AdKDD2020

Unbiased Lift-based Bidding System

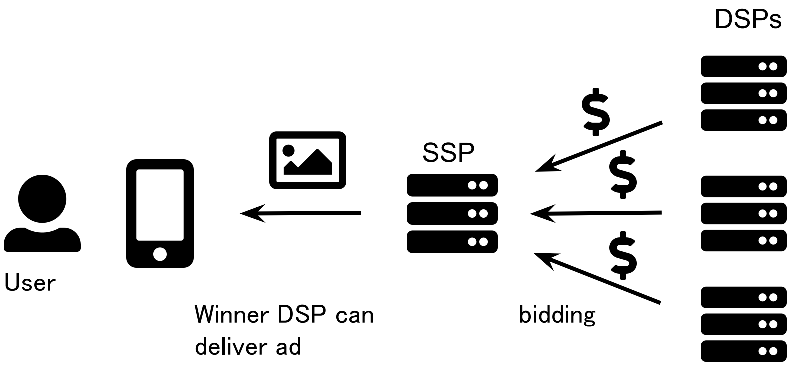
Daisuke Moriwaki, Yuta Hayakawa, Isshu Munemasa
CyberAgent, Inc.
{moriwaki_daisuke, hayakawa_yuta, munemasa_isshu}@cyberagent.co.jp

Yuta Saito
Tokyo Institute of Technology
saito.y.bj@m.titech.ac.jp

Akira Matsui
University of Southern California
amatui@usc.edu

Background

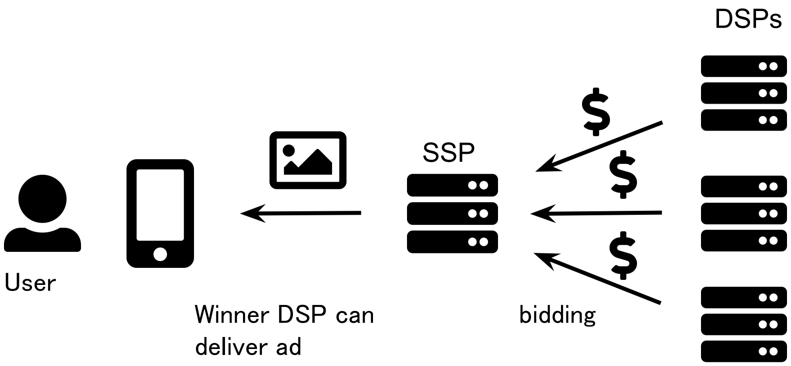
- DSPs (Demand-Side Platforms) Buy adslot through Real-Time Bidding



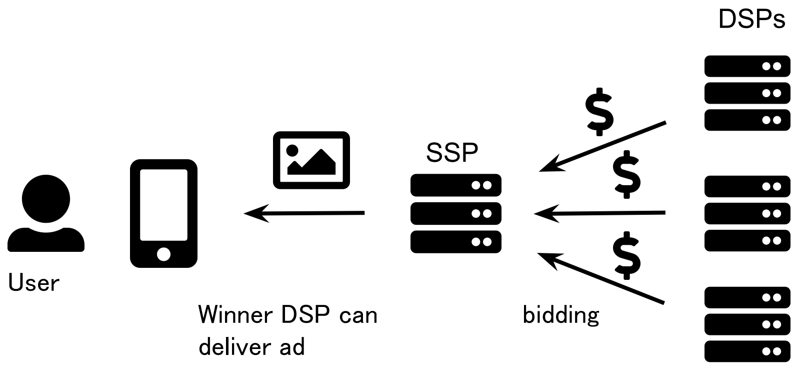
Background

- DSPs (Demand-Side Platforms) Buy adslot through Real-Time Bidding
- The prevalent bidding strategy is *Performance-based* or *Value-based* bidding

$$bid = CPC \cdot eCTR \cdot \alpha$$
or
$$bid = CPA \cdot eCTR \cdot eCVR \cdot \alpha$$



Background



- DSPs (Demand-Side Platforms) Buy adslot through Real-Time Bidding
- The prevalent bidding strategy is *Performance-based* or *Value-based* bidding

$$bid = CPC \cdot eCTR \cdot \alpha$$
or
$$bid = CPA \cdot eCTR \cdot eCVR \cdot \alpha$$

- Performance-based bidding is optimal under current billing system CPC, CPA because it maximizes DSP’s profit

