Online Algorithms for Covering and Packing Problems with Convex Objectives

Niv Buchbinder

Tel Aviv University

Based on:

Paper 1: N. Buchbinder, S. Chen, A. Gupta, V. Nagarajan, J. Naor

Paper 2: Y. Azar, I. R. Cohen, D. Panigrahi

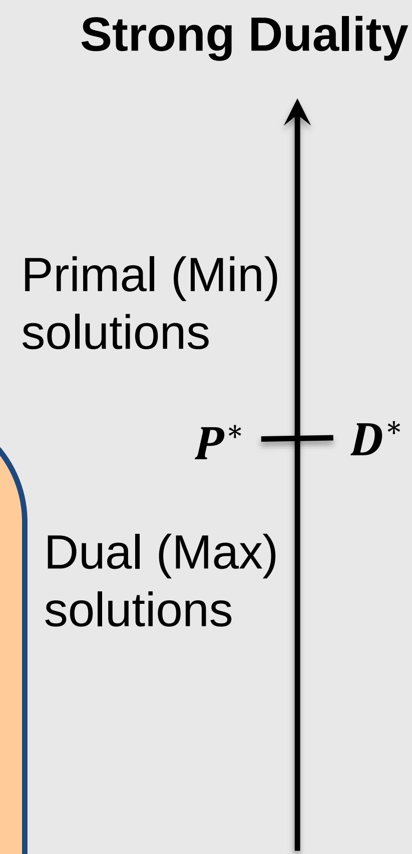
Paper 3: T.-H. H. Chan, Z. Huang, N. Kang

Road Map

- The online primal-dual framework
- A natural extension of the primal
- A natural extension of the dual
- Main results
- Algorithms and analysis ideas

Offline Covering/Packing Problems

Primal (covering)	Dual (Packing)
(P): Min $c'x$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq c'$ $y \geq 0$
$A \in R_+^{T \times n}, b \in R_+^T, c \in R_+^n$	



Captures many (relaxations) of combinatorial optimization problems:

- **Covering:** Covering problems (set-cover, facility location), connectivity/cut problems (steiner tree, shortest path), paging …
- **Packing:** knapsack , flow problems (Maximum multicommodity flow, matching), combinatorial auctions ....

Online Covering/Packing Problems

Primal (covering)	Dual (Packing)
(P): Min $c'x$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq c'$ $y \geq 0$
$A \in R_+^{T \times n}, b \in R_+^T, c \in R_+^n$	

- c is known in advance.

At time $t = 1,2, \ldots T$:

- The tth row of A is revealed (and a new dual y_t).

Covering: Variables x_j can only be increased to maintain a feasible solution.

- **Goal:** Minimize the total cost.

Packing: New dual variable y_t should be set immediately.

- **Goal:** Maintain a feasible solution, Max total profit.

Example 1: Online Set Cover

Primal (covering)	
(P): Min $\sum_s x_s$	Non negative objective function
$\sum_{s\mid e\in s} x_s \geq 1 \quad \forall e \in E$ $x \geq 0$	• x_s: Choose set s • Rows (=elements) arrive online • x_s can only be increased over time

Online set cover [Alon-Awerbuch-Azar-B-Naor03]:

- $E = \{1,2,\dots n\},\ S_i \subseteq E$ (m sets).
- Elements arrive one-by-one and should be covered **upon arrival**.
- Sets cannot be unchosen.

Goal: Minimize total cost of sets chosen.

Example 2: Virtual Circuits Routing

Dual (packing)	
(D): Max $\sum_i y_{r_i}$	Non-negative objective
$\sum_{r_i \mid e \in p_i} y_{r_i} \le c_e \quad \forall e \in E$ $y_{r_i} \le 1 \quad \forall r_i$ $y \ge 0$	• Packing constraints for all $e \in E$ • Variables y_r (= requests) arrive online. • Should be set upon arrival.

Online virtual circuits routing [Awerbuch-Azar-Plotkin93]:

- Graph $G = (V, E)$, capacities on edges c_e.
- Requests $\mathrm{r_i} = (s_i, t_i, p_i)$ arrive one-by-one.
- Should be connected using capacity 1, or rejected.
- Accepted requests cannot be rejected later.

Goal: Maximize number of requests accepted.

