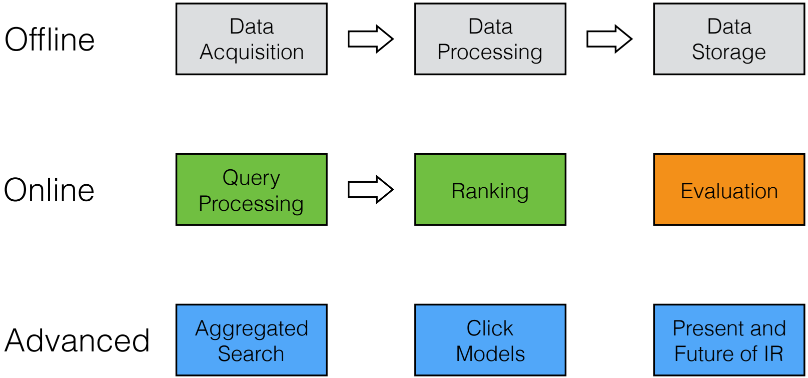
Information Retrieval

Click Models

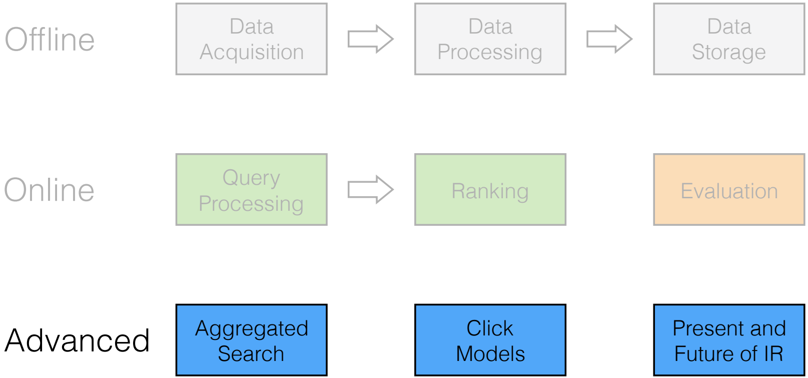
Ilya Markov
i.markov@uva.nl

University of Amsterdam

Course overview

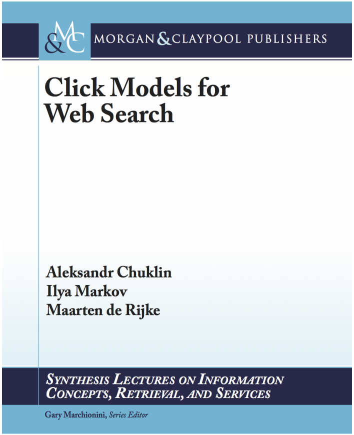


Advanced topics in IR



Outline

The book



http://clickmodels.weebly.com/the-book.html

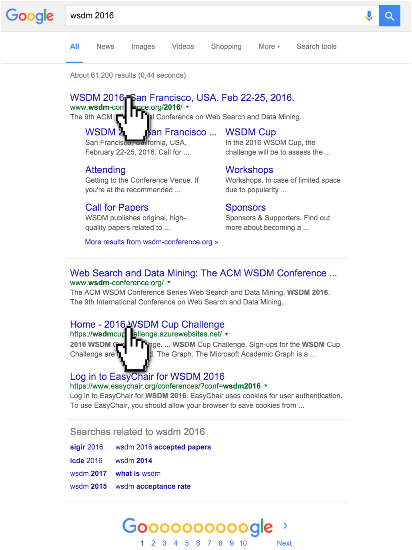
Tutorials

- SIGIR 2015, Santiago, Chile
- AINL-ISMW FRUCT 2015, St. Petersburg, Russia
- WSDM 2016, San Francisco, USA
- RuSSIR 2016, Saratov, Russia

`http://clickmodels.weebly.com/tutorials.html`

Outline

Why click models?



Why click models?

Scientific modelling is a scientific activity, the aim of which is to make a particular part or feature of the world easier to understand, define, quantify, visualize, or simulate by referencing it to existing and usually commonly accepted knowledge.

Wikipedia, Scientific modelling

Why click models?

Click models make user clicks in web search
easier to understand, define, quantify, visualize, or simulate
using (mostly) probabilistic graphical models.

