Frequency Capping in Online Advertising

(Extended Abstract)

Niv Buchbinder[1], Moran Feldman[2⋆], Arpita Ghosh[3], and Joseph (Seffi) Naor[2⋆⋆]

[1] Open University, Israel
niv.buchbinder@gmail.com
[2] Computer Science Dept., Technion, Haifa, Israel
{moranfe,naor}@cs.technion.ac.il
[3] Yahoo! Research, Santa Clara, CA
arpita@yahoo-inc.com

Abstract. We study the following online problem. Each advertiser a_i has a value v_i, demand d_i, and *frequency cap* f_i. Supply units arrive online, each one associated with a user. Each advertiser can be assigned at most d_i units in all, and at most f_i units from the same user. The goal is to design an online allocation algorithm maximizing total value. We first show a deterministic upper bound of 3/4-competitiveness, even when all frequency caps are 1, and all advertisers share identical values and demands. A competitive ratio approaching $1 - 1/e$ can be achieved via a reduction to a model with arbitrary decreasing valuations [GM07]. Our main contribution is analyzing two 3/4-competitive greedy algorithms for the cases of equal values, and arbitrary valuations with equal demands. Finally, we give a primal-dual algorithm which may serve as a good starting point for improving upon the $1 - 1/e$ ratio.

1 Introduction

Display advertising, consisting of graphic or text-based ads embedded in webpages, constitutes a large portion of the revenue from Internet advertising, totaling billions of dollars in 2008. Display, or brand, advertising is typically sold by publishers or ad networks on a pay-per-impression basis, with the advertiser specifying the total number of impressions she wants (the demand) and the price she is willing to pay per impression.[4]

Since display ads are sold on a pay-per-impression rather than on a pay-per-click or pay-per-action basis, *effective delivery* of display ads is very important to maximize advertiser value — each impression that an advertiser pays for must

⋆ Moran Feldman is a recipient of the Google Europe Fellowship in Market Algorithms, and this research is supported in part by this Google Fellowship.

⋆⋆ This work was supported in part by ISF grant 1366/07 and the Google Inter-university center for Electronic Markets and Auctions.

[4] In contrast, sponsored search advertisers typically pay per click or per action, and usually have budgets, rather than demands, or quotas, on the number of impressions.

be shown to *as valuable a user* as possible. One aspect of effectively delivering display ads, which has been widely studied, is good targeting — matching ads to users who are likely to be responsive to the content of the ad. Another very important, but less studied, aspect is *limiting user exposure* to an ad - displaying the same ad to a user multiple times diminishes value to the advertiser, since the incremental benefit from repeatedly displaying the same ad to a user is likely to be small (a user is unlikely to react to an ad after he has seen it a few times).

The notion of limiting the number of times a user is exposed to a particular ad is called *frequency capping* [19], and is often cited as a way to avoid banner ad burnout. That is, frequency capping prevents ads from being displayed repeatedly to the point where visitors are being overexposed and response drops.[5] Serving frequency capped ads is a very real requirement to maximize value delivered to display advertisers, particularly in the pay-per-impression structure of the display advertising marketplace. This is recognized by a number of publishers and ad networks (for instance, RightMedia, Google and Yahoo!) who already offer or implicitly implement frequency capping for their display advertisers.

Serving display ads subject to a frequency capping constraint poses an *online assignment* problem since the supply of users, or impressions, is not known to the ad server in advance. How should the ad server allocate impressions to advertisers in this setting? In this paper, we study the simplest abstractions of the assignment problems motivated by frequency capping.

Problem Statement. There are n advertisers. Advertiser i has value per impression v_i, which is the price she is willing to pay for an impression, and a demand d_i, which is the maximum number of impressions she is interested in. In addition, she also has a *frequency cap* f_i, which is the maximum number of times her ad can be displayed to the same user. That is, an advertiser pays v_i only for impressions from users who have not seen her ad more than f_i times. The set of advertisers, and their parameters, is known to the ad server in advance.

Impressions from users arrive online. We say an advertiser is *eligible* for an impression if she still has leftover demand, and has not yet exhausted her frequency cap for the user associated with this impression. When an impression arrives, the ad server must immediately decide which ad, among the set of eligible advertisers, to display for that impression. The total revenue obtained by an algorithm is the sum of the revenues from each impression it allocates. We want to design algorithms that are competitive against the optimal offline allocation, which knows the supply of impressions (with their associated users) in advance. (This problem is captured by the model in [14], see §1.1.)

