall (and extend) the combinatorial analysis of DST's with the help of generating functions.

2.1 Basic relations

As already introduced, $B_{n,k}$ denotes the (random) number of external nodes at level k in a digital search tree built over n strings generated by a memoryless source with parameter $q > p = 1-q$. We recall the initial conditions

$$B_{n,0} = \begin{cases} 1 & \text{for } n = 0, \\ 0 & \text{for } n \geq 1. \end{cases}$$

The probability generating function of the external profile, $P_{n,k}(u) = \mathbb{E}\Big(u^{B_{n,k}}\Big)$, satisfies the following recurrence relation (cf. [7])

$$P_{n+1,k}(u) = \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} P_{j,k-1}(u) P_{n-j,k-1}(u). \tag{2}$$

with initial conditions $P_{0,k}(u) = 1$ for $k \geq 1$, $P_{0,0}(u) = u$, $P_{n,0}(u) = 1$ for $n \geq 1$. The corresponding exponential generating function

$$G_k(x,u) = \sum_{n\geq 0} P_{n,k}(u) \frac{x^n}{n!}$$

fulfills the following functional recurrence

$$\frac{\partial}{\partial x} G_k(x,u) = G_{k-1}(px,u) G_{k-1}(qx,u), \qquad k \geq 1, \tag{3}$$

with initial conditions $G_0(x,u) = u + e^x - 1$ and $G_k(0,u) = 1$ $(k \geq 1)$. By taking derivatives with respect to u and setting $u = 1$ we obtain for the exponential generating function

$$E_k^{(1)}(x) = \sum_{n\geq 0} \mathbb{E}(B_{n,k}) \frac{x^n}{n!}$$

the following functional recurrence

$$E_k^{'(1)}(x) = e^{qx} E_{k-1}^{(1)}(px) + e^{px} E_{k-1}^{(1)}(qx), \tag{4}$$

with initial conditions $E_0^{(1)}(x) = 1$ and $E_k^{(1)}(0) = 0$ $(k \geq 1)$. The Poisson transform of $E_k^{(1)}(x)$, namely

$$\Delta_k^{(1)}(x) = e^{-x} \sum_{n \geq 0} \mathbb{E}(B_{n,k}) \frac{x^n}{n!} = e^{-x} E_k^{(1)}(x), \quad k \geq 0$$

translates recurrence (4) into

$$\Delta_k^{'(1)}(x) + \Delta_k^{(1)}(x) = \Delta_{k-1}^{(1)}(px) + \Delta_{k-1}^{(1)}(qx), \qquad k \geq 1, \tag{5}$$

with initial conditions $\Delta_0^{(1)}(x) = e^{-x}$ and $\Delta_k^{(1)}(0) = 0$ $(k \geq 1)$. Similarly, by taking second derivatives with respect to u and setting $u = 1$ we obtain for the exponential generating function

$$E_k^{(2)}(x) = \sum_{n \geq 0} \mathbb{E}(B_{n,k}^2) \frac{x^n}{n!}$$

the following functional recurrence

$$E_k^{'(2)}(x) = e^{qx} E_{k-1}^{(2)}(px) + e^{px} E_{k-1}^{(2)}(qx) + 2E_{k-1}^{(1)}(px) E_{k-1}^{(1)}(qx), \tag{6}$$

with initial conditions $E_0^{(2)}(x) = 1$ and $E_k^{(2)}(0) = 0$ $(k \geq 1)$. Furthermore, the Poisson transform of $E_k^{(2)}(x)$, namely

$$\Delta_k^{(2)}(x) = e^{-x} \sum_{n \geq 0} \mathbb{E}(B_{n,k}^2) \frac{x^n}{n!} = e^{-x} E_k^{(2)}(x), \qquad k \geq 0$$

translates recurrence (6) into

$$\Delta_k^{'(2)}(x) + \Delta_k^{(2)}(x) = \Delta_{k-1}^{(2)}(px) + \Delta_{k-1}^{(2)}(qx) + w_k(x), \qquad k \geq 1. \tag{7}$$

where $w_k(x) = 2\Delta_{k-1}^{(1)}(px)\Delta_{k-1}^{(1)}(qx)$, $\Delta_0^{(2)}(x) = e^{-x}$ and $\Delta_k^{(2)}(0) = 0$ $(k \geq 1)$.

By induction it is easy to prove that $\Delta_k^{(2)}(x)$ can be represented as finite linear combinations of functions of the form $e^{-p^{\ell_1} q^{\ell_2} x}$ with $\ell_1, \ell_2 \geq 0$, and of products of two of these functions. Hence, the Mellin transform of $\Delta_k^{(2)}(x)$ (see [5])

$$\Delta_k^{*(2)}(s) = \int_0^\infty \Delta_k^{(2)}(x) x^{s-1} dx.$$

exists for all s with $\Re(s) > 0$. Since $B_{n,k} = 0$ for $k > n$ it follows that $E_k^{(2)}(x) = \mathcal{O}(x^k)$ as $x \to 0$ which ensures that $\Delta_k^{*(2)}(s)$ actually exists for s with $\Re(s) > -k$.

