Toward Controlling Discrimination in Online Ad Auctions

L. Elisa Celis\textsuperscript{1}, Anay Mehrotra\textsuperscript{2}, Nisheeth K. Vishnoi\textsuperscript{1}

\textsuperscript{1}Yale University \textsuperscript{2}IIT Kanpur

Poster: Thursday, June 13\textsuperscript{th}, 6:30PM-9:00PM @ Pacific Ballroom #125
Online Advertising

Online advertising is a major source of revenue for many online platforms, contributing $100+ billion in revenue in 2018.
Discrimination in Online Advertising

On Facebook (with 52% women) a STEM job ad was shown to 20% more men than women (Lambrecht & Tucker 2018).

Also observed across race (Sweeney 2013) and in housing ads (Ali et al. 2019).
Discrimination in Online Advertising

On Facebook (with 52% women) a STEM job ad was shown to 20% more men than women (Lambrecht & Tucker 2018).

Also observed across race (Sweeney 2013) and in housing ads (Ali et al. 2019).

Can we develop a framework to mitigate this kind of discrimination?
Model and Preliminaries

• $n$ advertisers, $m$ types of users.

• For type $j \in [m]$, receiving bids $v_j \in \mathbb{R}_{\geq 0}^n$ as input, mechanism $\mathcal{M}$ decides an allocation $x(v_j) \in [0,1]^n$ and a price $p(v_j) \in \mathbb{R}^n$.

Choosing the mechanism $\mathcal{M}$, is a well studied problem.
Fairness Constraints

Coverage $q_{ij}$: Probability advertiser $i$ wins and user is of type $j$

For all $i \in [n], j \in [m]$

$$\ell_{ij} \leq \frac{q_{ij}}{\sum_{t=1}^{m} q_{it}} \leq u_{ij}.$$ 

Allows for

- constraints on some or all advertisers,
- across some or all sub-populations, and
- varying the fairness metric by varying the constraints.

Works for a wide class of fairness metrics; e.g., (Celis, Huang, Keswani and Vishnoi 2019).

Fairness Metric: *Equal Representation*
Constraints: $\ell_{ij} = \frac{1}{3}$ and $u_{ij} = \frac{1}{3}$
Infinite Dimensional Fair Advertising Problem

- For many platforms $\mathcal{M}$ is the 2\textsuperscript{nd} price auction.
- Myerson’s mechanism is the 2\textsuperscript{nd} price auction on virtual values,
  $$\phi(v) := v \cdot (1 - \text{cdf}(v))/\text{pdf}(v).$$
- Let $f_{ij}$ density function of $\phi_{ij}(v)$ of advertiser $i$ for type $j$, and $\mathcal{U}$ be the dist. of types.
- $x_{ij}$ are functions – infinite dimensional optimization problem.

**Input:** $\ell, u \in \mathbb{R}^{n \times m}$

**Output:** Set of allocation rules $x_{ij}: \mathbb{R}^n \to [0,1]^n$

\[
\begin{align*}
\max_{x_{ij}(\cdot) \geq 0} \quad & \text{rev}_\mathcal{M}(x_1, x_2, \ldots, x_m) \\
\text{s.t.,} \quad & q_{ij}(x_j) \geq \ell_{ij} \sum_{t=1}^{m} q_{it}(x_t) \quad \forall \, i \in [n], \, j \in [m] \\
& q_{ij}(x_j) \leq u_{ij} \sum_{t=1}^{m} q_{it}(x_t) \quad \forall \, i \in [n], \, j \in [m] \\
& \sum_{i=1}^{n} x_{ij}(\phi_j) \leq 1 \quad \forall j \in [m], \phi_j
\end{align*}
\]

*How can we find the optimal $x_{ij}$?*
Characterization Result

Assume:

- Bids are drawn from a regular distribution. (Equivalent to Myerson.)
Characterization Result

Assume:

- Bids are drawn from a regular distribution. (Equivalent to Myerson.)

Then:

**Theorem 4.1 (Informal)** There is a “shift” \( \alpha \in \mathbb{R}^{n \times m} \) such that

\[
x_{ij}(v_j, \alpha_j) := \mathbb{I}[i \in \arg\max_{\ell \leq n}(\phi_{ij}(v_{\ell j}) + \alpha_{\ell j})]
\]

is optimal.
Characterization Result

Assume:
- Bids are drawn from a regular distribution. (Equivalent to Myerson.)

Then:

**Theorem 4.1 (Informal)** There is a “shift” $\alpha \in \mathbb{R}^{n \times m}$ such that

$$x_{ij}(v_j, \alpha_j) := \mathbb{I}[i \in \arg\max_{\ell \in [n]} (\phi_{\ell j}(v_{\ell j}) + \alpha_{\ell j})]$$

is optimal.

**Infinite Dimensional Optimization $\rightarrow$ Finite Dimensional Optimization.**
Algorithmic Result

Assume:

\[ \forall i \in [n], j \in [m] \quad q_{ij} > \eta \]  
\text{(Minimum coverage)}

\[ \forall v \in \text{supp}(f_{ij}) \quad \mu_{\text{min}} \leq f_{ij}(v) \leq \mu_{\text{max}} \]  
\text{(Distributed Dist.)}

\[ \forall v_1, v_2 \in \text{supp}(f_{ij}) \quad |f_{ij}(v_1) - f_{ij}(v_2)| \leq L|v_1 - v_2| \]  
\text{(Lipschitz Cont. Dist.)}

\[ \forall i \in [n], j \in [m] \quad |\mathbb{E}[\phi_{ij}]| \leq \rho \]  
\text{(Bounded bid)}

Then:

**Theorem 4.3 (Informal)** There is an algorithm which solves (1) in

\[ \tilde{O} \left( n^7 \varepsilon^{-2} \log m \cdot \frac{(\mu_{\text{max}} \rho)^2}{(\mu_{\text{min}} \eta)^4} (L + n^2 \mu_{\text{max}}^2) \right) \] steps.
Empirical Results

Yahoo! A1 dataset; contains real bids from Yahoo! Online Auctions.

Keyword ↔ User type, consider “similar” keywords pairs.

Setting: $m = 2, u_{ij} = 1$, and $\#\text{auctions} = 3282$.
Vary: $\ell_{ij} = \ell \in [0,0.5]$
Measures:
- Fairness slift($F$) := $\min_{ij} q_{ij} / (1 - q_{ij})$, and
- Revenue ratio $\kappa_{M,F} := \text{rev}_M / \text{rev}_F$. 
Conclusion and Future Work

We give an optimal truthful mechanism which provably satisfies fairness constraints and an efficient algorithm to find it. We observe a minor loss to the revenue and change to advertiser distribution when using it.

• How does the mechanism affect user and advertiser satisfaction?
• Can we incorporate asynchronous campaigns?
• Can we extend our results to the GSP auctions?

Thanks!

Poster: Thursday, June 13th, 6:30PM-9:00PM @ Pacific Ballroom #125