# Online Ad Allocation, and Online Submodular Welfare Maximization

**Vahab Mirrokni**
**Google Research, New York**

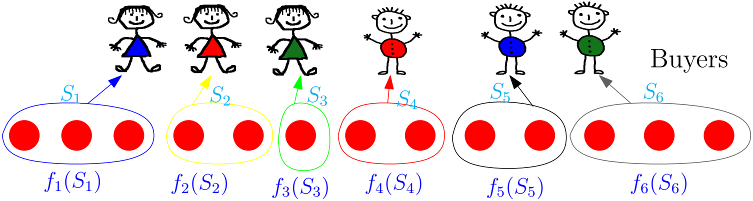
March 20, 2012

# Outline

- Problems: SWM and Online Ad Allocation
- Online Generalized Assignment (GAP)
- Page-based Allocation and SWM with online buyers
- Stochastic Settings:
  - Online Stochastic Matching: Primal Algorithms
  - Online Stochastic Packing: Dual Algorithms
  - Experimental Results
  - Simultaneous Stochastic and Adversarial Approximations

# Submodular Welfare Maximization(SWM): Offline

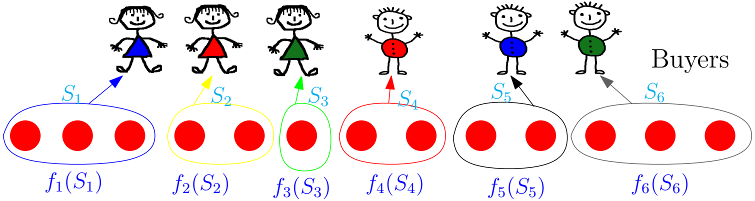
- $m$ buyers and $n$ items.
- Each buyers $i$ has a monotone submodular valuation $f_i$ on items.




- Goal: Partition items to maximize social welfare, i.e, $\sum_i f_i(S_i)$.
- Known Results:
  - There exists a $1 - \frac{1}{e}$-approximation for this problem. (Vondrak)
  - Achieving factor better than $1 - \frac{1}{e}$ needs exponential number of value queries. [M., Schapira, Vondrak]

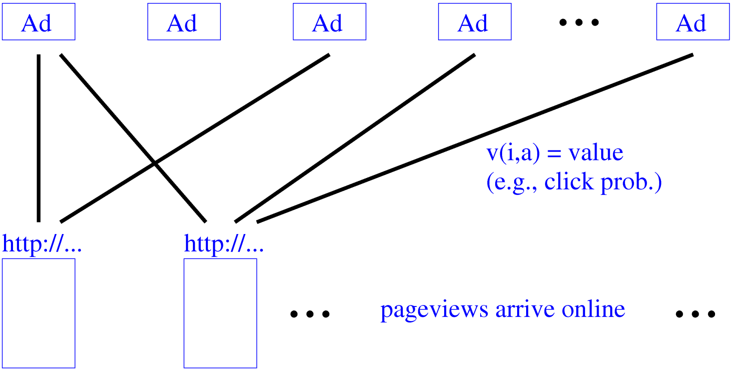
# Submodular Welfare Maximization(SWM): Online

- *m* buyers and *n* items.



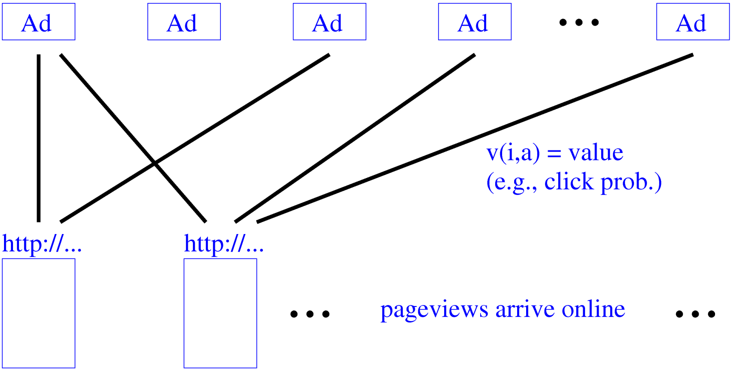

- Goal: Partition items to maximize social welfare, i.e, $\sum_i f_i(S_i)$.
- Online:
  - SWM with online items: items arrive online one by one
    - Greedy is a 1/2-approximation algorithm (NWF)
    - Will present improved algorithms for special cases.
  - SWM with online buyers with re-assignment: buyers arrive one by one.
    - Will present improved approximation algorithms.

# Online Ad Allocation



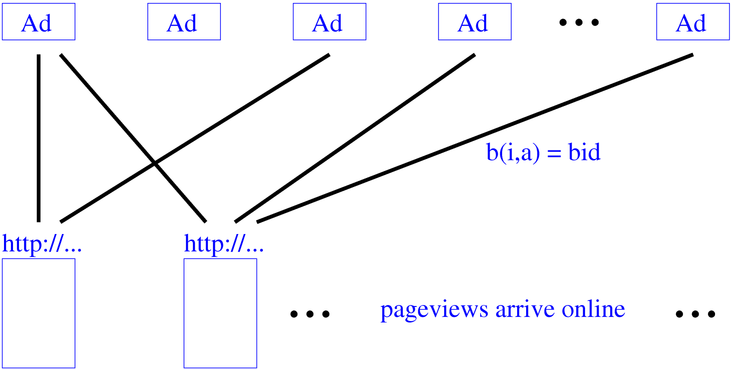

- When a page arrives, assign an eligible ad.
  - value of assigning page $i$ to ad $a$: $v_{ia}$