Online Covering/Packing problems

Primal (covering)	Dual (Packing)
(P): Min $c'x$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq c'$ $y \geq 0$
$A \in R_+^{T \times n}, b \in R_+^T, c \in R_+^n$	

Captures many (relaxations) of **online** combinatorial optimization problems:

- **Covering:** online set-cover, online connectivity/cut, facility location, (weighted) paging, Metrical task systems …
- **Packing:** routing, matching (ad-auctions), online knapsack, online combinatorial auctions.

Algorithm for the framework

Primal (covering)	Dual (Packing)
(P): Min $\sum_{j=1}^{n} c_j x_j$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq c'$ $y \geq 0$

Theorem [B-Naor05, Gupta-Nagarajan12]:

There is an algorithm that produces solutions x, y such that:

- x is $O(\log d)$-competitive,

  $d =$ Maximum row sparsity of A.

- y is $O\left(\log\left(d \cdot \frac{a_{max}}{a_{\min}}\right)\right)$-competitive,

  $a_{\max}/a_{\min}$- ratio of maximal to minimal (non-zero) entry in a column of A.

- Results are tight asymptotically.

A Natural Generalization of Covering

Primal (covering)
(P): Min $\sum_{j=1}^{n} c_j x_j$
$Ax \geq 1$ $x \geq 0$

⇒

Primal (covering)
(P): Min $f(x)$
$Ax \geq 1$ $x \geq 0$

- f is a convex monotone function.

(**Monotone:** $x \leq y \Rightarrow f(x) \leq f(y)$)

Offline: Problem is polynomially solvable.

Online (same setting):

- Rows of A arrive online.
- Variables should be monotonically increasing.

Example 1: L_p-norm Set Cover

Primal (covering)
(P): Min $\sum_{i=1}^{k}(c_i' x)^p$
$\sum_{s \mid e \in s} x_s \geq 1 \quad \forall e \in E$ $x \geq 0$

- Elements arrive one-by-one and should be covered **upon arrival**.
- Sets cannot be unchosen.
- $f(x) = \sum_{i=1}^{k}(c_i' x)^p$
- **Special case 1:** $f(x) = \sum_{j=1}^{n} c_s x_s$ $\quad (p = 1)$
- **Special case 2:** $f(x) = \max_{i=1}^{k}(c_i' x)$ $\quad (p \approx \log k)$

Motivation: combining multiple objectives, makespan, energy minimization.

The Dual Problem

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

- $f^*(z) = \sup_{x \geq 0}(z'x - f(x))$ (conjugate function)
- f^* always convex (even if f is not convex).

- [Nice function f]: if f is continuous, convex, monotone, differentiable and $f(0) = 0$

➔ f^* is convex, monotone, non-negative, $f^*(0) = 0$ and $f^{**} = f$.

The Dual Problem

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

- $f^*(z) = \sup_{x \geq 0}(z'x - f(x))$ (conjugate function)

Proof (weak duality): x, y solutions to primal/dual:

$$f(x) \geq f(x) - y'(Ax - 1)$$

$$= \sum_{t=1}^{T} y_t - \left(x'(A^T y) - f(x)\right)$$

$$\geq \sum_{t=1}^{T} y_t - \sup_{x \geq 0}\left(x'(A^T y) - f(x)\right) = \sum_{t=1}^{T} y_t - f^*(A^T y)$$

$\boldsymbol{y \geq 0, Ax \geq 1}$

$\boldsymbol{x \geq 0}$, **definition of** $\boldsymbol{f^*}$

Natural Extension of Dual Problem

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

Online setting (Dual):

- Primal constraint arrive at time t

  $\Rightarrow$ New dual variable y_t

- Value of y_t should be set immediately and cannot be changed later on.

Goal: Maximize profit $\sum_{t=1}^{T} y_t$ minus cost $f^*(A^T y)$.

Example 2: Virtual Circuits Routing

Dual (packing)
(D): Max $\sum_i y_{r_i} - f^*(z)$
$\sum_{r_i \mid e \in p_i} y_{r_i} = z_e \quad \forall e \in E$ $y_{r_i} \leq 1 \qquad \forall r_i$ $y \geq 0$

Online virtual circuits routing (with capacity costs):

- Requests arrive online as before and should be accepted/rejected immediately.
- Capacity should be bought at cost $f^*(z)$.
- **Special case 1:** $f^*(z) = \begin{Bmatrix} 0 & z_e \leq c_e \quad \forall e \in E \\ \infty & Otherwise \end{Bmatrix}$
- **Special case 2:** $f^*(z) = \sum_{e \in E} g_e(z_e)$

Extending the Basic Framework

Primal (covering)	Dual (Packing)
(P): Min $\sum_{j=1}^{n} c_j x_j$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq c'$ $y \geq 0$
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

- f(x): Non-negative monotone convex function (+ ∇f is monotone).

