



Click Modeling in Search Advertising: Challenges & Solutions

Jianchang (JC) Mao

Advertising Sciences

YAHOO! LABS

<http://labs.yahoo.com>



Disclaimers

- This talk presents the opinions of the author. It does not necessarily reflect the views of Yahoo! Inc or any other entity
- Algorithms, techniques, features, etc mentioned here might or might not be in use by Yahoo! or any other company



Sponsored Search Ads

YAHOO!

Web Images Video Local Shopping News More

Whistler ski resort

Search

Options

QuickApps

Search Pad

SearchScan - On

3,490,000 results for Whistler ski resort

Also try: [whistler ski resort map](#), [more...](#)

Whistler Ski Resort

5th Night Free **Ski Package Deal!** Whistler Blackcomb® Official Site.
www.WhistlerBlackcomb.com

Whistler Ski Resort

Great **Ski Packages** at Four Seasons **Resort Whistler**. Book Online Now!
FourSeasons.com/Whistler

Whistler Hotels

Get Our **Best Price Guarantee** on Any Hotel Reservation at Expedia.com.
www.expedia.com

Whistler Deals

Get Great Deals at Top BC **Resorts**. Get Your free 2011 BC **SKI** e-Guide.
HelloBC.com/Whistler

More Sponsors: [Whistler ski resort](#), [whistler](#)

Whistler Blackcomb - Official Ski Resort Website - Whistler ...

whistlerblackcomb.com - Official Site for **Whistler Blackcomb Ski Resort**. Plan and book **Whistler** vacations including lift tickets, **ski** rentals, snowboard rentals, and ...
www.whistlerblackcomb.com - [Cached](#)

[Web Cams](#)

[Trail Map](#)

[Weather Forecast](#)

[What's Open](#)

[Snow Report](#)

[Winter Lift Tickets](#)

[Live Temperatures](#)

[Photos](#)

[more results from whistlerblackcomb.com »](#)

Whistler Accommodations and Central Reservations Whistler.com ...

Tourism **Whistler** Accommodation, **Ski** Vacation Packages and complete **Whistler Blackcomb Resort** information. **Whistler.com** is your in **resort** connection to **Whistler BC**
www.whistler.com - [Cached](#)

North

ads

East

Sponsored Results

Sponsored Results

60 Hotels in Whistler

Book your hotel in **Whistler** online. No reservation costs. Great rates!
Booking.com/Whistler-Hotels-Canada

whistler ski resort

Whistler Resorts with Skiing. Up to 70% Off. Check Rates Now.
Whistler-Resorts.Tripzen.com

OwnerDirect Chalet Rental

Vacation Rental Homes & Condos. Rent Directly from Owner.
Ownerdirect.com

Marriott Whistler Hotels

Book Marriott's Best Rate Guarantee. Earn Rewards at Official Site.
Marriott.com/Canada

Whistler Vacation Chalets

Luxury Vacation Rental Near Slopes. Amazing View, Fireplace, Spa!
whistler.luxuryretreats.com

Whistler Lodging



Search Advertising - Problem

- General problem statement
 - Given a query by a user, select an optimal placement of **eligible** ads to maximize a total **utility** function that captures the expected revenue, user experience, and advertiser ROI
- Example utility: Expected Revenue
eCPM = Probability(good click | user, query, ad) * Bid
- Fundamental Problems
 1. Estimate relevance of an ad to the user query
 2. Estimate probability of clicks



Challenges in Click Modeling

- Biases
 - Position
 - Externality
 - Ad category
 - Selection
 - Sparsity
 - Missing data
 - Dynamic and seasonal effects
 - Noise, spam, etc.
-
- Modeling positional externalities*
 - Temporal Click Model
 - Dealing with sparsity*
 - Ad hierarchy
 - Query-segment
 - Dealing with missing data*
 - Mixture model

* Some of the recent work in Advertising Sciences, Y! Labs



Modeling User Click Behavior

- **Models**

- Cascade Model
- Dependent Click Model
- User Browsing Models
- Click Chain Model
- Session Utility Model
- Bayesian Browsing Model
- Dynamic Bayesian Browsing Model
- **Temporal Click Model***

- **Hypotheses**

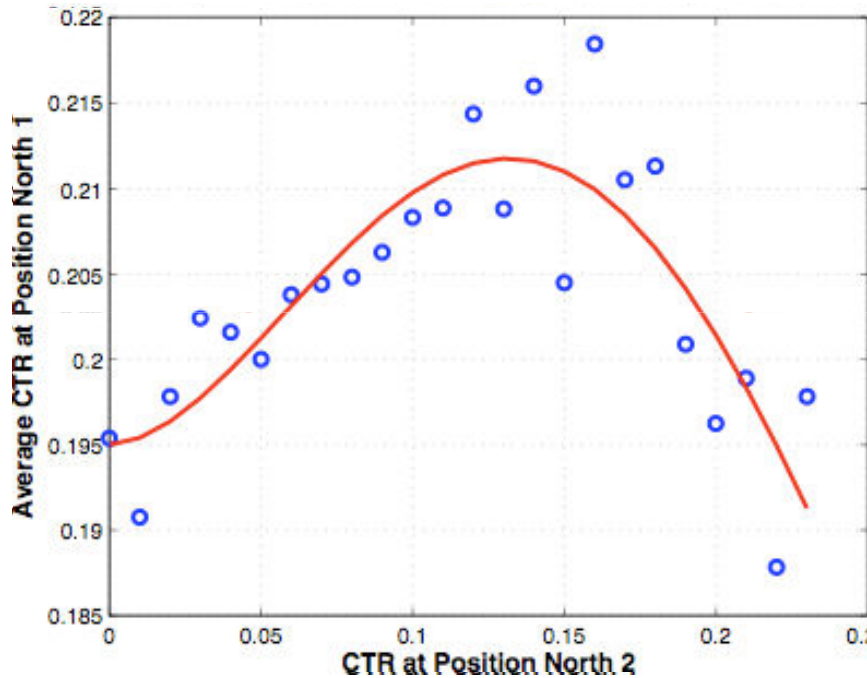
- Examination hypothesis
- Cascade hypothesis
- Rationality hypothesis
- **Positional rationality hypothesis***

