Functional Bid Landscape Forecasting for Display Advertising

Yuchen Wang[1] Kan Ren[1] Weinan Zhang[1]
Jun Wang[2] Yong Yu[1]

[1]Apex Data and Knowledge Management Lab
Shanghai Jiao Tong University

[2]University College London

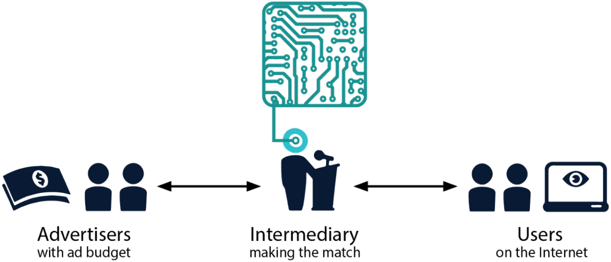
ECML-PKDD 2016

Outline

Outline

1. Background
   - Real-time Bidding
   - Bid Landscape Forecasting
2. Challenges
3. Related Work
4. Functional Bid Landscape Forecasting
   - Tree-based Mapping
   - Node Splitting
   - Survival Modeling
5. Experiments
6. Conclusion

Online Advertising

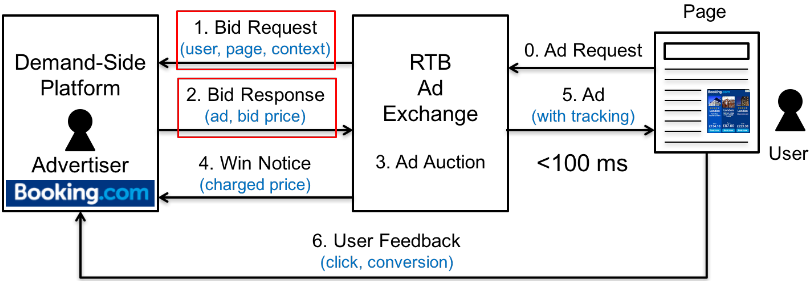


Goal of Computer

Address the **right user** with the **right message** in the **right context** and at the **right prices**.

Real-time Bidding (RTB)

in Display Ads Scenario



Outline

Terminologies

Market Price

The second highest bid price proposed by all the advertisers in the auction.

Terminologies

Market Price

The second highest bid price proposed by all the advertisers in the auction.

Bid Landscape Forecasting

To *forecast* the market price distribution (p.d.f.) of the specific auction.

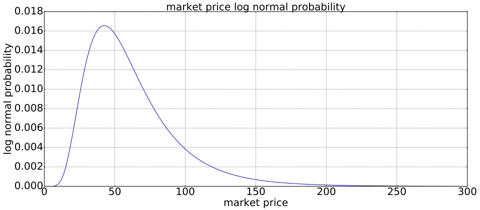
Bid Landscape Forecasting

Example

auction feature: weekday=*Friday*, city=*New York*, hour=*20*, ...

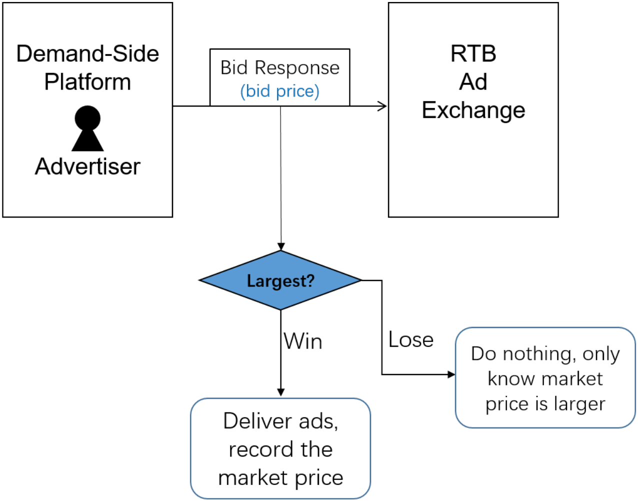
Goal

To forecast the market price distribution of the specific auction (impression level).



Modeling Right Censored Data

Losing and Winning

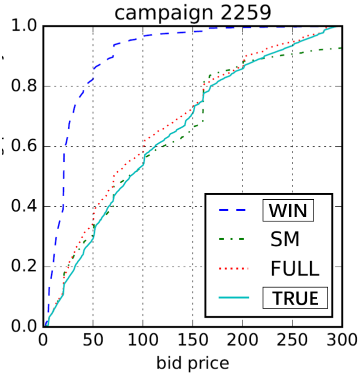


Modeling Right Censored Data

Right Censored

Right Censorship

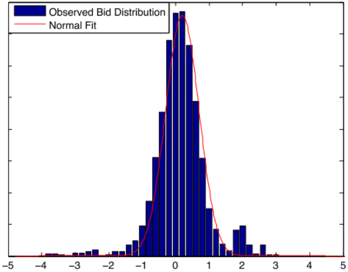
As in 2^{nd} price auction, if you *lose*, you only know that the *market price* is higher than your bidding price, which result in right censorship.



