Incremental Medians via Online Bidding

**Marek Chrobak · Claire Kenyon · John Noga ·
Neal E. Young**

Abstract In the k-median problem we are given sets of facilities and customers, and distances between them. For a given set F of facilities, the cost of serving a customer u is the minimum distance between u and a facility in F. The goal is to find a set F of k facilities that minimizes the sum, over all customers, of their service costs.

Following the work of Mettu and Plaxton, we study the *incremental* medians problem, where k is not known in advance. An incremental algorithm produces a nested sequence of facility sets $F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n$, where $|F_k| = k$ for each k. Such an algorithm is called *c-cost-competitive* if the cost of each F_k is at most c times the optimum k-median cost. We give improved incremental algorithms for the metric version of this problem: an 8-cost-competitive deterministic algorithm, a $2e \approx 5.44$-cost-competitive randomized algorithm, a $(24 + \epsilon)$-cost-competitive, polynomial-time deterministic algorithm, and a $6e + \epsilon \approx 16.31$-cost-competitive, polynomial-time randomized algorithm.

We also consider the competitive ratio with respect to *size*. An algorithm is *s-size-competitive* if the cost of each F_k is at most the minimum cost of any set of k facilities, while the size of F_k is at most sk. We show that the optimal size-competitive ratios for this problem, in the deterministic and randomized cases, are 4 and e. For polynomial-time algorithms, we present the first polynomial-time $O(\log m)$-

M. Chrobak · N.E. Young (✉)
Department of Computer Science, University of California, Riverside, CA 92521, USA
e-mail: algorithmica@neal.young.name

C. Kenyon
Computer Science Department, Brown University, Providence, RI 02912, USA

J. Noga
Department of Computer Science, California State University, Northridge, CA 91330, USA

size-approximation algorithm for the offline problem, as well as a polynomial-time $O(\log m)$-size-competitive algorithm for the incremental problem.

Our upper bound proofs reduce the incremental medians problem to the following *online bidding* problem: faced with some unknown threshold $T \in \mathbb{R}^+$, an algorithm must submit "bids" $b \in \mathbb{R}^+$ until it submits a bid $b \geq T$, paying the sum of all its bids. We present folklore algorithms for online bidding and prove that they are optimally competitive.

We extend some of the above results for incremental medians to approximately metric distance functions and to incremental fractional medians. Finally, we consider a restricted version of the incremental medians problem where k is restricted to one of two given values, for which we give a deterministic algorithm with a nearly optimal cost-competitive ratio.

1 Introduction and Summary of Results

The k-Median Problem   An instance of the k-median problem is specified by a finite set $\mathcal{C}$ of customers, a finite set $\mathcal{F}$ of facilities, and, for each customer u and facility f, a distance $d_{uf} \geq 0$ from u to f representing the cost of serving u from f. The cost of a set of facilities $F \subseteq \mathcal{F}$ is $cost(F) = \sum_{u \in \mathcal{C}} d_{uF}$, where $d_{uF} = \min_{f \in F} d_{uf}$. For a given k, the *offline k-median problem* is to compute a *k-median*, that is, a set $F \subseteq \mathcal{F}$ of cardinality k for which $cost(F)$ is minimum (among all sets of cardinality k). This minimum cost is denoted opt_k. An instance of the k-median problem is called *metric* if the distance function is a metric (the shortest u-to-f path has length d_{uf} for each $u \in \mathcal{C}$ and $f \in \mathcal{F}$).

The k-median problem is a well-known NP-hard facility location problem. Substantial work has been done on efficient approximation algorithms that, given k, find a set F_k of k facilities of approximately minimum cost [1–3, 7–9, 17–19, 26, 30]. In particular, for the metric version, Arya et al. [2, 3] show that, for any $\epsilon > 0$, a set F_k of cost at most $(3 + \epsilon)opt_k$ can be found in polynomial time.

Incremental Medians   The *incremental* medians problem is a version of the k-median problem where k is not specified in advance [27, 28]. Instead, authorizations for additional facilities arrive over time. Given an instance of the k-median problem, a (possibly randomized) algorithm produces a sequence $\bar{F} = (F_1, F_2, \ldots, F_n)$ of facility sets, where $F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n \subseteq \mathcal{F}$, and $|F_k| \leq k$ for all k.

In general, in an incremental solution, the F_k's cannot all simultaneously have minimum cost. An algorithm is said to be *c-cost-competitive*, or to have *cost-competitive ratio* of c, if it produces a (possibly random) sequence $\bar{F}$ of sets which is c-cost-competitive, that is, such that for each k, the set F_k has size at most k and (expected) cost at most $c \cdot opt_k$.

Alternatively, an algorithm is *s-size-competitive* if it produces a (possibly random) sequence $\bar{F}$ of sets which is s-size-competitive, that is, such that each set F_k has cost at most opt_k and (expected) *size* at most sk.

For offline solutions we use the term "approximate" instead of "competitive".

Online Bidding To analyze incremental medians, we reduce the various incremental medians problems to the following folklore *online bidding* problem. An algorithm repeatedly submits "bids" $b \in \mathbb{R}^+$, until it submits a bid b that is at least as large as some unknown threshold $T \in \mathbb{R}^+$. The algorithm's cost is the sum of the submitted bids. The algorithm is β-competitive if, for any $T \in \mathbb{R}^+$, its cost is at most βT (or, if the algorithm is randomized, its expected cost is at most βT). More generally, the algorithm may be given in advance a closed universe $\mathcal{U} \subseteq \mathbb{R}^+$, with a guarantee that the threshold T is in $\mathcal{U}$ and a requirement that all bids be in $\mathcal{U}$. (To handle the case when $\mathcal{U}$ includes arbitrarily small positive numbers, we allow the bidding sequence to be left-infinite—see Sect. 2 for a formal definition.)

In Sect. 2 we completely characterize optimal competitive ratios for deterministic and randomized algorithms for online bidding:

Theorem 1 (folklore)

(a) *The online bidding problem has a deterministic* 4-*competitive algorithm. If the universe $\mathcal{U}$ is finite, this algorithm runs in time polynomial in $|\mathcal{U}|$.*
(b) *The online bidding problem has a randomized e-competitive algorithm. If $\mathcal{U}$ is finite, this algorithm runs in time polynomial in $|\mathcal{U}|$.*

Throughout, $[n]$ denotes $\{1, 2, \ldots, n\}$.

Theorem 2 (folklore)

(a) *No deterministic algorithm for online bidding is less than* 4-*competitive, even when restricted to instances of the form $\mathcal{U} = [n]$ for some integer n.*
(b) *No randomized algorithm for online bidding is less than e-competitive, even when restricted to instances of the form $\mathcal{U} = [n]$ for some integer n.*

Portions of Theorems 1 and 2 are folklore. In particular, Theorem 1(a) uses a doubling algorithm that has been used previously in several papers, first in [21, 29] and later in [5, 6, 13, 15, 16]. Some of these papers also have the randomized upper bound. We include proofs of all bounds for completeness. Our main new contribution in this setting is Theorem 2(b), a randomized lower bound that matches the known upper bound.

Size-Competitiveness To our knowledge, size-competitive algorithms for incremental medians have not been studied, although other online problems have been analyzed in an analogous setting of *resource augmentation* (e.g. [10, 20, 23]). For unrestricted (possibly non-polynomial-time) algorithms, we completely characterize the optimal size-competitive ratios:

Theorem 3

(a) *The incremental medians problem has a* 4-*size-competitive deterministic algorithm.*
(b) *No deterministic incremental algorithm has size-competitive ratio smaller than* 4.

Theorem 4

(a) *The incremental medians problem has an e-size-competitive randomized algorithm.*
(b) *No randomized incremental algorithm has size-competitive ratio smaller than e.*

We stress that the upper and lower bounds in Theorems 3 and 4 are for unrestricted algorithms and hold for both the metric and non-metric problems.

Regarding polynomial-time algorithms, previously no polynomial-time size-approximation algorithms for k-medians were known, even for the offline problem. The best previous result for the offline problem is a bicriteria-approximation algorithm which finds a facility set of size $O(k \log(m + m/\epsilon))$ and cost at most $(1+\epsilon)opt_k$ [30] (improving on the previous work in [22, 25, 26, 30]). We first improve this result to obtain an offline size-approximation algorithm:

Theorem 5 *The offline k-median problem has a polynomial-time $O(\log m)$-size-approximation algorithm, where $m = |\mathcal{C}|$ is the number of customers.*

Note that this algorithm finds a true (not bicriteria) approximate solution: a facility set of size $O(k \log m)$ and cost at most opt_k. We use this result and a reduction to give a polynomial-time size-approximation algorithm for the incremental problem:

Theorem 6 *The incremental medians problem has a polynomial-time $O(\log m)$-size-competitive algorithm.*

The bounds in Theorems 5 and 6 are optimal in the sense that no polynomial-time algorithm (incremental or offline) is $o(\log m)$-size-competitive unless P=NP, even for the metric case. This follows from known results on the hardness of approximating set cover.

Theorems 3, 4, and 6 (proven in Sect. 3) imply the size-competitive ratios shown in Table 1.

Cost-competitive Incremental Medians   For incremental medians, Mettu and Plaxton [27, 28] give a c-cost-competitive linear-time algorithm with $c \approx 30$. We improve this result. The problem is difficult both because (i) the solution must be incremental, and (ii) even the offline problem is NP-hard. To study separately the effects of the two difficulties, we consider both polynomial-time and unrestricted algorithms, proving the following two theorems (see Sect. 4 for the proofs):

Table 1 Competitive ratios shown for incremental medians and online bidding. Ratios in bold are optimal

| problem | cost-competitive, metric | | size-competitive | | bidding |
time	polynomial	unrestricted	polynomial	unrestricted	polynomial
deterministic	$24 + \epsilon$	8	$\boldsymbol{O(\log m)}$	**4**	**4**
randomized	$6e + \epsilon < 16.31$	$2e < 5.44$	$\boldsymbol{O(\log m)}$	$\boldsymbol{e < 2.72}$	$\boldsymbol{e < 2.72}$

Theorem 7

(a) *The metric incremental medians problem has an 8-cost-competitive deterministic algorithm.*

(b) *Suppose that the offline metric k-median problem has a polynomial-time c-cost-approximation algorithm. Then the incremental medians problem has a polynomial-time 8c-cost-competitive deterministic algorithm.*

Theorem 8

(a) *The metric incremental medians problem has a 2e-cost-competitive randomized algorithm.*

(b) *Suppose that the offline metric k-median problem has a polynomial-time c-cost-approximation algorithm. Then the incremental medians problem has a polynomial-time 2ec-cost-competitive randomized algorithm.*

As it is known that there is a polynomial-time $(3 + \epsilon)$-cost-approximation algorithm for the metric k-median problem [2, 3], Theorems 7 and 8 imply the cost-competitive ratios shown in Table 1.