Let us now express $\Delta_k^{*(2)}(s)$ as

$$\Delta_k^{*(2)}(s) = \Gamma(s) F_k^{(2)}(s)$$

where $\Gamma(s)$ is the Euler gamma function. By definition we know that $F_k^{(2)}(s)$ is the finite linear combinations of functions a^{-s} (with certain values of a). Thus, $F_k^{(2)}(s)$ is an entire function. Furthermore (7) translates into

$$F_k^{(2)}(s) - F_k^{(2)}(s-1) = T(s) F_{k-1}^{(2)}(s) + H_k^{(2)}(s), \quad k \geq 0, \tag{8}$$

where

$$H_k^{(2)}(s) = \frac{1}{\Gamma(s)} \int_0^\infty w_k(x) x^{s-1} dx,$$

and $F_0^{(2)}(s) = 1$. Note that (8) does not only hold for $\Re(s) > -k$ where the Mellin transform exists. Since $F_k^{(2)}(s)$ continues analytically to an entire function, (8) holds for all s, too. The inhomogeneous part in (8) is very large compared to the order of magnitude of the homogeneous equation

$$F_k^{(1)}(s) - F_k^{(1)}(s-1) = T(s) F_{k-1}^{(1)}(s), \quad k \geq 0, \tag{9}$$

for the first moment (compare with [3]). Since $F_k^{(1)}(s)$ behaves geometrically as $T(s)^k$ it seems that the term $F_{k+1}^{(1)}(s-1)$ is negligible compared to the other two terms in (9) [3]. This phenomenon will also occur for $F_k^{(2)}(s)$ and $F_k^{(3)}(s)$ (Section 2.3). We introduce the so-called Poisson variance

$$V_k(x) := \Delta_k^{(2)}(x) - \left(\Delta_k^{(1)}(x)\right)^2$$

which should be a good approximation for the variance of the profile ([8]). By (5) and (7), $V_k(x)$ satisfies

$$V_k(x) + V_k'(x) = V_{k-1}(px) + V_{k-1}(qx) + \left(\Delta_k^{'(1)}(x)\right)^2, \tag{10}$$

with initial condition $V_0(x) = e^{-x}(1 - e^{-x})$ and $V_k(0) = 0$ $(k \geq 1)$. The Mellin transform of $V_k(x)$ is then given as

$$V_k^*(s) = \int_0^\infty V_k(x) x^{s-1} dx,$$

and again, we can use a factorization of the form

$$V_k^*(s) = \Gamma(s) F_k^{(3)}(s),$$

where $V_k^*(s)$ and $F_k^{(3)}(s)$ can be written in terms of $\Delta_k^{*(1)}(s)$, $\Delta_k^{*(2)}(s)$, $F_k^{(1)}(s)$, and $F_k^{(2)}(s)$ respectively. In particular, (10) translates into

$$F_k^{(3)}(s) - F_k^{(3)}(s-1) = T(s) F_{k-1}^{(3)}(s) + H_k^{(3)}(s), \quad k \geq 0, \tag{11}$$

where

$$H_k^{(3)}(s) = \frac{1}{\Gamma(s)} \int_0^\infty \left(\Delta_k^{'(1)}(x)\right)^2 x^{s-1} dx$$

and also $F_0^{(3)}(s) = 1 - 2^{-s}$. We also observe that $F_k^{(3)}(-r) = 0$ for $k > r$, since $\Gamma(s) F_k^{(3)}(s)$ is the Mellin transform of $V_k(x)$ that exists for $\Re(s) > -k$. We will use this property several times.

2.2 Analysis of $F_k^{(1)}(s)$

Before we will find a solution of (8) or (11) we recall (and extend) some facts of $F_k^{(1)}(s)$ (see [3]). We define the power series

$$f^{(i)}(s,w) = \sum_{k\ge 0} F_k^{(i)}(s)w^k, \quad i = 1,3. \tag{12}$$

Furthermore we introduce the function operator $\mathbf{A}$ as follows

$$\mathbf{A}[f](s) = \sum_{j\ge 0} f(s-j)T(s-j). \tag{13}$$

We also set $R_k(s) = \mathbf{A}^k[1](s)$ and

$$g(s,w) = \sum_{k\ge 0} R_k(s)w^k = (\mathbf{I} - w\mathbf{A})^{-1}[1](s). \tag{14}$$

By the way it is easy to compute $R_k(s)$ for a few small values of k. For example,

$$\begin{aligned}
R_0(s) &= 1,\\
R_1(s) &= \frac{p^{-s}}{1-p} + \frac{q^{-s}}{1-q},\\
R_2(s) &= \frac{p^{-2s}}{(1-p)(1-p^2)} + \frac{p^{-s}q^{-s}}{(1-p)(1-pq)} + \frac{p^{-s}q^{-s}}{(1-q)(1-pq)} + \frac{q^{-2s}}{(1-q)(1-q^2)}.
\end{aligned}$$

In general we have a representation of the form

$$R_k(s) = \sum_{i=0}^{k} \binom{k}{i} c_{k,i} \left(p^i q^{k-i}\right)^{-s}, \tag{15}$$

where the coefficients $c_{k,i}$ are uniformly bounded by

$$1 \le c_{k,i} \le C$$

for some constant C just depending on p and q.