# Online Ad Allocation




- When a page arrives, assign an eligible ad.
  - value of assigning page $i$ to ad $a$: $v_{ia}$
- Display Ads (DA) problem:
  - Maximize value of ads served: $\max \sum_{i,a} v_{ia} x_{ia}$
  - Capacity of ad $a$: $\sum_{i \in A(a)} x_{ia} \leq C_a$

# Online Ad Allocation




- When a page arrives, assign an eligible ad.
  - revenue from assigning page $i$ to ad $a$: $b_{ia}$
- "AdWords" (AW) problem:
  - Maximize revenue of ads served: $\max \sum_{i,a} b_{ia} x_{ia}$
  - Budget of ad $a$: $\sum_{i \in A(a)} b_{ia} x_{ia} \leq B_a$

# General Form of LP

$$\boxed{\begin{aligned} &\max \sum_{i,a} v_{ia} x_{ia} \\ \sum_a x_{ia} &\leq 1 && (\forall\ i) \\ \sum_i s_{ia} x_{ia} &\leq C_a && (\forall\ a) \\ x_{ia} &\geq 0 && (\forall\ i, a) \end{aligned}}$$

| | Online Matching: | Disp. Ads (DA): | AdWords (AW): |
|---|---|---|---|
| | $v_{ia} = s_{ia} = 1$ | $s_{ia} = 1$ | $s_{ia} = v_{ia}$ |

# General Form of LP

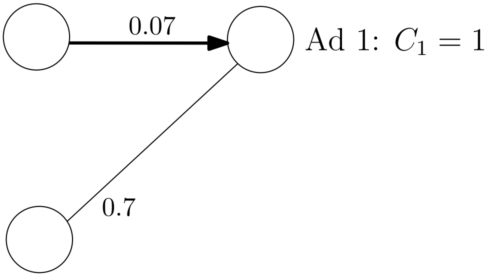
$$
\begin{array}{rcll}
\max \sum_{i,a} v_{ia} x_{ia} & & & \\
\sum_{a} x_{ia} & \leq & 1 & (\forall\ i) \\
\sum_{i} s_{ia} x_{ia} & \leq & C_a & (\forall\ a) \\
x_{ia} & \geq & 0 & (\forall\ i, a)
\end{array}
$$

| | Online Matching: $v_{ia} = s_{ia} = 1$ | Disp. Ads (DA): $s_{ia} = 1$ | AdWords (AW): $s_{ia} = v_{ia}$ |
|---|---|---|---|
| Worst-Case | Greedy: $\frac{1}{2}$, [KVV]: $1 - \frac{1}{e}$-aprx | | [MSVV,BJN]: $1 - \frac{1}{e}$-aprx if $B_a \gg b_{ia}$. |

# Outline

- Problems: SWM and Online Ad Allocation
- Online Generalized Assignment (GAP)
- Page-based Allocation and SWM with online buyers
- Stochastic Settings:
  - Online Stochastic Matching: Primal Algorithms
  - Online Stochastic Packing: Dual Algorithms
  - Experimental Results
  - Simultaneous Stochastic and Adversarial Approximations

# DA: Free Disposal Model




- Advertisers may not complain about extra impressions, but no bonus points for extra impressions, either.
- Value of advertiser = sum of values of top $C_a$ items she gets.

# Greedy Algorithm

Assign impression to an advertiser
maximizing Marginal Gain = (imp. value - min. impression value).

# Greedy Algorithm

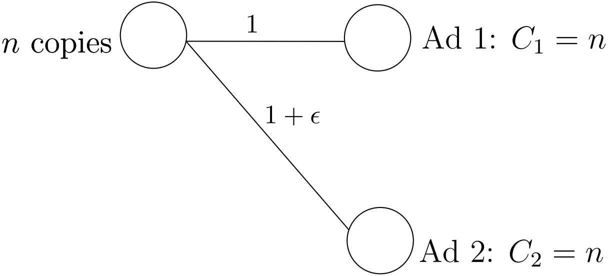
Assign impression to an advertiser
maximizing Marginal Gain = (imp. value - min. impression value).

- Competitive Ratio: 1/2. [NWF78]
  - Follows from submodularity of the value function.

# Greedy Algorithm

Assign impression to an advertiser
maximizing Marginal Gain = (imp. value - min. impression value).

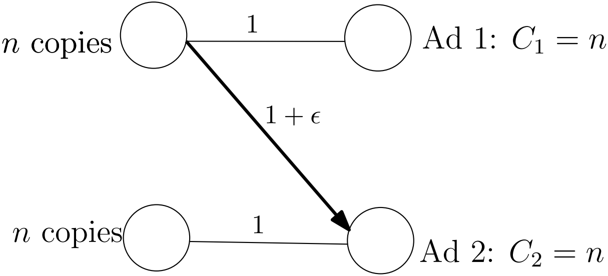
- Competitive Ratio: 1/2. [NWF78]
  - Follows from submodularity of the value function.



# Greedy Algorithm

Assign impression to an advertiser
maximizing Marginal Gain = (imp. value - min. impression value).

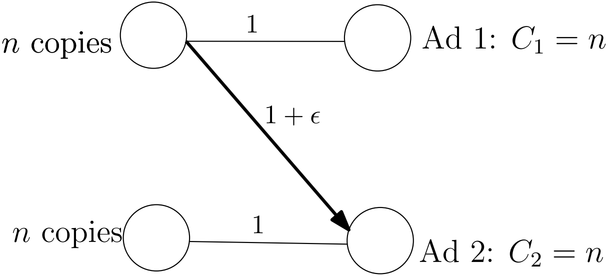
- Competitive Ratio: 1/2. [NWF78]
  - Follows from submodularity of the value function.



# Greedy Algorithm

Assign impression to an advertiser
maximizing Marginal Gain = (imp. value - min. impression value).

- Competitive Ratio: 1/2. [NWF78]
  - Follows from submodularity of the value function.



