

# Project: Probabilistic Analysis of Random Digital Trees

by

Michael Fuchs

## 1 General

This is the final report on the National Science Council project “Probabilistic Analysis of Random Digital Trees” with grant number NSC-99-2115-M-009-007-MY2.

We first summarize the main achievements of the project.

- We applied our approach from the project proposal to various cost measures in random symmetric digital search trees (see [7]).
- We applied our approach to partial match retrievals in  $k$ -dimensional bucket digital trees (see [4]).
- We applied our approach to approximate counting (see [6]).
- We applied our approach to the Wiener index in random digital trees (see [5]).
- We proposed a general framework to the variance of linear cost measures in tries and PATRICIA tries (this is the work [8] which is still in progress).
- We presented some of our results at the Workshop on Analytic Algorithmics and Combinatorics (ANALCO11), San Francisco, USA and at the 23rd International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, Montreal, Canada.
- Four master students and one Ph.-D. student were partially supported by the project (three master students already graduated).

## 2 Results

This project was concerned with probabilistic analysis of random digital trees. Before explaining our results, we briefly recall what digital trees are.

Consider  $n$  keys which are infinite 0-1 strings. The main three types of digital trees are constructed as follows.

- *Digital Search Trees*: The first key is placed into the root; all other keys are directed either to the left or right subtree according to whether their first bit is 0 or 1; subtrees are build recursively with the first bit from the keys removed.
- *Tries*: same construction principle as for digital search tree, but only leaves hold keys.
- *PATRICIA Tries*: same as for tries, but with one-way branching suppressed.

The above trees are still deterministic. In order to construct random trees, we have to fix a probability model. The model we considered in this project was the *Bernoulli model*, where we assumed that bits in the keys are generated independently with probability of generating 1 equal to  $p$  (consequently, the probability of generating 0 is  $q := 1 - p$ ). If  $p = 1/2$ , we called the tree *symmetric*, otherwise it was called *asymmetric*.

Now, we will discuss the main findings of the project.

## 2.1 Symmetric Digital Search Trees

In [7], we proposed a new approach to the variance of cost measures in symmetric digital search trees. Previous approaches were cumbersome and yielded very involved expressions for constants. Our approach greatly simplifies both the analysis and the final expressions. Roughly speaking, our approach consists of two major parts, namely, de-poissonization and a combination of Mellin and Laplace transform. The former is a classical tool in the analysis of algorithms; see [9]. However, we simplified the verification of the de-poissonization assumptions by proposing an admissibility notation and proving closure properties. As for the second part, the combination of Mellin and Laplace transform is entirely new in the analysis of algorithms. Previous methods used either Rice method or Euler transform combined with Mellin transform and singularity analysis. Our approach is computational easier and also gives simpler results.

The parameters to which our approach can be applied are so-called *additive cost measures*. We give some examples of parameters of this type which are analyzed in our paper.

- Internal path length: sum of distances of all nodes to the root.
- Peripheral path length: sum of sizes of subtrees rooted at the father of leaves.
- Number of leaves and more general patterns.
- Colless index: sum of absolute differences of sizes of left and right subtree over all nodes.

Moreover, our approach can also be applied to bucket digital search trees, where every node can hold up to  $b \geq 1$  keys. Here, there are two different notions of path length, namely, the sum of distances of all keys to the root (*key-wise path length*) and the sum of distances of all nodes to the root (*node-wise path length*). Interestingly, these two path length exhibit a quite different behavior; see our paper for details.

We finish this section by exemplifying our results. Therefore, we choose the number of leaves as guiding example. We use the notation  $X_n$  for the number of leaves in a random symmetric digital search tree of size  $n$ . The computation of the mean was a problem posed by Knuth which was solved by Flajolet and Sedgwick in [3]. They proved that, as  $n \rightarrow \infty$ ,

$$\mathbf{E}(X_n) = n(C_{mean} + P_1(\log_2 n)) + \mathcal{O}(n^{1/2})$$

with  $P_1(z)$  a one-periodic function with zero mean and computable Fourier coefficients and

$$C_{mean} = 1 + \sum_{k \geq 1} \frac{k}{Q_k 2^k} \sum_{1 \leq j \leq k} \frac{1}{2^j - 1} - \frac{1}{Q_\infty} \left( \frac{1}{\log 2} + \left( \sum_{k \geq 1} \frac{1}{2^k - 1} \right)^2 - \sum_{k \geq 1} \frac{1}{2^k - 1} \right),$$

