

# Unbounded Lower Bound for $k$ -Server against Weak Adversaries

Marcin Bienkowski

marcin.bienkowski@cs.uni.wroc.pl

Institute of Computer Science, University of Wrocław  
Poland

Christian Coester

christian.coester@cwi.nl

CWI  
Netherlands

Jarosław Byrka

jaroslaw.byrka@cs.uni.wroc.pl

Institute of Computer Science, University of Wrocław  
Poland

Łukasz Jeż

lukasz.jez@cs.uni.wroc.pl

Institute of Computer Science, University of Wrocław  
Poland

## ABSTRACT

We study the resource augmented version of the  $k$ -server problem, also known as the  $k$ -server problem against weak adversaries or the  $(h, k)$ -server problem. In this setting, an online algorithm using  $k$  servers is compared to an offline algorithm using  $h$  servers, where  $h \leq k$ . For uniform metrics, it has been known since the seminal work of Sleator and Tarjan (1985) that for any  $\epsilon > 0$ , the competitive ratio drops to a constant if  $k = (1 + \epsilon) \cdot h$ . This result was later generalized to weighted stars (Young 1994) and trees of bounded depth (Bansal et al. 2017). The main open problem for this setting is whether a similar phenomenon occurs on general metrics.

We resolve this question negatively. With a simple recursive construction, we show that the competitive ratio is at least  $\Omega(\log \log h)$ , even as  $k \rightarrow \infty$ . Our lower bound holds for both deterministic and randomized algorithms. It also disproves the existence of a competitive algorithm for the infinite server problem on general metrics.

## CCS CONCEPTS

• Theory of computation  $\rightarrow$  K-server algorithms.

## KEYWORDS

online algorithms,  $k$ -server, weak adversaries, resource augmentation

### ACM Reference Format:

Marcin Bienkowski, Jarosław Byrka, Christian Coester, and Łukasz Jeż. 2020. Unbounded Lower Bound for  $k$ -Server against Weak Adversaries. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC '20)*, June 22–26, 2020, Chicago, IL, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3357713.3384306>

## 1 INTRODUCTION

The  $k$ -server problem is one of the most well-studied and influential online problems in competitive analysis, defined in 1990 by Manasse et al. [17]. It generalizes many problems in which an algorithm

has to maintain a feasible state while satisfying a sequence of requests. Formally, the  $k$ -server problem is defined as follows. There are  $k$  servers in a metric space  $(\mathcal{X}, d)$  and a sequence  $r_1, r_2, r_3, \dots$  of requests to metric space points appears online. In response to a request  $r_i$ , an algorithm has to move its servers, so that one of them ends at point  $r_i$ . The goal is to minimize the cost defined as the total distance traveled by all servers.

### 1.1 From Uniform to General Metrics

The definition of the  $k$ -server problem is deceptively simple, but it has led to substantial progress in many branches of competitive analysis. Historically, the results were obtained first for the case where  $\mathcal{X}$  is a uniform metric space: the  $k$ -server problem is then equivalent to the paging problem with a cache of size  $k$  [20]. In particular, the competitive ratio for paging is  $k$  for deterministic algorithms and there is a lower bound of  $k$  that holds for arbitrary metric spaces of more than  $k$  points [17]. This led to the bold  *$k$ -server conjecture* [17] stating that this ratio is  $k$  for all metric spaces. After series of papers proving the upper bound of  $k$  for particular metrics (e.g., trees or lines), the conjecture has been positively resolved (in the asymptotic sense) by the celebrated  $2k - 1$  upper bound due to Koutsoupias and Papadimitriou [15]. For a more thorough treatment of the history of deterministic approaches, see a survey by Koutsoupias [14].