In the absence of the frequency capping constraint ($f_i = \infty$), the natural greedy algorithm, which assigns each arriving impression to the advertiser with the highest per-impression value v_i, is optimal. However, with the frequency capping constraint, the ad server faces a tradeoff between assigning an arriving impression to an advertiser with high v_i but large frequency cap (since the

[5] While it might be argued that multiple displays of an ad to a user reinforces its message, repeated display without an upper limit clearly diminishes value.

supply can stop anytime) and a lower value advertiser with a smaller frequency cap (since small f_i means this advertiser needs to be assigned to many distinct users). In fact, even when all advertisers have identical values (with arbitrary tie breaking), the greedy algorithm is not optimal, as the following example shows: there are two advertisers, the first with $v_1 = 1, f_1 = n$, and the second with $v_2 = 1 - \epsilon$ and $f_2 = 1$; both advertisers have demand n (the $1 - \epsilon$ is used for tie breaking). The sequence of users is $u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+1}$, where the last user appears n times (n impressions). The greedy allocation gets a value of $n+1$, whereas the optimal offline allocation gets $2n$.

As the next example shows, however, it is not even the different frequency caps that lead to the suboptimality of the greedy algorithm: suppose there are $n + 1$ advertisers each with $f_i = 1$. The first n advertisers have value 1 and demand 1, and the last advertiser has value $1 - \epsilon$ and demand n. With the same arrival sequence of users, a greedy allocation, again, gets a value of $n+1$, whereas the optimal value is $2n$. In fact, as we will show in §3, even when *all* values and demands are equal and *all* frequency caps are 1, no deterministic algorithm can have a competitive ratio better than 3/4.

Distinction from Online Matching. Finding a matching in a bipartite graph where one side is known and the other side is exposed one vertex at a time is known as *online matching*. While the problem of online allocation with frequency capping constraints appears to be similar to online matching, they are actually quite different. In the frequency capping problem, a-priori each impression can be assigned to any of the advertisers. Now, as the impressions arrive, in the language of online matching, the existence of an edge between an advertiser and an arriving impression *depends* on the previous assignments made by the algorithm because of the frequency capping constraint. Specifically, if the algorithm has already assigned enough impressions from user j to advertiser i, or has exhausted i's demand, there is no edge between advertiser i and a newly arrived impression; otherwise, there is an edge. This means that an adversary can no longer control the set of edges hitting each new impression; instead, the online algorithm determines the set of edges using indirect means. While we expect this property to translate into better competitive ratios for the frequency capping problem, taking advantage of the difference is not easy, a fact which is demonstrated by the involved analysis for the natural greedy algorithm for the problem.

Results. Our online assignment problem can be stated abstractly as follows: There are n agents, each with a total demand d_i, and a value v_i for items. Items of different types arrive one by one in an online fashion and must be allocated to an agent immediately. Agent i wants at most f_i copies of any single type of item. How should an online algorithm assign each arriving item to agents to maximize value? This abstract statement suggests the following simpler questions.

- *Equal values, arbitrary d_i, f_i:* Suppose agents (advertisers) have identical values for items (impressions), that is, $v_i = 1$ for all i. Now, the goal of the online algorithm is simply to assign as many items as possible. Our main technical contribution is the analysis of a novel greedy algorithm, proving that it is 3/4-competitive; this is optimal for a deterministic algorithm. The first step is to show that we can assume without loss of generality that every advertiser has frequency cap 1, i.e., wants no more than one impression from each user (the reduction is independent of advertisers having the same value, and also applies when advertisers have arbitrary values). This reduction is simple, yet crucial — for each of the cases we study, designing algorithms directly, with arbitrary frequency caps, turns out to be rather hard.

We then analyze our greedy algorithm, which assigns arriving impressions in decreasing order of *total* demand amongst eligible advertisers, for instances with unit frequency cap. (Assigning greedily according to maximum *residual demand* does not work; this algorithm cannot do better than 2/3.) The unit frequency cap means that an advertiser is eligible for an impression if she has leftover demand and has not yet been assigned to this user. We first prove that any *non-lazy* algorithm has competitive ratio 3/4 when all demands are equal (in addition to the equal value); then we build on this analysis to account for the fact that advertisers have unequal demands.