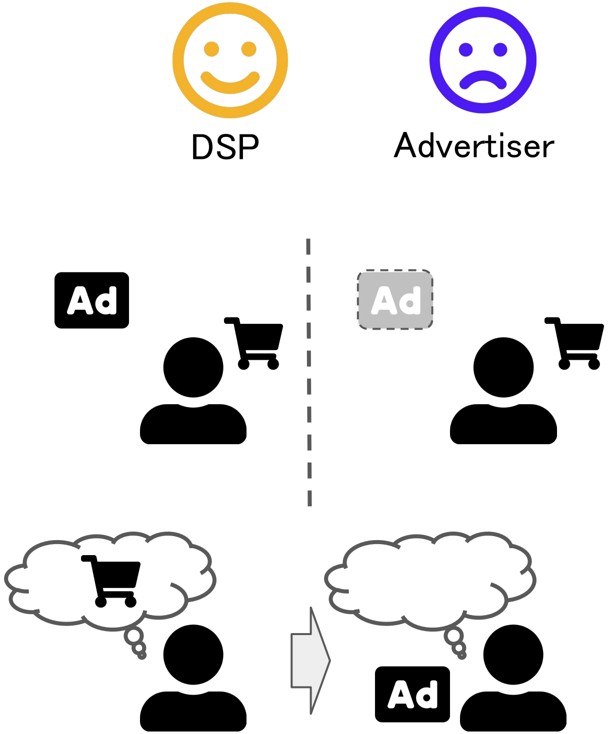
Background

- While the strategy is consistent with the DSP's objective, it is NOT always the optimal strategy for the *clients (advertisers)*

DSP Advertiser

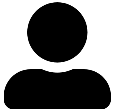
Background

- While the strategy is consistent with the DSP’s objective, it is NOT always the optimal strategy for the *clients (advertisers)*
- Example
  - A user may purchase a good without ad while the advertiser pay DSP fee for the ad
  - A user may stop buying a good because of an ad while the DSP is not given any penalty

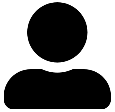


Background

- More specifically, suppose DSP bids adslot for two potential customers



$$\mathbb{E}[Y|ad] = 0.8$$
$$\mathbb{E}[Y|no\,ad] = 0.7$$

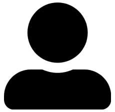


$$\mathbb{E}[Y|ad] = 0.2$$
$$\mathbb{E}[Y|no\,ad] = 0.0$$

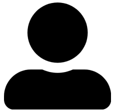
Performance-based bidder	High	Low
Optimal bidder	Low	High

Background

- More specifically, suppose DSP bids adslot for two potential customers
- Performance–based bidder bids high for user with high conversion rate given ad



$$\boxed{\mathbb{E}[Y|ad] = 0.8}$$
$$\mathbb{E}[Y|no\,ad] = 0.7$$



$$\boxed{\mathbb{E}[Y|ad] = 0.2}$$
$$\mathbb{E}[Y|no\,ad] = 0.0$$

Performance-based bidder	High	Low
Optimal bidder	Low	High

Background

- More specifically, suppose DSP bids adslot for two potential customers
- Performance-based bidder bids high for user with high conversion rate given ad
- Optimal bidder should bids high for user with high *lift*

$$\left.\begin{aligned}\mathbb{E}[Y|ad] &= 0.8\\ \mathbb{E}[Y|no\,ad] &= 0.7\end{aligned}\right\}\boxed{0.1} \qquad \left.\begin{aligned}\mathbb{E}[Y|ad] &= 0.2\\ \mathbb{E}[Y|no\,ad] &= 0.0\end{aligned}\right\}\boxed{0.2}$$

Performance-based bidder	High	Low
Optimal bidder	Low	High

Proposed Method

- *Lift-based bidding* is a bidding strategy that the bid price is based on the predicted lift τ

$$bid_{i,s(a)} = \mathrm{v}\tau(s(a) \mid \mathbf{x}_i)\alpha$$

Proposed Method

- *Lift-based bidding* is a bidding strategy that the bid price is based on the predicted lift τ

$$bid_{i,s(a)} = \boxed{\mathrm{v}}\tau(s(a) \mid \mathbf{x}_i)\alpha$$

- v is reward for a unit of conversion lift

Proposed Method

- *Lift-based bidding* is a bidding strategy that the bid price is based on the predicted lift τ

$$bid_{i,s(a)} = \mathrm{v}\tau(s(a) \mid \mathbf{x}_i)\boxed{\alpha}$$

- v is reward for a unit of conversion lift
- α is a parameter to adjust to RTB environment (competitiveness, # bid requests etc.)