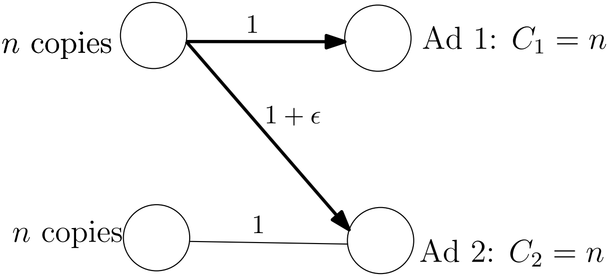

Evenly Split?

# A better algorithm?

Assign impression to an advertiser $a$
maximizing (imp. value - $\beta_a$),
where $\beta_a =$ average value of top $C_a$ impressions assigned to $a$.

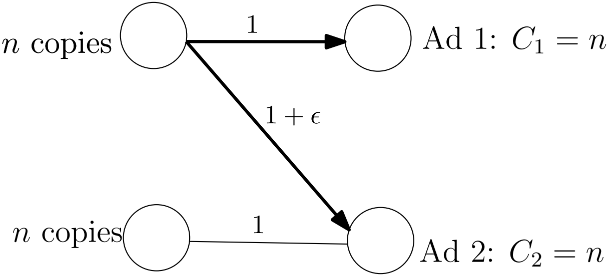
# A better algorithm?

Assign impression to an advertiser $a$
maximizing (imp. value - $\beta_a$),
where $\beta_a =$ average value of top $C_a$ impressions assigned to $a$.



# A better algorithm?

Assign impression to an advertiser $a$
maximizing (imp. value - $\beta_a$),
where $\beta_a =$ average value of top $C_a$ impressions assigned to $a$.




- Competitive Ratio: $\frac{1}{2}$ if $C_a >> 1$. [FKMMP09]
  - Primal-Dual Approach.

# An Optimal Algorithm

Assign impression to an advertiser $a$:
maximizing (imp. value - $\beta_a$),

- Greedy: $\beta_a =$ min. impression assigned to $a$.
- Better (pd-avg): $\beta_a =$ average value of top $C_a$ impressions assigned to $a$.

# An Optimal Algorithm

Assign impression to an advertiser $a$:
maximizing (imp. value - $\beta_a$),

- Greedy: $\beta_a =$ min. impression assigned to $a$.
- Better (pd-avg): $\beta_a =$ average value of top $C_a$ impressions assigned to $a$.
- Optimal (pd-exp): order value of edges assigned to $a$: $v(1) \geq v(2) \ldots \geq v(C_a)$:

$$\beta_a = \frac{1}{C_a(e-1)} \sum_{j=1}^{C_a} v(j)(1 + \frac{1}{C_a})^{j-1}.$$

# An Optimal Algorithm

Assign impression to an advertiser $a$:
maximizing (imp. value - $\beta_a$),

- Greedy: $\beta_a =$ min. impression assigned to $a$.
- Better (pd-avg): $\beta_a =$ average value of top $C_a$ impressions assigned to $a$.
- Optimal (pd-exp): order value of edges assigned to $a$: $v(1) \geq v(2) \ldots \geq v(C_a)$:

$$\beta_a = \frac{1}{C_a(e-1)} \sum_{j=1}^{C_a} v(j)(1 + \frac{1}{C_a})^{j-1}.$$

- Thm: pd-exp achieves optimal competitive Ratio: $1 - \frac{1}{e} - \epsilon$ if $C_a > O(\frac{1}{\epsilon})$. [Feldman, Korula, M., Muthukrishnan, Pal 2009]

# Online Generalized Assignment (with free disposal)

- Multiple Knapsack: Item $i$ may have different value ($v_{ia}$) and different size $s_{ia}$ for different ads $a$.
- DA: $s_{ia} = 1$, AW: $v_{ia} = s_{ia}$.

$$\begin{aligned} \max \sum_{i,a} v_{ia} x_{ia} & & & \\ \sum_{a} x_{ia} &\leq 1 & & (\forall\ i) \\ \sum_{i} s_{ia} x_{ia} &\leq C_a & & (\forall\ a) \\ x_{ia} &\geq 0 & & (\forall\ i, a) \end{aligned}$$

$$\begin{aligned} \min \sum_{a} C_a \beta_a + \sum_{i} z_i & & & \\ s_{ia}\beta_a + z_i &\geq v_{ia} & & (\forall i, a) \\ \beta_a, z_i &\geq 0 & & (\forall i, a) \end{aligned}$$

# Online Generalized Assignment (with free disposal)

- Multiple Knapsack: Item $i$ may have different value ($v_{ia}$) and different size $s_{ia}$ for different ads $a$.
- DA: $s_{ia} = 1$, AW: $v_{ia} = s_{ia}$.

$$\max \sum_{i,a} v_{ia} x_{ia}$$
$$\begin{aligned} \sum_a x_{ia} &\leq 1 && (\forall\ i) \\ \sum_i s_{ia} x_{ia} &\leq C_a && (\forall\ a) \\ x_{ia} &\geq 0 && (\forall\ i, a) \end{aligned}$$

$$\min \sum_a C_a \beta_a + \sum_i z_i$$
$$\begin{aligned} s_{ia}\beta_a + z_i &\geq v_{ia} && (\forall i, a) \\ \beta_a, z_i &\geq 0 && (\forall i, a) \end{aligned}$$

- Offline Optimization: $1 - \frac{1}{e} - \delta$-aprx[FGMS07,FV08].
- Thm[FKMMP09]: There exists a $1 - \frac{1}{e} - \epsilon$-approximation algorithm if $\frac{C_a}{\max s_{ia}} \geq \frac{1}{\epsilon}$.