Combinatorial analysis of online algorithms is usually done via a potential function argument which shows that at each step, the change in the potential function plus the algorithm's revenue are comparable to the gain of the optimal solution. Surprisingly, our analysis considers only the final assignment, disregarding the way in which it is reached. This allows us to avoid coming up with a potential function (which in many cases seems to come "out of nowhere"), and skip the tedious consideration of each possible step.

Our result is especially interesting in light of the known upper bounds for unweighted online matching: 0.5 and $1 - 1/e \approx 0.63$ for deterministic and randomized algorithms, respectively [16].

- *Arbitrary values, equal d_i/f_i*: The ideas used in the analysis of the equal values case can be extended to analyze the case where advertisers have different values, but the same ratio of demand to frequency cap. We show here that the natural greedy algorithm, which assigns in decreasing order of value, has a competitive ratio of 3/4; again, this is optimal in the sense that no deterministic algorithm can do better.
- *Arbitrary values, d_i and f_i, with targeting*: Finally, for the general case with arbitrary values, demands and frequency caps, we design a primal-dual algorithm whose competitive ratio approaches $1 - 1/e \approx 0.63$ when $d_i/f_i \gg 1$ [6]; we also show an upper bound of $1/\sqrt{2}$ for this case. Our online primal-dual algorithm has an interesting feature: it both increases and decreases primal variables during the execution of the algorithm. The same algorithm and competitive ratio also apply when advertisers have *target sets*, i.e., they have value v_i for impressions from a set S_i of users, and value of 0 for other impressions. For this case, we have a matching upper bound for deterministic online algorithms, using the upper bound on online b-matching [15]. (See §1.1 for a discussion regarding [14] and online primal dual algorithms.)

[6] The competitive ratio of $1 - 1/e$ in [14] is under an assumption similar to ours.

1.1 Related Work

Maximizing revenue in online ad auctions has received much attention in recent years [8, 7, 18, 17, 6, 11, 12]. The problem of designing online algorithms to maximize advertising revenue was introduced in the *adwords model* [18]: advertisers have budgets, and bids for different keywords. Keywords arrive online, and the goal is to match advertisers to keywords to maximize revenue, while respecting the advertisers' budget constraints. Goel and Mehta [14] extend the adwords model, allowing advertisers to specify bids for keywords which are decreasing functions of the number of impressions (of the keyword) already assigned to the advertiser. Our frequency capping problem is, in fact, a special case of the model of [14] (but not of the adwords model of [18]), where keywords correspond to users, and the decreasing function takes the form of a step function with cutoff f_i. Hence, the $(1-1/e)$-competitive online algorithm of [14] applies to our problem as well. On the other hand, the upper bounds in [14] do not apply to our problem since the model of [14] also captures online matching. Improving upon the ratio of $1-1/e$ in special cases is posed as an open problem in [14].

Our greedy algorithms in §3 and §4 obtain a ratio of $3/4$, improving upon this ratio of $1-1/e$. While the competitive ratio of our algorithm in §5 is the same as that in [14], the algorithms are quite different. Moreover, our model does not inherit the upper bound of $1-1/e$ [7], and in fact, the best upper bound for the case without target sets is $1/\sqrt{2}$. Also, while the most general problem we solve in this paper remains within the model of [14], the most general and realistic version of the frequency capping problem (§6) cannot be stated as a special case of the model of [14]. For this model the question of both a competitive algorithm and an upper bound (tighter than $1-1/e$) are open.

The primal dual framework for online problems, first introduced by Buchbinder and Naor [9], has been shown to be useful in many online scenarios including ad auctions, see [4, 5, 3, 2, 10, 11]. Unlike these primal-dual algorithms (e.g., [9, 11]), which simply update the primal variables monotonically in each round, our primal-dual algorithm is novel in that it reassigns primal variables several times during the execution of the algorithm; hence, the primal variables do not necessarily increase monotonically with each round of new supply.

Mirrokni et al. [13] consider frequency capping in a stochastic model, but they leave open the question of improving upon the $1-1/e$ ratio in this model. Finally, the work in [1] also addresses user fatigue in the context of sponsored search; however, the model and algorithms substantially differ from ours.