* Wanhong Xu, Eren Manavoglu and Erick Cantu-Paz. "Temporal Click Model for Sponsored Search", SIGIR 2010.

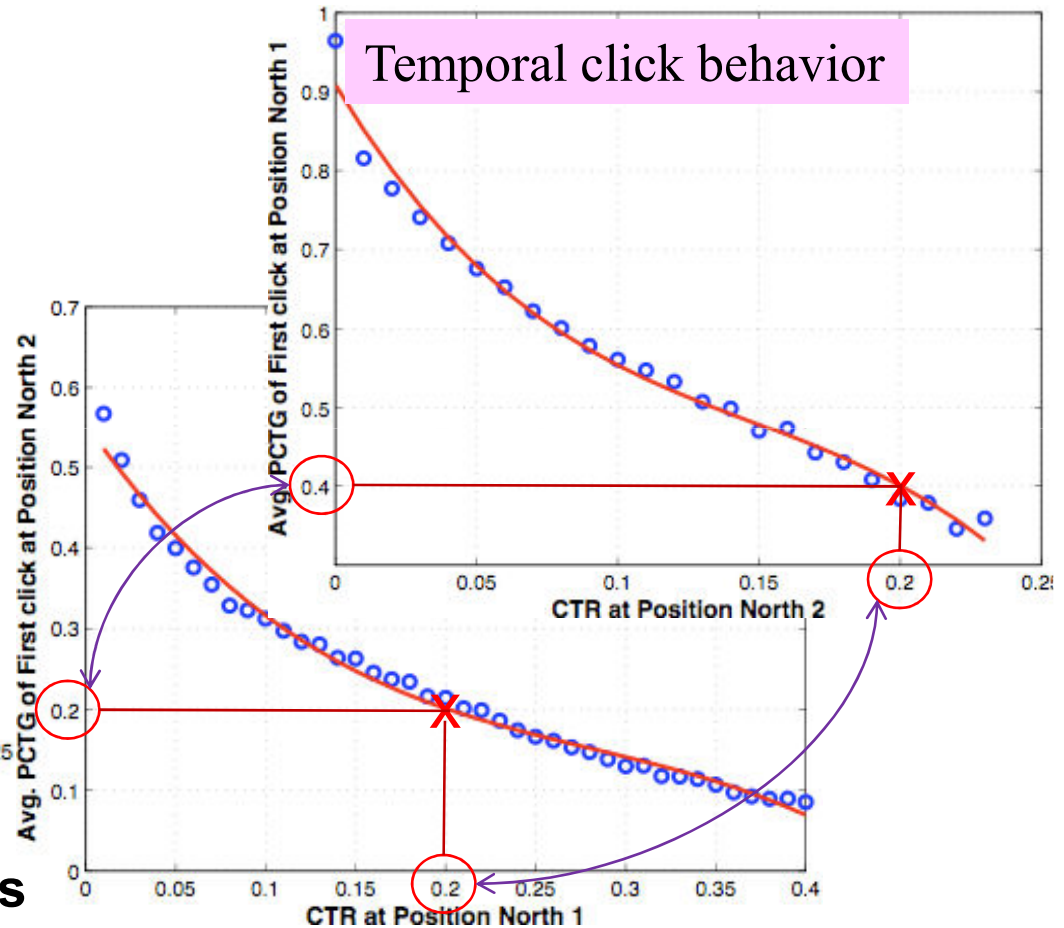


Positional Rationality Hypothesis (based on randomized ranking study)

