Weighted Paging

The problem  Page $i$ costs $w_i$ to fetch. We model this using the following variables.

$B(t)$: set of pages requested up to time $t$.

$p_t$: page requested at time $t$.

$t(i, j)$: times at which page $i$ is requested for the $j$th time.

$j(i, t)$: number of times page $i$ has been requested by time $t$.

$x(i, j)$: whether page $i$ has been evicted after its $j$th request and before the next request for $i$.

We write:

$$\min f = \sum_i \sum_j w_i x_{ij} \text{ s.t. } \sum_j \sum_{(i,j)} x_{ij} \geq |B(t)| - k \quad (\forall t)$$

This suggests the online fractional weighted paging problem, which is to solve this linear program in an online fashion as constraints arrive one by one over time. It can be shown that one can reduce the online weighted paging problem to the online fractional weighted paging problem, up to losing a constant factor in the competitive ratio.

To solve fractional weighted paging, we write the dual:

$$\max g = \sum_t (|B(t)| - k) y_t - \sum_{ij} z_{ij} \text{ s.t. } \sum_{t:(i,j)} y_t - z_{ij} \leq w_i \quad (\forall ij)$$

Here is the online algorithm upon arrival of page $p_t = i$ at time $t$, let $j = j(i, t)$.

$y_t = 0$

while not $(\sum_i \sum_j x_{ij} \geq |B(t)| - k))$ do

Increase $y_t$ by $dy$

if not $(\sum_{t:(i,j)} y_t - z_{ij} \leq w_i)$ then

if $x_{ij} = 0$ then

$x_{ij} \leftarrow \frac{1}{k}$

else if $1/k \leq x_{ij} < 1$ then

$x_{ij} \leftarrow (1/k) \exp[(\sum_{t:(i,j)} y_t - z_{ij} - w_i)/w_i]$

else

Increase $z_{ij}$ by $dy$ (and $x_{ij}$ does not change).

end if

end if

end while

To analyze the algorithm, we will prove that

1. $x$ is feasible: by construction: we keep increasing stuff until the new constraint is satisfied.

   We never get stuck because $x_{ij}$ only stops increasing once it has reached 1, and by the time all active $x_{ij}$'s have reached the value 1 the new constraint is definitely satisfied.

2. $y, z$ is feasible up to scaling: What is the maximum value of $\sum_{t:(i,j)} y_t - z_{ij} - w_i$? Since $x_{ij}$ is at most 1, this is at most $w_i \ln(k)$, so the left hand side of the dual constraint is bounded by $(1 + \ln k)w_i$, so $(y, z)$ is feasible up to scaling by $(1 + \ln k)$. 


3. Relating the increase in $f$ to the increase in $g$: that’s the hard part. We split the increase in $f$ into two parts:

$$f(x) = \sum_{ij: 0 < x_{ij}} w_i (1/k) + \sum_{ij: 0 < x_{ij}} w_i (x_{ij} - 1/k) = C_1 + C_2.$$ 

- First, $x_{ij} > 0$ implies that at some point we had $\sum_{t: (i,j)}$ active at time $t$, $y_t - z_{ij} \geq w_i$, and this remains true thereafter, so we can upper bound $w_i$ and write

$$C_1 \leq \sum_{ij: 0 < x_{ij}} \sum_{t: (i,j) \text{ active at time } t} y_t - z_{ij} (1/k) = C'_1.$$

We invert summations and rewrite

$$C'_1 = \sum_t \sum_{ij: 0 < x_{ij}, (i,j) \text{ active at time } t} (1/k) y_t - \sum_{ij: 0 < x_{ij}} z_{ij} (1/k).$$

We bound $C'_1$ by looking at its derivative. When $y_t$ increases by $dy$, $C'_1$ increases by

$$\sum_{ij: 0 < x_{ij}, (i,j) \text{ active at time } t, x_{ij} < 1} (1/k) dy$$

because each $(1/k)dy$ that comes from an $ij$ such that $x_{ij} = 1$ is cancelled by the $-(1/k)dy$ coming from the corresponding change in $z_{ij}$. Now, if $B'(t)$ denotes those $i$’s such that $x_{ij} = 1$ (for the $j = j(i,t)$ currently active), this is at most $(|B(t)| - 1 - |B'(t)|)(1/k)$, which is less than $|B(t)| - |B'(t)| - k$ since $|B(t)| - |B'(t)| \geq k + 1$ (this last inequality holds because otherwise the constraint would be satisfied.) Thus

$$dC'_1 \leq (|B(t)| - |B'(t)| - k) dy = (|B(t)| - k) dy - \sum_{ij: (i,j) \text{ active at time } t, x_{ij} = 1} dz_{ij}$$

because $dy = -dz_{ij}$ for each $i$ in $B'(t)$. Integrating yields $C'_1 \leq g(y, z)$, as desired.

- Second, we bound $C_2$ by looking at its derivative. At time $t$, $y_t$ increases by $dy_t$, and for $x_{ij} \in (0,1)$, $x_{ij}$ increases by $dy_t \cdot x_{ij}/w_i$. Thus

$$dC_2/dy \leq \sum_{i,j: \text{ active at time } t, x_{ij} < 1} w_i x_{ij}/w_i = \sum_{i,j: \text{ active at time } t, x_{ij} < 1} x_{ij}.$$ 

Then, since the constraint is not yet satisfied, $\sum_t \sum_{i,j} \text{ active at time } t x_{ij} \leq |B(t)| - k$, so $dC_2/dy \leq |B(t)| - k - |B'(t)|$. On the other hand it is easy to check that $dg/dy = |B(t)| - k - |B'(t)|$, hence $C_2 \leq g(y, z)$, as desired.