2 Preliminaries

We denote by $A(\sigma)$ the revenue of algorithm A on a sequence σ of arrivals of impressions, and by $OPT(\sigma)$ the revenue of the optimal offline algorithm, which

[7] Since the model of [14] captures the adwords model of [18], it inherits an upper bound of $1-1/e$ on the competitive factor. The frequency capping problem does not generalize the adwords model, and therefore, does not inherit this upper bound.

knows σ in advance. The goal is to design an online algorithm A that assigns each impression immediately upon arrival, and produce a feasible allocation whose total value $A(\sigma)$ is competitive against $OPT(\sigma)$ for any σ. The *natural greedy algorithm* for the problem, denoted by $GREEDY_V$, allocates each arriving impression to the eligible advertiser with the highest value (breaking ties arbitrarily, but consistently). The examples in the introduction show that the greedy algorithm at most 1/2-competitive. The next theorem shows that this is tight, due to space limitations, its proof is deferred to a full version of this paper.

Theorem 1. *The competitive ratio of* $GREEDY_V$ *is 1/2.*

We now establish a reduction from general frequency caps to unit frequency caps which greatly simplifies our algorithms. The following theorem allows us to assume $f_i = 1$ in the rest of the paper, its proof is also deferred to a full version.

Theorem 2 (Reduction to Unit Frequency Cap). *For every frequency capping instance there is an equivalent instance where all frequency caps are* 1*. Moreover, any solution to the equivalent instance can be transformed in an online fashion to an equivalent solution in the original model.*

3 Identical Valuations

In this section, we assume all advertisers have identical valuations, i.e., for each advertiser a_i, w.l.o.g., $v_i = 1$. The following theorem gives an upper bound on any deterministic online algorithm; due to space limitations, its proof is deferred to a full version of this paper.

Theorem 3. *No deterministic online algorithm is better than* 3/4*-competitive, even if all advertisers have identical values, demands, and frequency caps.*

We now turn to online algorithms. A natural greedy algorithm is one that assigns an arriving impression to an eligible advertiser with the maximum residual demand. However, assigning according to residual demand, breaking ties arbitrarily, cannot have a competitive ratio better than 2/3, as the following example shows. There are two advertisers, with $d_1 = 1$ and $d_2 = 2$, with ties broken in favor of a_1. The sequence of arrivals is u_2, u_1, u_2. The residual demand algorithm allocates only two impressions: the first impression to a_2 and then the second impression to a_1. The optimal assignment, however, can assign all 3 impressions.

We show that an alternative greedy algorithm, $GREEDY_D$, which assigns according to *total* demand, is 3/4-competitive. Hereby is algorithm $GREEDY_D$:

1. Sort advertisers $a_1, \ldots, a_n$ in a non-decreasing demand order ($d_1 \geq \ldots \geq d_n$).
2. Assign an arriving impression to the first eligible advertiser in this order.

We need the following notation. Let y_i denote the number of impressions assigned by $GREEDY_D$ to a_i, and let $y^* = \min_i y_i$. Let k denote the number of advertisers whose demand is exhausted by $GREEDY_D$. In §3.1, we analyze the case of equal demands, and in §3.2 we build on this analysis to deal with the case where demands are arbitrary. We include the proof of the equal demands case since it is simpler, yet gives some insight into the proof of the general case.

3.1 Equal Demand Case

Algorithm $GREEDY_D$ is *non-lazy*, i.e., it allocates every impression it receives, unless no advertiser is eligible for it. We show that any non-lazy algorithm, including $GREEDY_D$, is 3/4-competitive if all advertisers have equal demand d.

Theorem 4. *Let ALG be a non-lazy algorithm, and let σ be a sequence of impressions. Then, $ALG(\sigma)/OPT(\sigma) \geq 3/4$.*

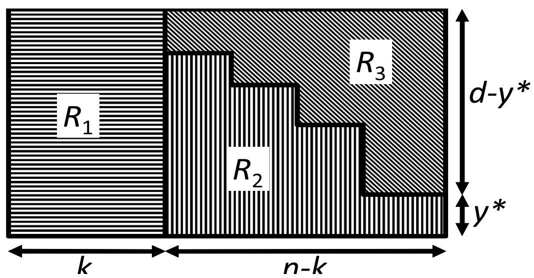

Fig. 1. Each column is an advertiser and each row corresponds to a unit demand.