Heuristic Form

Log-normal Form

$$p_z(z) = \frac{1}{z\sigma\sqrt{2\pi}} e^{\frac{-(\ln z - \mu)^2}{2\sigma^2}}, z > 0 \, .$$



- Y. Cui et al. Bid landscape forecasting in online ad exchange marketplace. KDD 2011

Forecasting
Regression Model

v_i as the predicted winning price,

$$v_i \approx \beta^T x_i + \epsilon_i \ ,$$

$$\text{minimize} \sum_{i \in W} -\log(\phi(\frac{w_i - \beta^T x_i}{\sigma})) \ .$$

- W. Wu et al. Predicting Winning Price in Real Time Bidding with Censored Data. KDD 2015

Censorship Handling

Mixture Model

$$\begin{aligned} v_i &= [P(v_i < b_i)\beta_{lm} + (1 - P(v_i < b_i))\beta_{clm}]^T x_i \\ &= \beta_{mix}^T x_i \ . \end{aligned}$$

- W. Wu et al. Predicting Winning Price in Real Time Bidding with Censored Data. KDD 2015

Outline

1. Background
   - Real-time Bidding
   - Bid Landscape Forecasting
2. Challenges
3. Related Work
4. Functional Bid Landscape Forecasting
   - Tree-based Mapping
   - Node Splitting
   - Survival Modeling
5. Experiments
6. Conclusion

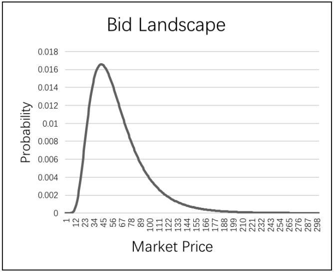
Tree-based Mapping

Goal

Given the auction feature $\boldsymbol{x}$, forecast the market price distribution $p_{\boldsymbol{x}}(z)$.

- Date: 20160320
- Hour: 14
- Weekday: 7
- IP: 119.163.222.*
- Region: England
- City: London
- Country: UK
- Ad Exchange: Google
- Domain: yahoo.co.uk
- URL: http://www.yahoo.co.uk/abc/xyz.html
- OS: Windows
- Browser: Chrome
- Ad size: 300*250
- Ad ID: a1890
- User tags: Sports, Electronics

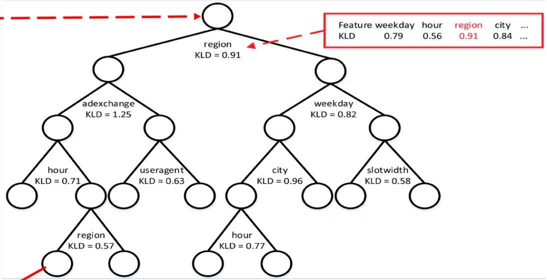
Mapping

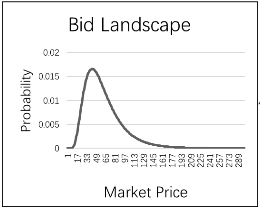


Tree-based Mapping

Methodology

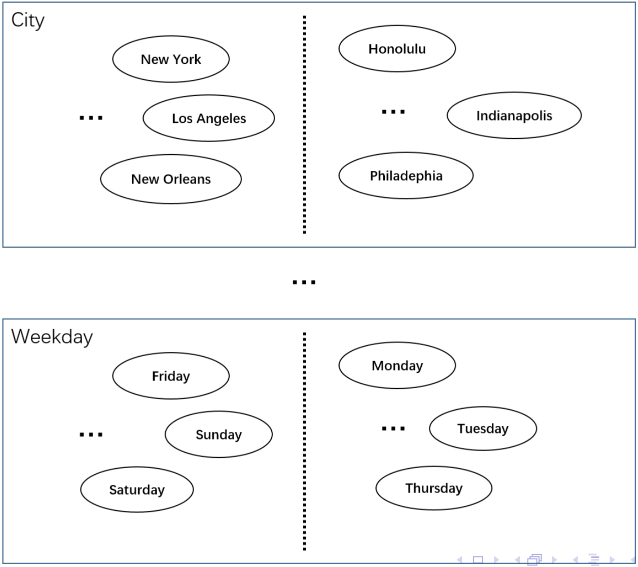
- Date: 20160320
- Hour: 14
- Weekday: 7
- IP: 119.163.222.*
- Region: England
- City: London
- Country: UK
- Ad Exchange: Google
- Domain: yahoo.co.uk
- URL: http://www.yahoo.co.uk/abc/xyz.html
- OS: Windows
- Browser: Chrome
- Ad size: 300*250
- Ad ID: a1890
- User tags: Sports, Electronics





Outline

Node Splitting



Node Splitting

KLD and Clustering

Kullback-Leibler Divergence (KLD)

A measure of the difference between two probability distributions P and Q.