Theorems 7 and 8 were recently and independently discovered by Lin, Nagarajan, Rajaraman and Williamson [24]. For polynomial-time algorithms, they improve the cost-competitive ratios further to 16 and 4e, in the deterministic and randomized cases, respectively. (The general approach in [24] is based on the doubling method similar to ours; the improvements were accomplished using a Lagrangian-multiplier-preserving approximation for facility location.) They also generalize the approach to incremental versions of k-MST, k-vertex cover, k-set cover, facility location, and to hierarchical k-median.

λ-*Relaxed Metrics*   Mettu and Plaxton show that their incremental medians algorithm also works in "λ-approximate" metric spaces, achieving cost-competitive-ratio $O(\lambda^4)$ [27, 28]. We get a similar result. Let $\lambda \geq 1$. We say that the cost function d is a λ-*relaxed metric* if $d_{fy} \leq \lambda(d_{fx} + d_{xg} + d_{gy})$ for any facilities f, g and customers x and y. (This condition is somewhat less restrictive than the one in [27, 28]. A related concept was studied in [14].) In Sect. 5, we prove that Theorems 7 and 8 generalize as follows:

Theorem 9

(a) *The incremental medians problem for λ-relaxed metrics has a deterministic algorithm with cost-competitive ratio $8\lambda^2$.*

(b) *If the offline k-median problem for λ-relaxed metrics has a polynomial-time c-cost-approximation algorithm, then the incremental version has a deterministic polynomial-time algorithm with cost-competitive ratio $8\lambda^2 c$.*

Theorem 10

(a) *The incremental medians problem for λ-relaxed metrics has a randomized algorithm with cost-competitive ratio $2e\lambda^2$.*

(b) *If the offline k-median problem for λ-relaxed metrics has a polynomial-time c-cost-approximation algorithm, then the incremental version has a randomized polynomial-time algorithm with cost-competitive ratio $2e\lambda^2 c$.*

Incremental Fractional Medians   A *fractional k-median* x is a solution to the standard linear program relaxation for the k-median problem (see Sect. 6). In this linear program, x_{uf} specifies how much of the demand from customer u is served by facility f; thus we have a constraint $\sum_{f \in \mathcal{F}} x_{uf} = 1$. For each $f \in \mathcal{F}$, the *capacity required* at f is $|x|_f = \max_{u \in \mathcal{C}} x_{uf}$, and the *total capacity* of x is $|x| = \sum_{f \in \mathcal{F}} |x|_f$. (Naturally, this corresponds to the cardinality of x in the integral case.) We require that $|x| \leq k$, and the objective is to minimize the cost of x, defined by $cost(x) = \sum_{u \in \mathcal{C}, f \in \mathcal{F}} d_{uf} x_{uf}$.

For two fractional medians x, x', we say that x *dominates* x' if $|x'|_f \leq |x|_f$ for each facility $f \in \mathcal{F}$. An *incremental fractional median* is defined by a sequence $(x^k)_k$ of fractional k-medians, one for each $k \in [n]$, where each x^{k+1} dominates x^k, for $k > n$. This sequence is *c-cost-competitive* if $cost(x^k) \leq c \cdot \overline{opt}_k$ for each k, where $\overline{opt}_k$ is the minimum cost of any (non-incremental) fractional k-median.

To prove the theorem below (see Sect. 6), we extend the proof of Theorem 8 to the fractional case, then note that the randomized algorithm for the fractional problem can be derandomized without increasing the competitive ratio.

Theorem 11 *If the distance function is metric then there is a deterministic polynomial-time algorithm that produces a 2e-cost-competitive incremental fractional median.*

One motivation for introducing fractional incremental medians (in addition to its own independent interest) is the possibility that the above theorem, or its improvements, could be used to improve the c-cost-competitive ratio for the integral case. Note that $\overline{opt}_k \leq opt_k$, so if we could somehow round a fractional c-competitive incremental median to an integral solution, giving up a factor of, say, 3 in the cost, then we would have a $3c$-cost-competitive deterministic algorithm for incremental medians. (See [4] for a similar approach for online problems.) However, the ratio $2e$ above is insufficient for this, in that even with $c = 2e$, the resulting ratio $3c = 6e$ is larger than the current best ratio of 16. Note that we are unlikely to lose less than a factor of 3 in rounding the fractional incremental median, as it is at least as hard as rounding a fractional k-median.

The kl-Medians Problem   A natural question to ask is whether better competitive ratios are possible if the number of facilities can take only some limited number of values. As shown in [27, 28], no algorithm can be better than 2-competitive even when there are only two possible numbers of facilities, either 1 or k, for some large k.

For any $1 \leq k < l \leq n$, we define the *kl-medians* problem as the restricted version of the incremental medians problem where the number of facilities is either k or l. In Sect. 7, we determine almost exactly the competitive ratio of the kl-medians problem in the deterministic case:

Theorem 12 *Let* $1 \le k < l \le n$. (a) *If the distance function is metric, then there is a deterministic kl-median algorithm with cost-competitive ratio* $2 - 1/l$, *and* (b) *no ratio better than* $2 - 1/(l - k + 1)$ *is possible.*

Bicriteria Approximations We say that an algorithm for the k-median problem is a *bicriteria* (c, s)-*approximation* algorithm if it approximates the cost within the ratio of c and the size within the ratio of s. Such bicriteria approximation algorithms for k-medians have been studied by many authors, [22, 25, 26, 30]. We remark without proof that, analogously to Theorem 7 and Theorem 3, one can transform any offline polynomial-time bicriteria (c, s)-approximation algorithm into a polynomial-time *incremental* algorithm whose bicriteria performance guarantee is either $(c, 4s)$ or $(8c, s)$. For example, for metric spaces, using the approximation results from [2, 3, 22, 25] one can obtain incremental polynomial-time algorithms with the following respective bicriteria ratios: $(3 + \epsilon, 4)$, $(2 + \epsilon, 4(1 + 2\epsilon^{-1}))$, and $(1 + \epsilon, 4(3 + 5\epsilon^{-1}))$, where $\epsilon > 0$.

Weighted Medians It is quite easy to see that all of the results in this paper extend to the version of k-medians where customers are given non-negative weights. Denote by w_u the weight of $u \in \mathcal{C}$. In this generalization, the cost of a facility set F is defined by $cost(F) = \sum_{u \in \mathcal{C}} w_u d_{uF}$. The results on size-competitive algorithms can be further extended to the case where each facility f is assigned a weight w_f. The definition of the cost function remains the same. The value k represents now an upper bound on the allowed total facility weight $\sum_{f \in F} w_f$.

2 Online Bidding

In this section we provide a complete analysis of online bidding by proving Theorems 1 and 2.

Throughout the paper, $\mathbb{R}^+$ denotes the set of non-negative reals, $\mathbb{Z}$ the set of integers, and $\mathbb{N}^+$ the set of positive integers. For $n \in \mathbb{N}^+$, let $[n] = \{1, 2, \ldots, n\}$.

Given a universe $\mathcal{U}$ which is a closed subset of $\mathbb{R}^+$, an *online bidding* algorithm outputs a *bid set* $\mathcal{B} \subseteq \mathcal{U}$. Against a particular threshold $T \in \mathcal{U}$, the algorithm's cost is

$$\sum \{b \in \mathcal{B} : b < T\} + \min\{b \in \mathcal{B} : b \ge T\}.$$

The bid set $\mathcal{B}$ is β-*competitive* if, for any $T \in \mathcal{U}$, this cost is at most βT.

Proof of Theorem 1(a) We give a deterministic 4-competitive algorithm.

First, consider the case $\mathcal{U} = \mathbb{R}^+$. Define the algorithm to produce the set of bids $\mathcal{B} = \{0\} \cup \{2^i : i \in \mathbb{Z}\}$. If the threshold T is zero, the algorithm pays zero. For any other threshold $T > 0$, let p be such that $T \in (2^{p-1}, 2^p]$. The algorithm pays $\sum_{i \le p} 2^i = 2^{p+1} \le 4T$, and thus its competitive ratio is at most 4.

Next, we reduce the general case to the case $\mathcal{U} = \mathbb{R}^+$. Let $\mathcal{B}$ be a 4-competitive bid set for $\mathcal{U} = \mathbb{R}^+$. For an arbitrary closed universe $\mathcal{U}' \subseteq \mathbb{R}^+$, the algorithm produces the bid set

$$\mathcal{B}' = \{\max(\mathcal{U}' \cap [0, b]) : b \in \mathcal{B}\}.$$

In other words, we replace each $b \in \mathcal{B}$ with the maximum element in $\mathcal{U}' \cap [0, b]$ (if any). The cost incurred against any threshold $T \in \mathcal{U}'$ is at most the cost incurred when using the bid set $\mathcal{B}$ against the same threshold T. Thus, the bid set $\mathcal{B}'$ is also 4-competitive.

Note that if $\mathcal{U}$ is finite then for the bid set $\mathcal{B}$ described in the previous paragraph, the corresponding bid set $\mathcal{B}'$ can be computed in time polynomial in $|\mathcal{U}|$. $\qquad\square$

Proof of Theorem 1(b) We give a randomized e-competitive algorithm.

First, we consider the case $\mathcal{U} = \mathbb{R}^+$. The algorithm picks a real number $\xi \in [0, 1)$ uniformly at random, then produces the set of bids $\mathcal{B} = \{0\} \cup \{e^\xi e^i : i \in \mathbb{Z}\}$.