It is also easy to verify that $R_k(s)$ satisfies the recurrence

$$R_k(s) - R_k(s-1) = T(s)R_{k-1}(s) \tag{16}$$

and satisfies $R_k(-\infty) = 0$ ($k \ge 1$). From (14) and (16) we obtain $g(s,w) = g(s-1,w)/(1-wT(s))$ and consequently

$$g(s,w) = \prod_{j\ge 0} \frac{1}{1-wT(s-j)}. \tag{17}$$

This shows that $w = 1/T(s)$ is the dominanting polar singularity of the mapping $w \mapsto g(s,w)$ (if s is sufficiently close to the real axis) and it also follows that

$$R_k(s) \sim \frac{T(s)^k}{\prod_{j\ge 1}(1 - T(s-j)/T(s))}$$

as $k \to \infty$. Actually this asymptotics is uniform for real $s \ge 0$. By (15) it also follows that

$$|R_k(s)| \le R_k(\Re(s)) = \mathcal{O}\left(T(\Re(s))^k\right).$$

Next we turn to $F_k^{(1)}(s)$ which satisfies the same recurrence (16) as $R_k(s)$. Hence we expect a similar representation for the generating function $f^{(1)}(s,w)$. However, we have different initial conditions, namely $F_k^{(1)}(0) = 0$ ($k \ge 1$) or $f^{(1)}(0,w) = 1$. Anyway, it follows (as above) that

$$f^{(1)}(s,w) = \sum_{k\ge 0} F_k^{(1)}(s) w^k = \frac{g(s,w)}{g(0,w)}. \tag{18}$$

Note that the mapping $w \mapsto 1/g(0,w)$ is an entire function. Hence, $w = 1/T(s)$ is again the dominating polar singularity (if s is sufficiently close to the real axis). Consequently

$$F_k^{(1)}(s) \sim \frac{\prod_{j\ge 0}(1 - T(-j)/T(s))}{\prod_{\ell\ge 1}(1 - T(s-\ell)/T(s))}\, T(s)^k.$$

This is also one of the main observations of [3].

It follows from (18) that $F_k^{(1)}(s)$ is a linear combination of $R_\ell(s)$, $0 \le \ell \le k$, which also shows that $F_k^{(1)}(-\infty)$ exists because $R_k(0) \ne 0$ for all k (see the proof of Theorem 3 in [3]). Furthermore, since the ratio $f^{(1)}(s,w)/g(s,w) = 1/g(0,w)$ is independent of s it follows that

$$\frac{f^{(1)}(0,w)}{g(0,w)} = \frac{f^{(1)}(-\infty,w)}{g(-\infty,w)}.$$

However, we have $f^{(1)}(0,w) = g(-\infty,w) = 1$ which implies that

$$f^{(1)}(-\infty,w) = \sum_{k\ge 0} F_k^{(1)}(-\infty) w^k = \frac{1}{g(0,w)} = \prod_{j\ge 0}(1 - wT(-j)).$$

This also shows that $f^{(1)}(s,w) = g(s,w) f^{(1)}(-\infty,w)$ or that we have the explicit representation

$$F_k^{(1)}(s) = \sum_{\ell=0}^{k} F_{k-\ell}^{(1)}(-\infty) R_\ell(s). \tag{19}$$

We also note that

$$|F_k^{(1)}(-\infty)| = \mathcal{O}(\eta^k) \tag{20}$$

for all (fixed) $\eta > 0$. This follows from the fact that the mapping $w \mapsto 1/g(0,w)$ is an entire function.

Remark 2 *The analysis of $F_k^{(1)}(s)$ that was given in [3] is not as direct as the approach given here. In [3] it was observed that the recurrence (9) translates to the relation*

$$F_k^{(1)}(s) = \boldsymbol{A}[F_{k-1}^{(1)}](s) - \boldsymbol{A}[F_{k-1}^{(1)}](0)$$

which can be translated to the power series representation (18). Both approaches use the fact that the limit $F_k^{(1)}(-\infty)$ exists.

2.3 Analysis of $F_k^{(3)}(s)$

The main difference between the recurrence (9) for $F_k^{(1)}(s)$ and the recurrence (11) for $F_k^{(3)}(s)$ is that (11) contains the inhomogeneous part

$$\begin{aligned} H_k^{(3)}(s) &= \frac{1}{\Gamma(s)} \int_0^\infty \left(\Delta_k^{\prime(1)}(x)\right)^2 x^{s-1}\,dx \\ &= \frac{1}{2\pi i\,\Gamma(s)} \int_{c-i\infty}^{c+i\infty} \Gamma(t)\Gamma(s-t)F_k^{(1)}(t-1)F_k^{(1)}(s-t-1)\,dt. \end{aligned} \tag{21}$$

The main problem with this inhomogeneous part is – as we will show in a moment – that the limit $H_k^{(3)}(-\infty)$ does not exist.

First we derive a proper representation for $H_k^{(3)}(s)$.

Lemma 1 *The function $H_k^{(3)}(s)$ can be represented by the sum*

$$H_k^{(3)}(s) = \sum_{\ell_1=0}^{k} \sum_{\ell_2=0}^{k} F_{k-\ell_1}^{(1)}(-\infty) F_{k-\ell_2}^{(1)}(-\infty) G_{\ell_1,\ell_2}(s), \tag{22}$$

where

$$\begin{aligned} G_{\ell_1,\ell_2}(s) &= \frac{1}{2\pi i\Gamma(s)} \int_{c-i\infty}^{c+i\infty} \Gamma(t)\Gamma(s-t)R_{\ell_1}(t-1)R_{\ell_2}(s-t-1)\,dt \qquad (23) \\ &= \sum_{i=0}^{\ell_1}\sum_{j=0}^{\ell_2} c_{\ell_1,i}c_{\ell_2,j}\binom{\ell_1}{i}\binom{\ell_2}{j}\left(p^i q^{\ell_1-i} + p^j q^{\ell_2-j}\right)^{-s+1} \end{aligned}$$

(with $c = \Re(s)/2$). Furthermore we have uniformly for $\Re(s) \in [a,b]$

$$G_{\ell_1,\ell_2}(s) = \mathcal{O}\left(T(\Re(s)/2-1)^{\ell_1+\ell_2}\right)$$

and

$$H_k^{(3)}(s) = \mathcal{O}\left(T(\Re(s)/2-1)^{2k}\right). \tag{24}$$

Proof: The representations (22) and (23) are immediate from (21) and from (19). We also can use the property that $(e^{-ax}e^{-bx})' = -(a+b)e^{-(a+b)x}$ and that its Mellin transform is given by $\Gamma(s)(a+b)^{-s+1}$.