where  $Q_k = \prod_{1 \leq j \leq k} (1 - 1/2^j)$  and  $Q_\infty = \lim_{k \rightarrow \infty} Q_k$ . Then, Kirschenhofer and Prodinger investigated the variance in [10] and obtained, as  $n \rightarrow \infty$ ,

$$\text{Var}(X_n) \sim n(C_{var} + P_2(\log_2 n))$$

with  $P_2(z)$  again a one-periodic function with zero mean. Moreover, their involved analysis yielded a very long and complicated expression for  $C_{var}$ . In our paper, we gave a much shorter proof of the above result. Another advantage of our approach is that the Fourier coefficients of  $P_2(z)$  are computable and our expression for  $C_{var}$  is much simpler. More precisely, our approach yields

$$C_{var} = \frac{1}{\log 2} \int_0^\infty \frac{s}{Q(-2s)} \int_0^\infty e^{-zs} \tilde{g}_2(z) dz ds,$$

where  $Q(s) = \prod_{j=1}^\infty (1 - z/2^j)$  and

$$\tilde{g}_2(z) = z \tilde{f}_1''(z)^2 + e^{-z} \left( 1 - e^{-z}(1+z) + 2z \tilde{f}_1'(z/2) - 4 \tilde{f}_1(z/2) \right)$$

with  $\tilde{f}_1(z) = e^{-z} \sum_{n \geq 0} \mathbf{E}(X_n) z^n / n!$ . Our paper also proposes a method of numerical evaluation of  $C_{var}$  and the first few digits are  $C_{var} = 0.034203 \dots$ .

## 2.2 $k$ -dimensional Bucket Digital Trees

We start by briefly explaining what  $k$ -dimensional bucket digital trees are. First,  $k$ -dimensional digital trees are build from keys which consist of  $k$  sequences of 0-1 sequences. The  $k$  sequences are suitable shuffled and then the digital tree is build as before. In the bucket version, nodes containing data can hold up to  $b \geq 1$  keys. Finally, the random model is also as before.

In [4], we considered partial match retrieval in  $k$ -dimensional bucket digital trees. Here, a partial match query  $\mathbf{q}$  is a  $k$ -tuple of specified and unspecified entries. This query is used to walk through the tree (starting from the root), where a specified entry means that we only proceed with one subtree, whereas an unspecified entry means that we proceed in both subtrees. The cost  $X_n$  is the number of nodes visited by such a partial match query. In [4], we derived the variance of  $X_n$  for partial match retrieval in bucket symmetric digital search trees as well as bucket symmetric tries and bucket symmetric PATRICIA tries. We briefly explain our results for symmetric digital search trees and refer the reader to [4] for the bucket version and other variants of digital trees.

First, Kirschenhofer and Prodinger derived the mean of  $X_n$  in [11]. They proved that, as  $n \rightarrow \infty$ ,

$$\mathbf{E}(X_n) \sim n^{u/k} P_1(\log_2 n^{1/k}),$$

where  $u$  is the number of unspecified entries in the partial match query and  $P_1(z)$  is a one-periodic function with computable Fourier coefficients. In [4], we derived the variance which satisfies, as  $n \rightarrow \infty$ ,

$$\text{Var}(X_n) \sim n^{u/k} P_2(\log_2 n^{1/k}),$$

where  $P_2(z)$  is again a one-periodic function with computable Fourier coefficients. As example, we state our expression for the 0-th Fourier coefficient of  $P_2(z)$  which is

$$\frac{1}{k(\log 2) Q_\infty \Gamma(1 + u/k)} \sum_{l=0}^{k-1} \delta_1 \dots \delta_l 2^{-ul/k} \sum_{j_1, j_2, j_3 \geq 0} \frac{(-1)^{j_1} \bar{\delta}_{\mathbf{q}^{(l)}, j_2} \bar{\delta}_{\mathbf{q}^{(l)}, j_3} 2^{-\binom{j_1}{2} + (1-u/k)j_1}}{2^{j_2+j_3} Q_{j_1} Q_{j_2} Q_{j_3}} \varphi(u/k; 2^{j_1-j_2} + 2^{j_1-j_3}),$$