We now give the analysis of this algorithm. We first observe that for any integer x, the set $\mathcal{L} = \{\ln(b) : b \in \mathcal{B} - \{0\}\}$ induces a uniform distribution in the interval $[x, x + 1)$, in the sense that $[x, x + 1)$ contains exactly one element of $\mathcal{L}$ and this element is uniformly distributed in $[x, x + 1)$. This immediately implies that the above property holds in fact for any real x.

Let $T > 0$ be the threshold. For any constant τ, by the paragraph above, we can equivalently describe the algorithm as producing the set of bids $\mathcal{B} = \{0\} \cup \{e^{\xi+\tau+i} : i \in \mathbb{Z}\}$ where ξ is distributed uniformly in $[0, 1)$. In particular, we can take $\tau = \ln(T)$.

Let b be the random variable equal to the largest bid paid by the algorithm, defined by $e^{\xi+\ln(T)+i-1} < T \le e^{\xi+\ln(T)+i} = b$, or, in other words, $e^{\xi+i-1} < e^0 \le e^{\xi+i} = b/T$. Thus $i = 0$ and b/T is distributed like e^ξ with ξ uniform in $[0, 1)$. It follows that the expectation of b is $T \int_0^1 e^z \, dz = T(e - 1)$, thus the expected total payment incurred by the algorithm is $E[\sum_{i=0}^\infty b e^{-i}] = E[b]e/(e - 1) = eT$, and the algorithm is e-competitive.

The general case, for an arbitrary closed universe, reduces to the case $\mathcal{U} = \mathbb{R}^+$ just as in the proof of Theorem 1(a) above. $\qquad\square$

Proof of Theorem 2(a) We now show a deterministic lower bound matching the upper bound.

In the simple case where $\mathcal{U} = \mathbb{N}^+$, assume (towards a contradiction) that there exists an online algorithm with competitive ratio $a < 4$. Let $\mathcal{B} = \{b_i\}_{i\geq1}$ be the bid set produced by the algorithm, $s_j = \sum_{i=1}^j b_i$ and $y_j = s_{j+1}/s_j$, for $j \geq 1$. Against the threshold $T = b_j + 1$, the algorithm pays s_{j+1}. By our assumption, $s_{j+1} \le a(b_j + 1) = a(s_j - s_{j-1} + 1)$ for all j. Rearranging, and using $s_j \geq j$, we get:

$$(y_{j+1} - y_j)y_j \le -y_j^2 + a(s_{j+1} - s_j + 1)y_j/s_{j+1}$$

$$= -y_j^2 + ay_j - a(1 - 1/s_j)$$

$$\le -y_j^2 + ay_j - a(1 - 1/j). \tag{1}$$

For $a < 4$ and j large enough, the discriminant $a^2 - 4a(1 - 1/j)$ is negative, thus the last expression is negative, and so the sequence (y_j) is ultimately decreasing. As it is bounded below by 1, it converges to some limit y which, by continuity, must satisfy $0 = (y - y)y \le -y^2 + ay - a$—a contradiction, since $-y^2 + ay - a < 0$ for $0 < a < 4$ and any y.

In the case when $\mathcal{U} = [n]$, for arbitrarily large n, the proof is similar but a bit more explicit. Let $\mathcal{B} = \{b_i\}_{i=1}^m$. We start out the same. Note that $s_{j+1} \leq a(b_j + 1)$ implies $s_{j+1} < 8s_j$, and so $y_j < 8$ for all $j = 1, \ldots, m-1$.

Let j_0 be the smallest positive integer such that $a^2 - 4a(1 - 1/j_0) = -\delta < 0$. Inequality (1) is valid for $j = j_0, \ldots, m-2$, so we have:

$$(y_{j+1} - y_j)y_j \leq \max_y(-y^2 + ay - a(1 - 1/j_0)) = -\delta/4.$$

Since $y_j \leq 8$, the above inequality implies that $y_{j+1} - y_j \leq -\delta/32$. Therefore, for $j = j_0, \ldots, m-1$, we have $1 \leq y_j \leq y_{j_0} - (j - j_0)\delta/32 \leq 8 - (j - j_0)\delta/32$, so, in particular, $m \leq j_0 + 1 + 224/\delta$. Since $b_m = n$ (otherwise, $\mathcal{B}$ would not be competitive at all), $b_1 \leq 8$, and $b_{j+1} \leq 8b_j$ for all $j = 1, \ldots, m-1$, we get $n \leq 8^{j_0 + 1 + 224/\delta}$, contradicting our assumption that n can be arbitrarily large. □

Proof of Theorem 2(b) We now show a randomized lower bound matching the upper bound. The proof consists of the two lemmas below.

Lemma 1 *Let $n \in \mathbb{N}^+$. Suppose that there are $\mu : [n] \to \mathbb{R}^+$ and $\pi : [n] \to \mathbb{R}^+$ that satisfy*

$$\sum_{T=t}^n \frac{1}{T}\pi(T) \geq \frac{1}{b}\sum_{T=t}^b \mu(T) \quad \forall b, t : 1 \leq t \leq b \leq n. \tag{2}$$

Then any randomized β-competitive online bidding algorithm for $\mathcal{U} = [n]$ must have

$$\beta \geq \frac{\sum_{T=1}^n \mu(T)}{\sum_{T=1}^n \pi(T)}. \tag{3}$$

Proof Consider a β-competitive randomized algorithm for $\mathcal{U}$, and let $\mathcal{B} = \{b_1, b_2, \ldots, b_m\}$ be the ordered (random) sequence of bids produced by the algorithm. Without loss of generality, $b_m = n$.

For $t \leq b$, let $X(t, b)$ be the characteristic function of the event that $t - 1$ and b are consecutive elements of $\{0\} \cup \mathcal{B}$. The algorithm pays $\sum_{t \leq T} \sum_b bX(t, b)$ against threshold T. Since it is β-competitive, this is at most βT in expectation. Since the algorithm only stops when reaching a bid greater than or equal to T, we must always have $\sum_{t,b:t \leq T \leq b} X(t, b) \geq 1$. Hence, together with β, the expectation $x(t, b)$ of $X(t, b)$ must form a feasible solution to the following linear program (LP):

$$\text{minimize}_{\beta, x} \ \beta \quad \text{subject to} \ \begin{cases} \beta - \sum_{b=1}^n \frac{b}{T}\sum_{t=1}^T x(t, b) \geq 0 & \forall T \in [n], \\ \sum_{b=T}^n \sum_{t=1}^T x(t, b) \geq 1 & \forall T \in [n], \\ x(t, b) \geq 0 & \forall t, b \in [n]. \end{cases}$$

Thus, the value of (LP) is a lower bound on the optimal competitive ratio of the randomized algorithm. To get a lower bound on the value of (LP), we use the dual (DLP) (where the dual variables $\pi(T)$ correspond to the first set of constraints and

$\mu(T)$ to the second set of constraints):

$$\text{maximize}_{\mu,\pi} \sum_{T=1}^{n} \mu(T) \quad \text{subject to} \quad \begin{cases} \sum_{T=1}^{n} \pi(T) \leq 1, \\ \sum_{T=t}^{b} \mu(T) - \sum_{T=t}^{n} \frac{b}{T}\pi(T) \leq 0 & \forall t, b \in [n], \\ \mu(T), \pi(T) \geq 0 & \forall T \in [n]. \end{cases}$$

Now, given any μ and π meeting the condition of the lemma, if we scale μ and π by dividing by $\sum_T \pi(T)$, we get a feasible dual solution whose value is $\sum_T \mu(T) / \sum_T \pi(T)$, and the lemma follows.   □

Lemma 2 *There exist $\mu : [n] \to \mathbb{R}^+$ and $\pi : [n] \to \mathbb{R}^+$ satisfying Condition* (2) *of Lemma* 1 *and such that* $\sum_{T=1}^{n} \mu(T) / \sum_{T=1}^{n} \pi(T) \geq (1 - o(1))e$ *as $n \to \infty$.*

Proof Fix U arbitrarily large and let $n = \lceil U^2 \log U \rceil$. Let $\alpha > 0$ be a constant to be determined later. We will choose α so that Condition (2) holds, and then show that the corresponding lower bound is $e(1 - o(1))$ as $U \to \infty$. Define

$$\mu(T) = \begin{cases} \alpha/T & \text{if } U \leq T \leq U^2, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \pi(T) = \begin{cases} 1/T & \text{if } U \leq T \leq U^2 \log U, \\ 0 & \text{otherwise.} \end{cases}$$

If $t > U^2$, then the right-hand side of Condition (2) has value 0, so the condition holds trivially. On the other hand, since $\pi(T)$ and $\mu(T)$ are zero for $T < U$, if the condition holds for $t = U$, then it also holds for $t < U$. So, we need only verify the condition for t in the range $U \leq t \leq U^2$. The expression on the left-hand side of (2) then has value

$$\sum_{T=t}^{U^2 \log U} \frac{1}{T^2} \geq \int_t^{1+U^2 \log U} \frac{1}{T^2} dT = \frac{1}{t} - \frac{1}{1 + U^2 \log U}$$

$$\geq \frac{1}{t}\left(1 - \frac{U^2}{1 + U^2 \log U}\right) \geq \frac{1}{t}(1 - o(1)).$$

In comparison, the expression on the right-hand side has value at most

$$\max_{b \geq t} \frac{1}{b} \sum_{T=t}^{b} \frac{\alpha}{T} \leq \alpha \max_{b \geq t} \frac{1}{b} \int_{t-1}^{b} \frac{1}{T} dT = \alpha \max_{b \geq t} \frac{1}{b} \ln \frac{b}{t-1} = \frac{\alpha}{e\, t\, (1 - o(1))}.$$

(The second equation follows by calculus, for the maximum occurs when $b = e(t - 1)$.) Thus, Condition (2) is met for $\alpha = (1 - o(1))e$. Denoting by H_j the jth harmonic number, we have $H_j = (1 + o(1)) \ln(j)$. Using this and the bounds derived above we get

$$\frac{\sum_T \mu(T)}{\sum_T \pi(T)} = \frac{\sum_{T=U}^{U^2} \alpha/T}{\sum_{T=U}^{U^2 \log U} 1/T}$$

$$= (1 - o(1))e \cdot \frac{H_{U^2} - H_{U-1}}{H_{U^2 \log U} - H_{U-1}}$$

$$\geq (1 - o(1))e \cdot \frac{\ln(U^2/(U-1))}{\ln((U^2 \log U)/(U-1))} \geq (1 - o(1))e$$

completing the proof of the lemma. $\square$

Using the functions μ and π from Lemma 2 in Lemma 1, we obtain a lower bound of $\beta \geq (1 - o(1))e$ on the competitive ratio, and Theorem 2 follows. $\square$

3 Incremental Size-Competitive Medians

In this section we prove Theorems 3, 4, 5, and 6. The proofs are based on the reduction shown in the next lemma. We show that from a β-competitive algorithm for online bidding, and a c-size-approximation algorithm for the offline k-median problem, we can construct a $c\beta$-size-competitive algorithm for the incremental medians problem. This reduction works even for the non-metric case. The reduction takes polynomial time, so, if the offline size-approximation algorithm runs in polynomial time, the reduction yields a polynomial-time size-competitive algorithm. Likewise, if the online bidding algorithm is randomized then the size-competitive algorithm will be randomized.