Let $p = \sup_{x \geq 0} \frac{\langle \nabla f(x), x \rangle}{f(x)}$ **(Intuition:** $f(x)$ is a polynomial of degree p)

- **Covering competitive ratio:** $O(p \cdot log\ d)^p$
- **Packing competitive ratio:** $O\left(p \cdot log\ \left(d \frac{a_{max}}{a_{\min}}\right)\right)^p$
- d – row sparsity of matrix A
- $a_{\max}/a_{\min}$- ratio of maximal to minimal (non-zero) entry in a column of A.

Our Results (cont.)

- Matches the best bounds for the linear case.

Theorem (lower bound):

There exists an instance with f =polynomial of degree p such that any online algorithm for the **primal** problem is $\Omega(p\log d)^p$-competitive.

Rounding (Integral solutions)

- **Example:** There exists a $(\frac{p^3}{\log p}\log d \log n)$-competitive algorithm for L_p-norm set cover

(n: num. of elements, d: max num. of sets containing an element)

- **Other applications:** scheduling, facility location …

Previous Results (Primal)

[Azar, Bhaskar, Fleischer, Panigrahi, 2013]

Online Mixed Packing and Covering

(P): Min $\mathrm{Max}_{\mathrm{i=1}}^{\mathrm{k}}\{c_i' x\}$

$$Ax \geq 1$$

$$x \geq 0$$

- $O\left(log\ k \cdot log\left(d \cdot \frac{a_{max}}{a_{min}} \cdot \frac{c_{max}}{c_{min}}\right)\right)$-competitive algorithm.

($a_{max}, a_{min}, c_{max}, c_{min}$: max / min (non-zero) coordinate)

Our result (for this case): $O(\log k\ \ \log d)$-competitive (best possible)

Previous Results (Dual)

[Blum, Gupta, Mansour, Sharma, 11], [Huang, Kim, 15]

Maximizing social welfare with (separable) production costs

- n item types, buyers arrive online. For each bundle S:
  - $v_{i,S}$: value of bundle S to buyer i
  - $a_{j,S}$: number of items of type j in bundle S
- $y_{i,S}$: allocate bundle S to buyer i
- z_j: how many items of type j to produce.

(D): Max $\sum_{i=1}^{m} \sum_S v_{i,S} \cdot y_{i,S} - \sum_{j=1}^{n} f_j^*(z_j)$

$$\sum_S y_{i,S} \leq 1 \quad for\ each\ buyer\ i$$

$$\sum_{i=1}^{m} \sum_S a_{j,S} y_{i,S} = z_j \quad for\ each\ item\ type\ j$$

$$y \geq 0$$

➔ $f^*$is separable (separate cost for each item type).

Algorithm for the framework

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

- Initially set $x = 0$.
- When t^{th} row of A arrives (and new y_t).
  - While t^{th} constraint is unsatisfied:
    - Increase y_t at rate δ

      (… depends on parameters of the problem).
    - Increase each x_j **with $\boldsymbol{a_{tj} > 0}$** with rate:

$$\frac{dx_j}{dy_t} = \frac{a_{tj}x_j + 1/d}{\nabla_j f(x)}$$

$d\ (\leq n) =$ Maximum row sparsity seen so far.

(**Intuition:** linear case, $\nabla_j f(x) = c_j$)

Algorithm for the framework

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

Theorem: Algorithm produces a **monotone** primal solution P and a **monotone** dual solution D such that:

$$P \leq O\left(p \log\left(d \frac{a_{\max}}{a_{\min}}\right)\right)^{p} \text{D}$$

➔(Weak duality):

$$P, D \text{ are } O\left(p \log\left(d \frac{a_{\max}}{a_{\min}}\right)\right)^{p} \text{-competitive.}$$

Analysis: Main Ideas

(P): Primal	(D): Dual
$Min_{Ax\geq 1, x\geq 0}\{f(x)\}$	$\text{Max}_{y\geq 0}\{\sum_{t=1}^{T} y_t - f^*(A^T y)\}$

Update rule:

Increase y_t at rate δ

$$\frac{dx_j}{dy_t} = \frac{a_{tj}x_j + 1/d}{\nabla_j f(x)}$$

Bounding the "profit":

Assume y_t is increased at rate 1 (not $\delta < 1$)

$$\frac{\partial f}{\partial y_t} = \sum_{j=1}^{n} \frac{\partial f(x)}{\partial x_j} \cdot \frac{\partial x_j}{\partial y_t} = \sum_{j|a_{tj}>0} \nabla_j f(x) \cdot \frac{a_{tj}x_j + \frac{1}{d}}{\nabla_j f(x)} = \sum_{j|a_{tj}>0} \left(a_{tj}x_j + \frac{1}{d}\right) \leq 2$$

$\Rightarrow f(\bar{x}) \leq \frac{2}{\delta} \sum_{t=1}^{T} y_t,$ $\bar{x}$: final value of x.