where  $\mathbf{q}^{(l)}$  is the  $l$ -th cyclic shift of  $\mathbf{q}$ ,  $\delta_i = 1$  if the  $(i \bmod k)$ -th entry of  $\mathbf{q}$  is specified and 2 otherwise,

$$\bar{\delta}_{\mathbf{q}, j} = \sum_{l \geq 0} \frac{(-1)^l 2^{-\binom{l+1}{2}}}{Q_l} \prod_{h=1}^{l+j} \delta_h,$$

and

$$\varphi(\omega; x) = \begin{cases} \pi(x^\omega - 1)/(\sin(\pi\omega)(x - 1)), & \text{if } x \neq 1; \\ \pi\omega/\sin(\pi\omega), & \text{if } x = 1. \end{cases}$$

## 2.3 Approximate Counting

In [6], we gave a new analysis of approximate counting. Approximate counting, an algorithm proposed by Morris in [12], is used for counting within a certain error tolerance a huge amount of objects with very limited space. The algorithm is very easy to describe: let  $C_n$  be the content of the counter after “counting  $n$  objects”. Then, a random decision determines whether the counter is increased or not when “counting the  $n + 1$ -st object”. More precisely, we have

$$C_{n+1} = \begin{cases} C_n + 1, & \text{with probability } 2^{-C_n}; \\ C_n, & \text{with probability } 1 - 2^{-C_n}. \end{cases}$$

Flajolet [2] was the first to give a detailed analysis of approximate counting. He proved that, as  $n \rightarrow \infty$ ,

$$\mathbf{E}(C_n) \sim \log_2 n + C_{mean} + P_1(\log_2 n),$$

where  $P_1(z)$  is a one-periodic function with zero mean and computable Fourier coefficients and

$$C_{mean} = \frac{\gamma}{\log 2} + \frac{1}{2} - \sum_{l \geq 1} \frac{1}{2^l - 1},$$

where  $\gamma$  is Euler’s constant. Moreover, for the variance, he proved that, as  $n \rightarrow \infty$ ,

$$\text{Var}(C_n) \sim C_{var} + P_2(\log_2 n),$$

where  $P_2(z)$  is a one-periodic function with zero mean and computable Fourier coefficients and

$$C_{var} = \frac{\pi^2}{6 \log^2 2} - \sum_{l \geq 1} \frac{1}{2^l - 1} - \sum_{l \geq 1} \frac{1}{(2^l - 1)^2} + \frac{1}{12} - \frac{1}{\log 2} \sum_{l \geq 1} \frac{1}{l \sinh(2l\pi^2/\log 2)}.$$

After Flajolet’s analysis, at least another half a dozen papers re-proved Flajolet’s results using all kind of different methods (Martingale theory, Euler transform, Rice method, probabilistic tools, etc.). In [6], we showed that  $C_n$  is related to the length of the left-path in a random symmetric digital search trees. Using this correspondence and our approach from Section 2.1, we again re-proved the above results of Flajolet. However, a new feature of our method is that it yields a very different expression for the Fourier coefficients of  $P_2(z)$  and  $C_{var}$ . We exemplify our result by stating the expression for  $C_{var}$  (notation as in Section 2.1)

$$C_{var} = \frac{Q_\infty}{\log 2} \sum_{h,l,j \geq 0} \frac{(-1)^j 2^{-h-l-\binom{j+1}{2}}}{Q_h Q_l Q_j} \psi(2^{-h-j} + 2^{-l-j}),$$

where

$$\psi(x) = \begin{cases} \log x/(x - 1), & \text{if } x \neq 1; \\ 1, & \text{if } x = 1. \end{cases}$$

Equating with the above expression of Flajolet gives an identity for which we provided a direct proof in [6] with tools from  $q$ -analysis.

Finally, we also considered variants of approximate counting such as several versions of approximate counting with  $m$ -counters and analysis of shape parameters in  $m$ -DSTs for which our approach can be used as well. For detailed results see our paper.

## 2.4 Wiener Index in Random Digital Trees

In [5], we applied our approach from Section 2.1 to the analysis of the Wiener index in random digital trees. The Wiener index is one of the most important topological indices in combinatorial chemistry and was investigated in numerous papers (see [1] for many references).

As for random trees, the Wiener index was investigated for simply generated random trees, non-plane unlabeled random trees and a huge subclass of random grid trees containing random binary search trees, random median-of- $(2k + 1)$  search trees, random  $m$ -ary search trees, random quadrees, etc. An important class of random trees which was left open were random digital trees. In [5], we closed this gap. In particular, we showed that random digital trees behave very different from other random grid trees (to which random digital trees belong as a subclass).