Before going into the proof of Theorem 4, consider the example depicted in Figure 1. The rectangle is divided into three areas: R_1 is the total allocation of advertisers who have exhausted their demand, R_2 is the total allocation of advertisers who have not exhausted their demand, and R_3 is "unused" demand. We use two bounds on $|OPT(\sigma)| - |ALG(\sigma)|$: $|R_3| \leq (d - y^*) \cdot (n - k) \leq |R_2| \cdot (d-y^*)/y^*$, and $k \cdot y^* \leq |R_1| \cdot y^*/d$ (note that $y^* > 0$ since an advertiser who has received no impressions can always be assigned at least one impression without violating the frequency cap constraint). The theorem follows from these bounds, and the observation $|ALG(\sigma)| = |R_1| + |R_2|$.

Let A be the set of impressions allocated by OPT, and let $B \subseteq A$ be of size $OPT(\sigma) - ALG(\sigma)$. Associate each impression of B with an advertiser, such that up to $d - y_i$ impressions of B are associated with each advertiser a_i. This is possible since $\sum_{i=1}^{n}(d - y_i) = nd - ALG(\sigma) \geq OPT(\sigma) - ALG(\sigma) = |B|$.

Lemma 1. $|B| = OPT(\sigma) - ALG(\sigma) \leq y^* k$.

Proof. Let a_{i^*} be an advertiser for which $y_{i^*} = y^*$. If $y^* = d$, then $ALG(\sigma) = nd = OPT(\sigma)$, so we can assume $y^* < d$. Thus, each impression ALG fails to allocate belongs to a user already having an impression allocated to a_{i^*} (else ALG could have assigned it to a_{i^*}). Hence, there are at most y^* users having unallocated impressions. Each such user u has at most k more impressions allocated by OPT than by ALG (if u has an unassigned impression, all $n - k$ advertisers with non-exhausted demands must have been assigned an impression of u).

We define two types of payments received by each impression $x \in B$. Suppose impression x is associated with advertiser a_i. The first payment x gets is $p_x = y_i/(d - y_i)$, and the second payment is $p'_x = d/y^*$.

Lemma 2. *The total payment received by all impressions of B is at most $ALG(\sigma)$.*

Proof. Let E denote the set of advertisers whose demand is not exhausted by ALG (i.e., $|E| = n - k$). Let $a_i \in E$. For each impression x associated with a_i, we have $p_x = y_i/(d - y_i)$ and the number of such impressions is at most $d - y_i$. Therefore, the first type of payment received by impressions associated with a_i sums up to at most y_i. Adding up over all advertisers of E, the sum of the first type payments to all impressions in B is at most $\sum_{a_i \in E} y_i$. Since payments of the second type all equal values, they add up to $|B| \cdot \frac{d}{y^*} \leq y^* k \cdot \frac{d}{y^*} = dk$. Note that $dk + \sum_{a_i \in E} y_i = ALG(\sigma)$, since $a_i \notin E \Rightarrow y_i = d$, completing the proof.

Lemma 3. *For each impression $x \in B$, $p_x + p'_x \geq 3$.*

Proof. Suppose x is associated with an advertiser a_i. The total payment received by x is: $\frac{y_i}{d-y_i} + \frac{d}{y^*} \geq \frac{y^*}{d-y^*} + \frac{d}{y^*} = \frac{{y^*}^2 + d(d-y^*)}{y^*(d-y^*)} = 3 + \frac{(2y^*-d)^2}{y^*(d-y^*)} \geq 3$.

Corollary 1. $ALG(\sigma) \geq 3|B|$.

The proof of Theorem 4 is now immediate:

$$\frac{ALG(\sigma)}{OPT(\sigma)} = \frac{ALG(\sigma)}{ALG(\sigma) + |B|} \geq \frac{3|B|}{3|B| + |B|} = \frac{3}{4} \ . \tag{1}$$

3.2 General Case

In this section we prove the main result of our paper. Unfortunately, the proof from the previous section does not readily generalize; the core of the difficulty is that it is no longer possible to sort the advertisers in non-decreasing demand order such that all exhausted advertisers appear before the non-exhausted advertisers. Instead, exhausted and non-exhausted advertisers might be interleaved in every non-decreasing demand ordering of the advertisers. Thus, it is hard to guarantee the extent to which impressions of exhausted advertisers can be charged. A simple approach to overcome this difficulty is to split the advertisers into blocks, making sure that within each block the exhausted advertisers appear before the non-exhausted ones. However, this fails since OPT and $GREEDY_D$ may place impressions in different blocks. To circumvent this problem we consider subsets of advertisers having demand above a given threshold. The proof then makes a connection between the difference in number of impressions allocated by OPT and $GREEDY_D$ to a subset of the advertisers and the number of exhausted advertisers in the subset, yielding a lower bound on the payment that can be extracted from the impressions of the exhausted advertisers.