Proposed Method

- *Lift-based bidding* is a bidding strategy that the bid price is based on the predicted lift τ

$$bid_{i,s(a)} = \mathrm{v}\boxed{\tau(s(a) \mid \mathbf{x}_i)}\alpha$$

- v is reward for a unit of conversion lift
- α is a parameter to adjust to RTB environment (competitiveness, # bid requests etc.)
- τ is calculated as the difference between two *potential outcomes*
  - s(a) : # ad exposure for ad *a*

$$\tau(s(a) \mid \mathbf{x}_i) = \underbrace{\mathbb{E}[y_i \mid s(a), \mathbf{x}_i]}_{\text{Expected outcome with additional ad}} - \underbrace{\mathbb{E}[y_i \mid s(a) - 1, \mathbf{x}_i]}_{\text{Expected outcome without additional ad}}$$

Proposed Method

- *Lift-based bidding* is a bidding strategy that the bid price is based on the predicted lift τ

$$bid_{i,s(a)} = \mathrm{v}\boxed{\tau(s(a) \mid \mathbf{x}_i)}\alpha$$

- v is reward for a unit of conversion lift
- α is a parameter to adjust to RTB environment (competitiveness, # bid requests etc.)
- τ is calculated as the difference between two *potential outcomes*
  - s(a) : # ad exposure for ad *a*

$$\tau(s(a) \mid \mathbf{x}_i) = \underbrace{\mathbb{E}[y_i \mid s(a), \mathbf{x}_i]}_{\text{Expected outcome with additional ad}} - \underbrace{\mathbb{E}[y_i \mid s(a) - 1, \mathbf{x}_i]}_{\text{Expected outcome without additional ad}}$$

- The basic idea is similar to Xu et al. (AAAI' 16)

Unbiased Lift-Effect Prediction

- Each potential outcome is predicted by the corresponding predictor f

$$\hat{\tau}(s(a) \mid \mathbf{x}_i) = f^{s(a)}(\mathbf{x}_i) - f^{s(a)-1}(\mathbf{x}_i)$$

Unbiased Lift-Effect Prediction

- Each potential outcome is predicted by the corresponding predictor f

$$\hat{\tau}(s(a) \mid \mathbf{x}_i) = f^{s(a)}(\mathbf{x}_i) - f^{s(a)-1}(\mathbf{x}_i)$$

- Challenges
  - Ideal estimation needs data from population distribution p(x, y, s(a))
  - In reality, one of the potential outcomes is available
  - Assignment of ads is biased due to past bidding strategy(e.g. user with higher bid price are more likely to see ad)
  - Naive ERM (Empirical Risk Minimization) leads to a biased estimator

Unbiased Lift-Effect Prediction

- Each potential outcome is predicted by the corresponding predictor f

$$\hat{\tau}\left(s(a) \mid \mathbf{x}_i\right) = f^{s(a)}\left(\mathbf{x}_i\right) - f^{s(a)-1}\left(\mathbf{x}_i\right)$$

- Challenges
  - Ideal estimation needs data from population distribution p(x, y, s(a))
  - In reality, one of the potential outcomes is available.
  - Assignment of ads is biased due to past bidding strategy (e.g. user with higher bid price are more likely to see ad)
  - Naive ERM (Empirical Risk Minimization) leads to biased estimator
- Solution
  - IPS (Inverse Propensity Score) estimation to debias the estimation

Unbiased Lift-Effect Prediciton

More specifically,

- Ideal loss function is defined as expectation of loss over population distribution

$$\mathcal{L}_{ideal}\left(f^{s(a)}\right) = \mathbb{E}_{p(\mathbf{x},y)}\left[l\left(y(s(a)), f^{s(a)}(\mathbf{x})\right)\right]$$

Unbiased Lift-Effect Prediciton

More specifically,

- Ideal loss function is defined as expectation of loss over population distribution

$$\mathcal{L}_{ideal}\left(f^{s(a)}\right) = \mathbb{E}_{p(\mathbf{x},y)}\left[l\left(y(s(a)), f^{s(a)}(\mathbf{x})\right)\right]$$

- A naive ERM estimation minimize loss over realized data which is biased since distribution of observed data is not the population distribution p(x,y|s(a)) ≠ p(x,y)

$$f^{s(a)}_{\mathrm{ERM}} = \arg\min_{f^{s(a)}} \hat{\mathcal{L}}_{\mathrm{ERM}}\left(f^{s(a)}\right) = \arg\min_{f^{s(a)}} \frac{1}{n_{s(a)}} \sum_{i \in \mathcal{D}_{s(a)}} l(y_i^{obs}, f^{s(a)}(\mathbf{x}_i))$$