Existence of ad externalities



Temporal click behavior

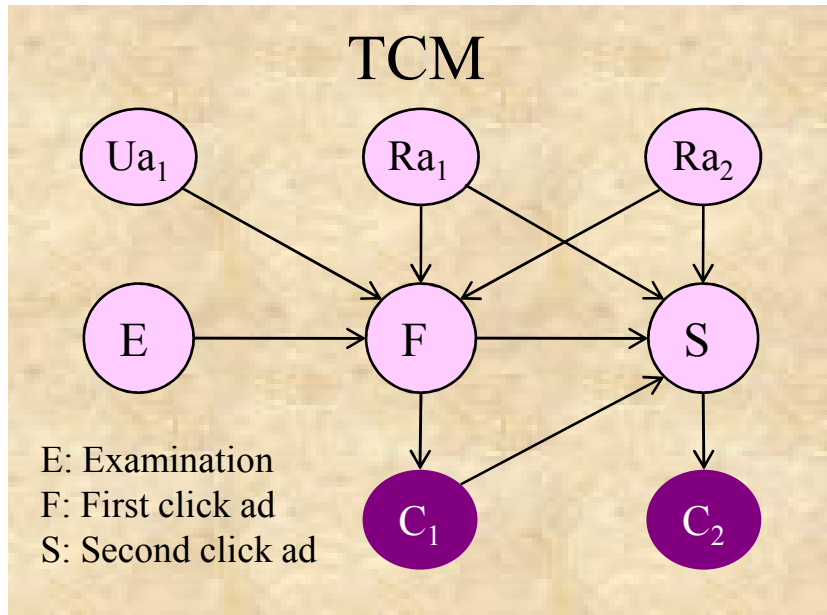
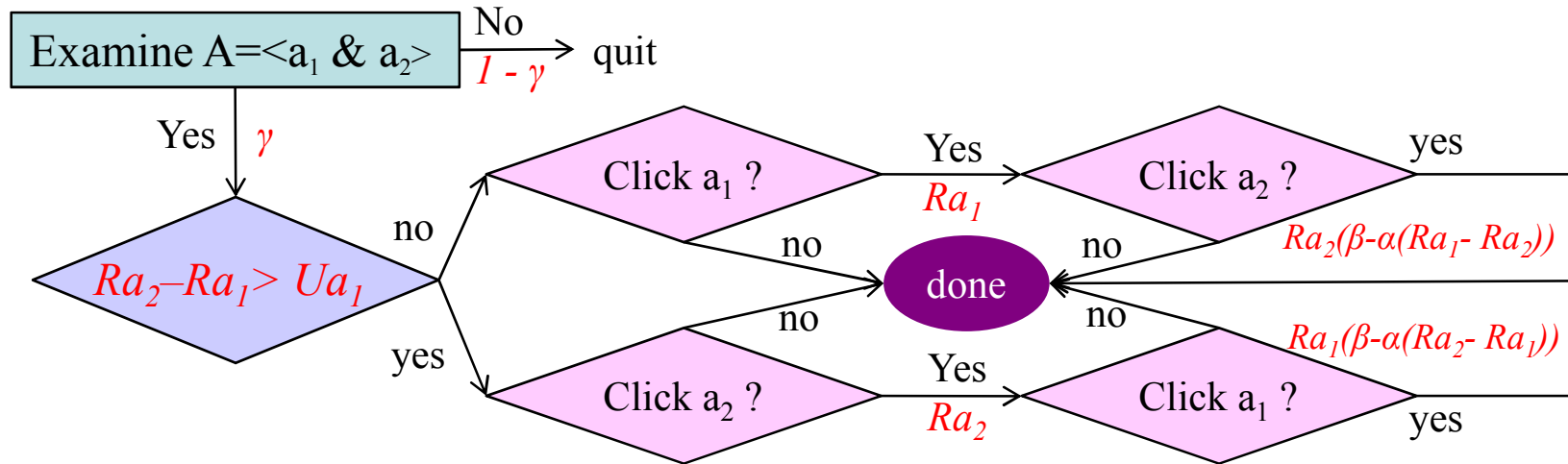


Positional Rationality Hypothesis

- Users examine both ads together to assess their qualities (Ra_1 & Ra_2)
- If the ad at position 2 is better than the ad at position 1 (by a positional preference threshold Ua_1), users would click the ad at position 2 first



Temporal Click Model



$$\begin{aligned}
 P(E = 1|A) &= \gamma \\
 P(F = a_2|E = 1, A) &= \mathbf{1}[Ra_2 - Ra_1 > Ua_1] \\
 P(F = a_1|E = 1, A) &= \mathbf{1}[Ra_2 - Ra_1 \leq Ua_1], \\
 P(c_1 = a_i|F = a_i) &= Ra_i \\
 P(c_1 = \square|F = a_i) &= 1 - Ra_i, i \in \{1, 2\}. \\
 P(S = \square|c_1 = \square) &= 1 \\
 P(S = a_{3-i}|c_1 = a_i, A) &= \beta - \alpha Ra_i - Ra_{3-i}, i \in \{1, 2\}. \\
 P(c_2 = a_j|S = a_j) &= Ra_j \\
 P(c_2 = \square|S = a_j) &= 1 - Ra_j, j \in \{1, 2\} \\
 P(F = \square|E = 0, A) &= 1 \\
 P(c_1 = \square|F = \square) &= 1 \\
 P(c_2 = \square|S = \square) &= 1
 \end{aligned}$$