In order to state our result, we consider the random symmetric digital tree. Denote by  $T_n$  the total path length (which was already investigated in [7]) and by  $W_n$  the Wiener index. Then, we proved in [5] that, as  $n \rightarrow \infty$ ,

$$\begin{aligned}\mathbf{E}(T_n) &= n \log_2 n + nP_1(\log_2 n) + \mathcal{O}(\log n), \\ \mathbf{E}(W_n) &= n^2 \log_2 n + n^2 P_1(\log_2 n) - n^2 + \mathcal{O}(n \log n),\end{aligned}$$

where  $P_1(z)$  is a one-periodic function with computable Fourier coefficients (see [7]). Moreover, we showed that for variances and covariances, as  $n \rightarrow \infty$ ,

$$\begin{aligned}\text{Var}(T_n) &= nP_2(\log_2 n) + \mathcal{O}(1), \\ \text{Cov}(T_n, W_n) &= n^2 P_2(\log_2 n) + \mathcal{O}(n \log n), \\ \text{Var}(W_n) &= n^3 P_2(\log_2 n) + \mathcal{O}(n^2 \log n),\end{aligned}$$

where  $P_2(z)$  is again a one-periodic function with computable Fourier coefficients (see [7]).

Using the above result for the variances and covariances, we also proved the following bivariate central limit theorem.

**Theorem 1.** *We have,*

$$\left( \frac{T_n - \mathbf{E}(T_n)}{\sqrt{\text{Var}(T_n)}}, \frac{W_n - \mathbf{E}(W_n)}{\sqrt{\text{Var}(W_n)}} \right) \xrightarrow{d} (X, X),$$

where  $X$  is a standard normal distributed random variable.

Moreover, we also obtained similar results for bucket digital search trees (where there are two notions of the Wiener index), tries (where there are again two notations of the Wiener index) and PATRICIA tries. See [5] for more details.

## 2.5 Tries and PATRICIA Tries

Our approach from Section 2.1. is also applicable to additive cost measures in tries and PATRICIA tries. Moreover, here the asymmetric version can be treated as well. In the work in progress [8], we have proposed a general framework to the variance of additive cost measures in random tries and PATRICIA tries. More precisely, for instance for tries, an additive cost measure can be interpreted as a sequence of random variables  $X_n$  satisfying the distributional recurrence

$$X_n \stackrel{d}{=} X_{I_n} + X_{n-I_n}^* + T_n,$$

where  $I_n = \text{Binomial}(n, p)$ ,  $X_n \stackrel{d}{=} X_n^*$  and  $X_n, X_n^*, I_n, T_n$  are independent. In [8], we have proved results which under some conditions on  $\mathbf{E}(T_n)$  and  $\mathbf{E}(T_n^2)$  give asymptotic expansions of the variance of  $X_n$ . Again, these asymptotic expansions involve periodic functions whose Fourier coefficients are computable. One new feature of this study compared to [7] is a new closure property of the Hadamard project which is needed for the depoissonization step.

We give an example to show the usefulness of our results. Let  $X_n$  be the number of internal nodes in tries (size of tries). Then, applying our results yields, as  $n \rightarrow \infty$ ,

$$\text{Var}(X_n) \sim \begin{cases} \frac{n}{-p \log p - q \log q} \sum_k G(-1 + \chi_k) n^{-\chi_k}, & \text{if } \log p / \log q \in \mathbb{Q}; \\ \frac{G(-1)}{-p \log p - q \log q} n, & \text{if } \log p / \log q \notin \mathbb{Q}, \end{cases}$$

where  $\chi_k = 2\rho k\pi i / \log p$  with  $\log p / \log q = \rho / \sigma$ ,  $\gcd(\rho, \sigma) = 1$  and  $G(\omega) = G_1(\omega) + H_1(\omega)$  with

$$G_1(\omega) = (\omega + 1)\Gamma(\omega) \left( 1 - \frac{\omega^2 + 4\omega + 8}{2^{\omega+3}} \right) + 2 \sum_{l \geq 1} \frac{(-1)^l l \Gamma(\omega + l + 1)}{(l + 1)! (1 - p^{l+1} - q^{l+1})} (p^{l+1} + q^{l+1}) (l(\omega + l + 1) - 1)$$