Lemma 3 *Assume that for each k we have a set of facilities F_k^* of size at most sk and cost at most opt_k. Let $\beta \geq 1$.*

Suppose that there exists a (*randomized*) *polynomial-time β-competitive algorithm for online bidding. Then, in* (*randomized*) *polynomial time, we can compute an* incremental *solution $\bar{F} = (F_1, \ldots, F_n)$ where each F_k has cost at most opt_k and* (*expected*) *size at most βsk.*

Proof We give the proof for the deterministic case. (The proof in the randomized setting is an easy extension, and we omit it.) Let $\mathcal{U} = [n]$ and take $\mathcal{B}$ to be the set of bids used by the β-competitive online bidding algorithm for universe $\mathcal{U}$. Let $\mathcal{B}_k$ be the set of bids issued against threshold $T = k$. We define $F_k = \bigcup_{b \in \mathcal{B}_k} F_b^*$. Note that $\mathcal{B}_j \subseteq \mathcal{B}_{j+1}$ for all $j < n$, and so $\bar{F}$ is indeed an incremental solution. Further, F_k contains F_b^* for some $b \geq k$, so $cost(F_k) \leq cost(F_b^*) \leq opt_b \leq opt_k$. Finally, $|F_k| \leq \sum_{b \in \mathcal{B}_k} |F_b^*| \leq \sum_{b \in \mathcal{B}_k} sb \leq s\beta k$ since $\mathcal{B}$ is β-competitive. $\square$

Proof of Theorems 3(a) and 4(a) By Theorems 1(a), and 2(a), there are deterministic and randomized algorithms for online bidding with competitive ratios of 4 and e, respectively. Using these online bidding algorithms, and taking each F_k^* to be the optimal k-median, the reduction in Lemma 3 gives a 4-size-competitive deterministic algorithm and an e-size-competitive randomized algorithm for incremental medians. $\square$

Next, we turn our attention to the proof of Theorem 3(b), our lower bound on the size-competitiveness of unrestricted algorithms for incremental medians. We do

this by showing the converse of the reduction in Lemma 3. That is, we show that competitive online bidding reduces to size-competitive incremental medians:

Lemma 4 *Let $s \geq 1$ and assume that, for incremental medians (metric or not), there is a (possibly randomized) s-size-competitive algorithm. Then, for any integer n, there is a (randomized) s-competitive algorithm for online bidding with $\mathcal{U} = [n]$.*

Proof We give the proof in the deterministic setting. (The proof in the randomized setting is an easy extension, and we omit it.) For any arbitrarily large m, we construct a set $\mathcal{C}$ of customers, a set $\mathcal{F}$ of facilities, and a metric distance function d_{uf}, for $u \in \mathcal{C}$ and $f \in \mathcal{F}$. The facility set $\mathcal{F}$ will be partitioned into disjoint sets $M_1, M_2, \ldots, M_m$, where $|M_k| = k$ for each k, with the following property:

(∗) For all k, and for every set F of facilities, if $cost(F) \leq cost(M_k)$ then there exists $\ell \geq k$ such that $M_\ell \subseteq F$.

Notice that the condition (∗) implies that $cost(M_k) > cost(M_{k+1})$ for $k < m$, and that each M_k is the unique optimum k-median.

Assume for the moment that there exists such a metric space, and consider an s-size-competitive incremental median $\bar{F} = (F_1, \ldots, F_m)$ for it. This means that for each k we have $|F_k| \leq \beta k$ and that F_k satisfies condition (∗).

Let $\mathcal{B} = \{k : M_k \subseteq F_k\}$. We show that $\mathcal{B}$ is an s-competitive bid set for universe $\mathcal{U} = [m]$. For $k = m$, (∗) implies that $M_m \subseteq F_m$, and thus $\mathcal{B} \neq \emptyset$. Against any threshold $T \in [m]$, the total of the bids paid will be

$$C = \sum \{k : k < T, M_k \subseteq F_k\} + \min\{\ell : \ell \geq T, M_\ell \subseteq F_\ell\}.$$

Now, $\sum \{k : k < T, M_k \subseteq F_k\} \leq \sum \{k : k < T, M_k \subseteq F_T\}$ since $\bar{F}$ is a nested sequence. Similarly, we have $\min\{\ell : \ell \geq T, M_\ell \subseteq F_\ell\} \leq \min\{\ell : \ell \geq T, M_\ell \subseteq F_T\}$. Note that by (∗), $M_\ell \subseteq F_T$ for some $\ell \geq T$, so the minimum on the right is well-defined for $T \in [m]$. Thus we can bound the cost of $\mathcal{B}$ as follows:

$$C \leq \sum \{k : k < T, M_k \subseteq F_T\} + \min\{\ell : \ell \geq T, M_\ell \subseteq F_T\}$$

$$= \sum \{|M_k| : k < T, M_k \subseteq F_T\} + \min\{|M_\ell| : \ell \geq T, M_\ell \subseteq F_T\}$$

$$\quad \text{because } |M_k| = k$$

$$\leq \sum \{|M_k| : M_k \subseteq F_T\}$$

$$\leq |F_T| \quad \text{since the } M_k\text{'s are disjoint}$$

$$\leq sT \quad \text{since } \bar{F} \text{ is } s\text{-size-competitive.}$$

Thus, the bid set $\mathcal{B}$ is s-competitive for universe $\mathcal{U} = [m]$.

We now present the construction of the metric space satisfying condition (∗). Let $\mathcal{C} = [1] \times [2] \times \ldots \times [m]$, that is, $\mathcal{C}$ is the set of integer vectors $\bar{u} = (u_1, u_2, \ldots, u_m)$ where $u_\ell \in [\ell]$ for all $\ell = 1, 2, \ldots, m$. For each $\ell \in [m]$, introduce a set $M_\ell = \{f_{\ell,1}, f_{\ell,2}, \ldots, f_{\ell,\ell}\}$, and for each node $\bar{u}$ in $\mathcal{C}$, connect $\bar{u}$ to f_{ℓ,u_ℓ} with an edge of

length $\delta_\ell = 1 + (m!)^{-\ell}$. The set of facilities is $\mathcal{F} = \bigcup_{\ell=1}^{m} M_\ell$. All distances between points in $\mathcal{C} \cup \mathcal{F}$ other than those specified above are determined by shortest-path lengths. Since $1 < \delta_\ell \leq 2$ for all ℓ, the resulting distance function satisfies the triangle inequality.

We have $cost(M_j) = m!\delta_j$ for each $j \in [m]$. We prove ($*$) by contradiction. Fix some index j and consider a set $F \subseteq \mathcal{F}$ that does not contain M_ℓ for any $\ell \geq j$: for each $\ell \geq j$ there is $i_\ell \leq \ell$ such that $f_{\ell,i_\ell} \notin F$. Define a customer $\bar{u}$ as follows: $u_i = 1$ for $\ell = 1, \ldots, j-1$ and $u_i = i_\ell$ for $\ell = j, \ldots, m$. Then the facility $f_{\ell,i} \in F$ serving this $\bar{u}$ must have $\ell < j$ or $i \neq i_\ell$. Either way, it is at distance at least δ_{j-1} from $\bar{u}$. Since every other customers pays strictly more than 1, we get $cost(F) > m! - 1 + \delta_{j-1} = m!\delta_j = cost(M_j)$—a contradiction. □

Proof of Theorems 3(b) and Theorem 4(b) Lemma 4 and the lower bounds for online bidding in Theorem 2 immediately imply Theorems 3(b) and 4(b). □

Next we turn attention to polynomial-time algorithms. First, we prove Theorem 5—that there exists a polynomial-time $O(\log m)$-size-approximation algorithm for the offline problem. We will give a polynomial-time algorithm that, given k and a problem instance, finds a facility set of size $O(k \log m)$ and cost at most opt_k. (Here $m = |\mathcal{C}|$ is the number of customers.)

Proof of Theorem 5 Without loss of generality, we assume that each customer has distance 0 to its closest facility. (Otherwise, for each customer u we can subtract $d_{u\mathcal{F}}$ from the distance d_{uf} to each facility. This decreases the cost of each facility set by a uniform amount, so any solution having optimal cost in the modified instance will also have optimal cost in the original instance.)

The algorithm first runs the bicriteria approximation algorithm from [30] that, in time $O(k(m+n)\log(m/\epsilon))$, finds a facility set of size $O(k\log(m+m/\epsilon))$ and cost at most $(1+\epsilon)opt_k$. This algorithm is run with $\epsilon = 1/m$. As a result, we obtain a facility set F of size $O(k\log m)$ and cost at most $(1+1/m)opt_k$. We then greedily add to F a single facility f that minimizes $cost(F \cup \{f\})$. The algorithm returns $F \cup \{f\}$.