Similarly, randomized competitive solutions for uniform metrics [1, 11, 18] showed that the achievable competitive ratio is exactly  $H_k = \Theta(\log k)$  and led to the analogous *randomized  $k$ -server conjecture*, stating that the randomized competitive ratio is  $\Theta(\log k)$  on arbitrary metrics. Some cornerstone results towards resolving this conjecture deserve closer attention. On the lower bound side, Bartal et al. [6] used Ramsey-type phenomena for metric spaces to show that the randomized competitive ratio is at least  $\Omega(\log k / \log \log k)$  for any metric space.<sup>1</sup> On the algorithmic side, a major breakthrough (building on a long line of results for particular metrics) was obtained by Bansal et al. [2], who constructed an algorithm of ratio poly-logarithmic in the number of metric space points, based on HST embeddings (hierarchically separated trees) and the so-called fractional allocation problem. It was recently improved by Bubeck et al. [8], who used mirror descent dynamics with multi-scale entropic regularization to obtain an  $O(\log^2 k)$ -competitive algorithm on HSTs and an  $O(\log^2 k \log n)$ -competitive

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
STOC '20, June 22–26, 2020, Chicago, IL, USA

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-6979-4/20/06...\$15.00  
<https://doi.org/10.1145/3357713.3384306>

<sup>1</sup>In the description of all lower bounds on the competitive ratio for the  $k$ -server problem, we silently assume that the metric space in question has more than  $k$  points.

algorithm on general  $n$ -point metrics. Based on this, Lee [16] proposed a dynamic embedding technique to achieve a competitive ratio poly-logarithmic in  $k$  on arbitrary metrics.

## 1.2 Weak Adversaries

A way to compensate for the online algorithm's lack of knowledge of the future is to assume that the algorithm has more "resources" than the offline optimum it is compared to. This natural concept, called *resource augmentation*, has led to spectacular successes for online scheduling problems (see e.g. [12, 19]). It can be a way to overcome pessimistic worst-case bounds of the original setting. In the context of the  $k$ -server problem, it is also known as the *weak adversaries* model [5, 13] or the  $(h, k)$ -server problem: an online algorithm with  $k$  servers is compared to an optimal algorithm (an adversary) with  $h \leq k$  servers. For a metric space  $X$ , let  $\mathcal{D}_X(h, k)$  and  $\mathcal{R}_X(h, k)$  denote the best competitive ratios of deterministic and randomized algorithms, respectively, for the  $(h, k)$ -server problem on  $X$ .

Again, the first results for the  $(h, k)$ -server problem were developed for uniform metrics: Sleator and Tarjan [20] gave an exact answer of  $\mathcal{D}_X(h, k) = k/(k - h + 1)$ , with the upper bound being achieved by the LRU (least recently used) paging strategy. This implies that having  $k = (1 + \epsilon) \cdot h$  servers suffices to attain a constant competitive ratio. It is natural to ask whether such phenomenon extends to other metrics. This question was raised already by Manasse et al [17] when they introduced the  $k$ -server problem.

Formally, we study the following questions.

**Strong  $(h, k)$ -server hypothesis:** For any metric space  $X$  and any  $\epsilon > 0$ ,  $\mathcal{D}_X(h, k) = O_\epsilon(1)$  whenever  $k \geq (1 + \epsilon) \cdot h$ .

**Weak  $(h, k)$ -server hypothesis:** For any metric space  $X$  and any  $h \in \mathbb{N}$ ,  $\mathcal{D}_X(h, k) = O(1)$  as  $k \rightarrow \infty$ .

Generalizing the result for uniform metrics, the same competitive ratio of  $k/(k - h + 1)$  was later also obtained for weighted star metrics [22]. More recently, Bansal et al. [4] confirmed the strong  $(h, k)$ -server hypothesis also for trees of bounded depth. Using randomization, tight bounds of  $\mathcal{R}_X(h, k) = \Theta(\log(1/\epsilon))$  were obtained for uniform metrics [21] and weighted stars [3] when  $k = (1 + \epsilon) \cdot h$ . The recent results by Bubeck et al. [8] and Buchbinder et al. [9] for the  $k$ -server problem extend also to the  $(h, k)$ -server setting, implying that  $\mathcal{R}_X(h, k) = O(D \cdot \log(1/\epsilon))$  for HSTs of depth  $D$  when  $k = (1 + \epsilon) \cdot h$ .<sup>2</sup>

Surprisingly, the performance of some classical algorithms can slightly degrade when additional online servers are available. Bansal et al. [4, 5] showed that this can occur for both the Work Function algorithm and the Double Coverage algorithm. On the positive side, Koutsoupias [13] showed that the Work Function algorithm obtains a competitive ratio of at most  $2h$  simultaneously for all  $h \leq k$ . The algorithm of [4] confirming the  $(h, k)$ -server hypothesis on bounded depth trees is actually a variant of the Double Coverage algorithm.