Next, we use the estimate $|F_k^{(1)}(s)| \le CT(\Re(s))^k$ to obtain (with $c = \Re(s)/2$)

$$\begin{aligned} \left|G_{\ell_1,\ell_2}(s)\right| &\le \frac{C^2}{2\pi i|\Gamma(s)|}\int_{c-i\infty}^{c+i\infty} |\Gamma(t)||\Gamma(s-t)|\, T(\Re(t)-1)^{\ell_1} T(\Re(s-t)-1)^{\ell_2} dt \\ &\le C'(s)\, T(\Re(s)/2-1)^{\ell_1+\ell_2} \end{aligned} \tag{25}$$

Finally by using (20) and the explicit representation (22) we also obtain (24). □

The next lemma is crucial for the solution of our problem.

Lemma 2 *There exists a function $D_0(s,w)$ such that*

$$D_0(s,w) - D_0(s-1,w) = \frac{1}{g(s-1,w)} \sum_{k\ge1} H_k^{(3)}(s)\, w^k \tag{26}$$

and that the mapping $w \mapsto D_0(s,w)$ is analytic for $|w| < 1/T(\Re(s)/2-1)^2$.

Proof: Let $L_k(s)$ be defined by

$$\sum_{k\ge0} L_k(s) w^k = \frac{1}{g(s-1,w)} = \prod_{j\ge1} \left(1 - wT(s-j)\right).$$

By definition it is clear that $L_k(s)$ is a linear combination of terms of the form $(p^{k_1}q^{k_2})^{-s}$, where all coefficients have the same sign. Furthermore, since $1/g(s,w)$ is an entire function, we have $L_k(s) = \mathcal{O}(\eta^k)$ for every (fixed) $\eta > 0$.

We will show that for every $k \ge 1$ there exists a function $D_k(s)$ with

$$\begin{aligned} D_k(s) - D_k(s-1) &= \sum_{m=1}^{k} L_{k-m}(s) H_m^{(3)}(s) \\ &= \sum_{m=1}^{k} L_{k-m}(s) \sum_{\ell_1=0}^{m}\sum_{\ell_2=0}^{m} F_{m-\ell_1}^{(1)}(-\infty) F_{m-\ell_2}^{(1)}(-\infty) G_{\ell_1,\ell_2}(s) \end{aligned} \tag{27}$$

and with an upper bound of the form

$$D_k(s) = \mathcal{O}\left(\min\{1,|s|\}\, T(\Re(s)/2-1)^{2k}\right). \tag{28}$$

By Lemma 1 and by the definition of $L_k(s)$ it follows that $L_{k-m}(s)G_{\ell_1,\ell_2}(s)$ is a linear combination of functions of the form a^{-s} (for certain positive numbers a), where all coefficients have the same sign. Furthermore, there is only a bounded number of real numbers a appearing there with $a \ge 1$. Recall that a is of the form

$$p^{k_1}q^{k_2}\left(p^i q^{\ell_1-i} + p^j q^{\ell_2-j}\right).$$

This bound is also independent of k, m, ℓ_1, ℓ_2. Next observe that

$$a^{-s} = \frac{a^{-s}}{1-a} - \frac{a^{-(s-1)}}{1-a}$$

for $a \neq 1$ and that

$$a^{-s} = 1 = s - (s-1)$$

for $a = 1$. Hence, each term of the form a^{-s} can be written as a difference of the form $d(s) - d(s-1)$. Of course, this implies the existence of a function $D_k(s)$ satisfying (27).

Let us consider one of the terms $L_{k-m}(s)G_{\ell_1,\ell_2}(s)$. Recall that we already know that $L_{k-m}(s)G_{\ell_1,\ell_2}(s) = \mathcal{O}(\eta^{k-m}T(\Re(s)/2-1)^{\ell_1+\ell_2})$, where $\eta > 0$ is arbitrary. We split it up into a sum

$$L_{k-m}(s)G_{\ell_1,\ell_2}(s) = A(s) + B(s),$$

where $A(s)$ contains all terms of the form a^{-s} with $a < 1$ and $B(s)$ the remaining ones. Since all coefficients have the same sign, $A(s)$ and $B(s)$ can be bounded from the above as $A(s) + B(s)$. By the above observation we can represent $A(s)$ as $A(s) = d_1(s) - d_1(s-1)$, where

$$d_1(s) = \mathcal{O}(A(\Re(s))) = \mathcal{O}(\eta^{k-m}T(\Re(s)/2-1)^{\ell_1+\ell_2}).$$

Similarly we have $B(s) = d_2(s) - d_2(s-1)$ with

$$d_2(s) = \mathcal{O}(B(\Re(s))) = \mathcal{O}(\min\{1,|s|\}\,\eta^{k-m}T(\Re(s)/2-1)^{\ell_1+\ell_2}).$$

Note also that the constants that are implied by the $\mathcal{O}$-notation are uniform, they only depend on p and q. Putting all these estimates together we immediately deduce (28).