Unbiased Lift-Effect Prediciton

More specifically,

- Ideal loss function is defined as expectation of loss over population distribution

$$\mathcal{L}_{ideal}\left(f^{s(a)}\right) = \mathbb{E}_{p(\mathbf{x},y)}\left[l\left(y(s(a)), f^{s(a)}(\mathbf{x})\right)\right]$$

- A naive ERM estimation minimize loss over realized data which is biased since distribution of observed data is not the population distribution p(x,y|s(a)) ≠ p(x,y)

$$f^{s(a)}_{\mathrm{ERM}} = \arg\min_{f^{s(a)}} \hat{\mathcal{L}}_{\mathrm{ERM}}\left(f^{s(a)}\right) = \arg\min_{f^{s(a)}} \frac{1}{n_{s(a)}} \sum_{i \in \mathcal{D}_{s(a)}} l(y_i^{obs}, f^{s(a)}(\mathbf{x}_i))$$

- IPS $e_{s(a)}$(x_i) *corrects for* the bias of observed data by reweighting loss

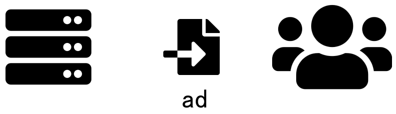
$$\hat{\mathcal{L}}_{\mathrm{IPS}}\left(f^{s(a)}\right) = \frac{1}{n} \sum_{i \in \mathcal{D}_{s(a)}} \frac{1}{e_{s(a)}(\mathbf{x}_i)} l(y_i^{obs}, f^{s(a)}(\mathbf{x}_i))$$

Implementation

- A unique DSP/DMP for O2O marketing provided by CyberAgent

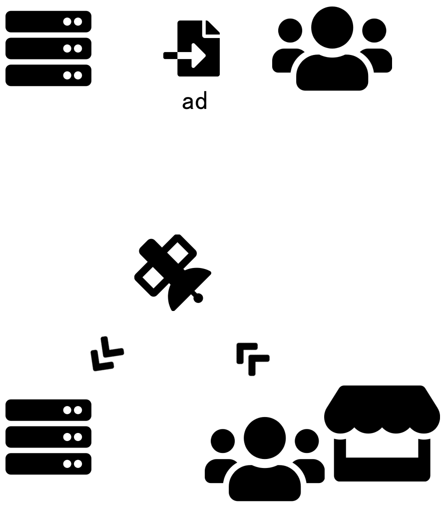
Implementation

- A unique DSP/DMP for O2O marketing provided by CyberAgent
  - Target potential customers with behavior and deliver ads



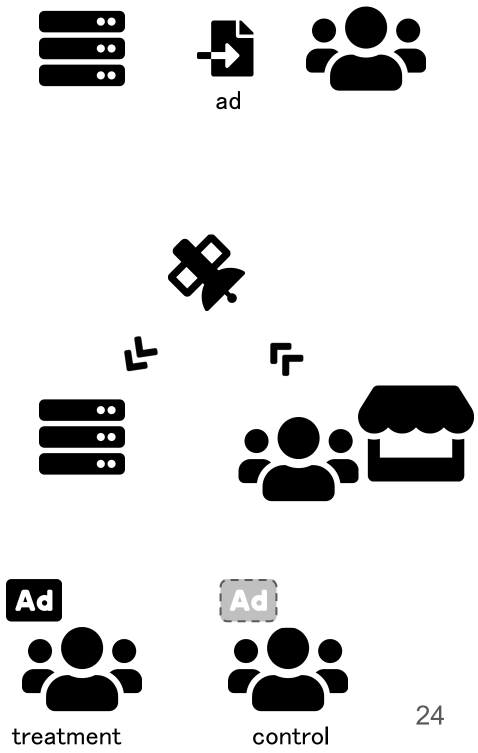
Implementation

- A unique DSP/DMP for O2O marketing provided by CyberAgent
  - Target potential customers with behavior and deliver ads
  - Tracking offline store visits using location data (e.g. GPS)