In [10], the infinite server problem (denoted  $\infty$ -server problem here) has been introduced as a possible way to resolve the question on general metrics. This is the variant of the  $k$ -server problem where  $k = \infty$ , and all infinitely many servers initially reside at the same

point of the metric space. The existence of an  $O(1)$ -competitive algorithm for the  $\infty$ -server problem was shown to be equivalent to an affirmative resolution of the weak  $(h, k)$ -server hypothesis.

In terms of lower bounds, it is known that — unlike in the case of uniform and weighted star metrics — the ratio  $\mathcal{D}_X(h, k)$  does *not* converge to 1 on general metrics even as  $k \rightarrow \infty$ . Namely, Bar-Noy and Schieber [7, page 175] showed that  $\mathcal{D}_X(2, k) = 2$  for all  $k$  when  $X$  is the line metric. For large  $h$ , the lower bound on  $\mathcal{D}_X(h, k)$  was improved to 2.41 [4] using depth-2 trees and later to 3.14 [10] by a reduction from the  $\infty$ -server problem. In the absence of any super-constant lower bounds, the  $(h, k)$ -server hypothesis continued to seem plausible. In fact, Bansal et al. [4] argued that it would be very surprising if  $\mathcal{D}_X(h, k) = \omega(1)$  (even for a sufficiently large  $k$ ).

## 1.3 Our Results

Our main result is the refutation of both versions of the  $(h, k)$ -server hypothesis:

**THEOREM 1.** *There exists a tree metric  $\mathcal{T}$  such that  $\mathcal{R}_{\mathcal{T}}(h, k) = \Omega(\log \log h)$ , even for arbitrarily large  $k$ .*

Since  $\mathcal{D}_{\mathcal{T}}(h, k) \geq \mathcal{R}_{\mathcal{T}}(h, k)$ , the lower bound obviously extends to deterministic algorithms. The underlying construction is simple. It is based on recursively combining Young's lower bound for randomized  $(h, k)$ -paging [21] along many scales. At higher scales, the construction is applied to groups of servers rather than individual servers.

Due to the connection between the  $(h, k)$ -server problem and the  $\infty$ -server problem [10], a direct consequence of Theorem 1 is that there is no competitive algorithm for the  $\infty$ -server problem on general metrics. In fact, we first found the lower bound by analyzing the  $\infty$ -server problem.

**COROLLARY 2.** *The competitive ratio of the  $\infty$ -server problem on trees of depth  $D$  is  $\Omega(\log D)$ . In particular, no algorithm for the  $\infty$ -server problem on general metrics has a constant competitive ratio.*

## 1.4 Preliminaries

An online algorithm  $\text{ALG}$  is called  $\rho$ -competitive if

$$\text{ALG}(\sigma) \leq \rho \cdot \text{OPT}(\sigma) + C$$

for all request sequences  $\sigma$ , where  $\text{ALG}(\sigma)$  and  $\text{OPT}(\sigma)$  denote the cost of  $\text{ALG}$  and the optimal (offline) cost for  $\sigma$ , respectively, and  $C \geq 0$  is a constant independent of  $\sigma$ . The competitive ratio of a problem is the infimum of all  $\rho$  such that a  $\rho$ -competitive algorithm exists. In the case of randomized algorithms,  $\text{ALG}(\sigma)$  is replaced by its expectation. Note that for the  $(h, k)$ -server problem,  $\text{OPT}$  denotes the optimal solution using  $h$  servers, while  $\text{ALG}$  uses  $k$  servers.

An algorithm is *fractional* if it is allowed to move an arbitrary fraction of a server, paying the same fractions of the distance moved, but it is still required to bring "a total mass" of at least one server to the requested point. A fractional algorithm can be derived from a randomized one by setting the server mass at each point to the

<sup>2</sup>For general trees of depth  $D$ , they obtain a *fractional* algorithm achieving the same competitive ratio.

expected number of servers; clearly, the cost of the fractional algorithm is at most the expected cost of the randomized algorithm.<sup>3</sup>

All metric spaces constructed in this paper are trees with a distinguished root, and we assume that servers reside initially at the root. We will charge cost (to both the online and offline algorithms) only for traversing edges in direction *away* from the root. Since movement away from the root is within a factor 2 of the total movement, the error due to this is absorbed in the asymptotic notation of our results.