To finish, we show that the facility set $F \cup \{f\}$ has size $O(k\log m)$ and cost at most opt_k. The size bound is immediate from the size bound on F. To show the cost bound, note that if we add to F, for the customer contributing the maximum amount to the current cost, the closest facility f to that customer, the cost for that customer would decrease to 0. (Recall our assumption that $d_{u\mathcal{F}} = 0$ for each customer u.) Thus, adding this f decreases the overall cost by at least the current cost times $1/m$, and we get $cost(F \cup \{f\}) \leq (1-1/m)cost(F) \leq (1-1/m)(1+1/m)opt_k \leq opt_k$. □

By the reduction in Lemma 3, this gives a polynomial-time size-approximation algorithm for the incremental problem:

Proof of Theorem 6 By Theorem 1(a), there is a deterministic polynomial-time algorithm for online bidding with competitive ratio $O(1)$. Using this online bidding algorithm, and using Theorem 5 to compute an $O(\log m)$-size-approximate k-median F_k^* for each k, the reduction in Lemma 3 gives an $O(\log m)$-size-competitive deterministic polynomial-time algorithm for incremental medians. □

4 Incremental Cost-Competitive Medians

In this section we prove Theorems 7 and 8.

In the analysis we will use the following fact, whose proof can be found in [12] and is also implicit in [19]. (See also Sects. 5 and 6 for generalizations with proofs.) Given two sets of facilities A and B, let $\Gamma(A, B)$ be a set $C \subseteq B$ of cardinality at most $|A|$ defined as follows. For each facility $g \in A$, choose a single customer $r(g) \in C$ to be defined shortly, and let f be the facility of B serving $r(g)$. We define C to be the set of such f's as g spans A, where $r(g)$ is chosen so as to minimize $d_{gr(g)} + d_{r(g)f}$.

It will be convenient to introduce distances between facilities: given two $f, g \in \mathcal{F}$, let $d_{fg} = \min_{x \in \mathcal{C}}(d_{fx} + d_{xg})$. For $f \in \mathcal{F}$ and $F \subseteq \mathcal{F}$, define the distance from f to F by $d_{fF} = \min_{g \in F} d_{fg}$. In other words, in order to define C, for each $a \in A$ we add to C an element of B closest to a.

Fact 1 *Let $1 \le h < k \le n$, and assume that the distance function is metric. Consider an h-median A and a k-median B, and let $C = \Gamma(A, B)$. Then $|C| \le h$, and C is a subset of B such that for every customer u we have $d_{uC} \le 2d_{uA} + d_{uB}$.*

This implies $cost(C) \le cost(B) + 2cost(A)$.

To prove Theorems 7 and 8, the argument is based on another reduction from incremental medians to online bidding, presented in the lemma below. This implies the theorems.

Lemma 5 *Consider an instance of the metric medians problem. Assume that for each k we have a set of facilities F_k^* of size at most k and cost at most $c \cdot opt_k$. Let $\beta \ge 1$. Assume that we have a (randomized) polynomial-time β-competitive algorithm for online bidding. Then in (randomized) polynomial time we can compute an incremental solution $\bar{F} = (F_1, \ldots, F_n)$, where each F_k has size at most k and (expected) cost at most $2\beta c \cdot opt_k$.*

Proof Without loss of generality $cost(F_k^*) \ge cost(F_{k+1}^*)$ for all $k < n$. The algorithm constructs the incremental solution $\bar{F}$ from $F_1^*, \ldots, F_n^*$ in several steps. First, fix some indices $1 = \kappa(1) < \kappa(2) < \cdots < \kappa(m)$ by a method to be described later, and let $\mathcal{K} \subseteq [n]$ denote the set of indices.

Next, compute the sets $F_{\kappa(i)}$ as follows. $F_{\kappa(m)} = F_{\kappa(m)}^*$. For $i = m - 1, m - 2, \ldots, 1$, inductively define $F_{\kappa(i)}$ to contain the facilities within $F_{\kappa(i+1)}$ that are "closest" to $F_{\kappa(i)}^*$ in the following sense. With the $\Gamma(\cdot, \cdot)$ notation seen above, we define $F_{\kappa(i)} = \Gamma(F_{\kappa(i)}^*, F_{\kappa(i+1)})$. Directly from the definition, $|F_{\kappa(i)}| \le |F_{\kappa(i)}^*| \le \kappa(i)$.

Finally, for indices $k \notin \mathcal{K}$, define $F_k = F_{\kappa(i)}$, where i is maximum such that $\kappa(i) \le k$ (this i is well defined, since $\kappa(1) = 1$.) Obviously, $|F_k| \le \kappa(i) \le k$. To complete the construction, it remains to describe how to compute $\mathcal{K}$, which we momentarily defer.

To analyze the cost, for a given k, let i be maximum such that $\kappa(i) \le k$. From the definition of $F_{\kappa(i)}$ and Fact 1 summed over all $u \in \mathcal{C}$, we have: $cost(F_{\kappa(i)}) \le$

$2cost(F^*_{\kappa(i)}) + cost(F_{\kappa(i+1)})$. Applying this inequality repeatedly gives:

$$cost(F_k) = cost(F_{\kappa(i)}) \leq 2 \sum_{j=i}^{m} cost(F^*_{\kappa(j)}).$$

To complete the proof of the lemma, it suffices to define $\mathcal{K}$ so that

$$\sum_{j=i}^{m} cost(F^*_{\kappa(j)}) \leq \beta cost(F^*_k) \tag{4}$$

since this will imply $cost(F_k) \leq 2\beta cost(F^*_k) \leq 2\beta c \cdot opt_k$.

We now prove (4). Let $\mathcal{U} = \{cost(F^*_k) : 1 \leq k \leq n\}$ and take $\mathcal{B}$ to be the set of bids used by the β-competitive online bidding algorithm for universe $\mathcal{U}$. We define $\mathcal{K} = \{\kappa : cost(F^*_\kappa) \in \mathcal{B}\}$, with ties broken in favor of smaller indices (ties may happen if several sets of facilities have the same cost). Note that with this tie-breaking rule, we always have $1 \in \mathcal{K}$. (If the online bidding algorithm is randomized, then $\mathcal{B}$ and therefore $\mathcal{K}$ are random.) Then the left-hand side of (4) is exactly the sum of the bids paid by the online bidding algorithm for threshold $T = cost(F^*_k)$. Since $\mathcal{B}$ is β-competitive, this cost is at most $\beta cost(F^*_k)$, so (4) holds. This completes the proof. □

The algorithm of Theorem 7(a) is obtained by combining the reduction used in the proof of Lemma 5 with the deterministic 4-competitive online bidding algorithm given in the proof of Theorem 1(a). The resulting algorithm is detailed in the displayed Algorithm 1.
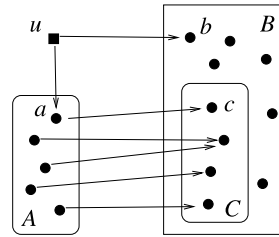
Algorithm 1 The 8-cost-competitive deterministic incremental median algorithm

---

for all $k = 1, \ldots, n$, **do**
    compute a k-median solution F^*_k
end for
$\mathcal{K} \leftarrow \{1\}$
for all $k = 2, \ldots, n$, **do**
    if $cost(F^*_k) = 0 < cost(F^*_{k-1})$ then add k to $\mathcal{K}$
    if $cost(F^*_k) > 0$ then if $\lceil \log cost(F^*_k) \rceil < \lceil \log cost(F^*_{k-1}) \rceil$ then add k to $\mathcal{K}$
end for
$\ell \leftarrow \max(\mathcal{K})$; $F_\ell \leftarrow F^*_\ell$
for all $k = \ell + 1, \ldots, n$ do $F_k \leftarrow F_\ell$
for each $k \in \mathcal{K} - \{\max(\mathcal{K})\}$, in decreasing order **do**
    $F_k \leftarrow \Gamma(F^*_k, F_\ell)$
    for all $k' = k + 1, \ldots, \ell - 1$ do $F_{k'} \leftarrow F_k$
    $\ell \leftarrow k$
end for

---

5 λ-Relaxed Metrics

In this section we prove Theorem 9, namely the upper bounds for cost-competitive ratios when the distance function is a λ-relaxed metric. For $\lambda \geq 1$, by a λ-relaxed metric, we mean a distance function that satisfies $d_{yf} \leq \lambda(d_{xf} + d_{xg} + d_{yg})$ for all $f, g \in \mathcal{F}$ and $x, y \in \mathcal{C}$.

We start with a generalization of Fact 1 from the previous section. Recall that $\Gamma(A, B)$ is the set of up to $|A|$ elements of B that are closest to the elements of A (see the definition in previous section.)

Lemma 6 *Assume that the distance function is a λ-relaxed metric. Given two sets of facilities A and B, let $C = \Gamma(A, B)$. Then for every customer $u \in \mathcal{C}$ we have $d_{uC} \leq 2\lambda d_{uA} + \lambda d_{uB}$.*

Proof For a given $u \in \mathcal{C}$, choose $a \in A$ such that $d_{uA} = d_{ua}$, $b \in B$ such that $d_{uB} = d_{ub}$, and $c \in C$ such that $d_{aB} = d_{ac}$. (See Fig. 1.) By the definition of the λ-relaxed metric, for every customer $x \in \mathcal{C}$ we have $d_{uc} \leq \lambda(d_{ua} + d_{xa} + d_{xc})$. There is $x \in \mathcal{C}$ for which $d_{ac} = d_{xa} + d_{xc}$, and choosing this x we get $d_{uc} \leq \lambda(d_{ua} + d_{ac})$. Thus:

$$d_{uC} \leq d_{uc} \leq \lambda(d_{ua} + d_{ac}) \leq \lambda(d_{ua} + d_{ab}) \leq \lambda(d_{ua} + (d_{ua} + d_{ub}))$$
$$= 2\lambda d_{uA} + \lambda d_{uB},$$

completing the proof.   □

Proof of Theorem 9 The algorithm used to prove Theorem 9 is the same as in the metric case (see the proof of Lemma 5), except for the definition of $\mathcal{B}$. Let $\mathcal{U} = \{cost(F_k^*) : 1 \leq k \leq n\}$.