# Proof Idea: Primal-Dual Analysis [BJN]

$$\max \sum_{i,a} v_{ia} x_{ia}$$

$$\begin{aligned}
\sum_a x_{ia} &\leq 1 && (\forall\ i) \\
\sum_i s_{ia} x_{ia} &\leq C_a && (\forall\ a) \\
x_{ia} &\geq 0 && (\forall\ i, a)
\end{aligned}$$

$$\min \sum_a C_a \beta_a + \sum_i z_i$$

$$\begin{aligned}
s_{ia}\beta_a + z_i &\geq v_{ia} && (\forall i, a) \\
\beta_a, z_i &\geq 0 && (\forall i, a)
\end{aligned}$$

# Proof Idea: Primal-Dual Analysis [BJN]

$$\max \sum_{i,a} v_{ia} x_{ia}$$

$$\begin{aligned} \sum_a x_{ia} &\leq 1 && (\forall\ i) \\ \sum_i s_{ia} x_{ia} &\leq C_a && (\forall\ a) \\ x_{ia} &\geq 0 && (\forall\ i, a) \end{aligned}$$

$$\min \sum_a C_a \beta_a + \sum_i z_i$$

$$\begin{aligned} s_{ia}\beta_a + z_i &\geq v_{ia} && (\forall i, a) \\ \beta_a, z_i &\geq 0 && (\forall i, a) \end{aligned}$$

- Proof:
  1. Start from feasible primal and dual ($x_{ia} = 0$, $\beta_a = 0$, and $z_i = 0$, i.e., Primal=Dual=0).
  2. After each assignment, update $x, \beta, z$ variables and keep primal and dual solutions.
  3. Show $\Delta(\text{Dual}) \leq (1 - \frac{1}{e})\Delta(\text{Primal})$.

# SWM with online items?

Special Cases:

| | Online Matching: $v_{ia} = s_{ia} = 1$ | Disp. Ads (DA): $s_{ia} = 1$ | AdWords (AW): $s_{ia} = v_{ia}$ |
|---|---|---|---|
| Worst-Case | Greedy: $\frac{1}{2}$, [KVV]: $1 - \frac{1}{e}$-aprx | Free Disposal [FKMMP09]: $1 - \frac{1}{e}$-aprx $C_a \gg \max s_{ia}$ | [MSVV,BJN]: $1 - \frac{1}{e}$-aprx if $B_a \gg b_{ia}$. |

- Open Problem 1: What about small budgets ($B_a$) or small capacities ($C_a$)?
- Open Problem 2: How to generalize large budgets ($B_a$) and large capacities ($C_a$) for online SWM with online items, and get a $1 - 1/e$-approximation?

# Outline

- Problems: SWM and Online Ad Allocation
- Online Generalized Assignment (GAP)
- Page-based Allocation and SWM with online buyers
- Stochastic Settings:
  - Online Stochastic Matching: Primal Algorithms
  - Online Stochastic Packing: Dual Algorithms
  - Experimental Results
  - Simultaneous Stochastic and Adversarial Approximations

# Page-based Ad Allocation

- Each page can be assigned multiple ads.
- Feasible configurations of ads:
  - Exclusion Constraints: Nike and Adidas ads should not appear on the same page?
  - All-or-nothing Constraints: Either all ads on the page are from Ford or none.
  - Diversity Constraints: at most one ad from one advertiser.

# Page-based Ad Allocation

- Each page can be assigned multiple ads.
- Feasible configurations of ads:
  - Exclusion Constraints: Nike and Adidas ads should not appear on the same page?
  - All-or-nothing Constraints: Either all ads on the page are from Ford or none.
  - Diversity Constraints: at most one ad from one advertiser.
- Dependent-value model based on value sharing: $v_p(C, a)$ = value of ad $a$ in configuration $C$ on a page $p$.
- Assume $v_p(C, a)$ is cross-monotonic, i.e.,

$$\sum_{a' \neq a} v_p((C \backslash a), a') \geq \sum_{a' \neq a} v_p(C, a').$$

# Page-based Ad Allocation: LP and Algorithm

$$\begin{array}{rcl}
\text{maximize} & \sum_{p,C\in\mathcal{C}_p,a} v_p(C,a)\cdot x_{p,C,a} & \text{(Primal)} \\
\forall p,a: & \sum_{C\in\mathcal{C}_p} x_{p,C,a} \le 1 & [z_{p,a}] \\
\forall a: & \sum_{p,C\in\mathcal{C}_p} |C_a|\cdot x_{p,C,a} \le n_a & [\beta_a] \\
\forall p, C\in\mathcal{C}_p, a: & x_{p,C,a} \ge 0 &
\end{array}$$

$$\begin{array}{rcl}
\text{minimize} & \sum_{p,a} z_{p,a} + \sum_a n_a\cdot\beta_a & \text{(Dual)} \\
\forall p, C\in\mathcal{C}_p, a: & z_{p,a} + |C_a|\cdot\beta_a \ge v_p(C,a) & [x_{p,C,a}] \\
\forall p,a: & z_{p,a}\ge 0, \beta_a \ge 0 &
\end{array}$$

# Page-based Ad Allocation: LP and Algorithm

$$\begin{array}{rcl} \text{maximize} & \sum_{p,C\in\mathcal{C}_p,a} v_p(C,a)\cdot x_{p,C,a} & \text{(Primal)} \\ \forall p,a: & \sum_{C\in\mathcal{C}_p} x_{p,C,a} \le 1 & [z_{p,a}] \\ \forall a: & \sum_{p,C\in\mathcal{C}_p} |C_a|\cdot x_{p,C,a} \le n_a & [\beta_a] \\ \forall p, C\in\mathcal{C}_p, a: & x_{p,C,a} \ge 0 & \end{array}$$

$$\begin{array}{rcl} \text{minimize} & \sum_{p,a} z_{p,a} + \sum_a n_a\cdot\beta_a & \text{(Dual)} \\ \forall p, C\in\mathcal{C}_p, a: & z_{p,a} + |C_a|\cdot\beta_a \ge v_p(C,a) & [x_{p,C,a}] \\ \forall p,a: & z_{p,a}\ge 0, \beta_a \ge 0 & \end{array}$$

1. Initially, $\beta_a = 0$ for each advertiser $a$.
2. For every arriving page, do the following:
   2.1 Choose feasible allocation $C$ to maximize the discounted value $\sum_{a\in C} v_p(C,a) - |C_a|\cdot\beta_a$.
   2.2 Allocate according to $C$.
   2.3 Recalculate $\beta_a$ as defined as the exp-avg scoring.