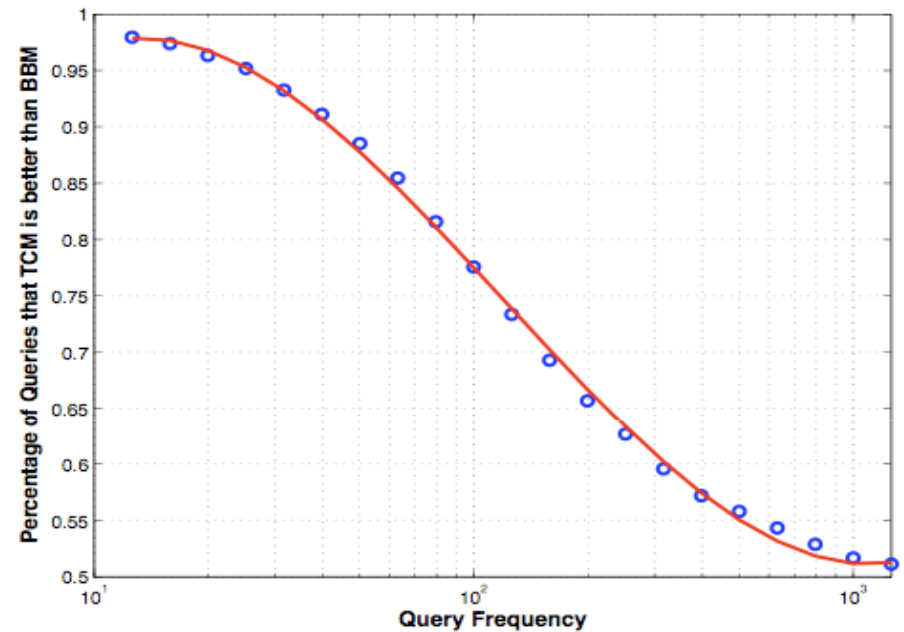
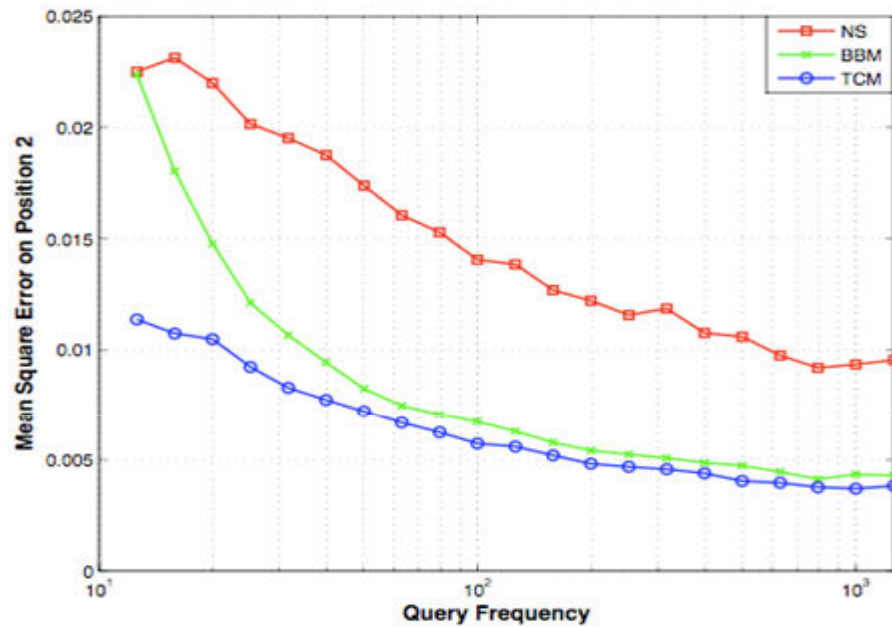


TCM Algorithms

- Posterior Distribution of user-perceived relevance (Ra_i) and ad position advantage (Ua_i) can be obtained in a closed form
- Parameters (α, β, γ) are estimated using maximum likelihood principle
 - γ in a closed form
 - α, β by a barrier method combined with Newton's method
- Click Through Rate (CTR) prediction in a closed form
 - CTR of a click sequence
 - CTR of an ad
- Map-Reduce implementation to process a large volume of search log



Experimental Results



Conclusions

- Temporal click sequences do provide implicit quality feedback
- TCM incorporates positional bias, externality, and user perceived relevance into a combined model
- TCM can be used to improve the CTR estimates. It outperforms BBM
- Positional Rationality Hypothesis may be too strong for higher frequency queries



Challenges in Click Modeling

- Biases
 - Position
 - Externality
 - Ad category
 - Selection
 - Sparsity
 - Missing data
 - Dynamic and seasonal effects
-
- Modeling positional externalities
 - Temporal Click Model
- Dealing with sparsity
 - Ad hierarchy
 - Query-segment
-
- Dealing with missing data
 - Mixture model



Predictive Click Modeling

- Model probability of click per position

$P(\text{click}|q, a, u, \text{pos})$: q =query, a =ad, u =user, pos =position

- Maximum Entropy model trained using regularized maximum likelihood estimation
- Generate a position normalized click score

$$P(\text{click}|q, a, u) = \sum_{\text{pos}} P(\text{click}|q, a, u, \text{pos}) * P(\text{pos} | q, a, u)$$

*Dustin Hillard, Eren Manavoglu, Hema Raghavan, Erick Cantu-Paz, Chris Leggetter, Rukmini Iyer. "The Sum of Its Parts: Reducing Sparsity in Click Estimation with Query Segments". To appear in Journal of Information Retrieval, Special Topic Issue on Web Mining for Search



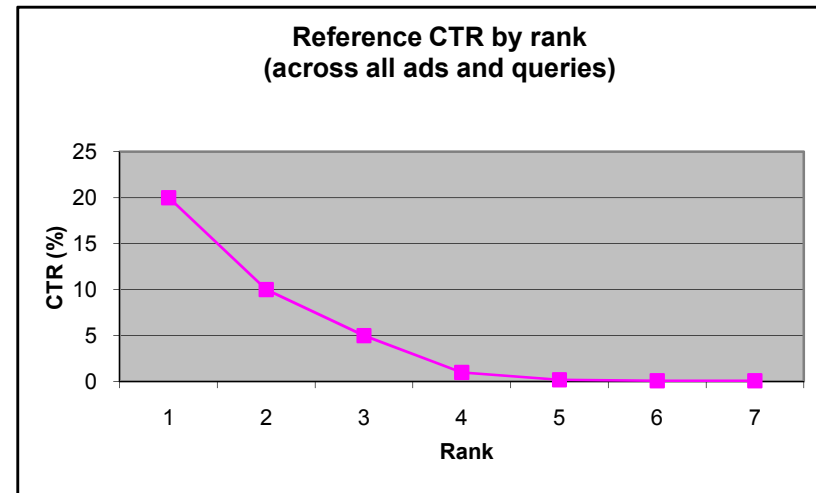
Click Prediction - Features

- Query features
- Query-Ad text matching features
- Time features
- Presentation features
- Click feedback features