$\mathcal{B}$ contains 0 iff $\mathcal{U}$ does, plus the following elements: in the deterministic case, for every $i \in \mathbb{Z}$, $\mathcal{B}$ contains the maximum element in the set $\mathcal{U} \cap [0, (2\lambda)^i]$ (if it is non-empty); in the randomized case, pick a real number $\xi \in (0, 1]$ uniformly at random, and for every $i \in \mathbb{Z}$, $\mathcal{B}$ contains the maximum element in the set $\mathcal{U} \cap [0, e^\xi (e\lambda)^i]$ (if it is non-empty).

The analysis is similar to the proof of Theorem 7. Let $\mathcal{K} = \{\kappa(i) : 1 \leq i \leq |\mathcal{B}|\}$ be defined from $\mathcal{U}$ and $\mathcal{B}$ as in the proof of Theorem 7. Choose i to be the maximum index such that $\kappa(i) \leq k$. From Lemma 6 summed over $u \in \mathcal{C}$, we have: $cost(F_{\kappa(i)}) \leq$

$2\lambda cost(F^*_{\kappa(i)}) + \lambda cost(F_{\kappa(i+1)})$. Applying this inequality repeatedly gives

$$cost(F_k) = cost(F_{\kappa(i)}) \leq 2\lambda \sum_{j=i}^{m} \lambda^{j-i} cost(F^*_{\kappa(j)}).$$

To continue the analysis in the deterministic case, let p be such that $(2\lambda)^{p-1} < cost(F^*_{\kappa(i)}) \leq (2\lambda)^p$. By the definition of $\kappa(i)$, we have:

$$cost(F_k) \leq 2\lambda \sum_{j \geq i} \lambda^{j-i}(2\lambda)^{p-j+i} = 2(2\lambda)(2\lambda)^p \leq 8\lambda^2 cost(F^*_k),$$

since $k < \kappa(i+1)$ and so $cost(F^*_k) \geq (2\lambda)^{p-1}$. Hence our algorithm is $8\lambda^2$-cost-competitive when F^*_k is the optimal k-median and is $8\lambda^2 c$-cost-competitive when F^*_k is a c-approximation.

In the randomized case, let p such that $e^\xi (e\lambda)^{p-1} < cost(F^*_{\kappa(i)}) \leq e^\xi (e\lambda)^p$. By the definition of $\kappa(i)$, we have:

$$cost(F_k) \leq 2\lambda e^\xi \sum_{j \geq i} \lambda^{j-i}(e\lambda)^{p-j+i} = \frac{1}{1-1/e} 2\lambda e^\xi (e\lambda)^p.$$

As in the proof of Theorem 1(b), the expected value of $e^\xi (e\lambda)^p / cost(F^*_k)$ is distributed like $e^\xi \lambda$ where ξ is uniform in $(0, 1]$, so

$$E[cost(F_k)] \leq \frac{1}{1-1/e}(2\lambda) cost(F^*_k)\lambda(e-1) = 2e\lambda^2 cost(F^*_k),$$

and the theorem follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

6 Incremental Fractional Medians

In this section we prove Theorem 11.

A fractional k-median is given by a feasible solution to the following linear program relaxation of the standard k-median integer linear program:

$$\text{minimize}_{x,y} \sum_{f \in \mathcal{F}, u \in \mathcal{C}} d_{uf} x_{uf} \quad \text{subject to} \begin{cases} \sum_{f \in \mathcal{F}} x_{uf} = 1 & \forall u \in \mathcal{C}, \\ x_{uf} \leq y_f & \forall u \in \mathcal{C}, f \in \mathcal{F}, \\ \sum_{f \in \mathcal{F}} y_f \leq k, \\ x_{uf} \geq 0 & \forall u \in \mathcal{C}, f \in \mathcal{F}. \end{cases}$$

In this section we let x denote a fractional median. The *capacity required by x at f* is defined as $|x|_f = \max_{u \in \mathcal{C}} x_{uf}$ (essentially, the value of y_f in the linear program). The total *capacity required by x* is $|x| = \sum_{f \in \mathcal{F}} |x|_f$, and the *cost of x* is the objective function value $cost(x) = \sum_{f \in \mathcal{F}, u \in \mathcal{C}} d_{uf} x_{uf}$. Recall that for two fractional medians x, x', we say that x *dominates* x' if $|x'|_f \leq |x|_f$ for all $f \in \mathcal{F}$.

Lemma 7 *Let $1 \leq h \leq k \leq n$, and assume that the distance function is metric. Consider two fractional medians, a fractional k-median x and a fractional h-median z. There exists a fractional h-median x' dominated by x such that $cost(x') \leq cost(x) + 2 \cdot cost(z)$.*

Proof Construct x' as follows. For each facility $g \in \mathcal{F}$, choose (by a method to be described later) a single customer $r(g) \in \mathcal{C}$ "responsible" for g. For each $u \in \mathcal{C}$ and $f \in \mathcal{F}$, take $x'_{uf} = \sum_{g \in \mathcal{F}} z_{ug} x_{r(g)f}$.

For each $u \in \mathcal{C}$, applying the first constraint of the above linear program, we have $\sum_{f \in \mathcal{F}} x'_{uf} = \sum_{f,g \in \mathcal{F}} z_{ug} x_{r(g)f} = \sum_{g \in \mathcal{F}} (\sum_{f \in \mathcal{F}} x_{r(g)f}) z_{ug} = \sum_{g \in \mathcal{F}} z_{ug} = 1$, so x' is a valid fractional median (for some capacity value).

Next we prove that $|x'|$ is a fractional h-median. The intuition is that "routing" x' through z (as described later) ensures this. To prove that x' is a h-median, we bound the total capacity required by x':

$$|x'| = \sum_{f \in \mathcal{F}} \max_{u \in \mathcal{C}} \sum_{g \in \mathcal{F}} z_{ug} x_{r(g)f} \leq \sum_{g,f \in \mathcal{F}} \max_{u \in \mathcal{C}} z_{ug} x_{r(g)f}$$

$$= \sum_{g \in \mathcal{F}} \left(\sum_{f \in \mathcal{F}} x_{r(g)f} \right) \max_{u \in \mathcal{C}} z_{ug} = \sum_{g \in \mathcal{F}} \max_{u \in \mathcal{C}} z_{ug} = |z| \leq h.$$

We now claim that the fractional median x' is dominated by x. Indeed, fixing an $f \in \mathcal{F}$, for any $u \in \mathcal{C}$ we have $x'_{uf} = \sum_{g \in \mathcal{F}} z_{ug} x_{r(g)f} \leq \sum_{g \in \mathcal{F}} z_{ug} \cdot \max_{w \in \mathcal{C}} x_{wf} \leq \max_{w \in \mathcal{C}} x_{wf} = |x|_f$; therefore $|x'|_f \leq |x|_f$, as claimed.

To finish the proof, we show that we can choose the responsible customers so that $cost(x') \leq 2cost(z) + cost(x)$. Consider choosing $r(g)$ randomly for each $g \in \mathcal{F}$ so that $\Pr[r(g) = w] = z_{wg} / \sum_{u \in \mathcal{C}} z_{ug}$ (if the denominator is zero, choose $r(g)$ arbitrarily). Now bound the expected cost of x' as follows. Imagine "routing" one unit of weight from each customer $u \in \mathcal{C}$ to the facilities in stages: (stage 1) send z_{ug} units from each $u \in \mathcal{C}$ to each $g \in \mathcal{F}$; (stage 2) from each $g \in \mathcal{F}$, send all arriving weight to $r(g) \in \mathcal{C}$; (stage 3) from each $w \in \mathcal{C}$, split all arriving weight and send an x_{wf} fraction to each $f \in \mathcal{F}$.

For each $u \in \mathcal{U}$ and $f \in \mathcal{F}$, an easy calculation shows that for every choice of $r(g)$, x'_{uf} units of the weight that leaves u at the start end up at f at the end.

Since $d_{uf} \leq d_{ug} + d_{r(g)g} + d_{r(g)f}$, the cost of x' is at most the sum of the costs of the stages, where the cost of sending weight between any $u \in \mathcal{C}$ and any $f \in \mathcal{F}$ in a stage is d_{uf} per unit.

The first stage costs $\sum_{u \in \mathcal{C}, g \in \mathcal{F}} d_{ug} z_{ug} = cost(z)$.

In the second stage, for each $g \in \mathcal{F}$, the total weight to be sent is $\sum_{u \in \mathcal{C}} z_{ug}$. Therefore, with the random choice of $r(g)$, the expected weight sent to any given $w \in \mathcal{C}$ is $(\sum_{u \in \mathcal{C}} z_{ug}) \Pr[r(g) = w] = z_{wg}$ (using the definition of $r(g)$). Thus, the expected cost of the second stage is $\sum_{g \in \mathcal{F}, w \in \mathcal{C}} d_{wg} z_{wg} = cost(z)$.

In the third stage, for each $w \in \mathcal{C}$, the expected weight to be split and sent on is $\sum_{g \in \mathcal{F}} z_{wg} = 1$ (using from above that the expected weight sent from $g \in \mathcal{C}$ to w in the second stage is z_{wg}). The fraction of this weight sent to each $f \in \mathcal{F}$ is x_{wf}, so the expected weight sent from w to f is x_{wf}. Thus, the expected cost of the third stage is $\sum_{w \in \mathcal{C}, f \in \mathcal{F}} d_{wf} x_{wf} = cost(x)$.

In sum, the expected total cost of the stages is at most $2cost(z) + cost(x)$. Since the cost of the stages is an upper bound on the expectation of $cost(x')$, we conclude $E[cost(x')] \leq 2cost(z) + cost(x)$. So there is some way to choose the responsible customers so that $cost(x') \leq 2cost(z) + cost(x)$. □

Proof of Theorem 11 With the lemma in place, the proof of the theorem is essentially the same as the proof of Theorems 7 and 8, along with a minor observation about fractional strategies being closed under randomization. More precisely, recall that a c-cost-competitive incremental fractional median gives, for every integer $k \in [n]$, a fractional k-median x^k with $cost(x^k) \leq c \cdot \overline{opt}_k$, such that x^k is dominated by x^{k+1} for all $k < n$.