and

$$H_1(-1 + \chi_k) = pq \left( \sum_{l \geq 1} \frac{(-1)^{l-1} \Gamma(l + \chi_k + 1)}{(l - 1)! (1 - p^{1+l} - q^{1+l}) (1 - p^{1-l} - q^{1-l})} (p^l - q^l) (p^{-l} - q^{-l}) - \sum_{\kappa_l} \frac{\Gamma(\kappa_l + 1) \Gamma(-\kappa_l + \chi_k + 1)}{(1 - p^{1-\kappa_l} - q^{1-\kappa_l}) (p^{1+\kappa_l} \log p + q^{1+\kappa_l} \log q)} (p^{-\kappa_l} - q^{-\kappa_l}) (p^{\kappa_l} - q^{\kappa_l}) \right),$$

where  $\kappa_l$  are all zeros of  $1 - p^{1+\tau} - q^{1+\tau} = 0$  with  $\Re(\kappa_l) < 0$ .

### 3 Papers and Preprints

The following papers and preprints were (partially) written during the lifetime of this project (all of them are available from my webpage <http://www.math.nctu.edu.tw/~mfuchs/>).

1. H.-K. Hwang, M. Fuchs, V. Zacharovas (2010). Asymptotic variance of random symmetric digital search trees, *Discrete Math. Theor. Comput. Sci.*, **12**, 103–166.
2. M. Fuchs (2010). The variance for partial match retrievals in k-dimensional bucket digital trees, *Discrete Math. Theor. Comput. Sci. Proc.*, Proceedings of the 21st International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, 261–276.
3. M. Fuchs (2011). The subtree size profile of plane-oriented recursive trees, Proceedings of the Eighth Workshop on Analytic Algorithmics and Combinatorics (ANALCO), 85–92.
4. M. Fuchs (2011). A note on simultaneous Diophantine approximation in positive characteristic, *Acta Arith.*, **147**, 161–171.

5. C.-H. Chen and M. Fuchs (2011). On the moment-transfer approach for random variables satisfying a one-sided distributional recurrence, *Electron. J. Probab.*, **16**, 903–928.
6. M. Fuchs (2012). Limit theorems for subtree size profiles of increasing trees, *Combin. Probab. Comput.*, **21**, 412–441.
7. M. Fuchs, C.-K. Lee, H. Prodinger (2012). Approximate counting via the Poisson-Laplace-Mellin method, *Discrete Math. Theor. Comput. Sci. Proc.*, Proceedings of the 23rd International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, 13–28.
8. S.-Y. Chen and M. Fuchs. A higher-dimensional Kurzwil theorem for formal Laurent series over finite fields, 10 pages, *Finite Fields Appl.*, in press.
9. M. Fuchs and C.-K. Lee. The Wiener index of random digital trees, 28 pages, submitted.

The papers [2,3,7] are conference papers which were submitted and presented at the following meetings.

- 21st International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, Vienna, Austria ([2]).
- Workshop on Analytic Algorithmics and Combinatorics (ANALCO11), San Francisco, USA ([3]).
- 23rd International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, Montreal, Canada ([7]).

From these three papers, only [3] is an extended abstract (the journal version of this paper is [6]), whereas the other two are complete papers which will not appear elsewhere.

## 4 Students

Four master students and one Ph.-D. students were partially supported by the project.

- Shu-Yi Chen, master student, graduated on June 7, 2012. (Thesis: “Kurzwil’s Theorem in the Field of Formal Laurent Series”.)
- Yen-Ting Kuo, master student, graduated on June 27, 2012. (Thesis: “Limit Theorems for the Cost of Splitting Random Cayley Trees”.)
- Yi-Wen Lee, master student, graduated on June 27, 2012. (Thesis: “Group Patterns of Social Animals under the Neutral Model”.)
- Chi-Wei Chiang, master student, expected to graduate in November, 2012. (Thesis: “Random Polynomials over Finite Fields via Analytic Combinatorics”.)
- Chung-Kuei Lee, Ph.D. student, since 2011.

## 5 Conclusion and Future Work

In this project, we proposed a new approach to the asymptotic of the variance of cost measures in digital trees. Our approach was already outlined in the project proposal and it was claimed that this approach has many applications. We confirmed this claim by applying it to various problems such as cost measures in random digital trees, analysis of partial match retrievals in  $k$ -dimensional digital trees, approximate counting, analysis of the Wiener index in random digital trees, etc. Moreover, we presented our findings at international meetings. Finally, four master students and one Ph.-D. student were partially supported by this project (three master students already graduated).