Or, $\sum_{t=1}^{T} y_t \geq \frac{\delta}{2} f(\bar{x})$ (➔ Profit is large compared to primal cost)

Analysis: Main Ideas

(P): Primal	(D): Dual
$Min_{Ax\geq 1, x\geq 0}\{f(x)\}$	$\text{Max}_{y\geq 0}\{\sum_{t=1}^{T} y_t - f^*(A^T y)\}$

Update rule:

Increase y_t at rate δ

$$\frac{dx_j}{dy_t} = \frac{a_{tj}x_j + 1/d}{\nabla_j f(x)}$$

Bounding the "production cost":

Claim: $x_j \geq \frac{1}{\text{d}\cdot\max_{t\in S_j}\{a_{tj}\}}\left[exp\left(\frac{\sum_{t\in S_j} a_{tj}\cdot y_t}{\delta \nabla_j f(\bar{x})}\right) - 1\right]$, $S_j \subseteq \{t \mid a_{tj} > 0\}$

Proof: Solving differential equation of update rule + ∇f is monotone.

$x_j \leq \frac{1}{\min_{t\in S_j}\{a_{tj}\}}$ (at this value all constraints are feasible)

➔ $(A^T y)_j = \sum_{t\in S_j} a_{tj} \cdot y_t \leq \delta \nabla_j f(\bar{x}) \cdot O\left(\log\left(d\frac{a_{max}}{a_{min}}\right)\right)$

➔ (By prop. of f^*+ bound on "convexity" of f) bound on $f^*(A^T y)$

Finally, optimizing the value δ.

Analysis: Main Ideas

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

Theorem: The algorithm produces a **monotone** primal solution P and a **monotone** dual solution D such that:

$$P \leq O\left(p \log\left(d \frac{a_{\max}}{a_{\min}}\right)\right)^p \text{D}$$

How to remove the a_{max}/a_{min} term?

There exists a feasible solution D' such that:

$$P \leq O(p \log(d))^p \cdot D'$$

D' is not monotone!

Maintaining D’: The Linear Case

Primal (covering)	Dual (Packing)
(P): Min $\sum_{j=1}^{n} x_j$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq 1$ $y \geq 0$

Why we must decrease dual variables?

Primal (covering)		Dual (Packing)
(P): Min x_1		(D): Max $y_1 + My_2 + M^2 y_3 + \ldots$
$x_1 \geq 1$ $x_1 \geq M$ $x_1 \geq M^2$ $\ldots$	$(M \gg 1)$	$y_1 + y_2 + y_3 + \ldots \leq 1$ $y \geq 0$

Aiming towards constant competitive ratio ‘c’:

$$y_1 = 1/c,\ y_2 = 1/c,\ y_3 = 1/c\ \ldots$$

But dual constraint should be satisfied!

Maintaining D’: The Linear Case

Primal (covering)	Dual (Packing)
(P): Min $\sum_{j=1}^{n} x_j$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq 1$ $y \geq 0$

- When t^{th} row of
  - While t^{th} co

Primal update: Inc

$$\frac{dx_j}{dy_t} = \ldots 1/d$$

Observation: D’ is feasible.

The change in the dual constraint at most:

$$a_{tj}\delta - a_{t_j^*} \cdot \frac{a_{tj}}{a_{t_j^*}}\delta = 0$$

Dual update for D’: Increase y_t at rate δ.

If for $j = 1, \ldots, n$ the dual constraint $\sum_{t'=1}^{t} a_{t'j} y_{t'} = 1$:

- Let $t_j^* = \operatorname{argmax}_{t' \leq t}\{a_{t'j} | y_{t'} > 0\}$
- Decrease $y_{t_j^*}$ at rate $-\frac{a_{tj}}{a_{t_j^*}}\delta$.