- Expected Clicks:
$$EC = \sum_{r=1}^n refCTR_r * imps_r$$

- Clicks over EC: $COEC = clicks / EC$

- Pair-wise conjunctions
- Continuous features are quantized
 - K-means clustering per feature
 - Special bin for the missing value
 - Allows non-linearity



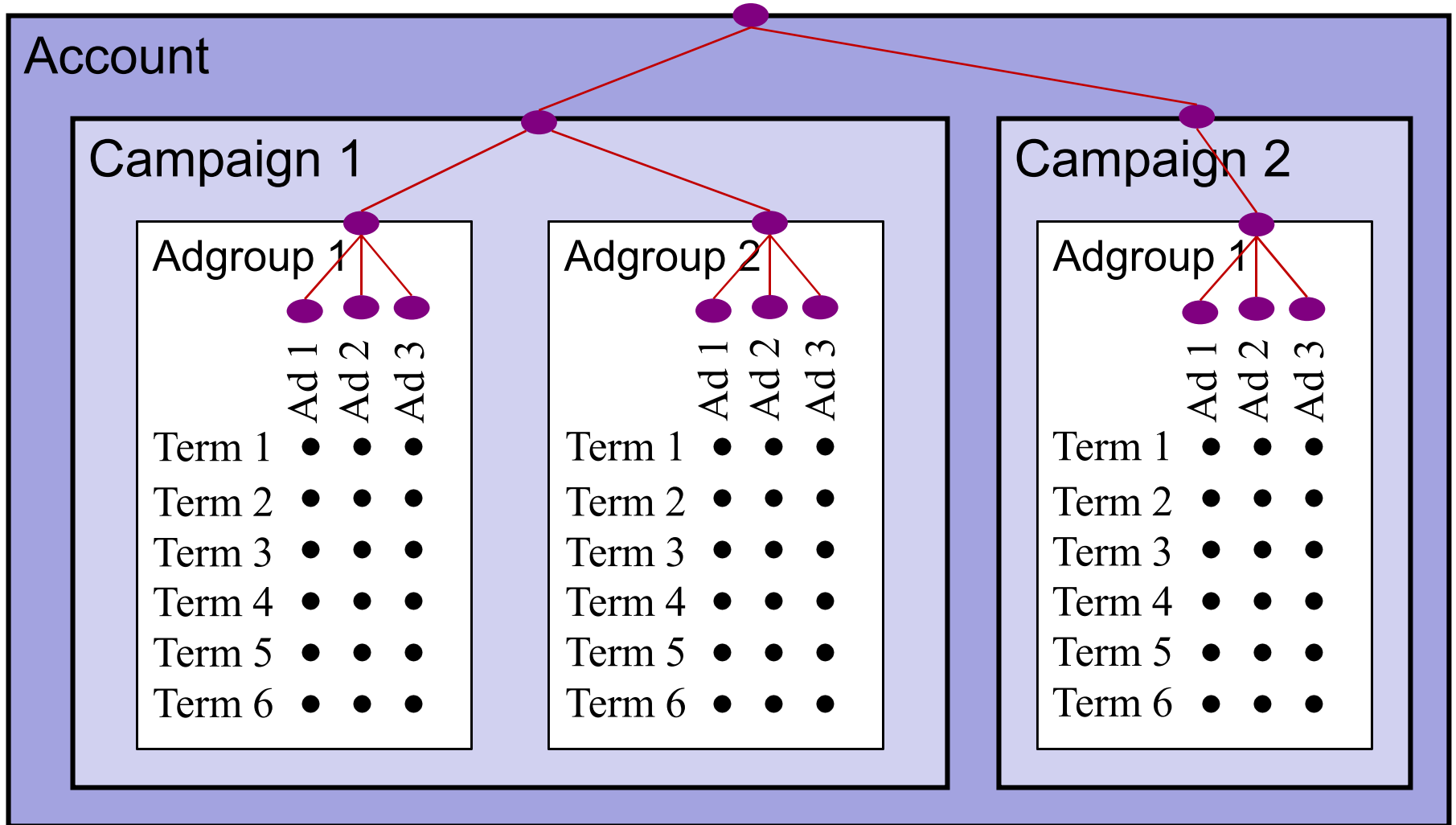


Deal with Sparsity - Leverage Data Hierarchy

- Past performance of ads for specific queries is a good predictor of future performance.
- But using only query-ad level history features is problematic
 - Data is sparse, tail is long
 - There are many new queries, and many new ads
- We can aggregate click history at coarser granularity in data hierarchies