Finally, we set

$$D_0(s,w) = \sum_{k\geq 0} D_k(s)w^k$$

which is analytic for $|w| < 1/T(\Re(s)/2-1)^2$. □

We are now ready to state and prove a proper representation for $f^{(3)}(s,w)$.

Lemma 3 *The function $f^{(3)}(s,w)$ can be represented as*

$$f^{(3)}(s,w) = D(s,w)g(s,w),$$

where $D(s,w)$ is analytic for $|w| < 1/T(\Re(s)/2-1)^2$.

We recall that the dominating singularity of $g(s,w)$ is $w = 1/T(s)$. Since we have the relation

$$T\left(\frac{\sigma}{2}-1\right)^2 < T(\sigma)$$

the function $D(s,w)$ is analytic in a region containing this singularity. Hence $g(s,w)$ dominates the behaviour of $f^{(3)}(s,w)$ as was the case with $f^{(1)}(s,w)$. Therefore we can expect that $F_k^{(3)}(s)$ behaves similarly as $F_k^{(1)}(s)$ which turns out to be true, see Section 3.

Proof: From (11) it follows that $f^{(3)}(s,w)$ satisfies

$$f^{(3)}(s,w)(1-wT(s)) = f^{(3)}(s-1,w) + h(s,w),$$

where
$$h(s,w) = \sum_{k\geq 1} H_k^{(3)}(s) w^k.$$

Setting $D(s,w) = f^{(3)}(s,w)/g(s,w)$ and by using the relation $g(s,w) = g(s-1,w)/(1-wT(s))$ it follows that
$$D(s,w) - D(s-1,w) = \frac{h(s,w)}{g(s-1,w)}.$$

Hence, $D(s,w)$ and $D_0(s,w)$ differ by a function $C(s,w)$ that is periodic in s (with period 1). However, since $\Delta_k^{(1)}(x)$ and $\Delta_k^{(2)}(x)$ are linear combinations of functions of the form $e^{-(p^i q^j)x}$ it follows that the Mellin transform of $V_k(x) = \Delta_k^{(2)}(x) - \left(\Delta_k^{(1)}(x)\right)^2$ has no periodic parts. Consequently, we have

$$D(s,w) = D_0(s,w) + K(w)$$

for some function $K(w)$.

Next we observe that for every non-negative integer m the function $f^{(3)}(-m,w)$ is a polynomial in w. This follows from the fact that $F_k^{(3)}(-m) = 0$ for $k > m$. Thus,

$$f^{(3)}(s,w) = \left(D_0(s,w) - D_0(-m,w) + \frac{f^{(3)}(-m,w)}{g(-m,w)}\right) g(s,w).$$

We now choose $-m \leq \Re(s)$. Then this representation implies that $D(s,w)$ is analytic for $|w| < 1/T(\Re(s)/2-1)^2$. □

3 Asymptotic Analysis

We now prove Theorem 1 by establishing the asymptotic behavior of the variance of the profiles. The plan of the proof is as follows. First we concentrate on the singularity analysis of $f^{(3)}(s,w)$ and obtain an asymptotic expansion for $F_k^{(3)}(s)$. Second we invert the Mellin transform $V_k^*(s)$ by a proper saddle point method of the dominant part $T(s)^k x^{-s} = \exp\Big(k \log T(s) - s\log x\Big)$ of the integrand in the inverse Mellin integral. Finally we show $\mathbb{V}ar(B_{n,k}) \sim V_k(n)$ by standard depoissonization methods [18].

3.1 Singularity Analysis of $f^{(3)}(s,w)$

In order to obtain asymptotic information for $F_k^{(3)}(s)$ we will analyze the generating function $f^{(3)}(s,w)$ that by Lemma 3 is given by
$$f^{(3)}(s,w) = D(s,w)g(s,w).$$

Since we will apply the inverse Mellin transform to $V_k^*(s) = \Gamma(s)F_k^{(3)}(s)$ we have to know the behaviour in a strip $\Re(s) \in [a,b]$ for given real numbers a, b.

Lemma 4 *For every real interval $[a,b]$ there exist k_0, $\gamma > 0$ and $\varepsilon > 0$ such that*

$$F_k^{(3)}(s) = f(s)T(s)^k \left(1 + \mathcal{O}\left(e^{-\gamma k}\right)\right) \tag{29}$$

uniformly for all s with $\Re(s) \in [a,b]$, $|\Im(s) - 2\ell\pi \log(q/p)| \le \varepsilon$ for some integer ℓ and $k \ge k_0$, where

$$f(s) = D(s, 1/T(s))g(s-1, 1/T(s))$$

is an analytic function that satisfies $f(-r) = 0$ for $r = 1, 2, \ldots$ and is bounded in this region.

Furthermore, if $|\Im(s) - 2\ell\pi \log(q/p)| > \varepsilon$ for for all integers ℓ then we have

$$F_k^{(3)}(s) = \mathcal{O}\left(T(\Re(s))^k \, e^{-\gamma k}\right). \tag{30}$$

uniformly for $\Re(s) \in [a,b]$.