The next theorem shows that $GREEDY_D$ is 3/4-competitive also for arbitrary demands; due to space limitations, its proof is deferred to a full version.

Theorem 5. *For any sequence σ of input impressions, $\frac{GREEDY_D(\sigma)}{OPT(\sigma)} \geq 3/4$.*

4 Equal Demands/Arbitrary Valuations

In this section, we assume advertisers have different values, but equal ratio of demand to frequency cap (this can happen, e.g., when each advertiser has frequency cap f_i and wants to advertise to the same number of distinct users u, i.e., $d_i = f_i u$). The reduction to unit frequency caps makes this equivalent to assuming all demands are equal and all frequency caps are 1. The following theorem shows that the natural greedy algorithm $GREEDY_V$, assigning in decreasing order of value, is 3/4-competitive. Note that by Theorem 3, this ratio is optimal.

Theorem 6. *For any sequence* σ *of input impressions,* $\frac{GREEDY_V(\sigma)}{OPT(\sigma)} \geq 3/4$*, under the above assumptions.*

The proof builds on the ideas developed in Theorem 5, and due to lack of space, it is deferred to a full version of this paper.

5 Arbitrary Valuations

We now consider arbitrary valuations v_i. We first prove an improved upper bound for this case. Due to space imitations, the proofs of the theorems of this section are deferred to a full version of this paper.

Theorem 7. *No deterministic algorithm is better than* $1/\sqrt{2} \approx 0.707$*-competitive.*

5.1 A Primal-Dual Algorithm

In order to apply the primal-dual approach to the problem, we first formulate the **offline** allocation problem as a linear program as following. Let A be the set of advertisers. Let B be the set of users. Finally, for each user $j \in B$, let $K(j)$ be the number of impressions of user j. We define variables $y(i,j,k)$ indicating that the k-th impression of user j is assigned to advertiser a_i.

$$\begin{aligned}
\max \quad & \sum_{a_i \in A} v_i \sum_{j \in B} \sum_{k=1}^{K(j)} y(i,j,k) & & (D) \\
\text{s.t.} \quad & \sum_{j \in B} \sum_{k=1}^{K(j)} y(i,j,k) \leq d_i & \forall a_i \in A & \\
& \sum_{k=1}^{K(j)} y(i,j,k) \leq f_i & \forall a_i \in A, j \in B & \\
& \sum_{a_i \in A} y(i,j,k) \leq 1 & \forall j \in B, k \in \{1,2\ldots,K(j)\} & \\
& y(i,j,k) \geq 0 & &
\end{aligned}$$

The first set of constraints guarantees that at most d_i impressions are assigned to advertiser a_i. The second set of constraints guarantees the frequency cap of each advertiser. Finally, the last set of constraints guarantees that each impression is assigned only once. For consistency with previous work [9], we refer to the maximization problem as the dual problem. We now define the primal problem. We have variable $x(i)$ for each advertiser a_i, a variable $w(i,j)$ for each pair of advertiser a_i and user j and variable $z(j,k)$ for the k-th impression of user j.

$$\begin{aligned} \min \sum_{a_i \in A} d_i x(i) + \sum_{a_i \in A, j \in B} f_i w(i,j) + \sum_{j \in B, k} z(j,k) \qquad (P) \\ \text{s.t. } x(i) + w(i,j) + z(j,k) \geq v_i \; \forall a_i \in A, j \in B, k \\ x, w, z \geq 0 \end{aligned}$$

The allocation algorithm is as follows. We assume that the reduction to the case where the frequency cap of each advertiser is 1 has already been applied.