We first show that a c-cost-competitive fractional median with the minimum ratio c can be computed in polynomial time using linear programming as follows. For each $k \in [n]$, compute $\overline{opt}_k$ (using linear programming). Then, set up a linear program that has the following constraints: (i) the fractional k-median constraints for each x^k, (ii) the incrementality constraints saying that each x^{k+1} dominates x^k for $k < n$, and (iii) the cost-competitiveness constraints saying that $cost(x^k) \leq c \cdot \overline{opt}_k$, for each k. The objective function is to minimize c. The solution of this linear program is a c-cost-competitive incremental fractional median with minimum c.

It remains to show that $c \leq 2e$, by proving that there exists a $2e$-cost-competitive fractional incremental median. We first show that $c \leq 8$, using the existence of a deterministic 4-competitive online bidding algorithm.

First, fix some indices $1 = \kappa(1) < \kappa(2) < \cdots < \kappa(m)$ by a method to be described later, and let $\mathcal{K} \subseteq [n]$ denote this set of indices.

Next, compute the fractional medians x^k as follows. Let z^k denote an optimal fractional k-median. We initialize $x^{\kappa(m)} = z^{\kappa(m)}$. Then, for $i = m - 1, m - 2, \ldots, 1$, inductively define $x^{\kappa(i)}$ to be a minimum-cost fractional $\kappa(i)$-median among those dominated by $x^{\kappa(i+1)}$. Finally, for indices $k \notin \mathcal{K}$, define $x^k = x^{\kappa(i)}$, where i is maximum such that $\kappa(i) \leq k$. To complete the construction, it remains to describe how to compute $\mathcal{K}$, which we momentarily defer.

To analyze the required capacity, note that for $\kappa(i) \leq k < \kappa(i + 1)$ we have $|x^k| = |x^{\kappa(i)}| \leq \kappa(i) \leq k$, and thus x^k is indeed a fractional k-median.

To analyze the cost, for a given k, let i be maximum such that $\kappa(i) \leq k$. From Lemma 7 we have $cost(x^{\kappa(i)}) \leq cost(x^{\kappa(i+1)}) + 2 \cdot \overline{opt}_{\kappa(i)}$. Applying this inequality repeatedly gives:

$$cost(x^k) = cost(x^{\kappa(i)}) \leq 2 \sum_{j=i}^{m} \overline{opt}_{\kappa(j)}. \tag{5}$$

To complete the proof, it suffices to define $\mathcal{K}$ so that

$$\sum_{j=i}^{m} cost(x^{\kappa(j)}) \leq \beta cost(z^k) \tag{6}$$

since this will imply $cost(x^k) \leq 2\beta cost(z^k) \leq 2\beta c \cdot opt_k$.

We will now prove (6) for $\beta = 4$. Let $\mathcal{U} = \{\overline{opt}_k : 1 \leq k \leq n\}$ and take $\mathcal{B}$ to be the set of bids used by the 4-competitive online bidding algorithm for universe $\mathcal{U}$.

Let $\mathcal{K} = \{\kappa(i) : 1 \le i \le m\}$ be a minimal set (containing 1) such that $\mathcal{B} = \{\overline{opt}_{\kappa(i)} : 1 \le i \le m\}$ (breaking ties in favor of smaller indices). Then the left hand side of (6) is exactly the sum of the bids paid by the online bidding algorithm for threshold $T = cost(z^k)$. Since $\mathcal{B}$ is β-competitive, this cost is at most $\beta cost(z^k)$, so (6) holds. This completes the proof of 8-competitiveness.

To improve this ratio to $2e$, consider carrying out the above construction using the e-competitive randomized online bidding algorithm. That algorithm generates a random bidding set $\mathcal{B}$ for our universe $\mathcal{U}$. Applying the construction above gives us a random incremental fractional solution $(x^k)_k$. For each k, f and u, take $\bar{x}^k_{uf}$ to be the expectation of x^k_{uf} for this random solution. This gives an incremental fractional solution which, for each k, has $|\bar{x}| \le k$, and whose cost is equal to the expected cost of the random solution. Thus, $(\bar{x}^k)_k$ is $2e$-cost-competitive.

Summarizing, we showed that there exists a $2e$-cost-competitive fractional incremental median. As the algorithm given earlier computes a c-cost-competitive fractional incremental median that minimizes c, the theorem follows. $\square$

7 Incremental Algorithms for kl-Medians

In this section we prove Theorem 12. Recall that in the kl-median problem, for given $1 \le k < l \le n$, we need to compute two facility sets $F_k \subseteq F_l$ with $|F_k| = k$ and $|F_l| = l$, minimizing the cost-competitive ratio

$$c = \max \left\{ \frac{cost(F_k)}{opt_k}, \frac{cost(F_l)}{opt_l} \right\}.$$

We now prove that the optimal ratio c for this problem is between $2 - 1/(l - k + 1)$ and $2 - 1/l$.

Proof of Theorem 12(a) We start with the upper bound proof. Our method here is very different from the previous bounds in this paper and it does not rely on online bidding. Let F and G denote, respectively, the optimum k-median and the optimum l-median. Without loss of generality, we can assume that $F \cap G = \emptyset$, for otherwise we can duplicate the facilities in $F \cap G$. Our algorithm chooses the better of two options below (the one with the better ratio):

(i) $F_k = F$, and F_l is a set with minimum $cost(F_l)$, such that $|F_l| = l$ and $F \subseteq F_l \subseteq \mathcal{F}$, or

(ii) F_k is a set with minimum $cost(F_k)$ such that $|F_k| = k$ and $F_k \subseteq G$, and $F_l = G$.

We now show that this algorithm's competitive ratio is at most $2 - 1/l$. It is sufficient to show that there exists a k-element set $X \subseteq G$ such that

$$cost(X) + cost(F \cup G - X) \le (2 - 1/l)[cost(F) + cost(G)]. \qquad (7)$$

Indeed, Inequality (7) implies that at least one of the following two options must hold: either $cost(X) \le (2 - 1/l)cost(F)$, or $cost(F \cup G - X) \le (2 - 1/l)cost(G)$. In

the first case our algorithm can choose option (ii), and then we will have $cost(F_k) \leq cost(X) \leq (2 - 1/l)cost(F)$. In the second case it can choose option (i), and then $cost(F_l) \leq cost(F \cup G - X) \leq (2 - 1/l)cost(G)$. In both cases, the algorithm's ratio is at most $2 - 1/l$.

It remains to show that there is a k-element set $X \subseteq G$ that satisfies (7). Our proof is based on a probabilistic argument. We start with some notation. For each $f \in F$ and $g \in G$, denote by C_{fg} the set of customers that are served by f when F is the facility set and by g when G is the facility set. Let w_{fg} be the cardinality of C_{fg}. By a_{fg} (resp. b_{fg}) we denote the average distance between a customer $x \in C_{fg}$ and f (resp. g.) Formally, $a_{fg} = \sum_{u \in C_{fg}} d_{uf}/|C_{fg}|$ and $b_{fg} = \sum_{u \in C_{fg}} d_{ug}/|C_{fg}|$. It is convenient to think of C_{fg} as a single point with weight w_{fg} whose distances to f and g are a_{fg} and b_{fg}, respectively. The costs of F and G can then be written as $cost(F) = \sum_{f \in F, g \in G} w_{fg} a_{fg}$, and $cost(G) = \sum_{f \in F, g \in G} w_{fg} b_{fg}$.

We now define a probability distribution on k-element subsets of G. Let $w_{fG} = \sum_{g \in G} w_{fg}$. For each $f \in F$ and $g \in G$, define $\bar{w}_{fg} = w_{fg}(w_{fG} - w_{fg})^{-1}$. (Assume for now that $w_{fg} > 0$ for all f, g. We will explain later how to extend the argument to the general case.) Choose a random mapping $\pi : F \to G$ as follows: for any $f \in F$ set $\pi(f) = g$ with probability $\bar{w}_{fg}/\sum_{h \in G} \bar{w}_{fh}$. For any such mapping π let X_π be the k-element subset of G that consists of $\pi(F)$ and arbitrary $k - |\pi(F)|$ elements of $G - \pi(F)$. (Intuitively, we would like to take our random k-set $X \subseteq G$ to be $\pi(F)$, but then $F \cup G - \pi(F)$ may have cardinality larger than l and will not be a valid l-median. This is why we add these additional elements to X_π.)

Consider some $f \in F$ and $h \in G$. The cost of serving C_{fh} from $F \cup G - X_\pi$ is at most $w_{fh} a_{fh}$. If $h \notin X$, we can also bound this cost by $w_{fh} b_{fh}$. We also want to estimate the cost of serving C_{fh} from X_π. If $h \in X_\pi$ then this cost is at most $w_{fh} b_{fh}$. If $h \in G - X_\pi$ then, for any $g \in X_\pi$, the cost of serving C_{fh} from X_π is bounded by the cost of serving C_{fg} from g, and thus (by the triangle inequality) it is at most $w_{fh} a_{fh} + w_{fh} \min_{u \in C_{fg}}(d_{uf} + d_{ug}) \leq w_{fh}(a_{fh} + a_{fg} + b_{fg})$. Using $g = \pi(f)$ and summing over all $f \in F$ and $h \in G$, we get

$$
\begin{aligned}
cost(X_\pi) + cost(F \cup G - X_\pi) \\
\leq \sum_{f \in F} \Bigg[\sum_{h \in G - X_\pi} w_{fh}(a_{fh} + a_{f\pi(f)} + b_{f\pi(f)} + b_{fh}) \\
+ \sum_{h \in X_\pi} w_{fh}(a_{fh} + b_{fh}) \Bigg] \\
= cost(F) + cost(G) + \sum_{f \in F} \sum_{h \in G - X_\pi} w_{fh}(a_{f\pi(f)} + b_{f\pi(f)}) \\
\leq cost(F) + cost(G) + \sum_{f \in F}(w_{fG} - w_{f\pi(f)})(a_{f\pi(f)} + b_{f\pi(f)}).
\end{aligned}
$$

Now, by the linearity of expectation, we have

$$E[cost(X_\pi) + cost(F \cup G - X_\pi)]$$

$$\leq cost(F) + cost(G) + \sum_{f \in F} \sum_{g \in G} \frac{\bar{w}_{fg}}{\sum_{h \in G} \bar{w}_{fh}} (w_{fG} - w_{fg})(a_{fg} + b_{fg})$$

$$= cost(F) + cost(G) + \sum_{f \in F} \frac{1}{\sum_{h \in G} \bar{w}_{fh}} \sum_{g \in G} w_{fg}(a_{fg} + b_{fg})$$

$$\leq (2 - 1/l)[cost(F) + cost(G)].$$

The last inequality holds because, for each $f \in F$, $\sum_{h \in G} \bar{w}_{fh}$ is minimized when $w_{fh} = w_{fG}/l$ for all $h \in G$, and thus $(\sum_{h \in G} \bar{w}_{fh})^{-1} \leq 1 - 1/l$. We conclude that there exists a set X_π that satisfies $(*)$, as claimed.