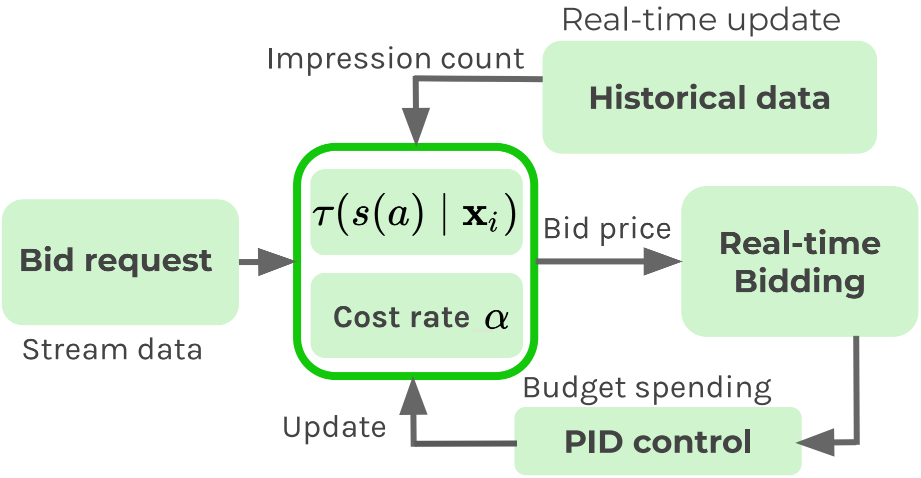
Implementation

- A unique DSP/DMP for O2O marketing provided by CyberAgent
  - Target potential customers with behavior and deliver ads
  - Tracking offline store visits using location data (e.g. GPS)
  - Measuring causal effect by A/B testing



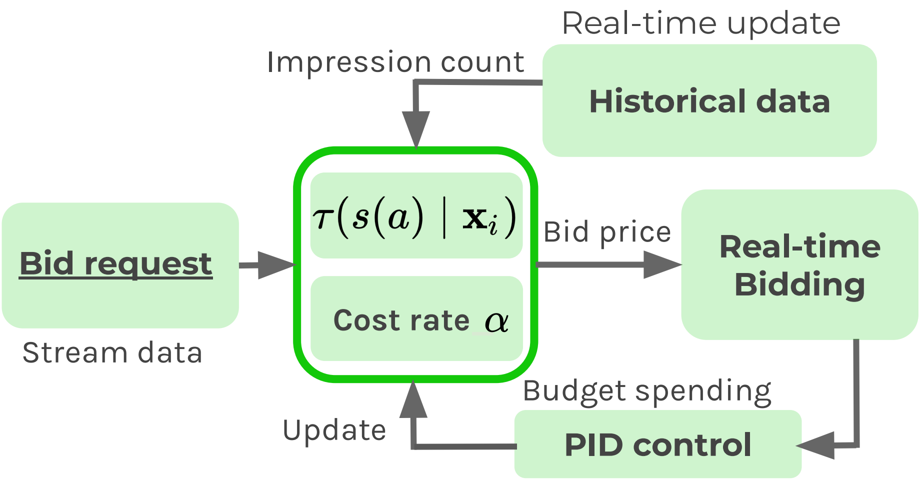
Unbiased Lift-based bidding system

- Bid request with device id and feature x
- Check past impression record
- Set pre-calculated τ
- Calculate bid price = v $\cdot\ \tau \cdot \alpha$
- Updated α to pace budget



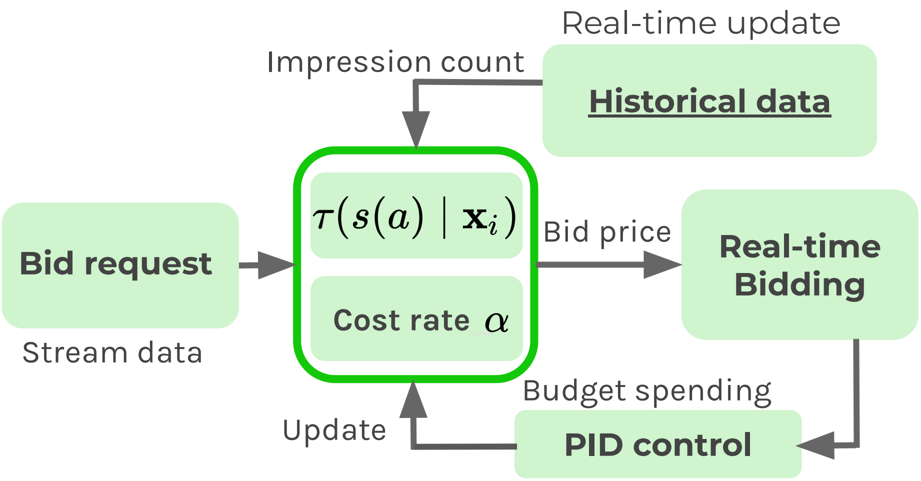
Unbiased Lift-based bidding system

- Bid request with device id and feature x
- Check past impression record
- Set pre-calculated τ
- Calculate bid price = v$\cdot \tau \cdot \alpha$
- Updated α to pace budget