For an infinite request sequence  $\sigma$ , we denote its prefix of the first  $m$  requests by  $\sigma_m$ .

## 2 PROOF OF THE LOWER BOUND

Below we state the main lemma towards proving Theorem 1.

**LEMMA 3.** *Fix arbitrary  $\rho \geq 1$ ,  $\delta > 0$  and an integer  $i \geq 0$ . Let  $b = \lceil \exp(3\rho) \rceil$ ,  $h_i = b^i$ ,  $k_i = b^i \cdot (1 + i/(2b))$ . There exists a tree  $T_i$  of depth  $i$  such that, for any fractional online  $k_i$ -server algorithm  $ALG$ , there exists an infinite request sequence  $\sigma$  on  $T_i$  satisfying two properties:*

- (a)  $ALG(\sigma_m) \geq \rho \cdot OPT_{h_i}(\sigma_m) - \delta$  for all  $m \in \mathbb{N}$ , where  $OPT_{h_i}(\sigma_m)$  denotes the optimal cost for serving  $\sigma_m$  using  $h_i$  servers.
- (b) If  $i \geq 1$ , then  $ALG(\sigma_m) \rightarrow \infty$  as  $m \rightarrow \infty$ .

**PROOF.** We prove the lemma by induction on  $i$ . For  $i = 0$ , the tree  $T_0$  is simply a single node, and all requests are given at this node. Clearly, the lemma holds here.

For the inductive step, we fix any  $i \geq 0$ . We will show that the lemma properties for  $i + 1$  hold for *some*  $\delta$ . By scaling all distances by a small multiplicative constant, this implies that  $\delta$  can be made arbitrarily close to 0, yielding the lemma statement for  $i + 1$  and arbitrary  $\delta$ .

Let  $T_i$  be the tree induced by the induction hypothesis for  $\delta = 1/2$ . The root of  $T_{i+1}$  has infinitely many children at distance 1; all the subtrees rooted at these children are copies of  $T_i$ . We will assume that the server mass in each subtree never exceeds  $k_i$ ; this assumption will be justified later. It allows us to invoke the induction hypothesis on the subtrees. If the mass inside a subtree is  $k_i - c$  for some  $c \geq 0$ , we interpret this as mass  $c$  sitting at the root of the subtrees. Note that the sub-algorithms for the different subtrees are not independent of each other, as a request in one subtree can trigger movement towards the root in another subtree. However, we construct the request sequence in an online manner where each request is independent of decisions of the algorithm for future requests, and thus we can analyze the sub-algorithms independently of each other.

Let  $\epsilon > 0$  be some small constant. The request sequence  $\sigma$  consists of phases numbered from 1. In each phase,  $b$  subtrees  $T_i$  will be marked and among them  $b - 1$  subtrees were marked in the previous phase. For this definition, we assume that right before phase 1, in an artificial phase 0 containing no requests,  $b$  arbitrary subtrees were marked. All phases proceed as follows:

- Mark a fresh subtree  $T_i$  that has never received any requests before.

<sup>3</sup>On weighted stars and HST metrics, the converse is also true: Any fractional algorithm can be rounded online to a randomized integral one while increasing its cost by at most a multiplicative constant [2, 3]. It is unknown whether this also holds for general metrics.

- While the server mass in the fresh subtree is at most  $k_i - \epsilon$ , issue requests in it according to the induction hypothesis.
- For  $j = 1, \dots, b - 1$ :
  - Among the subtrees that were marked in the last phase but have not been marked (yet) in the current phase, mark the one with the least server mass.
  - While there exists a subtree marked in the current phase where the server mass is at most  $k_i - \epsilon$ , issue requests in this subtree according to the induction hypothesis.

The request sequence satisfies Property (b): if  $i = 0$ , then each request incurs at least cost  $\epsilon$ , and if  $i \geq 1$  this follows by the induction hypothesis. We now prove that Property (a) also holds.