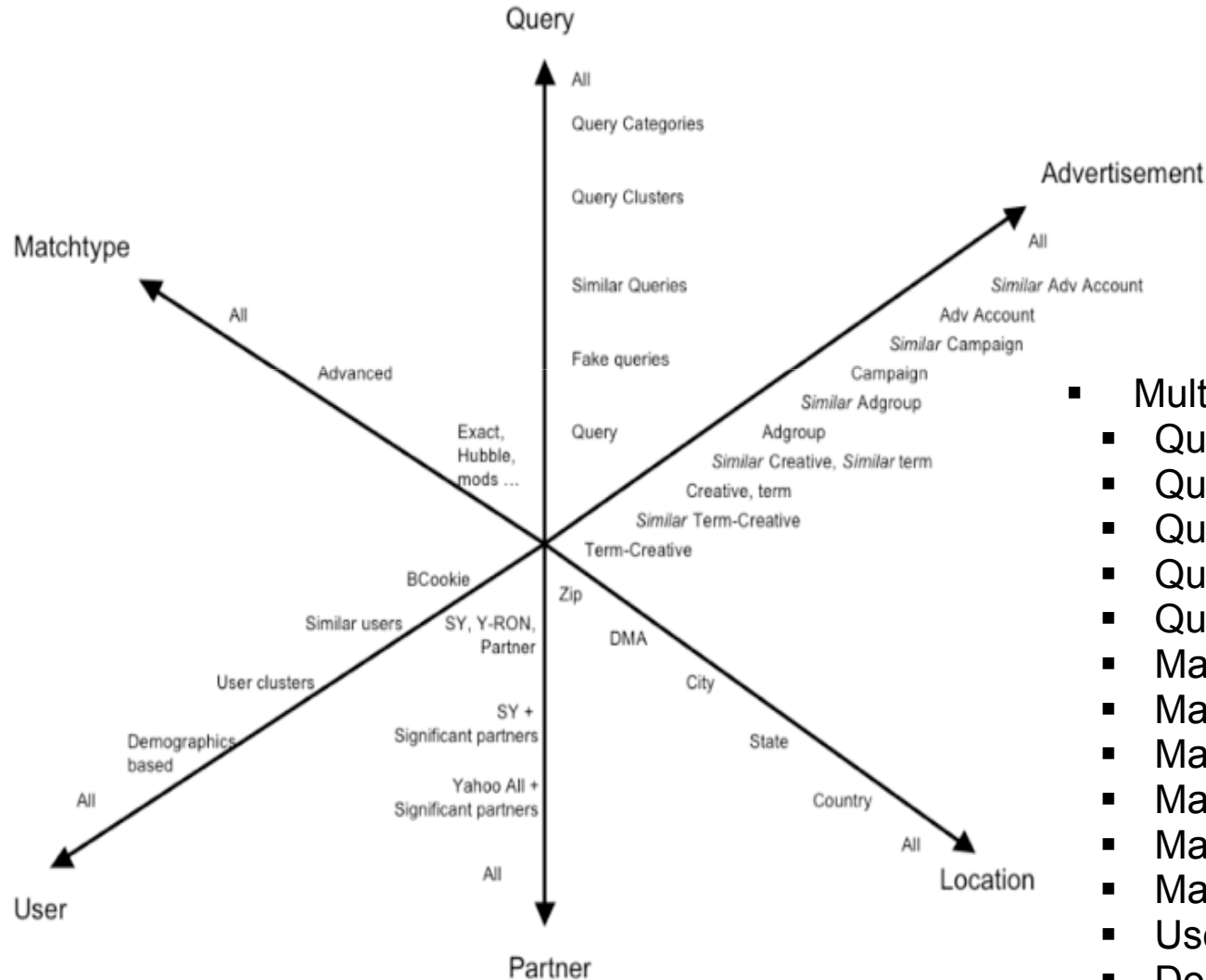


Natural Ad Data Hierarchy





Hierarchies Exist in other Dimensions



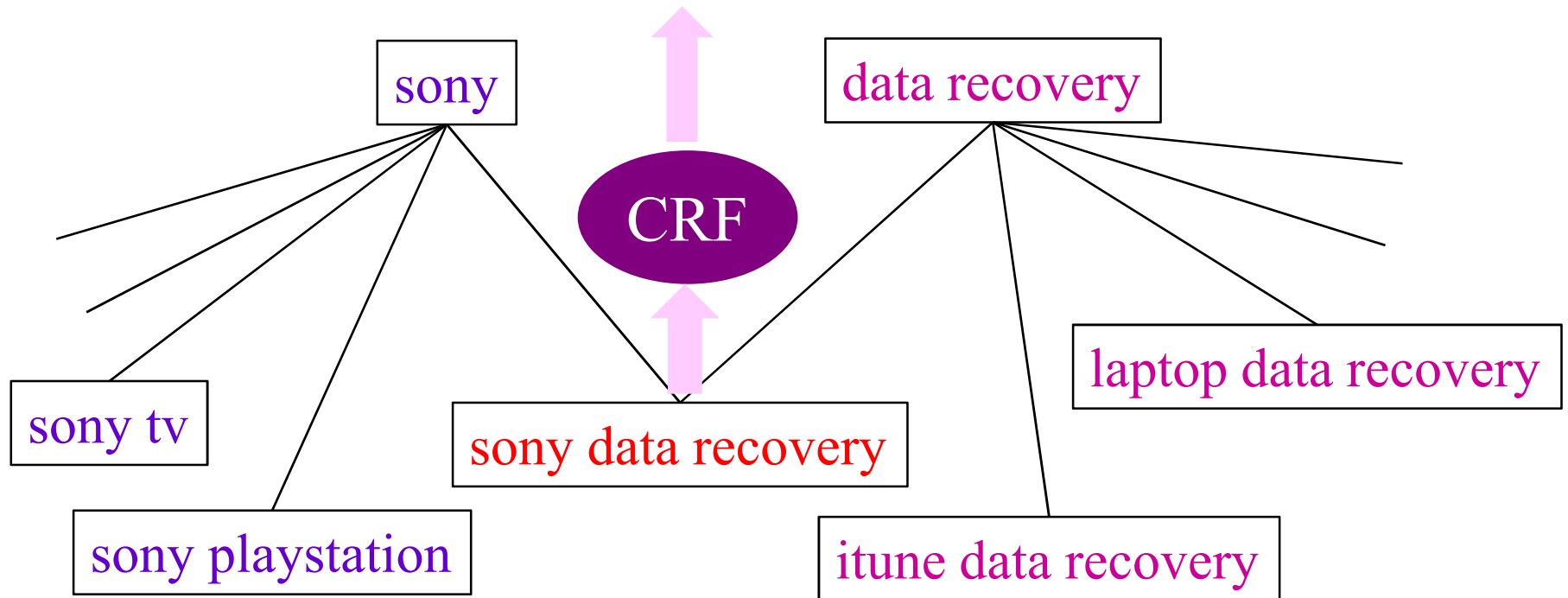
- Multiple levels
 - Query X Term X Creative
 - Query X Ad Group
 - Query X User Location
 - Query X Domain
 - Query X Bidded Term
 - Matchtype X Term X Creative
 - Matchtype X Term
 - Matchtype X Creative
 - Matchtype X Ad Group
 - Matchtype X Campaign
 - Matchtype X Account
 - User Location X Domain
 - Domain