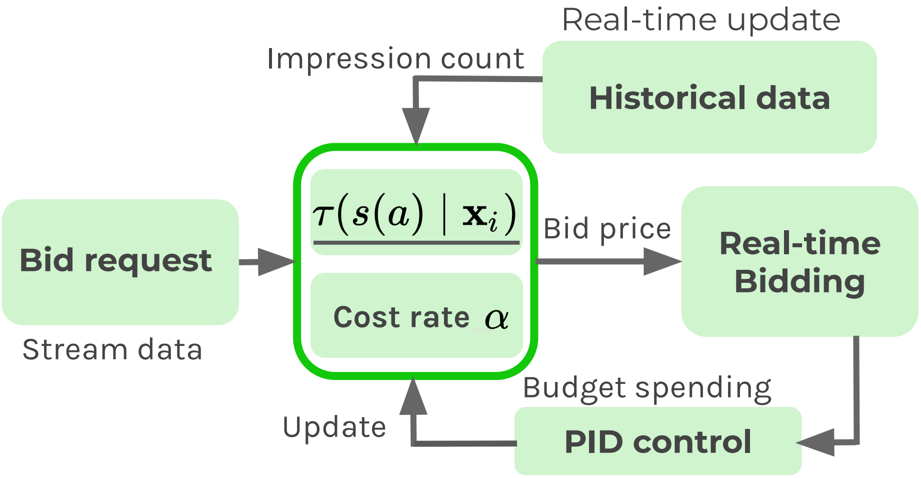
Unbiased Lift-based bidding system

- Bid request with device id and feature x
- Check past impression record
- Set pre-calculated τ
- Calculate bid price = v$\cdot\tau\cdot\alpha$
- Updated α to pace budget



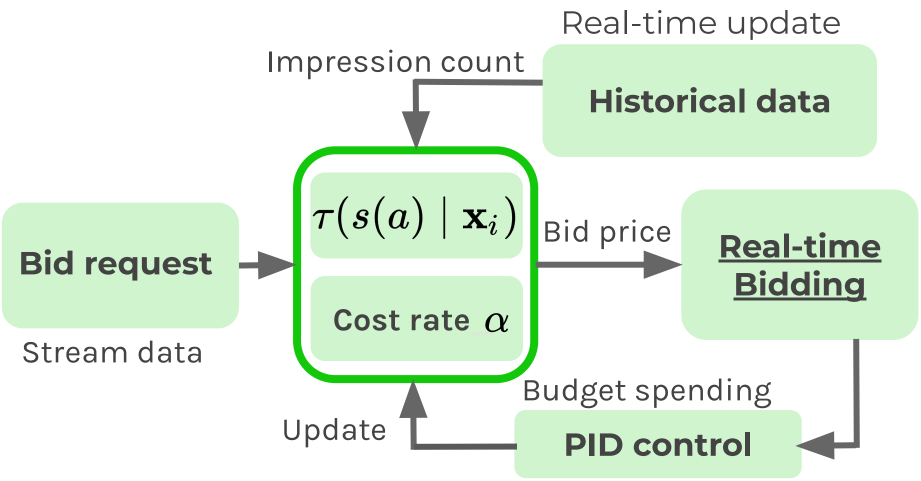
Unbiased Lift-based bidding system

- Bid request with device id and feature x
- Check past impression record
- Set pre-calculated τ
- Calculate bid price = v $\cdot\ \tau \cdot \alpha$
- Updated α to pace budget