Node Splitting

KLD and Clustering

Kullback-Leibler Divergence (KLD)

A measure of the difference between two probability distributions P and Q.

Node Splitting (one step)

Divide all the category (including in this node) values into two sets, maximizing KLD between the resulted two sets.

Node Splitting

KLD and Clustering

Kullback-Leibler Divergence (KLD)

A measure of the difference between two probability distributions P and Q.

Node Splitting (one step)

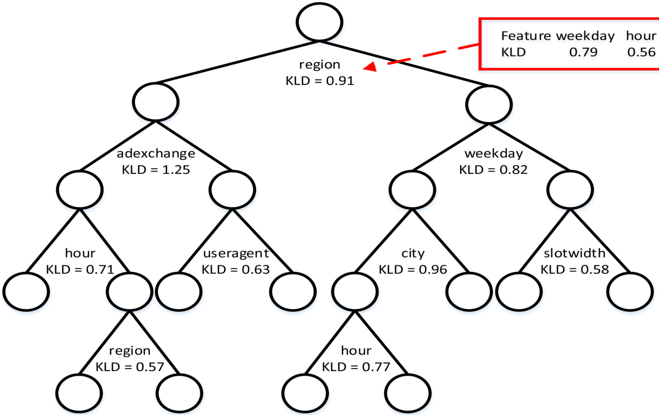
Divide all the category (including in this node) values into two sets, maximizing KLD between the resulted two sets.

Algorithm

Using K-Means Clustering according to KLD values.

Node Splitting

KLD and Clustering



Feature	weekday	hour	region	city	...
KLD	0.79	0.56	0.91	0.84	...

Outline

Handling Censorship

Survival Model

- For winning auctions: We have the true market price value.
- For lost auctions: We only know our proposed bid price and know that the true market price is higher than that.

Intuition

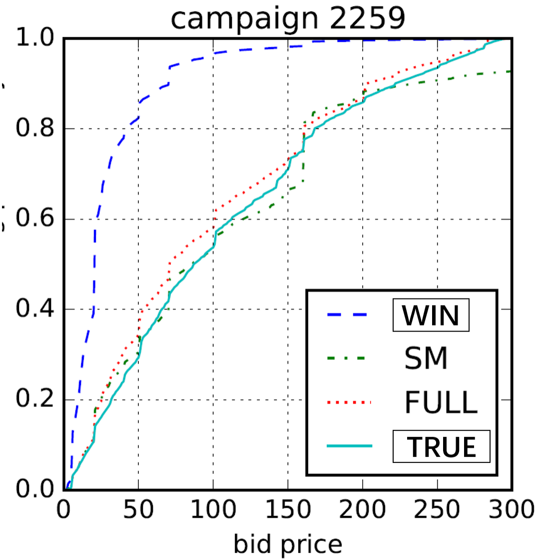
Most related works focus only on the winning auctions without considering the lost auction, which contains the information to infer the true distribution.

$$(b_i, w_i, m_i)_{i=1,2,\cdots,M} \quad \longrightarrow \quad (b_j, d_j, n_j)_{j=1,2,\cdots,N}$$

$b_j < b_{j+1}$, d_j is number of winning auctions by $b_j - 1$, n_j is number of lost auctions by $b_j - 1$. So

$$w(b_x) = 1 - \prod_{b_j < b_x} \frac{n_j - d_j}{n_j}, \quad p(z) = w(z+1) - w(z). \tag{1}$$

Survival Model



Exp. Setup

- Dataset
  - iPinYou: real world RTB data including 64.7M auction samples.
- Exp. Flow: split the bidding records into two sets
  - Winning set W: remain the same as the original samples.
  - Lost set L: hide the true market price
  - Create a simulating environment for the compared models to forecast the bid landscape of each bid request.

Example

Original samples:

bid_price=90, market_price=86, feature ...

bid_price=101, market_price=112, feature ...

Exp. Setup

- Dataset
  - iPinYou: real world RTB data including 64.7M auction samples.
- Exp. Flow: split the bidding records into two sets
  - Winning set W: remain the same as the original samples.
  - Lost set L: hide the true market price
  - Create a simulating environment for the compared models to forecast the bid landscape of each bid request.