# Page-based Ad Allocation: Algorithm and Result

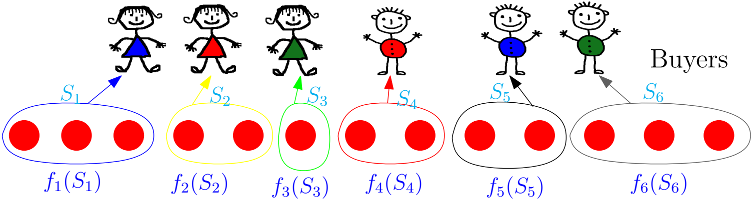
1. Initially, $\beta_a = 0$ for each advertiser $a$.
2. For every arriving page, do the following:
   2.1 Choose feasible allocation $C$ to maximize the discounted value $\sum_{a \in C} v_p(C, a) - |C_a| \cdot \beta_a$.
   2.2 Allocate according to $C$.
   2.3 Recalculate $\beta_a$ as defined as the exp-avg scoring.

[Exp-Avg Scoring] Let $w_1 \geq w_2 \geq \cdots \geq w_n$ be the top $n$ weights assigned to an advertiser $a$ with capacity $n$, and let $d \in \{1, \ldots, n\}$, $\beta_a = \frac{1}{\hat{n}_a \cdot (e_{n/d} - 1)} \cdot \sum_{i=1}^{n} \alpha^{i-1} \cdot w_i$, where $\alpha_a = (1 + \frac{d}{n_a})^{\frac{1}{d}}$.

Theorem (Korula, M., Yan)

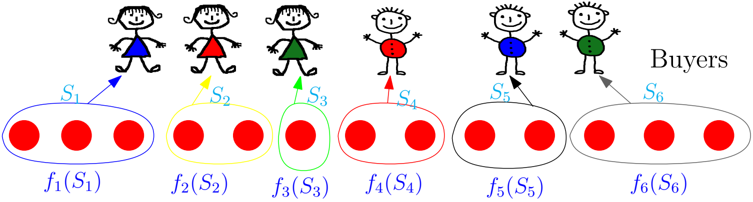
*For the page-based ad allocation problem with cross-monotonic value-sharing, this algorithm gives a $(1 - \frac{1}{e} - \epsilon)$-approximation for large capacities. For small capacities, the approximation ratio is $\frac{1}{2}$.*

# SWM with Online buyers




- Goal: Partition items to maximize social welfare, i.e, $\sum_i f_i(S_i)$.
- SWM with online buyers with re-assignment: buyers arrive one by one, and we can re-assign items from older buyers to new buyers (but not vice versa).

# SWM with Online buyers



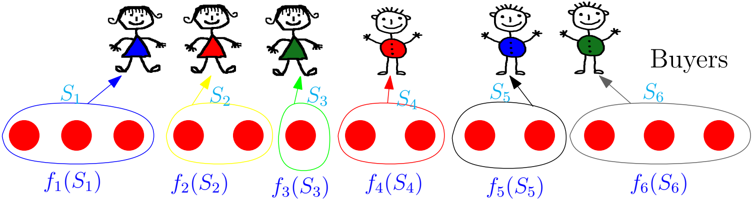

- Goal: Partition items to maximize social welfare, i.e, $\sum_i f_i(S_i)$.
- SWM with online buyers with re-assignment: buyers arrive one by one, and we can re-assign items from older buyers to new buyers (but not vice versa).

**Theorem (Korula, M., Yan)**

*The SWM with online buyers admits a $(1 - \frac{1}{e} - \epsilon)$- and $(\frac{1}{2})$-approximations for large and small multiplicity of items, respectively. The algorithm uses a demand oracle access to the submodular function.*

- Proof Technique: reduce SWM with online buyers to the page-based allocation with cross-monotonic value-sharing.

# SWM with Online buyers




**Theorem** (Korula, M., Yan)

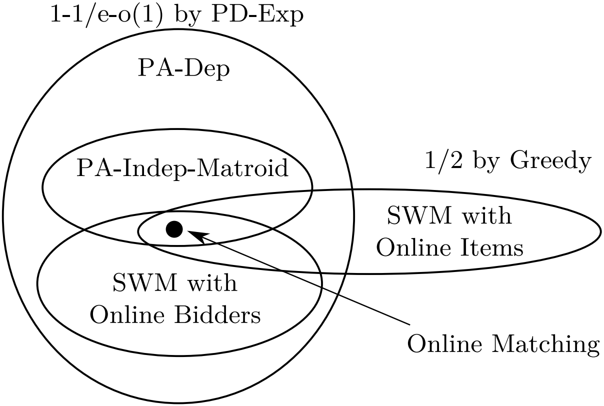
*The SWM with online buyers admits a* $(1 - \frac{1}{e} - \epsilon)$*- and* $(\frac{1}{2})$*-approximations for large and small multiplicity of items, respectively. The algorithm uses a demand oracle access to the submodular function.*

- Proof Technique: reduce SWM with online buyers to the page-based allocation with cross-monotonic value-sharing.
- Item $a$ corresponds to advertiser $a$ with capacity one.
- Pages corresponds to buyers, and

$$v_p(S, a) = f_p(\{1, \ldots, a\} \cap S) - f_p(\{1, \ldots, a-1\} \cap S).$$

# SWM with Online buyers: Summary

PA-Dep: Page-based Ad Allocation with value-sharing.
PA-Indep-Matroid: Page-based Ad Allocation with separable vlauations and matroid constraints.