Proof: Since $D(s,w)$ is analytic for $|w| < T(\Re(s)/2-1)^{-2}$ it is also analytic for $|w| < \Big(T(\Re(s))-\eta\Big)^{-1}$ for some $\eta > 0$ that can be chosen to be uniform for $\Re(s) \in [a,b]$.

Furthermore, by the representation (17) it is clear that $w = 1/T(s)$ is the dominant (and polar) singularity of $g(s,w)$ if s is sufficiently close to the real axis, that is, $|\Im(s)| \le \varepsilon$ for some $\varepsilon > 0$ that can be be chosen to be uniform for $\Re(s) \in [a,b]$. Since

$$T(s + 2\ell\pi i \log(q/p)) = e^{-2\pi i \log(p)/\log(q/p)}\, T(s).$$

this is also true if $|\Im(s) - 2\ell\pi \log(q/p)| \le \varepsilon$ for some integer ℓ.

Next we apply Cauchy's formula for a contour of integration on the circle $|w| = e^\gamma/T(s)$ (for some sufficiently small $\gamma > 0$) and the residue theorem it follows that

$$\begin{aligned}
[w^k]f^{(3)}(s,w) &= \frac{1}{2\pi i}\int_{|w|=\, e^\gamma/T(s)} f^{(3)}(s,w)\frac{dw}{w^{k+1}} \\
&= -\mathrm{Res}\left[f^{(3)}(s,w)w^{-k-1}; w = \frac{1}{T(s)}\right] + \mathcal{O}\left(|T(s)e^{-\gamma}|^k\right) \\
&= f(s)T(s)^k + \mathcal{O}\left(|T(s)e^{-\gamma}|^k\right),
\end{aligned}$$

where

$$f(s) = D(s, 1/T(s))g(s-1, 1/T(s))$$

These estimates are uniform for s contained in a compact interval $[a,b]$. However, note that $f(-r) = 0$ for non-negative integer r, since $F_k^{(3)}(-r) = 0$. Hence, if the interval $[a,b]$ contains a non-positive integer then we multiply by $\Gamma(s)$ and do completely the same calulations which give (after all) the expansion (29).

Finally, if $|\Im(s) - 2\pi i\ell/\log(q/p))| > \varepsilon$ for some integer ℓ, then there exists $\gamma > 0$ such that $|T(s)| < e^{-2\gamma}|T(\Re(s)|$. Hence it follows that $f^{(3)}(s,w)$ is regular for $|w| < e^{2\gamma}/T(\Re(s))$. Thus, for a contour of integration on the circle $|w| = e^\gamma/T(\Re(s))$ in Cauchy's formula we obtain

$$[w^k]f^{(3)}(s,w) = \mathcal{O}\left(T(\Re(s))^k \, e^{-\gamma k}\right).$$

This completes the proof of Theorem 1. □

3.2 Saddle Point Analysis

By the discussion of Lemma 4, we see that $F_k^{(3)}(s)$ and consequently $V_k^*(s)$ behave asymptotically as $T(s)^k$ like $F_k^{(1)}(s)$ and $\Delta_k^{*(1)}(s)$. Thus we are in the same situation as in the analysis of the average profile presented in [3]. Our asymptotic analysis will be therefore similar to that of [3] with a new function $A(s)$ (here called $f(s)$) that is also equal to zero for $-1, -2, -3, \ldots$.

We start with a very short outline of the proof. By applying the inverse Mellin transform in case $x = n$,

$$V_k(n) = \frac{1}{2\pi i}\int_{\rho-i\infty}^{\rho+i\infty} V_k^*(s)n^{-s}ds. \tag{31}$$

it is natural to choose $\rho = \rho_{n,k}$ as the saddle point of the function

$$T(s)^k n^{-s} = \exp\Big(k\log T(s) - s\log n\Big).$$

Note also that on the line $\Re(s) = \rho$ there will be infinitely many saddle points

$$s_j = \rho + \frac{2\pi i j}{\log\frac{p}{q}}$$

since $T(s_j) = e^{-2\pi i j(\log p)/(\log p/q)}T(\rho)$, consequently the behaviour of $T(s)^k z^{-s}$ around $s = s_j$ is almost the same that of $T(s)^k z^{-s}$ around $s = \rho$. This phenomenon gives a periodic leading factor in the asymptotics of $V_k(x)$. By Lemma 4, we can safely replace $F_k^{(3)}(s)$ by $f(s)T(s)^k$ since the error term is of order $\mathcal{O}(|T(s)|^k e^{-\gamma k})$ and leads to an exponentially small contribution compared to the asymptotic leading term.

We describe the following lemma for the Poisson variance $V_k(n)$ and in the next section we will show that $\mathbb{V}ar(B_{n,k}) \sim V_k(n)$ in the given range.

