Information Processing Letters 114 (2014) 38–41

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Fitness levels with tail bounds for the analysis of randomized search heuristics

Carsten Witt *

DTU Compute, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 12 August 2013
Received in revised form 19 September 2013
Accepted 24 September 2013
Available online 1 October 2013
Communicated by B. Doerr

Keywords:
Randomized algorithms
Randomized search heuristics
Running time analysis
Fitness-level method
Tail bounds

ABSTRACT

The fitness-level method, also called the method of f-based partitions, is an intuitive and widely used technique for the running time analysis of randomized search heuristics. It was originally defined to prove upper and lower bounds on the expected running time. Recently, upper tail bounds were added to the technique; however, these tail bounds only apply to running times that are at least twice as large as the expectation.

We remove this restriction and supplement the fitness-level method with sharp tail bounds, including lower tails. As an exemplary application, we prove that the running time of randomized local search on OneMax is sharply concentrated around $n \ln n - 0.1159\ldots n$.

1. Introduction

The running time analysis of randomized search heuristics, including evolutionary algorithms, ant colony optimization and particle swarm optimization, is a vivid research area where many results have been obtained in the last 15 years. Different methods for the analysis were developed as the research area grew. For an overview of the state of the art in the area see the books by Auger and Doerr [1], Neumann and Witt [8] and Jansen [4].

The fitness-level method, also called the method of f-based partitions, is a classical and intuitive method for running time analysis, first formalized by Wegener [11]. It applies to the case that the total running time of a search heuristic can be represented as (or bounded by) a sum of geometrically distributed waiting times, where the waiting times account for the number of steps spent on certain levels of the search space. Wegener [11] presented both upper and lower bounds on the running time of randomized search heuristics using the fitness-level method. The lower bounds relied on the assumption that no level was allowed to be skipped. Sudholt [10] significantly relaxed this assumption and presented a very general lower-bound version of the fitness-level method that allows levels to be skipped with some probability.

Only recently, the focus in running time analysis turned to tail bounds, also called concentration inequalities. Zhou, Luo, Lu, Han [12] were the first to add tail bounds to the fitness-level method. Roughly speaking, they prove w.r.t. the running time T that $\Pr(T > 2E(T) + 2\delta h) = e^{-\delta}$ holds, where h is the worst-case expected waiting time over all fitness levels and $\delta > 0$ is arbitrary. An obvious open question was whether the factor 2 in front of the expected value could be "removed" from the tail bound, i.e., replaced with 1; Zhou et al. [12] only remark that the factor 2 can be replaced with 1.883.

In this article, we give a positive answer to this question and supplement the fitness-level method also with lower tail bounds. Roughly speaking, we prove in Section 2 that $\Pr(T \leqslant E(T) - \delta) \leqslant e^{-\delta^2/(2s)}$ and $\Pr(T \geqslant E(T) + \delta) \leqslant e^{-\frac{\delta}{4}\cdot\min\{\frac{\delta}{s}, h\}}$, where s is the sum of the squares of the waiting times over all fitness levels. We apply the technique to a classical benchmark problem, more precisely to the

* Tel.: +45 45253722.
E-mail address: cawi@dtu.dk.

0020-0190/$ – see front matter
http://dx.doi.org/10.1016/j.ipl.2013.09.013

running time analysis of randomized local search (RLS) on OneMax in Section 3, and prove a very sharp concentration of the running time around $n \ln n - 0.1159\ldots n$. We finish with some conclusions and a pointer to related work.

2. New tail bounds for fitness levels

Miscellaneous authors [6] on the Internet discussed tail bounds for a special case of our problem, namely the coupon collector problem (Motwani and Raghavan [7, Chapter 3.6]). Inspired by this discussion, we present our main result in Theorem 1 below. It applies to the scenario that a random variable (e.g., a running time) is given as a sum of geometrically distributed independent random variables (e.g., waiting times on fitness levels). A concrete application will be presented in Section 3.

Theorem 1. *Let X_i, $1 \leqslant i \leqslant n$, be independent random variables following the geometric distribution with success probability p_i, and let $X := \sum_{i=1}^{n} X_i$. If $\sum_{i=1}^{n}(1/p_i^2) \leqslant s < \infty$ then for any $\delta > 0$*

$$\Pr\big(X \leqslant E(X) - \delta\big) \leqslant e^{-\frac{\delta^2}{2s}}.$$

For $h := \min\{p_i \mid i = 1, \ldots, n\}$,

$$\Pr\big(X \geqslant E(X) + \delta\big) \leqslant e^{-\frac{\delta}{4}\cdot\min\{\frac{\delta}{s}, h\}}.$$

For the proof, the following two simple inequalities will be used.

Lemma 1.