Allocation Algorithm: Upon arrival of impression k of user j:

- Let $S(j)$ be those advertisers not yet assigned impressions of user j, and let $\overline{S(j)} = A \setminus S(j)$.
- Let $m_1 \in S(j)$ be the advertiser that maximizes $v_i - x(i)$. Let $m_2 \in S(j) \setminus m_1$ be the advertiser that maximizes $v_i - x(i)$.[a]

1. Assign impression k to advertiser m_1.
2. For each advertiser $i \in \overline{S(j)} \cup m_1$ set: $w(i,j) \leftarrow \max\{0, (v_i - x(i)) - (v_{m_2} - x(m_2))\}$.
3. For each advertiser $i \in S(j) \setminus m_1$ set: $w(i,j) \leftarrow 0$.
4. For each impression $\ell \leq k$ of user j set: $z(j,\ell) \leftarrow v_{m_2} - x(m_2)$.
5. For advertiser m_1: $x(m_1) \leftarrow x(m_1)\left(1 + \frac{1}{d_i}\right) + \frac{v_{m_1}}{c \cdot d_i}$ (c is a constant to be determined later).

[a] If $\max_{S(j)}(v_i - x(i)) \leq 0$, or $S(j) = \emptyset$, no assignment is made and no variables are updated. If there is no m_2, we view $v_{m_2} - x(m_2)$ as equal to 0.

Notice that this algorithm differs from the standard online primal-dual approach because it both increases and decreases primal variables.

Theorem 8. *The algorithm is* $(1-(c+1)^{-1})$*-competitive, for* $c = (1+\frac{1}{d_{\min}})^{d_{\min}} - 1$*, where* $d_{\min}$ *is the minimum demand of any advertiser.*

Targeting constraints. We assumed thus far that advertisers valued all users equally. In practice, however, when buying display ad space, advertisers can provide *targeting* information, specifying which subset of impressions is acceptable. That is, advertisers have value v_i for acceptable impressions that meet the targeting constraints and value of 0 for others (contracts for display ads typically specify a single price-per-impression that does not vary across the set of acceptable impressions, i.e., v_i does not take on different non-zero values).

Suppose targeting information is user-dependent only, i.e., an advertiser may value only a subset of users with certain characteristics (age, gender, location, etc.), but does not distinguish between different impressions (e.g., when visiting different webpages) from the same user. In this case, advertiser values have the following form: $v(i,j)$ is either v_i or 0 (i.e., a_i finds users with $v(i,j) = v_i$ *acceptable*, and the rest unacceptable). We observe that the above algorithm also works for this more general setting. The only change is that the sets $S(j)$ and $\overline{S(j)}$ include only advertisers that accept user j. This implies the following.

Theorem 9. *For $c = (1 + \frac{1}{d_{\min}})^{d_{\min}} - 1$, the algorithm remains $(1-(c+1)^{-1})$-competitive, when $v(i,j) \in \{0, v_i\}$ for all i,j.*

Theorem 10. *With targeting constraints, no deterministic algorithm has a competitive ratio higher than $1 - 1/e$, even when demand are large.*

6 Further Directions

The frequency capping problem is an important practical problem which imposes interesting algorithmic challenges. Here are two main directions for further work.

- *Improving $1 - 1/e$ for arbitrary valuations:* There is a gap between the best upper bound of $1/\sqrt{2}$ and the best algorithm $(1 - 1/e)$ for the case of arbitrary valuations without targeting constraints, discussed in §5. The targeting constraints are to be blamed for the "matching" aspects, leading to the upper bound of $1 - 1/e$ in Theorem 10. By removing these constraints, the difference between our problem and online matching resurfaces, and the upper bound of $1 - 1/e$ does not hold anymore. We believe that our primal-dual algorithm is an excellent starting point for a future online algorithm for frequency capping with arbitrary values that will go beyond $1 - 1/e$.
- *Content-based targeting specifications:* Targeting specifications may be not only user-based, but also depend on the webpage's content. For instance, an advertiser might want to display her ads only to males (user targeting) when they browse a sports related webpage (content targeting); targeting constraints are often of this form. So, advertisers now have valuations of the form $v(i,j,k) \in \{0, v_i\}$, i.e., the value of the k-th impression of the j-th user to advertiser i is either v_i or 0 depending on what page the user was surfing on his k-th impression. Note that the model of [14] does not capture this problem, which entangles a matching aspect with frequency capping. The questions of designing a good online algorithm and finding the smallest upper bound (of course, $1 - 1/e$ is a trivial upper bound since this problem generalizes arbitrary valuations with targeting) are both open.