We compare  $ALG$  against an offline algorithm  $ADV$  with  $h_{i+1} = b^{i+1}$  servers that always has  $h_i = b^i$  servers at each marked subtree of the current phase, and uses servers optimally within the subtrees.

Consider some phase. Denote by  $ALG_\ell$  and  $ALG_{\leq \ell}$  the cost of  $ALG$  incurred during the phase along edges of level  $\ell$  and along edges of level at most  $\ell$ , respectively. We define  $ADV_\ell$  and  $ADV_{\leq \ell}$  analogously. Here, we use the convention that edges incident to the leaves have level 1 and edges incident to the root have level  $i + 1$ .

Consider the case that the phase under consideration is complete. We analyze first the cost along edges incident to the root.  $ALG$  pays at least  $k_i - \epsilon$  to move server mass  $k_i - \epsilon$  to the fresh subtree. At the beginning of iteration  $j$  of the for-loop,  $ALG$  has server mass at least  $k_i - \epsilon$  in each of the  $j$  subtrees that were marked during the current phase. Note that  $j - 1$  of them were marked in the previous phase. Thus, the average amount of server mass in the  $b - (j - 1)$  subtrees that were marked in the last phase but not yet in the current phase is at most  $(k_{i+1} - j \cdot (k_i - \epsilon)) / (b - j + 1)$ . In effect, the cost to move mass to the subtree that is marked in the  $j$ th iteration is at least

$$k_i - \epsilon - \frac{k_{i+1} - j \cdot (k_i - \epsilon)}{b - j + 1} = \frac{(k_i - \epsilon)(b + 1) - k_{i+1}}{b - j + 1}.$$

Hence, as  $\epsilon \rightarrow 0$ , the total cost of moving server mass to the marked subtrees of the phase is at least

$$\begin{aligned} ALG_{i+1} &\geq k_i - o(1) + \sum_{j=1}^{b-1} \frac{k_i(b+1) - k_{i+1}}{b-j+1} \\ &= b^i \cdot \left( 1 + \frac{i}{2b} + \sum_{j=1}^{b-1} \frac{\left(1 + \frac{i}{2b}\right)(b+1) - b \left(1 + \frac{i+1}{2b}\right)}{b-j+1} \right) - o(1) \\ &= b^i \cdot \left( \frac{1}{2} + \left(\frac{1}{2} + \frac{i}{2b}\right) + \sum_{j=1}^{b-1} \frac{\frac{1}{2} + \frac{i}{2b}}{b-j+1} \right) - o(1) \\ &\geq b^i \cdot \left( \frac{1}{2} + \frac{\ln b}{3} \right) \\ &\geq b^i \rho + b^i / 2. \end{aligned}$$

In contrast, the offline cost during the phase along edges incident to the root is only

$$ADV_{i+1} = b^i$$

because the offline algorithm moves only  $b^i$  servers from the subtree that was marked in the last but not in the current phase to the fresh subtree of the current phase.

For the cost within the subtrees, the induction hypothesis of (a) (applied to  $b$  marked subtrees with  $\delta = 1/2$ ) yields

$$\text{ALG}_{\leq i} \geq \rho \cdot \text{ADV}_{\leq i} - b/2.$$

Therefore, for the total cost during a complete phase, we obtain

$$\begin{aligned} \text{ALG}_{\leq i+1} &= \text{ALG}_{i+1} + \text{ALG}_{\leq i} \\ &\geq \rho \cdot \text{ADV}_{i+1} + b^i/2 + \rho \cdot \text{ADV}_{\leq i} - b/2 \\ &\geq \rho \cdot \text{ADV}_{\leq i+1}. \end{aligned}$$

In the last phase, which may be incomplete, we have

$$\begin{aligned} \text{ALG}_{\leq i+1} &\geq \text{ALG}_{\leq i} \\ &\geq \rho \cdot \text{ADV}_{\leq i} - b/2 \\ &\geq \rho \cdot \text{ADV}_{\leq i+1} - \rho b^i - b/2. \end{aligned}$$

Now if we set  $\delta'$  to be  $\rho b^i + b/2$  plus the cost of  $\text{ADV}$  to bring servers to the marked subtrees of phase 0, then we obtain Property (a) for  $i + 1$  and a fixed  $\delta'$ . Recall that this yields the same property for arbitrary  $\delta'$  by scaling all distances by a small factor.