Creating Query Segment Hierarchy

Segments	Probability
sony data recovery	0.742
sony data recovery	0.208
sony data recovery	0.038

- CRF Segmentation
 - Trained on human labeled query phrases
 - Better than HMM and other methods in P/R & Phrase-WER





Segment-level Aggregates

- New click feedback features
 - QuerySegment
 - QuerySegment X Ad Id
 - QuerySegment X Domain
- Query: super mario bros 3 game online

Click feedback

for segments

for segments X domain

conf	Query Segment	Clicks	EC	Query Segment	Domain	Clicks	EC
0.31	online	200k	160k	online	gamenet	none	none
0.24	super mario bros 3	30	38	super mario bros 3	gamenet	4.4	5
0.20	game	171k	142k	game	gamenet	520	370
0.11	super mario bros 3 game	10	7	super mario bros 3 game	gamenet	none	none
0.04	game online	821	711	game online	gamenet	2	2.5



Click Model w/ Segment-level Aggregates

- Approach:
 - Provide query segment click history as features
 - Combine query segments, weighted by confidence
 - For missing ad, back-off to query segment level
- Test on two weeks of search traffic
 - Neutral impact on traffic with sufficient history
 - Significant gains across multiple low history slices

Traffic Slice	Slice Coverage	Precision/Recall AUC
No North Ads	29%	+15%
South Ads	25%	+12%
No Account History	0.3%	+9%
No Query History	25%	+9%
No Query/Domain History	49%	+6%
All Advanced Ads	49%	+5%
Query \geq 4 words	20%	+3%



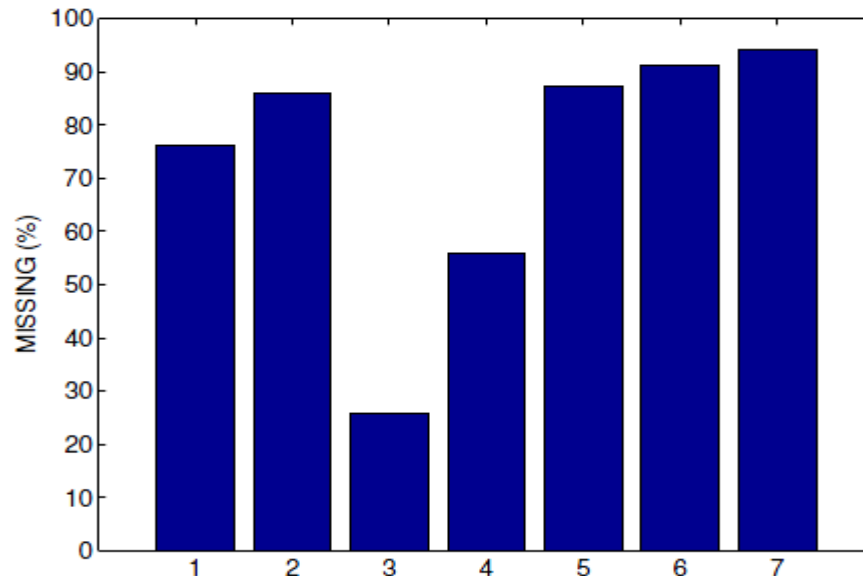
Challenges in Click Modeling

- Biases
 - Position
 - Externality
 - Ad category
 - Selection
 - Sparsity
 - Missing data
 - Dynamic and seasonal effects
-
- Modeling positional externalities
 - Temporal Click Model
 - Dealing with sparsity
 - Ad hierarchy
 - Query-segment
-
- Dealing with missing data
 - Mixture model



Missing Click Data

- Missing click feedback for new ads, tail queries, etc.
- One analysis
 - 1.6 Billion random samples in a 3-week period
 - Missing click data when $EC < 200$ (std dev of CTR estimate < 0.1)
 - Only about 5.7% of samples have all the 7 aggregated clicks available



*Ozgur Cetin, Kannan Achan, Erick Cantu-Paz, and Rukmini Iyer. "Missing Click History in Sponsored Search: A Generative Modeling Solution". AdKDD 2010.