Lemma 5 *Let $V_k(x)$ denote the Poisson variance of profile in unbalanced digital search trees with underlying probabilities $0 < p < q$. Let k and n be positive integers such that $k/\log n$ satisfies $(\log\frac{1}{p})^{-1} + \varepsilon < k/\log n < (\log\frac{1}{q})^{-1} - \varepsilon$. Then uniformly*

$$V_k(n) = L\left(\rho_{n,k}, \log_{p/q} p^k n\right)\frac{(p^{-\rho_{n,k}} + q^{-\rho_{n,k}})^k n^{-\rho_{n,k}}}{\sqrt{2\pi\beta(\rho_{n,k})k}}\left(1 + \mathcal{O}\left(\frac{1}{\sqrt{k}}\right)\right),$$

where the periodic function $L(\rho, x)$ has the following representation

$$L(\rho, x) = \sum_{j\in Z} f(\rho + it_j)\Gamma(\rho + it_j)e^{-2j\pi i x}, \quad t_j = 2j\pi/\log(p/q). \tag{32}$$

Furthermore, for $x = ne^{i\theta}$, where $|\theta| \le \pi/2 - \varepsilon$ (for some $\varepsilon > 0$), we still have uniformly

$$\begin{aligned} V_k(ne^{i\theta}) &= \frac{T(\rho_{n,k})^k}{\sqrt{2\pi\beta(\rho_{n,k})k}}\sum_{|j|\le j_0} f(\rho_{n,k} + it_j)\Gamma(\rho_{n,k} + it_j)(ne^{i\theta})^{-\rho_{n,k} - it_j}p^{-ikt_j} \\ &\times \left(1 + \mathcal{O}\Big(\frac{1}{\sqrt{k}}\Big)\right), \end{aligned} \tag{33}$$

where j_0 abbreviates $j_0 = \sqrt{\log n}\log(p/q)/(2\pi)$.

Proof: The proof is quite identical to that of Lemma 5 in [3] for a new function $f(s)$, see [3] in detail. □

3.3 Proof of Theorem 1

In this section, we show that the Poisson variance is asymptotically equal to the variance of the profile. i.e., $V_k(n) \sim \mathbb{V}ar(B_{n,k})$. For this goal, we use the same technique of [17, 18]. First we prove the following Lemma that is necessary for proving the our main result.

Lemma 6 *Let $f(x) := \sum_{n\geq 0} a_n x^n/n!$ be an entire function, where a_n is a given sequence, and $\widetilde{f}(x) := e^{-x}f(x)$. If $x = re^{i\theta}$ and*

$$|f(x)| \leq c_1\sqrt{\log r}\, f(r)e^{-rc_2\theta^2}$$

holds uniformly for $r \geq 1$, $c_1, c_2 > 0$, and $|\theta| \leq \pi$, where $f(r) \geq 0$, and

$$\widetilde{f}^{(\ell)}(ne^{i\theta}) = \mathcal{O}\Big(\delta(n)^\ell \widetilde{f}(n)\Big), \qquad \ell = 0, 1, 2, ..., \tag{34}$$

uniformly for $|\theta| \leq \theta_1$, where $\theta_1 \geq n^{-1/2+\varepsilon}$ and $\delta(n) = o(n^{-1/2})$, then for any $\ell_0 \geq 2$,

$$a_n = \sum_{0\leq \ell < \ell_0} \frac{\widetilde{f}^{(\ell)}(n)}{\ell!}\tau_\ell(n) + \mathcal{O}\Big(n^{\ell_0/2}\delta(n)^{\ell_0}\widetilde{f}(n)\Big), \tag{35}$$

where $\tau_\ell(n) = \ell![x^\ell](1+x)^n e^{-nx}$. In particular

$$a_n = \widetilde{f}(n) - \frac{n}{2}\widetilde{f}\,''(n) + \mathcal{O}\Big(n^2\delta^4(n)\widetilde{f}(n)\Big).$$

Proof: By the Cauchy formula and upper bound of $f(x)$,

$$a_n = \frac{n!}{2\pi i}\int_{|x|=n, |\arg(x)|\leq n^{-2/5}} x^{-n-1}e^x\widetilde{f}(x)dx + \mathcal{O}\Big(n!n^{-n}\sqrt{\log n}\, f(n)\int_{n^{-2/5}}^{\infty} e^{-c_2 n\theta^2}d\theta\Big).$$

By Stirling's formula the $\mathcal{O}$-term is equal to

$$\mathcal{O}\Big(n^{1/2}\sqrt{\log n}\,\widetilde{f}(n)n^{-1/2}e^{-c_2 n^{1/5}}\Big) = \mathcal{O}\Big(\sqrt{\log n}e^{-c_2 n^{1/5}}\widetilde{f}(n)\Big),$$

which is negligible in comparison to the main term $\widetilde{f}(n)$. The proof is completed similarly to [18, Proposition 1]. □

Proof: (Theorem 1) By [3, Lemma 6],

$$|e^x\Delta_k^{(1)}(x)| \leq c_1 e^r\Delta_k^{(1)}(r)\sqrt{\log r}e^{-rc_2\theta^2}$$

for $r \geq 1$, $c_1, c_2 > 0$, and $|\theta| \leq \pi$. We can obtain such upper bound for $|e^x V_k(x)|$. By [18, Lemma 4] and Lemma 5 (above) for $\delta(n) = \rho/n$,

$$\Delta_k^{(1)(\ell)}(ne^{i\theta}) = \mathcal{O}\Big(\rho^\ell n^{-\ell}\Delta_k^{(1)}(n)\Big), \qquad V_k^{(\ell)}(ne^{i\theta}) = \mathcal{O}\Big(\rho^\ell n^{-\ell}V_k(n)\Big).$$

Since $\Delta_k^{(1)}(x)$ satisfies the above estimate, we have

$$\frac{\partial^\ell}{\partial x^\ell}\Big(\Delta_k^{(1)}(x)\Big)^2\Big|_{x=ne^{i\theta}} = \mathcal{O}\left(\rho^\ell n^{-\ell}\Big(\Delta_k^{(1)}(n)\Big)^2\right), \qquad \ell = 0, 1, 2, \ldots . \tag{36}$$