Acknowledgments. We are extremely grateful to Ning Chen for several helpful discussions, and for first suggesting the total demand algorithm.

References

1. Abrams, Z. Vee, E.: Personalized ad delivery when ads fatigue: An approximation algorithm. In: 3rd International Workshop On Internet And Network Economics, pp. 535–540. Springer-Verlag, Heidelberg (2007)

2. Alon, N., Awerbuch, B., Azar, Y., Buchbinder, N., and Naor, J.: A general approach to online network optimization problems. ACM Transactions on Algorithms 2(4), 640–660 (2006)
3. Alon, N., Awerbuch, B., Azar, Y., Buchbinder, N., and Naor, J.: The online set cover problem. SIAM J. Comput. 39(2), 361–370 (2009)
4. Bansal, N., Buchbinder, N., Naor, J.: A primal-dual randomized algorithm for weighted paging. In: 48th Annual IEEE Symposium on Foundations of Computer Science, pp. 507–517. IEEE Computer Society, Washington, DC (2007)
5. Bansal, N., Buchbinder, N., Naor, J.: Randomized competitive algorithms for generalized caching. In: 40th ACM Symposium on Theory of Computer Science, pp. 235–244. ACM, New York, NY (2008)
6. Bansal, N., Chen, N., Cherniavsky, N., Rudra, A., Schieber, B., Sviridenko, M.: Dynamic pricing for impatient bidders. In: 18th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 726–735. Society for Industrial and Applied Mathematics, Philadelphia, PA (2007)
7. Blum, A., Hartline, J.: Near-optimal online auctions. In: 16th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1156–1163. Society for Industrial and Applied Mathematics, Philadelphia, PA (2005)
8. Blum, A., Kumar, V., Rudra, A., Wu, F.: Online learning in online auctions. Theor. Comput. Sci. 324(2–3), 137–146 (2004)
9. Buchbinder, N., Naor, J.: Online primal-dual algorithms for covering and packing. Math. Oper. Res. 34(2), 270–286 (2009)
10. Buchbinder, N., Naor, J.: Improved bounds for online routing and packing via a primal-dual approach. In: 47th annual ieee Symposium on Foundations of Computer Science, pp. 293–304. IEEE Computer Society, Washington, DC (2006)
11. Buchbinder, N., Jain, K., Naor, J.: Online primal-dual algorithms for maximizing ad-auctions revenue. In: 15th Annual European Symposium on Algorithms, pp. 253–264. Springer-Verlag, Heidelberg (2007)
12. Feldman, J., Korula, N., Mirrokni, V., Muthukrishnan, S., Pál, M.: Online ad assignment with free disposal. In: 5th International Workshop on Internet and Network Economics, pp. 374–385. Springer-Verlag, Heidelberg (2009)
13. Feldman, J., Mehta, A., Mirrokni, V.S., Muthukrishnan, S.: Online stochastic matching: Beating $1 - 1/e$. In: 50th Annual IEEE Symposium on Foundations of Computer Science, pp. 117–126. IEEE Computer Society, Washington DC (2009)
14. Goel, G., Mehta, A.: Adwords auctions with decreasing valuation bids. In: 3rd International Conference on Internet and Network Economics, pp. 335–340. Springer-Verlag, Heidelberg (2007)
15. Kalyanasundaram, B., Pruhs, K.R.: An optimal deterministic algorithm for online b-matching. Theor. Comput. Sci. 233(1–2),319–325 (2000)
16. Karp, R.M., Vazirani, U.V., Vazirani, V.V.: An optimal algorithm for on-line bipartite matching. In: 22nd Annual ACM Symposium on Theory of Computing, pp 352–358. ACM, New York, NY (1990)
17. Mahdian, M., Saberi, A.: Multi-unit auctions with unknown supply. In: 7th ACM conference on Electronic Commerce, pp. 243–249. ACM, New York, NY (2006)
18. Mehta, A., Saberi, A., Vazirani, U., Vazirani, V.: Adwords and generalized online matching. J. ACM 54(5), 22 (2007)
19. Marketing Terms.com, http://www.marketingterms.com/dictionary/frequency_cap