Finally, it remains to justify the assumption that the server mass in each subtree  $T_i$  never exceeds  $k_i$ , which was necessary to allow invoking the induction hypothesis. Suppose after serving a request to a leaf  $u$ , the algorithm ends up with server mass  $k_i + c$  in the subtree that contains  $u$ , for some  $c > 0$ . Call this subtree  $S$ . Without loss of generality, the distance from the root of  $S$  to any leaf in  $S$  is at most 1 (we can scale all subtrees  $T_i$  down to achieve this). Upon serving the request at leaf  $u$ , at least mass  $c + \epsilon$  traveled to leaf  $u$  across the root  $r$  of  $T_{i+1}$ . Consider an alternative algorithm  $\text{ALG}'$  that stores an amount  $c$  of this mass at  $r$  and brings mass  $c$  from another vertex in  $S$  to leaf  $u$  instead. Upon serving this request,  $\text{ALG}'$  saves cost  $c$  compared to  $\text{ALG}$ , since it does not need to bring this mass from  $r$  to the root of  $S$ . The next time that  $\text{ALG}$  would take mass from  $S$  to another subtree  $T_i$ , we first take it from the mass  $c$  that is stored at the root of  $T_{i+1}$ . Notice that there will be no request in  $S$  until after the stored mass  $c$  at  $r$  has been used up. Once it has been used up,  $\text{ALG}'$  reorganizes its server mass so that its configuration is the same as that of  $\text{ALG}$  again. This requires at most cost  $c$  for moving this much mass within  $S$ , which is the cost that  $\text{ALG}'$  had saved before. Thus, we have a new algorithm whose cost is less than that of  $\text{ALG}$ , which never exceeds mass  $k_i$  in any subtree (by repeating this idea), and for which our lower bound holds.  $\square$

We obtain the main result by combining the trees guaranteed by Lemma 3.

**PROOF OF THEOREM 1.** For  $i \in \mathbb{N}$ , let  $b_i = \lfloor \sqrt{i} \rfloor$  and  $\rho_i = \frac{1}{3} \ln b_i$ . The lower bound holds on the following tree  $\mathcal{T}$ : It contains as subtrees, for each  $i \in \mathbb{N}$ , infinitely many copies of the tree  $T_i$  guaranteed by Lemma 3 for  $\rho = \rho_i$  and  $\delta = b_i^i$ . The roots of the subtrees  $T_i$  are connected to the root of  $\mathcal{T}$  by edges of length 1.

Let  $h \leq k$  be the numbers of offline and online servers respectively. Let  $i_h = \lfloor \sqrt{\ln h} \rfloor$ . The adversarial sequence uses only subtrees of type  $T_{i_h}$ . It consists of epochs: In each epoch, select a subtree of type  $T_{i_h}$  whose online server mass is zero. Requests are issued in this subtree as induced by Lemma 3. As soon as the online server mass in the subtree exceeds  $b_{i_h}^{i_h} \cdot (1 + i_h/(2b_{i_h}))$ , the epoch ends and a new epoch begins.

At the start of each epoch, the offline algorithm brings  $b_{i_h}^{i_h} \leq i_h^{i_h/2} \leq \exp(i_h^2) \leq h$  servers to the subtree of that epoch. For a given epoch, denote by  $\text{ALG}$  and  $\text{ADV}$ , respectively, the online and offline cost suffered *within* the active subtree of the epoch. By Lemma 3,

$$\text{ALG} \geq \rho_{i_h} \cdot \text{ADV} - b_{i_h}^{i_h}.$$

If the epoch runs indefinitely (because the algorithm never brings the required number of servers to the subtree), then the cost within the active subtree dominates the competitive ratio. Since  $\rho_{i_h} = \Omega(\log b_{i_h}) = \Omega(\log i_h) = \Omega(\log \log h)$ , the theorem follows.