As for future work, there are a lot of possible directions.

**Extension of the Wiener Index and Correlation in Random Trees.** In [5], we studied the Wiener index in random symmetric digital trees. Our approach also applies to asymmetric digital trees. Moreover, extensions of the Wiener index, such as the  $k$ -Steiner distance as well as other topological indices from combinatorial chemistry can be studied as well.

One main feature of our study of the Wiener index was the correlation between the Wiener index and the total path length which lead to a bivariate central limit theorem. Correlations of other parameters in random trees could be studied as well. For instance, an interesting example is the correlation between the space requirement of an  $m$ -ary search tree and its total path length. Since the limit law of the space requirement is known to be normal for  $m \leq 26$  and does not exist for  $m > 26$ , whereas the limit law of the total path length always exist, it is interesting to ask what happens with the bivariate limit law? We will investigate such and similar questions in one of our forthcoming projects.

**Other Cost Measures in Random Digital Trees.** Apart from the cost measures, we considered above, our approach can also be applied to many other cost measures. Some examples include the number of  $k$ -protected nodes, node-sorts and size of PATRICIA trie, etc.

**A General Framework for Logarithmic Cost Measures.** Most of the cost measures, we considered so far are either linear or superlinear. The only logarithmic cost measure, we considered is the left-path length in random digital search trees (which was needed for the analysis of approximate counting). Since there are many other logarithmic cost measures, it would be nice to have a more general framework like the one we proposed for tries and PATRICIA tries.

**From Asymptotic Expansions to Identities.** One of our main contribution in [7] was the use of “poissonized variances”. By suitable refining this notation, one does not only get asymptotic expansions but even identities (of an asymptotic nature). We have already worked out this for some easy examples, but the full usefulness of such identities still remains to be investigated.

**Limit Laws, Local Limit Theorems, Rates of Convergence.** In our original project proposal, we claimed that once the variance is well-understood, we plan to propose a general framework for deriving more refined results such as limit laws, local limit theorems, rates of convergence, etc. However, due to too many interesting problems related to the variance, we had no time to work on this second stage of the project yet. We will postpone such a study (which certainly is of great importance) to a future project.



## References

- [1] A. A. Dobrynin, R. Entinger, I. Gutman (2001). Wiener index of trees: Theory and applications, *Acta Appl. Math.*, **66**, 211–249.
- [2] P. Flajolet (1985). Approximate counting: a detailed analysis, *BIT*, **25**, 113–134.
- [3] P. Flajolet and R. Sedgewick (1986). Digital search trees revisited, *SIAM J. Comp.*, **15**, 748–767.
- [4] M. Fuchs (2010). The variance for partial match retrievals in k-dimensional bucket digital trees, *Discrete Math. Theor. Comput. Sci. Proc.*, Proceedings of the 21st International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, 261–276.
- [5] M. Fuchs and C.-K. Lee. The Wiener index of random digital trees, submitted.
- [6] M. Fuchs, C.-K. Lee, H. Prodinger (2012). Approximate counting via the Poisson-Laplace-Mellin method, *Discrete Math. Theor. Comput. Sci. Proc.*, Proceedings of the 23rd International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms, 13–28.
- [7] H.-K. Hwang, M. Fuchs, V. Zacharovas (2010). Asymptotic variance of random symmetric digital search trees, *Discrete Math. Theor. Comput. Sci.*, **12**, 103–166.
- [8] H.-K. Hwang, M. Fuchs, V. Zacharovas. Asymptotic variance of random tries and related structures, in preparation.
- [9] P. Jacquet and W. Szpankowski (1998). Analytic de-Poissonization and its applications, *Theor. Comput. Sci.*, **201**, 1–62.
- [10] P. Kirschenhofer and H. Prodinger (1988). Eine Anwendung der Theorie der Modulfunctionen in der Informatik, *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II*, **197**, 339–366.
- [11] P. Kirschenhofer and H. Prodinger (1994). Multidimensional digital searching - alternative data structures, *Random Struc. Algorithms*, **5**, 123–134.
- [12] R. Morris (1978). Counting large numbers of events in small registers, *Comm. ACM*, **21**, 840–842.