1. *For $x \geqslant 0$ it holds $\frac{e^x}{1+x} \leqslant e^{x^2/2}$.*
2. *For $0 \leqslant x \leqslant 1$ it holds $\frac{e^{-x}}{1-x} \leqslant e^{x^2/(2-2x)}$.*

Proof. We start with the first inequality. The series representation of the exponential function yields

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!} \leqslant \sum_{i=0}^{\infty} (1+x)\frac{x^{2i}}{(2i)!}$$

since $x \geqslant 0$. Hence,

$$\frac{e^x}{1+x} \leqslant \sum_{i=0}^{\infty} \frac{x^{2i}}{(2i)!}.$$

Since $(2i)! \geqslant 2^i i!$, we get

$$\frac{e^x}{1+x} \leqslant \sum_{i=0}^{\infty} \frac{x^{2i}}{2^i i!} = e^{x^2/2}.$$

To prove the second inequality, we omit all negative terms except for $-x$ from the series representation of e^{-x} to get

$$\frac{e^{-x}}{1-x} \leqslant \frac{1 - x + \sum_{i=1}^{\infty}\frac{x^{2i}}{(2i)!}}{1-x} = 1 + \sum_{i=1}^{\infty} \frac{x^{2i}}{(1-x)\cdot(2i)!}.$$

For comparison,

$$e^{x^2/(2-2x)} = 1 + \sum_{i=1}^{\infty} \frac{x^{2i}}{2^i(1-x)^i i!},$$

which, as $x \leqslant 1$, is clearly not less than our estimate for $e^{-x}/(1-x)$. $\square$

Proof of Theorem 1. Both the lower and upper tail are analyzed similarly, using the exponential method (see, e.g., the proof of the Chernoff bound in Motwani and Raghavan [7, Chapter 3.6]). We start with the lower tail. Let $d := E(X) - \delta = \sum_{i=1}^{n}(1/p_i) - \delta$. Since for any $t \geqslant 0$

$$X \leqslant d \quad \Longleftrightarrow \quad -X \geqslant -d \quad \Longleftrightarrow \quad e^{-tX} \geqslant e^{-td},$$

Markov's inequality and the independence of the X_i yield that

$$\Pr(X \leqslant d) \leqslant \frac{E(e^{-tX})}{e^{-td}} = e^{td} \cdot \prod_{i=1}^{n} E\big(e^{-tX_i}\big).$$

Note that the last product involves the moment-generating functions (mgf's) of the X_i. Given a geometrically distributed random variable Y with parameter p, its moment-generating function at $r \in \mathbb{R}$ equals $E(e^{rY}) = \frac{pe^r}{1-e^r(1-p)} = \frac{1}{1-(1-e^{-r})/p}$ for $r < -\ln(1-p)$. We will only use negative values for r, which guarantees existence of the mgf's used in the following. Hence,

$$\Pr(X \leqslant d) \leqslant e^{td} \cdot \prod_{i=1}^{n} \frac{1}{1-(1-e^{t})/p_i} \leqslant e^{td} \cdot \prod_{i=1}^{n} \frac{1}{1+t/p_i},$$

where we have used $e^x \geqslant 1 + x$ for $x \in \mathbb{R}$. Now, by writing the numerators as $e^{t/p_i} \cdot e^{-t/p_i}$, using $\frac{e^x}{1+x} \leqslant e^{x^2/2}$ for $x \geqslant 0$ (Lemma 1) and finally plugging in d, we get

$$\Pr(X \leqslant d) \leqslant e^{td} \cdot \left(\prod_{i=1}^{n} e^{t^2/(2p_i^2)} e^{-t/p_i}\right)$$
$$= e^{td} e^{(t^2/2)\sum_{i=1}^{n}(1/p_i)^2} e^{-tE(X)} \leqslant e^{-t\delta + (t^2/2)s}.$$

The last exponent is minimized for $t = \delta/s$, which yields

$$\Pr(X \leqslant d) \leqslant e^{-\frac{\delta^2}{2s}}$$

and proves the lower tail inequality.

For the upper tail, we redefine $d := E(X) + \delta$ and obtain

$$\Pr(X \geqslant d) \leqslant \frac{E(e^{tX})}{e^{td}} = e^{-td} \cdot \prod_{i=1}^{n} E\big(e^{tX_i}\big).$$

Since now positive arguments will be used for the moment-generating functions of the X_i, their existence is not trivial. In the following, we assume $t \leqslant \min\{p_i \mid i = 1, \ldots, n\}/2 = h/2$. Since $p_i \leqslant -\ln(1-p_i)$, our assumption implies $t \leqslant -\ln(1-p_i)$ for any i. The factor $1/2$ will be crucial later.

Estimating the moment-generating functions similarly as above, we get

$$\Pr(X \geqslant d) \leqslant e^{-td} \cdot \left(\prod_{i=1}^{n} \frac{e^{-t/p_i}}{1-t/p_i} \cdot e^{t/p_i}\right).$$