Otherwise, the online algorithm pays at least  $b_{i_h}^{i_h} \cdot (1 + i_h/(2b_{i_h}))$  to bring as many servers to the subtree, whereas the offline algorithm pays only  $b_{i_h}^{i_h}$  to move servers to the subtree. Thus, the ratio of the total online to offline cost during each epoch is at least

$$\begin{aligned} \frac{\rho_{i_h} \cdot \text{ADV} + b_{i_h}^{i_h} \cdot \frac{i_h}{2b_{i_h}}}{\text{ADV} + b_{i_h}^{i_h}} &\geq \min \left\{ \rho_{i_h}, \frac{i_h}{2b_{i_h}} \right\} = \Omega(\log i_h) \\ &= \Omega(\log \log h). \quad \square \end{aligned}$$

The theorem holds also if instead of a single infinite tree  $\mathcal{T}$ , there is a *finite* tree  $\mathcal{T}_k$  for each  $k$ . The trees only need to be large enough so that, whenever we want to choose an empty subtree, we can instead find a subtree with negligibly small server mass.

**PROOF OF COROLLARY 2.** Consider the same tree as in the proof of Theorem 1, except that it contains the subtrees  $T_i$  for only one value of  $i = i_h$ . By the identical arguments as in the proof of Theorem 1, we obtain a lower bound of  $\Omega(\log i_h)$  for trees of depth  $i_h + 1$ . If the subtrees  $T_i$  are included for all  $i$ , we obtain a metric space with no competitive algorithm for the  $\infty$ -server problem.  $\square$

### 3 CONCLUSIONS

We have refuted the  $(h, k)$ -server hypothesis by proving that, on trees of sufficient depth,  $\mathcal{R}_{\mathcal{T}}(h, k) = \Omega(\log \log h)$  even when  $k$  is arbitrarily large. When expressed in terms of the depth  $D$  of the tree, the lower bound amounts to  $\Omega(\log D)$  and applies also to the  $\infty$ -server problem.

The construction of our lower bound is inherently fractional: On higher scales, even if an algorithm is deterministic, it can move only a fraction of a group of servers. It would be interesting to show a lower bound for deterministic algorithms that is substantially larger than the randomized one.

Intriguing gaps remain between the lower and upper bounds. The upper bound that would follow from the randomized  $k$ -server conjecture when disabling the  $k - h$  extra servers,  $O(\log h)$ , is exponentially larger than our lower bound. For deterministic algorithms, the gap is even doubly exponential.

### ACKNOWLEDGMENTS

Supported by Polish National Science Centre grants 2015/18/E/ST6/00456, 2016/22/E/ST6/00499, 2016/21/D/ST6/02402, and the NWO VICI grant 639.023.812.