Handling Missing Click Data

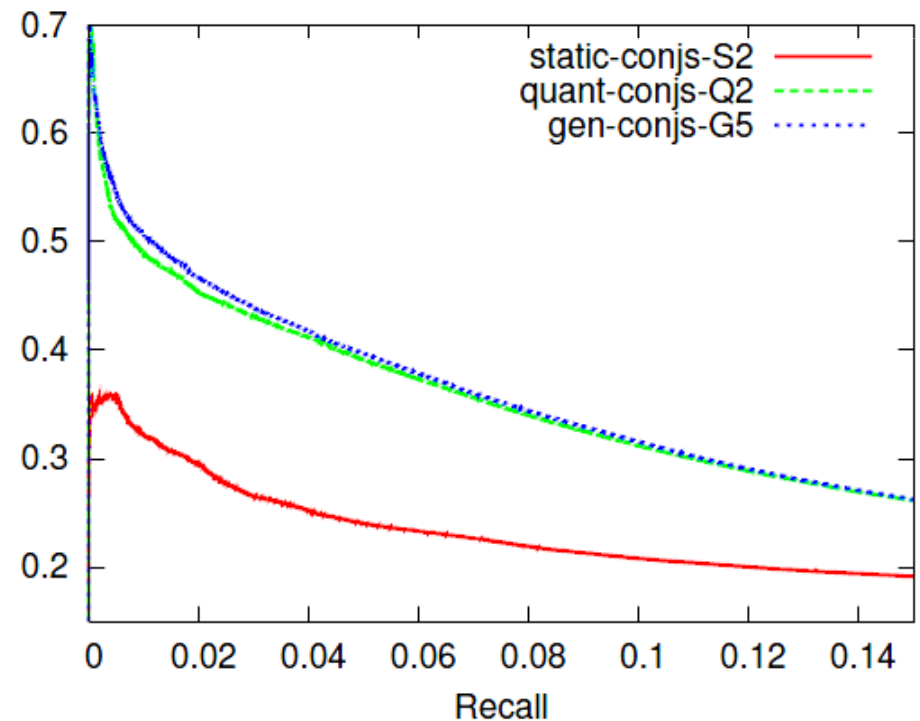
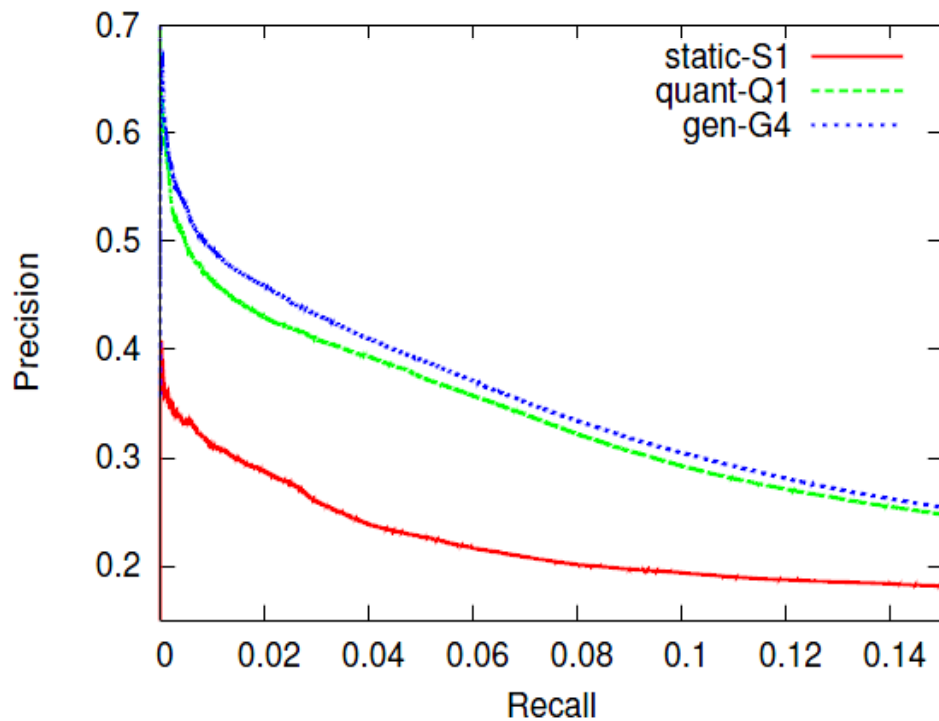
- Mixture of Gaussians model for click feedback features
- Parameter estimation
 - Use EM algorithm
- Probabilistic inference
 - Posterior distribution of the missing data conditioned on the observed ones
- Derive features from latent structure in the mixture model
 - Posterior probability vector
- Handle missing data in Maximum Entropy Model
 - Add a special bin for indication of a missing feature
 - Use imputed feature mean and variance when missing
 - Add posterior probability vector as additional features



Precision-Recall Curves

	Features	Params
S1	static	147
Q1	static + quant (click feats)	294
G4	static + click feats + imputed	338

	Features	Params
S2	static + conj's	3583
Q2	static + quant (click feats) + conj's	14944
G5	static + click feats+imputed+conj's	4896





Concluding Remarks

- Click prediction is a central problem in Search Advertising
- Click modeling is challenging because of various biases, sparsity, missing data, and the dynamic nature of clicks and marketplace
- Machine learning techniques can be employed to deal with some of those challenging problems
- Computational Advertising is a rich interdisciplinary field for applying machine learning techniques



Thank you!

Acknowledgement:

Eren Manavoglu

Erick Cantu-paz

Dustin Hillard

Chris Leggetter

Kannan Achan

Ashvin Kannan

Deepak Agarwal

Andrei Broder

Sponsored Search Science Team