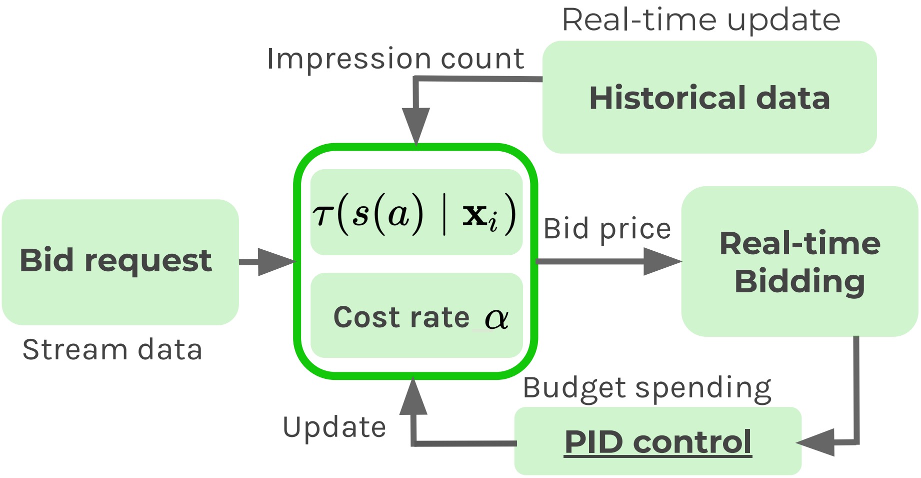
Unbiased Lift-based bidding system

- Bid request with device id and feature x
- Check past impression record
- Set pre-calculated τ
- Calculate bid price = $v \cdot \tau \cdot \alpha$
- Updated α to pace budget



Unbiased Lift-based bidding system

- Bid request with device id and feature x
- Check past impression record
- Set pre-calculated τ
- Calculate bid price = v $\cdot$ τ $\cdot$ α
- <u>Updated α to pace budget</u>



Model Training

- Data extraction
  - Ad exposure
  - Feature
  - response

Model Training

- Data extraction
    - Ad exposure
    - Feature
    - response
- User−level data
    - Use all the data
    - Consistent with debias

User-level learning

user	impression	conversion
A	3	2
B	1	0

Model Training

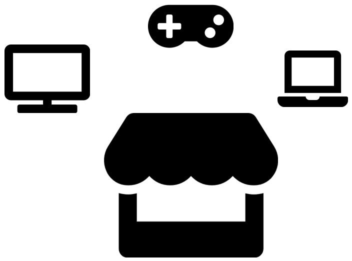
- Data extraction
  - Ad exposure
  - Feature
  - response
- User–level data
  - Use all the data
  - Consistent with debias
- Models
  - First stage: IPS estimator $e(x)$
  - Second stage: outcome predictors fs
  - Both models use XGBoost

User-level learning

user	impression	conversion
A	3	2
B	1	0

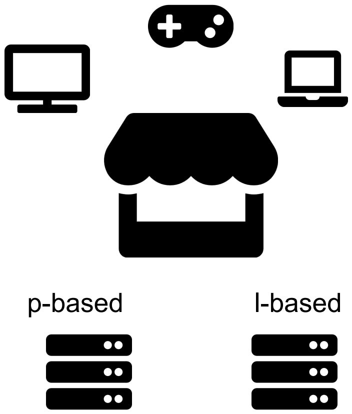
Online Experiment

- Evaluate models in real advertising campaign
  - An electronics retailer (like Best Buy) in Japan promotes app downloads and visits to real stores with display ads



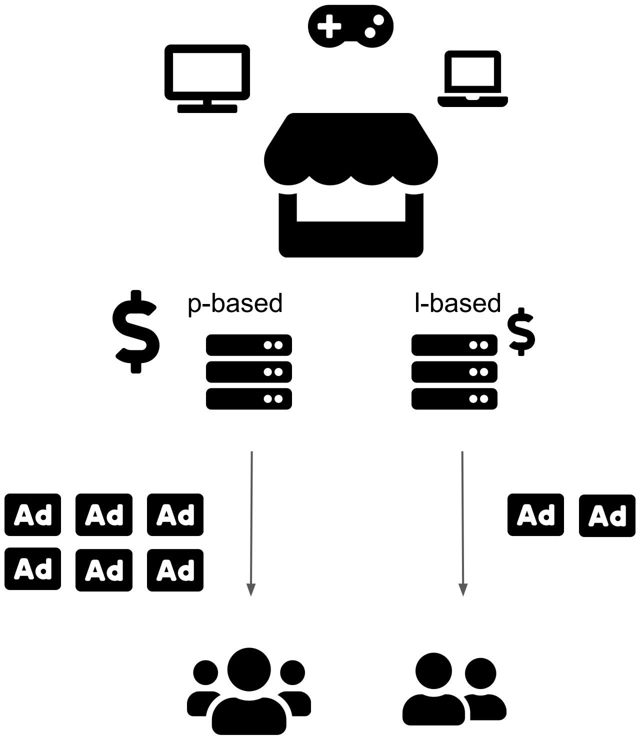
Online Experiment

- Evaluate models in real advertising campaign
  - An electronics retailer (like Best Buy) in Japan promotes app downloads and visits to real stores with display ads
- Lift-based bidder vs. Existing bidder
  - Existing bidder maximizes performance



Online Experiment

- Evaluate models in real advertising campaign
  - An electronics retailer (like Best Buy) in Japan promotes app downloads and visits to real stores with display ads
- Lift-based bidder vs. Existing bidder
  - Existing bidder maximizes performance
- Split users and budget into two groups
  - assign 88% to existing algorithm for business reason