Thus $\Delta_k^{(2)}(x) = V_k(x) + \Big(\Delta_k^{(1)}(x)\Big)^2$ also satisfies conditions of Lemma 6 and

$$\mathbb{E}(B_{n,k}^2) = \Delta_k^{(2)}(n) - \frac{n}{2}\Delta_k^{(2)\,\prime\prime}(n) + \mathcal{O}\Big(\rho^4 n^{-2}\Delta_k^{(2)}(n)\Big).$$

Also

$$\begin{aligned}\Big(\mathbb{E}(B_{n,k})\Big)^2 &= \Big(\Delta_k^{(1)}(n) - \frac{n}{2}\Delta_k^{(1)\,\prime\prime}(n) + \mathcal{O}\Big(\rho^4 n^{-2}\Delta_k^{(1)}(n)\Big)\Big)^2 \\ &= \Big(\Delta_k^{(1)}(n)\Big)^2 - n\Delta_k^{(1)\,\prime\prime}(n)\Delta_k^{(1)}(n) + \mathcal{O}\Big(\rho^4 n^{-2}(\Delta_k^{(1)}(n))^2\Big).\end{aligned}$$

Therefore

$$\mathbb{V}ar(B_{n,k}) = \mathbb{E}(B_{n,k}^2) - \Big(\mathbb{E}(B_{n,k})\Big)^2 = V_k(n)\left(1 + \mathcal{O}\Big(\rho^2 n^{-1}\mathbb{E}(B_{n,k})\Big)\right),$$

namely, $\mathbb{V}ar(B_{n,k}) \sim V_k(n)$. □

Acknowledgements

The first author would like to thank Professor Michael Drmota for suggesting this topic of research. He had the pleasure of meeting him in Vienna. The authors would like to thank the anonymous referees for their valuable comments leading to the improvement of our manuscript.

References

[1] J. E. Coffman and J. Eve, File structures using hashing functions, *Communications of the ACM*, Vol. 13, 427-432, 1970.

[2] M. Drmota, *Random Trees, An Interplay Between Combinatorics and Probability*, Springer, Wien-New York, 2009.

[3] M. Drmota and W. Szpankowski, The Expected Profile of Digital Search Trees, *Journal of Combinatorial Theory, Series A*, Vol 118, 1939-1965, 2011.

[4] P. Flajolet, X. Gourdon, and P. Dumas, Mellin Transforms and Asymptotics: *harmonic sums, Theoret. Comput. Sci*. 144, 3-58, 1995

[5] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2008.

[6] P. Jacquet and W. Szpankowski, Analysis of Digital Tries with Markovian Dependency, IEEE Trans. *Information Theory* 37, 1470-1475, 1991.

[7] P. Jacquet, and W. Szpankowski, Asymptotic Behavior of the Lempel-Ziv Parsing Scheme and Digital Search Trees, *Theoretical Computer Science*, 144, 161-197, 1995.

[8] P. Jacquet, and W. Szpankowski, Analytical Depoissonization and Its Applications, *Theoretical Computer Science*, 201, 1-62, 1998.

[9] P. Jacquet and M. RÂ´egnier, Trie Partitioning Process: Limiting Distributions, Lecture Notes in Comput. Sci. 214, 196-210, Springer, Berlin, 1986.

[10] C. Knessl, and W. Szpankowski, Asymptotic Behavior of the Height in a Digital Search Tree and the Longest Phrase of the Lempel-Ziv Scheme, SIAM *J. Computing*, 30, 923-964, 2000.

[11] C. Knessl and W. Szpankowski, On the Average Profile of Symmetric Digital Search Trees, *Analytic Combinatorics*, 4, article 6, 2009.

[12] D. Knuth, *The Art of Computer Programming. Sorting and Searching*, Vol. 3, Second Edition, Addison-Wesley, Reading, MA, 1998.

[13] M. Lothaire, *Applied Combinatorics on Words*. Cambridge University Press, New York, 2005

[14] G. Louchard, Exact and Asymptotic Distributions in Digital and Binary Search Trees, RAIRO *Theoretical Inform. Applications*, 21, 479-495, 1987.

[15] G. Louchard and W. Szpankowski, Average Profile and Limiting Distribution for a Phrase Size in the Lempel-Ziv Parsing Algorithm, IEEE Trans. *Information Theory*, 41, 478-488, 1995.

[16] H. Mahmoud, Evolution of Random Search Trees, John Wiley and Sons Inc., New York, 1992.

[17] G. Park, Profile of Tries, Ph.D. Thesis, Purdue University, 2006.

[18] G. Park, H. K. Hwang, P. Nicodeme, and W. Szpankowski, Profile of Tries, *SIAM J. Computing*, 8, 1821-1880, 2009.

[19] B. Pittel, Asymptotic Growth of a Class of Random Trees, *Annals of Probability*, 18, 414-427, 1985.

[20] H. Prodinger, Digital Search Trees and Basic Hypergeometric Functions, *Bulletin of the EATCS*, 56, 1995.

[21] R. Sedgewick, *Algorithms in C: Fundamental Algorithms, Data Structures, Sorting, Searching*, Addison-Weseley, 1997.

[22] W. Szpankowski, A Characterization of Digital Search Trees From the Successful Search Viewpoint, *Theoretical Computer Science*, 85, 117-134, 1991.

[23] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, John Wiley, New York, 2001.