To complete the proof, we still need to extend the argument to the general case, when some weights w_{fg} are zero. Suppose first that we allow arbitrary positive weights. Choose an arbitrarily small $\epsilon > 0$ and set all zero weights w_{fg} to ϵ instead. The earlier argument implies that for each ϵ there is X that satisfies $(*)$. Since there are finitely many choices for X, there is X that satisfies $(*)$ for infinitely many values of ϵ. The continuity of $cost(X) + cost(F \cup G - X)$ with respect to ϵ implies that $(*)$ holds for $\epsilon = 0$ as well.

In our case the weights are integer, so we cannot use arbitrarily small weights. Instead, we create a large number of copies of each customer (same for each), and then add one customer to each empty set C_{fg}. Then the same asymptotic argument as above applies.   □

Proof of Theorem 12(b) Our lower bound is a slight refinement of the one in [27, 28]. Let $k = 1$ and $1 < l \leq n$. Consider the metric space M with customers $v_1, \ldots, v_l$ and facilities $f_1, g_1, g_2, \ldots, g_l$. Each customer v_j is connected to facility g_j by an edge of length $\delta = 1/l$. All customers are also connected to facility f_1 with edges of length 1. (See Fig. 2.) All other distances are measured along the above defined edges.

Let $G = \{g_1, \ldots, g_l\}$. We have $cost(f_1) = l$ and $cost(G) = l\delta$. Further, for each i, we have $cost(g_i) = \delta + (l - 1)(2 + \delta)$ and $cost(G - \{g_i\} \cup \{f_1\}) = (l - 1)\delta + 1$. Thus, substituting $\delta = 1/l$, we get $cost(g_i)/cost(f_1) = 2 - 1/l$ and $cost(G - \{g_i\} \cup \{f_1\})/cost(G) = 2 - 1/l$ for all i. So, if an incremental algorithm chooses $F_1 = \{g_i\}$, for some i, the ratio is at least $2 - 1/l$. On the other hand, if it chooses $F_1 = \{f_1\}$, then $F_l = G - \{g_i\} \cup \{f_1\}$, for some i, and the ratio is again at least $2 - 1/l$.

This completes the lower bound proof for $k = 1$. For $1 < k < l$, we use the above construction with $l' = l - k + 1$ instead of l. We add $k - 1$ facilities $f_2, \ldots, f_k$ at a very large distance from the above space and each other, and $k - 1$ customers

Fig. 2 The metric space in the lower bound. Facilities are represented by squares and customers by circles
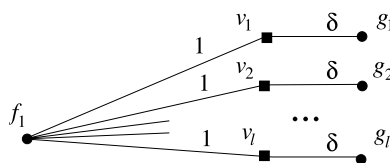
$v_{l+1} = f_2, \ldots, v_{l+k-1} = f_k$. Any k-median must include $f_2, \ldots, f_k$, and thus the argument above applies to this new space. □

References

1. Archer, A., Rajagopalan, R., Shmoys, D.B.: Lagrangian relaxation for the k-median problem: new insights and continuity properties. In: Proc. 11th European Symp. on Algorithms (ESA), pp. 31–42 (2003)
2. Arya, V., Garg, N., Khandekar, R., Meyerson, A., Munagala, K., Pandit, V.: Local search heuristic for k-median and facility location problems. In: Proc. 33rd Symp. Theory of Computing (STOC), pp. 21–29. ACM, New York (2001)
3. Arya, V., Garg, N., Khandekar, R., Meyerson, A., Munagala, K., Pandit, V.: Local search heuristics for k-median and facility location problems. SIAM J. Comput. **33**(3), 544–562 (2004)
4. Buchbinder, N., Naor, J.: Improved bounds for online routing and packing via a primal-dual approach. In: Proc. 46th Symp. Foundations of Computer Science (FOCS), pp. 293–304 (2006)
5. Chakrabarti, S., Phillips, C.A., Schulz, A.S., Shmoys, D.B., Stein, C., Wein, J.: Improved scheduling algorithms for minsum criteria. In: Automata, Languages and Programming, pp. 646–657 (1996)
6. Charikar, M., Chekuri, C., Feder, T., Motwani, R.: Incremental clustering and dynamic information retrieval. In: Proc. 29th Symp. Theory of Computing (STOC), pp. 626–635. ACM, New York (1997)
7. Charikar, M., Guha, S.: Improved combinatorial algorithms for the facility location and k-median problems. In: Proc. 40th Symp. Foundations of Computer Science (FOCS), pp. 378–388. IEEE, New York (1999)
8. Charikar, M., Guha, S.: Improved combinatorial algorithms for facility location problems. SIAM J. Comput. **34**(4), 803–824 (2005)
9. Charikar, M., Guha, S., Tardos, E., Shmoys, D.B.: A constant-factor approximation algorithm for the k-median problem. In: Proc. 31st Symp. Theory of Computing (STOC), pp. 1–10. ACM, New York (1999)
10. Chekuri, C., Goel, A., Khanna, S., Kumar, A.: Multi-processor scheduling to minimize flow time with ϵ-resource augmentation. In: Proc. 36th Symp. Theory of Computing (STOC), pp. 363–372. ACM, New York (2004)
11. Chrobak, M., Kenyon, C., Noga, J., Young, N.: Online medians via online bidding. In: Proc. 7th Latin American Theoretical Informatics Symp. (LATIN). Lecture Notes in Computer Science, vol. 3887, pp. 311–322. Springer, Berlin (2006)
12. Chrobak, M., Kenyon, C., Young, N.E.: The reverse greedy algorithm for the k-median problem. Inf. Process. Lett. **97**, 68–72 (2006)
13. Dasgupta, S., Long, P.M.: Performance guarantees for hierarchical clustering. J. Comput. Syst. Sci. **70**(4), 555–569 (2005)
14. Fagin, R., Stockmeyer, L.: Relaxing the triangle inequality in pattern matching. Int. J. Comput. Vis. **30**, 219–231 (1998)
15. Goemans, M., Kleinberg, J.: An improved approximation ratio for the minimum latency problem. In: Proc. 7th Symp. on Discrete Algorithms (SODA), pp. 152–158. ACM/SIAM, New York (1996)
16. Goemans, M., Kleinberg, J.: An improved approximation ratio for the minimum latency problem. Math. Program. **82**(1), 111–124 (1998)
17. Jain, K., Mahdian, M., Markakis, E., Saberi, A., Vazirani, V.V.: Greedy facility location algorithms analyzed using dual fitting with factor-revealing lp. J. ACM **50**, 795–824 (2003)
18. Jain, K., Mahdian, M., Saberi, A.: A new greedy approach for facility location problems. In: Proc. 34th Symp. Theory of Computing (STOC), pp. 731–740. ACM, New York (2002)
19. Jain, K., Vazirani, V.V.: Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation. J. ACM **48**, 274–296 (2001)
20. Kalyanasundaram, B., Pruhs, K.: Speed is as powerful as clairvoyance. J. ACM **47**, 214–221 (2000)
21. Kao, M.-Y., Reif, J.H., Tate, S.R.: Searching in an unknown environment: An optimal randomized algorithm for the cow-path problem. Inf. Comput. **131**(1), 63–80 (1996). Preliminary version appeared in the Proceedings of the Symp. on Discrete Algorithms, Austin, TX, January 1993

22. Korupolu, M.R., Greg Plaxton, C., Rajaraman, R.: Analysis of a local search heuristic for facility location problems. J. Algorithms **37**, 146–188 (2000)
23. Koutsoupias, E.: Weak adversaries for the k-server problem. In: Proc. 40th Symp. Foundations of Computer Science (FOCS), pp. 444–449. IEEE, New York (1999)
24. Lin, G., Nagarajan, C., Rajamaran, R., Williamson, D.P.: A general approach for incremental approximation and hierarchical clustering. In: Proc. 17th Symp. on Discrete Algorithms (SODA). ACM/SIAM, New York (2006)
25. Lin, J.-H., Vitter, J.S.: Approximation algorithms for geometric median problems. Inf. Process. Lett. **44**, 245–249 (1992)
26. Lin, J.-H., Vitter, J.S.: ϵ-approximations with minimum packing constraint violation (extended abstract). In: Proc. 24th Symp. Theory of Computing (STOC), pp. 771–782. ACM, New York (1992)
27. Mettu, R.R., Greg Plaxton, C.: The online median problem. In: Proc. 41st Symp. Foundations of Computer Science (FOCS), pp. 339–348. IEEE, New York (2000)
28. Mettu, R.R., Greg Plaxton, C.: The online median problem. SIAM J. Comput. **32**, 816–832 (2003)
29. Motwani, R., Phillips, S., Torng, E.: Nonclairvoyant scheduling. Theor. Comput. Sci. **130**(1), 17–47 (1994)
30. Young, N.E.: K-medians, facility location, and the Chernoff–Wald bound. In: Proc. 11th Symp. on Discrete Algorithms (SODA), pp. 86–95. ACM/SIAM, New York (2000)