# Outline

- Problems: SWM and Online Ad Allocation
- Online Generalized Assignment (GAP)
- Page-based Allocation and SWM with online buyers
- Stochastic Settings:
  - Online Stochastic Matching: Primal Algorithms
  - Online Stochastic Packing: Dual Algorithms
  - Experimental Results
  - Simultaneous Stochastic and Adversarial Approximations

# Ad Allocation: Problems and Models

| | Online Matching: $v_{ia} = s_{ia} = 1$ | Disp. Ads (DA): $s_{ia} = 1$ | AdWords (AW): $s_{ia} = v_{ia}$ |
|---|---|---|---|
| Worst Case | Greedy: $\frac{1}{2}$, [KVV]: $1 - \frac{1}{e}$-aprx | Free Disposal [FKMMP09]: $1 - \frac{1}{e}$-aprx | [MSVV,BJN]: $1 - \frac{1}{e}$-aprx if $B_a \gg b_{ia}$. |
| Stochastic (random arrival order) | [FMMM09,MOS11] 0.703-aprx | [HMZ11] 0.66-aprx i.i.d with known distribution [FHKMS10,AWY]: $1 - \epsilon$-aprx, if $\text{OPT} \gg \max v_{ia}$ and $C_a \gg \max s_{ia}$ | [DH09]: $1 - \epsilon$-aprx, if $\text{OPT} \gg \max v_{ia}$ |

# Primal Algorithm: "Two-suggested-matchings"

$$\text{“ALG is } \alpha\text{-approximation?” if w.h.p.,} \frac{\mathrm{ALG}(H)}{\mathrm{OPT}(H)} \geq \alpha$$

Simple Primal Algorithm:

- Find one matching in expected graph $G$ offline, and try to apply it online.
- Tight $1 - \frac{1}{e}$-approximation.

# Primal Algorithm: "Two-suggested-matchings"

"ALG is $\alpha$-approximation?" if w.h.p., $\frac{\mathrm{ALG}(H)}{\mathrm{OPT}(H)} \geq \alpha$

Simple Primal Algorithm:

- Find one matching in expected graph $G$ offline, and try to apply it online.
- Tight $1 - \frac{1}{e}$-approximation.

Better Algorithm: Two-Suggested-Matchings

- Offline: Find two disjoint matchings, blue(B) and red(R), on the expected graph $G$.
- Online: try the blue matching first, then if that doesn't work, try the red one.

# Primal Algorithm: "Two-suggested-matchings"

"ALG is $\alpha$-approximation?" if w.h.p., $\frac{\mathrm{ALG}(H)}{\mathrm{OPT}(H)} \geq \alpha$

Simple Primal Algorithm:

- Find one matching in expected graph $G$ offline, and try to apply it online.
- Tight $1 - \frac{1}{e}$-approximation.

Better Algorithm: Two-Suggested-Matchings

- Offline: Find two disjoint matchings, blue(B) and red(R), on the expected graph $G$.
- Online: try the blue matching first, then if that doesn't work, try the red one.
- Thm: Tight $\frac{1-2/e^2}{4/3-2/3e} \geq 0.67$

(Feldman, M., M., Muthukrishnan, 2009).

# Analysis: Two-suggested-matching Algorithm

- Proof Ideas: Balls-into-Bins concentration inequalities, structural properties of min-cuts, etc.

# Analysis: Two-suggested-matching Algorithm

- Proof Ideas: Balls-into-Bins concentration inequalities, structural properties of min-cuts, etc.
- Bounding ALG: Classify $a \in A$ based on its neighbors in the blue and red matchings: $A_{BR}, A_{BB}, A_B, A_R$

$$ALG \geq \left(1 - \frac{1}{e^2}\right)|A_{BB}| + \left(1 - \frac{2}{e^2}\right)|A_{BR}| + \left(1 - \frac{3}{2e}\right)(|A_B| + |A_R|)$$

# Analysis: Two-suggested-matching Algorithm

- Proof Ideas: Balls-into-Bins concentration inequalities, structural properties of min-cuts, etc.
- Bounding ALG: Classify $a \in A$ based on its neighbors in the blue and red matchings: $A_{BR}, A_{BB}, A_B, A_R$

$$ALG \geq \left(1 - \frac{1}{e^2}\right)|A_{BB}| + \left(1 - \frac{2}{e^2}\right)|A_{BR}| + \left(1 - \frac{3}{2e}\right)(|A_B| + |A_R|)$$

- Bounding $\text{OPT}$: Find min-cut in augmented expected graph $G$, and use it min-cut in $G$ as a "guide" for cut in each scenario.

# Primal Algorithms: Two Offline Solutions

- Online stochastic matching: 0.67-approximation[FMMM09]
- Improved to 0.702-approximation[MOS11]
- Improve to 0.703-approximation using 3 matchings[HMZ11]
- Online stochastic weighted matching: 0.66-approximation [HMZ11].

# Online Stochastic Weighted Matching

"ALG is $\alpha$-approximation?" if $\frac{\mathrm{E}[\mathrm{ALG}(H)]}{\mathrm{E}[\mathrm{OPT}(H)]} \geq \alpha$

Power of Two Choices:

- Offline:
  1. Find an optimal fractional solution $x_e$ to a discounted matching LP, where $x_e \leq 1 - \frac{1}{e}$.
  2. Sample a matching $M$ from $x$.
  3. Let $M_0 = M_1 \backslash M$, where $M_1$ is the maximum weighted matching.
- Online: try the edges in $M$ first, and if it does not work, try $M_0$.
- Thm: Approximation factor is better than 0.66. (Haeupler, M., ZadiMoghaddam, 2011).