## REFERENCES

- [1] Dimitris Achlioptas, Marek Chrobak, and John Noga. 2000. Competitive analysis of randomized paging algorithms. *Theoretical Computer Science* 234, 1–2 (2000), 203–218. [https://doi.org/10.1016/S0304-3975\(98\)00116-9](https://doi.org/10.1016/S0304-3975(98)00116-9)
- [2] Nikhil Bansal, Niv Buchbinder, Aleksander Madry, and Joseph Naor. 2015. A Polylogarithmic-Competitive Algorithm for the k-Server Problem. *J. ACM* 62, 5 (2015), 40:1–40:49. <https://doi.org/10.1145/2783434>
- [3] Nikhil Bansal, Niv Buchbinder, and Joseph Naor. 2012. A Primal-Dual Randomized Algorithm for Weighted Paging. *J. ACM* 59, 4 (2012), 19:1–19:24. <https://doi.org/10.1145/2339123.2339126>
- [4] Nikhil Bansal, Marek Eliáš, Lukasz Jez, and Grigorios Koumoutsos. 2019. The  $(h, k)$ -Server Problem on Bounded Depth Trees. *ACM Transactions on Algorithms* 15, 2 (2019), 28:1–28:26. <https://doi.org/10.1145/3301314>
- [5] Nikhil Bansal, Marek Eliáš, Lukasz Jez, Grigorios Koumoutsos, and Kirk Pruhs. 2018. Tight Bounds for Double Coverage Against Weak Adversaries. *Theory of Computing Systems* 62, 2 (2018), 349–365. <https://doi.org/10.1007/s00224-016-9703-3>
- [6] Yair Bartal, Béla Bollobás, and Manor Mendel. 2006. Ramsey-type theorems for metric spaces with applications to online problems. *J. Comput. System Sci.* 72, 5 (2006), 890–921. <https://doi.org/10.1016/j.jcss.2005.05.008>
- [7] Allan Borodin and Ran El-Yaniv. 1998. *Online Computation and Competitive Analysis*. Cambridge University Press.
- [8] Sébastien Bubeck, Michael B. Cohen, Yin Tat Lee, James R. Lee, and Aleksander Madry. 2018. k-server via multiscale entropic regularization. In *Proc. 50th ACM Symp. on Theory of Computing (STOC)*. 3–16. <https://doi.org/10.1145/3188745.3188798>
- [9] Niv Buchbinder, Anupam Gupta, Marco Molinaro, and Joseph (Seffi) Naor. 2019. k-Servers with a Smile: Online Algorithms via Projections. In *Proc. 30th ACM-SIAM Symp. on Discrete Algorithms (SODA)*. 98–116. <https://doi.org/10.1137/1.9781611975482.7>
- [10] Christian Coester, Elias Koutsoupias, and Philip Lazos. 2017. The Infinite Server Problem. In *Proc. 44th Int. Colloq. on Automata, Languages and Programming (ICALP)*. 14:1–14:14. <https://doi.org/10.4230/LIPIcs.ICALP.2017.14>
- [11] Amos Fiat, Richard M. Karp, Michael Luby, Lyle A. McGeoch, Daniel D. Sleator, and Neal E. Young. 1991. Competitive paging algorithms. *Journal of Algorithms* 12, 4 (1991), 685–699. [https://doi.org/10.1016/0196-6774\(91\)90041-V](https://doi.org/10.1016/0196-6774(91)90041-V)
- [12] Bala Kalyanasundaram and Kirk Pruhs. 2000. Speed is as powerful as clairvoyance. *J. ACM* 47, 4 (2000), 617–643. <https://doi.org/10.1145/347476.347479>
- [13] Elias Koutsoupias. 1999. Weak Adversaries for the k-Server Problem. In *Proc. 40th IEEE Symp. on Foundations of Computer Science (FOCS)*. 444–449. <https://doi.org/10.1109/SFFCS.1999.814616>
- [14] Elias Koutsoupias. 2009. The k-server problem. *Computer Science Review* 3, 2 (2009), 105–118. <https://doi.org/10.1016/j.cosrev.2009.04.002>
- [15] Elias Koutsoupias and Christos H. Papadimitriou. 1995. On the k-Server Conjecture. *J. ACM* 42, 5 (1995), 971–983. <https://doi.org/10.1145/210118.210128>
- [16] James R. Lee. 2018. Fusible HSTs and the Randomized k-Server Conjecture. In *Proc. 59th IEEE Symp. on Foundations of Computer Science (FOCS)*. 438–449. <https://doi.org/10.1109/FOCS.2018.00049>
- [17] Mark S. Manasse, Lyle A. McGeoch, and Daniel D. Sleator. 1990. Competitive algorithms for server problems. *Journal of Algorithms* 11, 2 (1990), 208–230. [https://doi.org/10.1016/0196-6774\(90\)90003-W](https://doi.org/10.1016/0196-6774(90)90003-W)
- [18] Lyle A. McGeoch and Daniel D. Sleator. 1991. A Strongly Competitive Randomized Paging Algorithm. *Algorithmica* 6, 6 (1991), 816–825. <https://doi.org/10.1007/BF01759073>
- [19] Cynthia A. Phillips, Clifford Stein, Eric Torng, and Joel Wein. 2002. Optimal Time-Critical Scheduling via Resource Augmentation. *Algorithmica* 32, 2 (2002), 163–200. <https://doi.org/10.1007/s00453-001-0068-9>
- [20] Daniel D. Sleator and Robert E. Tarjan. 1985. Amortized efficiency of list update and paging rules. *Commun. ACM* 28, 2 (1985), 202–208. <https://doi.org/10.1145/2786.2793>
- [21] Neal E. Young. 1991. On-line caching as cache size varies. In *Proc. 2nd ACM-SIAM Symp. on Discrete Algorithms (SODA)*. 241–250.
- [22] Neal E. Young. 1994. The k-Server Dual and Loose Competitiveness for Paging. *Algorithmica* 11, 6 (1994), 525–541. <https://doi.org/10.1007/BF01189992>