Example

Original samples:
bid_price=90, market_price=86, feature ...
bid_price=101, market_price=112, feature ...

Example

Winning sample: *bid_price=90, market_price=86, feature ...*
Lost sample: *bid_price=101, market_price=NULL, feature ...*

Compared Models

- NM: Normal Model, only make statistics on the winning auctions.
- SM: Survival Model, additionally utilize the lost auctions.
- MM: Mixture Model, implemented as KDD 2015 paper.
    - W. Wu et al. Predicting Winning Price in Real Time Bidding with Censored Data. KDD 2015
- NTM: Normal Tree Model, forecast the bid landscape using only tree model.
- STM: Survival Tree Model, the proposed tree Model with survival analysis.

Evaluation Measures

Objective

Measure the error between the forecasted market price distribution and the true one.

- Average Negtive Log Probability (ANLP)

$$P_{\text{nl}} = \sum_{i=1}^{k} \sum_{j=1}^{z_{\max}} (-\log P_{ij}) N_{ij}, \tag{2}$$

$$N = \sum_{i=1}^{k} \sum_{j=1}^{z_{\max}} N_{ij}, \quad \bar{P}_{\text{nl}} = P_{\text{nl}} / N, \tag{3}$$

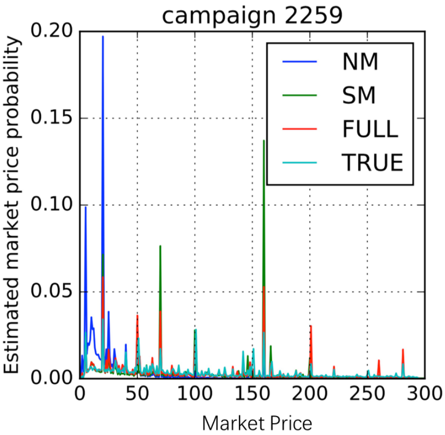
- KL-Divergence (KLD) between forecasted distribution and the true one.

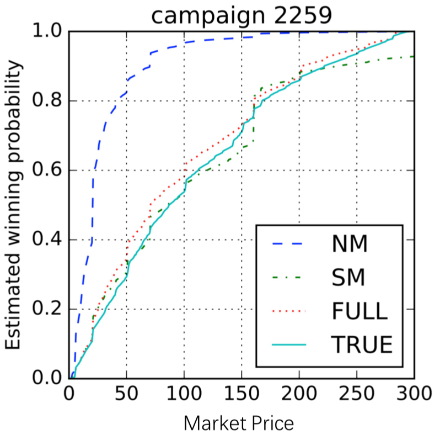
Results

Table: Performance illustration. Average negative probability of five compared settings. ANLP: the smaller, the better. KLD: the smaller, the better.

	ANLP					KLD				
Campaign	MM	NM	SM	NTM	STM	MM	NM	SM	NTM	STM
1458	5.7887	5.3662	4.7885	4.7160	4.3308	0.7323	0.7463	0.2367	0.6591	0.2095
2259	7.3285	6.7686	5.8204	5.4943	5.4021	0.8264	0.9633	0.3709	0.8757	0.1668
2261	7.0205	5.5310	5.1053	4.4444	4.3137	1.0181	0.4029	0.2943	0.3165	0.1222
2821	7.2628	6.5508	5.6710	5.4196	5.3721	0.7816	0.9671	0.3562	0.6170	0.2880
2997	6.7024	5.3642	5.1411	5.1626	5.0944	0.7450	0.4526	0.1399	0.3312	0.1214
3358	7.1779	5.8345	5.2771	4.8377	4.6168	1.4968	0.8367	0.5148	0.8367	0.3900
3386	6.1418	5.2791	4.8721	4.6698	4.2577	0.8761	0.6811	0.3474	0.6064	0.2236
3427	6.1852	4.8838	4.6453	4.1047	4.0580	1.0564	0.3247	0.1478	0.3247	0.1478
3476	6.0220	5.2884	4.7535	4.3516	4.2951	0.9821	0.6134	0.2239	0.5650	0.2238
overall	6.5520	5.6635	5.0997	4.7792	**4.6065**	0.9239	0.6898	0.2927	0.5834	**0.2160**

Forecasted Bid Langscape Comparison





Conclusion

- Model
  - Function mapping from bid request features to bid landscape.
  - Clustering-based node splitting with KL-Divergence objective.
  - Survival analysis to handle censorship in learning problems.
- Significant improvement of forecasting performance over baselines and state-of-the-art models in various metrics.
- Future work
  - Embed bid landscape forecasting into utility (click-through rate, conversion rate) estimation model.
    - K. Ren et al. User Response Learning for Directly Optimizing Campaign Performance in Display Advertising. CIKM 2016