# Online Stochastic Weighted Matching

"ALG is $\alpha$-approximation?" if $\frac{\mathrm{E}[\mathrm{ALG}(H)]}{\mathrm{E}[\mathrm{OPT}(H)]} \geq \alpha$

Power of Two Choices:

- Offline:
  1. Find an optimal fractional solution $x_e$ to a discounted matching LP, where $x_e \leq 1 - \frac{1}{e}$.
  2. Sample a matching $M$ from $x$.
  3. Let $M_0 = M_1 \backslash M$, where $M_1$ is the maximum weighted matching.
- Online: try the edges in $M$ first, and if it does not work, try $M_0$.
- Thm: Approximation factor is better than 0.66. (Haeupler, M., ZadiMoghaddam, 2011).
- Open Problem 3: Generalize this algorithm to online stochastic SWM and get better than $1 - 1/e$ with extra assumptions.

# Ad Allocation: Problems and Models

| | Online Matching: $v_{ia} = s_{ia} = 1$ | Disp. Ads (DA): $s_{ia} = 1$ | AdWords (AW): $s_{ia} = v_{ia}$ |
|---|---|---|---|
| Worst Case | Greedy: $\frac{1}{2}$, [KVV]: $1-\frac{1}{e}$-aprx | ? | [MSVV,BJN]: $1-\frac{1}{e}$-aprx if $B_a \gg b_{ia}$. |
| Stochastic (i.i.d.) | [FMMM09,MOS11] 0.703-aprx i.i.d with known distribution | [HMZ11] 0.66-aprx i.i.d with known distribution [FHKMS10,AWY]: $1-\epsilon$-aprx, if $\mathrm{OPT} \gg \max v_{ia}$ and $C_a \gg \max s_{ia}$ | [DH09]: $1-\epsilon$-aprx, if $\mathrm{OPT} \gg \max v_{ia}$ |

random order = i.i.d. model with unknown distribution

# Stochastic DA: Dual Algorithm

$$\max \sum_{i,a} v_{ia} x_{ia}$$

$$\begin{aligned}
\sum_a x_{ia} &\leq 1 && (\forall\ i) \\
\sum_i x_{ia} &\leq C_a && (\forall\ a) \\
x_{ia} &\geq 0 && (\forall\ i, a)
\end{aligned}$$

$$\min \sum_a C_a \beta_a + \sum_i z_i$$

$$\begin{aligned}
z_i &\geq v_{ia} - \beta_a && (\forall i, a) \\
\beta_a, z_i &\geq 0 && (\forall i, a)
\end{aligned}$$

Algorithm:

- Observe the first $\epsilon$ fraction sample of impressions.
- Learn a dual variable for each ad $\beta_a$, by solving the dual program on the sample.
- Assign each impression $i$ to ad a that maximizes $v_{ia} - \beta_a$.

# Stochastic DA: Dual Algorithm

$$\max \sum_{i,a} v_{ia} x_{ia}$$

$$\begin{aligned}
\sum_a x_{ia} &\leq 1 && (\forall\ i) \\
\sum_i x_{ia} &\leq C_a && (\forall\ a) \\
x_{ia} &\geq 0 && (\forall\ i, a)
\end{aligned}$$

$$\min \sum_a C_a \beta_a + \sum_i z_i$$

$$\begin{aligned}
z_i &\geq v_{ia} - \beta_a && (\forall i, a) \\
\beta_a, z_i &\geq 0 && (\forall i, a)
\end{aligned}$$

Algorithm:

- Observe the first $\epsilon$ fraction sample of impressions.
- Learn a dual variable for each ad $\beta_a$, by solving the dual program on the sample.
- Assign each impression $i$ to ad a that maximizes $v_{ia} - \beta_a$.

# Stochastic DA: Dual Algorithm

Feldman, Henzinger, Korula, M., Stein 2010

Thm[FHKMS10,AWY]: W.h.p, this algorithm is a $(1-O(\epsilon))$-aprx, as long as each item has low value ($v_{ia} \leq \frac{\epsilon \mathrm{OPT}}{m \log n}$), and large capacity ($C_a \leq \frac{m \log n}{\epsilon^3}$)

# Stochastic DA: Dual Algorithm

Feldman, Henzinger, Korula, M., Stein 2010

Thm[FHKMS10,AWY]: W.h.p, this algorithm is a $(1-O(\epsilon))$-aprx, as long as each item has low value ($v_{ia} \leq \frac{\epsilon \mathrm{OPT}}{m \log n}$), and large capacity ($C_a \leq \frac{m \log n}{\epsilon^3}$)

Fact: If optimum $\beta_a^*$ are known, this alg. finds $\mathrm{OPT}$

- Proof: Comp. slackness. Given $\beta_a^*$, compute $x^*$ as follows: $x_{ia}^* = 1$ if $a = \mathrm{argmax}(v_{ia} - \beta_a^*)$.

# Stochastic DA: Dual Algorithm

Feldman, Henzinger, Korula, M., Stein 2010

Thm[FHKMS10,AWY]: W.h.p, this algorithm is a $(1-O(\epsilon))$-aprx, as long as each item has low value ($v_{ia} \le \frac{\epsilon \mathrm{OPT}}{m \log n}$), and large capacity ($C_a \le \frac{m \log n}{\epsilon^3}$)

Fact: If optimum $\beta_a^*$ are known, this alg. finds $\mathrm{OPT}$

- Proof: Comp. slackness. Given $\beta_a^*$, compute $x^*$ as follows: $x_{ia}^* = 1$ if $a = \operatorname{argmax}(v_{ia} - \beta_a^*)$.