Maintaining D': The Linear Case

Primal (covering)	Dual (Packing)
(P): Min $\sum_{j=1}^{n} x_j$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq 1$ $y \geq 0$

Primal update as before:

Increase each x_j **with** $\boldsymbol{a_{tj} > 0}$ with rate:

$$\frac{dx_j}{dy_t} = a_{tj}x_j + 1/d$$

Change in the primal objective function:

$$\frac{\partial P}{\partial y_t} = \sum_{j|\ a_{tj}>0} \left(a_{tj}x_j + \frac{1}{d} \right) \leq 2$$

Main question: Does the Dual D' increase enough?

Maintaining D': The Linear Case

Primal (covering)	Dual (Packing)
(P): Min $\sum_{j=1}^{n} x_j$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq 1$ $y \geq 0$

Dual update: Increase y_t at rate δ.

If for $j = 1, \dots, n$ the dual constraint $\sum_{t'=1}^{t} a_{t'j} y_{t'} = 1$:

- Let $t_j^* = \operatorname{argmax}_{t' \leq t} \{ a_{t'j} | y_{t'} > 0 \}$. Decrease $y_{t_j^*}$ at rate $-\frac{a_{tj}}{a_{t_j^*}} \delta$.

Change in the dual objective function:

$$\frac{\partial D}{\partial y_t} = \delta - \sum_{dual\ of\ j\ is\ tight} \frac{a_{tj}}{a_{t_j^*}} \delta = \delta \cdot \left(1 - \sum_{dual\ of\ j\ is\ tight} \frac{a_{tj}}{a_{t_j^*}} \right)$$

Final claim: $\sum_{j\ is\ tight} \frac{a_{tj}}{a_{t_j^*}} \leq \frac{1}{2}$ (so dual increase ≥$\delta/2$)

Maintaining D': The Linear Case

Primal (covering)	Dual (Packing)
(P): Min $\sum_{j=1}^{n} x_j$	(D): Max $\sum_{t=1}^{T} y_t$
$Ax \geq 1$ $x \geq 0$	$y'A \leq 1$ $y \geq 0$

Final claim: $\sum_{j \text{ is tight}} \frac{a_{tj}}{a_{t_j^*}} \leq \frac{1}{2}$

- **Claim:** $x_j \geq \frac{1}{\mathrm{d} \cdot \max_{t \in S_j}\{a_{tj}\}} \left[exp\left(\frac{\sum_{t \in S_j} a_{tj} \cdot y_t}{\delta} \right) - 1 \right]$, $S_j \subseteq \{t \mid a_{tj} > 0\}$
- $S_j = \{t \mid a_{tj} > 0, y_t > 0\}$: $x_j \geq \frac{1}{d \cdot a_{t_j^*}} \left[exp\left(\frac{\sum_{t \in S_j} a_{tj} \cdot y_t}{\delta} \right) - 1 \right]$
- $\sum_j a_{tj} x_j \leq 1$ + dual constraints of variables j are tight

➔ $\sum_{j \text{ is tight}} a_{tj} \left(\frac{1}{d \cdot a_{t_j^*}} \left[exp\left(\frac{1}{\delta}\right) - 1 \right] \right) \leq \sum_j a_{tj} x_j \leq 1$

➔ (Plugging $\delta = 1/(\log(1+2d))$): $\sum_j \frac{a_{tj}}{a_{t_j^*}} \leq \frac{1}{2}$

Questions

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

- f(x): Non-negative monotone convex function (+ ∇f is monotone).

Let $p = \sup_{x \geq 0} \frac{\langle \nabla f(x), x \rangle}{f(x)}$ **(Intuition:** $f(x)$ is a polynomial of degree p)

- **Covering competitive ratio:** $O(p \cdot log\ d)^p$
- **Packing competitive ratio:** $O\left(p \cdot log\ \left(d \frac{a_{max}}{a_{\min}}\right)\right)^p$
- d – row sparsity of matrix A
- $a_{\max}/a_{\min}$- ratio of maximal to minimal (non-zero) entry in a column of A.

Questions

Primal (covering)	Dual (Packing)
(P): Min $f(x)$	(D): Max $\sum_{t=1}^{T} y_t - f^*(A^T y)$
$Ax \geq 1$ $x \geq 0$	$y \geq 0$

- Is $'d'$ (row sparsity) the right parameter? Is there a more refined parameter?

  (adding ϵ noise doesn't change problem, but makes $d = n$)
- More applications.
- Additional extensions of the framework.
- Handling non-covering constraints (paying for changing x).
- Connections to learning.

Thank you