Results: ad performance

- Impressions and reach rate (# users with impression/#users) increased
- Clicks decreased
- No significant difference for conversion rate (CVR)

	Lift-based	Performance-based	Diff
impression-per-user	1.28	1.0	0.28***
clicks-per-user	0.54	1.0	-0.46**
reach rate	1.71	1.0	0.71***

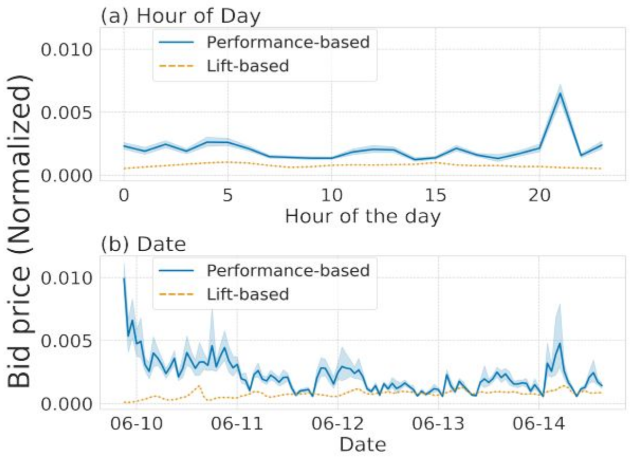
Results: Business KPIs

- Besides the same CVR, cost has been remarkably decreased
- Lift-based bidding archived the same results with less than 30% of the money that existing bidder spent!

	Lift-based	Performance-based	business impact
Cost-per-impression	0.27	1.0	73%saved!
Cost-per-reach	0.24	1.0	76% saved!
Cost-per-visit (CPA)	0.3	1.0	70% saved!

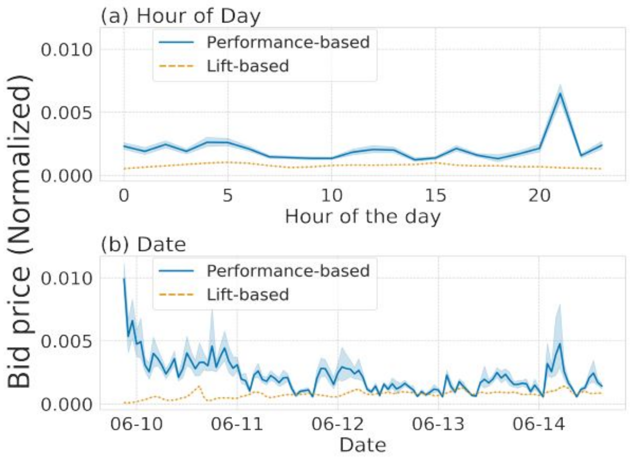
Discussion

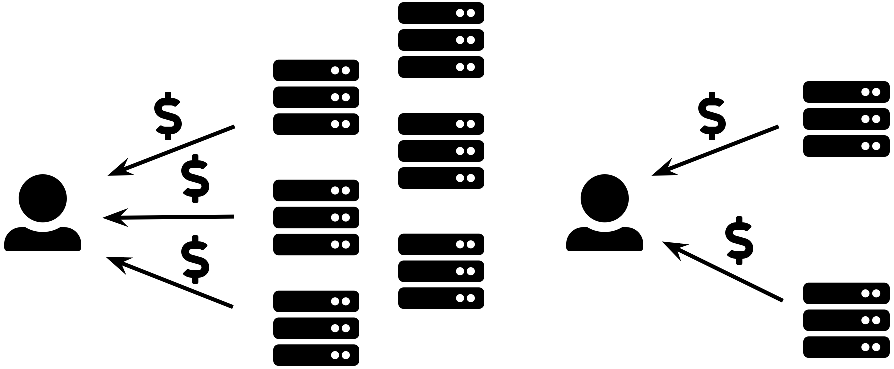
- Lift-based algorithm
  - bid much lower bid price.
  - very stable



Discussion

- Lift-based algorithm
  - bid much lower bid price.
  - very stable
- Possible reasons
  - Other DSPs also follow performance-based algorithm. very competitive auctions
  - In contrast, lift-based bidder participates less competitive auctions. Win easily





Conclusion and Future Study

- Propose unbiased lift-based bidding system
- Debias lift-effect prediction with IPS
- Deployed in real environment
- Online experiment
- Future work includes
  - More real-world experiments
  - Post-hoc analysis

Thank you!