Lemma: In the random order model, W.h.p., the sample $\beta_a'$ are close to $\beta_a^*$.

- Extending DH09.

# Ad Allocation: Problems and Models

| | Online Matching: $v_{ia} = s_{ia} = 1$ | Disp. Ads (DA): $s_{ia} = 1$ | AdWords (AW): $s_{ia} = v_{ia}$ |
|---|---|---|---|
| Worst Case | Greedy: $\frac{1}{2}$, [KVV]: $1 - \frac{1}{e}$-aprx | Free Disposal [FKMMP09]: $1 - \frac{1}{e}$-aprx | [MSVV,BJN]: $1 - \frac{1}{e}$-aprx if $B_a \gg b_{ia}$. |
| Stochastic (random arrival order) | [FMMM09,MOS11] 0.703-aprx | [HMZ11] 0.66-aprx i.i.d with known distribution [FHKMS10,AWY]: $1 - \epsilon$-aprx, if $\text{OPT} \gg \max v_{ia}$ and $C_a \gg \max s_{ia}$ | [DH09]: $1 - \epsilon$-aprx, if $\text{OPT} \gg \max v_{ia}$ |

# Outline

- Problems: SWM and Online Ad Allocation
- Online Generalized Assignment through Primal-Dual Analysis
- Page-based Allocation and SWM with online bidders
- Stochastic Settings:
  - Online Stochastic Matching: Primal Algorithms
  - Online Stochastic Packing: Dual Algorithms
  - Experimental Results
  - Simultaneous Stochastic and Adversarial Approximations

# Experiments: setup

- Real ad impression data from several large publishers
- 200k - 1.5M impressions in simulation period
- 100 - 2600 advertisers
- Edge weights = predicted click probability
- Efficiency: free disposal model
- Algorithms:
  - greedy: maximum marginal value
  - pd-avg, pd-exp: pure online primal-dual from [FKMMP09].
  - dualbase: training-based primal-dual [FHKMS10]
  - hybrid: convex combo of training based, pure online.
  - lp-weight: optimum efficiency

# Experimental Evaluation: Summary

| Algorithm | Avg Efficiency% |
|---|---|
| opt | 100 |
| greedy | 69 |
| pd-avg | 77 |
| pd-exp | 82 |
| dualbase | 87 |
| hybrid | 89 |

- pd-exp & pd-avg outperform greedy by 9% and 14% (with more improvements in *tight* competition.)
- dualbase outperforms pure online algorithms by 6% to 12%.
- Hybrid has a mild improvement of 2% (up to 10%).
- pd-avg performs much better than the theoretical analysis.

# In Production

- Algorithms inspired by these techniques are in use at Google display ad serving system, delivering billions of ads per day.

# In Production

- Algorithms inspired by these techniques are in use at Google display ad serving system, delivering billions of ads per day.
- Smooth Delivery of Display Ads (Bhalgat, Feldman, M.)
- Display Ad Allocation with Ad Exchange (Belsairo, Feldman, M., Muthukrishnan)

# Simultaneous worst-case & stochastic optimization

- Dealing with Traffic Spikes: Primal and Dual techniques fail in the adversarial models.

# Simultaneous worst-case & stochastic optimization

- Dealing with Traffic Spikes: Primal and Dual techniques fail in the adversarial models.
- Assuming $C_a \gg \max s_{ia}$, are there algorithms that achieve good approximation factors for both adversarial and stochastic models simultaneously?

# Simultaneous worst-case & stochastic optimization

- Dealing with Traffic Spikes: Primal and Dual techniques fail in the adversarial models.
- Assuming $C_a \gg \max s_{ia}$, are there algorithms that achieve good approximation factors for both adversarial and stochastic models simultaneously?
- Yes for unweighted edges!(M.,OveisGharan, ZadiMoghaddam)
  - Balance algorithm achieves $1-\epsilon$-approximation for random order and $1-\frac{1}{e}$-approximation for the adversarial model.

# Simultaneous worst-case & stochastic optimization

- Dealing with Traffic Spikes: Primal and Dual techniques fail in the adversarial models.
- Assuming $C_a \gg \max s_{ia}$, are there algorithms that achieve good approximation factors for both adversarial and stochastic models simultaneously?
- Yes for unweighted edges!(M.,OveisGharan, ZadiMoghaddam)
  - Balance algorithm achieves $1-\epsilon$-approximation for random order and $1-\frac{1}{e}$-approximation for the adversarial model.
- No for weighted edges! (M.,OveisGharan,ZadiMoghaddam)
  - Impossible: $1-\frac{1}{e}$-approximation for adversarial and better than 0.97-approximation for random order.
  - PD-EXP achieves achieves 0.76-approximation for random order and $1-\frac{1}{e}$-approximation for the adversarial model.

# Online SWM: Interesting Problems

- Adversarial:
  - Open Problem 1: Get better than $1/2$-approximation for online budgeted allocation with small budgets ($B_a$) or small capacities ($C_a$)?
  - Open Problem 2: How to generalize large budgets ($B_a$) and large capacities ($C_a$) assumptions for online SWM with online items, and get a $1 - 1/e$-approximation?
- Primal Techniques:
  - Open Problem 3: Generalize the two-offline-matching algorithm to online stochastic SWM and get better than $1 - 1/e$ *with extra assumptions* for the iid model.
  - Vondrak: MSV'08 implies that getting better than $1 - 1/e$ is impossible without extra assumptions.
- Dual Techniques:
  - Open Problem 4: Generalize the dual-based algorithm to online stochastic SWM and get better than $1 - \epsilon$ with extra assumptions for random order model.