Click log

0	0	Q	8	0	7	103	51	92	43	12	73	69	27	105
0	36	Q	174	0	1625	1627	1623	1626	1624	1622	1619	1621	1620	1618
0	50	Q	227	0	2094	2091	2087	2089	2093	2088	2090	2092	2095	2086
0	515	Q	174	0	1625	1627	1623	1626	1624	1622	1619	1621	1620	1618
0	524	Q	1974	0	17562	1627	1626	1623	2091	17559	17563	17558	17561	17560
0	527	C	17562											
0	528	C	1627											
0	529	C	1626											
1	0	Q	9	0	13	70	66	94	50	104	29	21	89	85
1	20	C	104											
1	123	C	21											
1	291	Q	1324	0	11807	11805	11812	11813	11804	11809	11806	11811	11808	11810
1	301	C	11813											
1	8605	C	11808											
1	8737	C	11810											
1	8884	C	11811											
2	0	Q	7	0	77	93	55	86	64	67	76	98	18	54
2	11	C	18											
2	1122	Q	4088	0	35554	35561	35562	35556	35557	35567	35550	35566	35568	35553
2	1127	C	35561											
2	1645	Q	5863	0	36505	36514	36508	36509	50480	36510	36507	50482	50483	50481
2	1646	C	36505											

Yandex Relevance Prediction Challenge

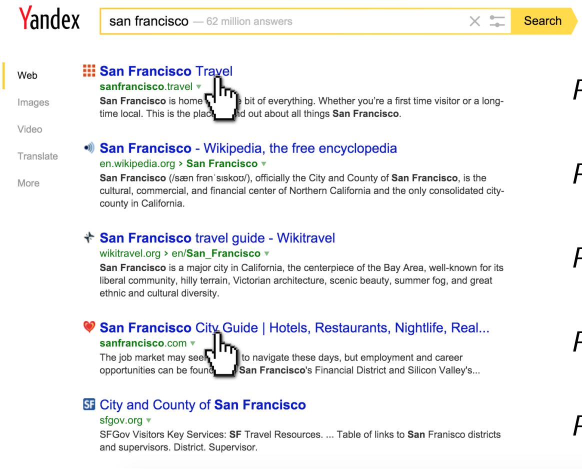
`http://imat-relpred.yandex.ru/en`

Outline

1. Introduction
2. Basic click models
   - Random click model
   - Position-based model
   - Cascade model
   - Click probabilities
   - Evaluation
   - Parameter estimation
3. Applications
4. Advanced models

Outline

Random click model



P_{click}

P_{click}

P_{click}

P_{click}

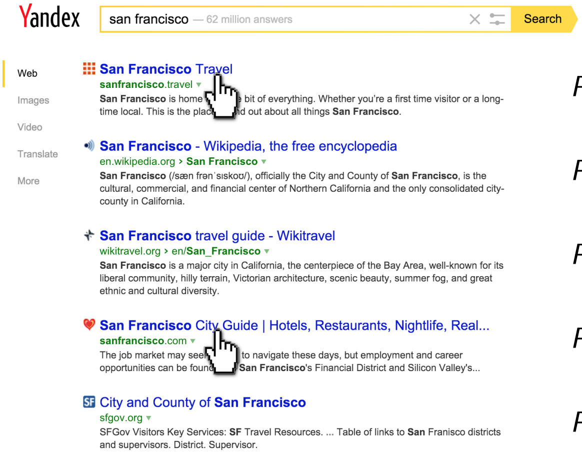
P_{click}

Random click model

- Terminology
  - C_u – binary random variable denoting a click on document u
- Random click model (RCM)
  - Any document can be clicked with the same (fixed) probability

$$P(C_u = 1) = const = \rho$$

Random click model



$$P(C_{u_1} = 1) = \rho$$

$$P(C_{u_2} = 1) = \rho$$

$$P(C_{u_3} = 1) = \rho$$

$$P(C_{u_4} = 1) = \rho$$

$$P(C_{u_5} = 1) = \rho$$

$$\rho = \frac{\#\ \text{clicks}}{\#\ \text{shown docs}}$$

CTR models

Random click model (global CTR):

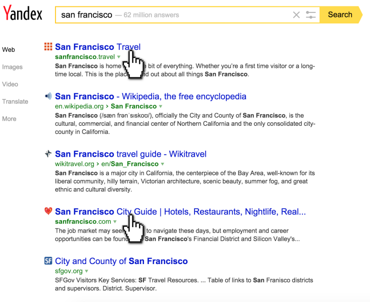
$$P(C_u = 1) = \rho$$

Rank-based CTR:

$$P(C_{u_r} = 1) = \rho_r$$

Query-document CTR:

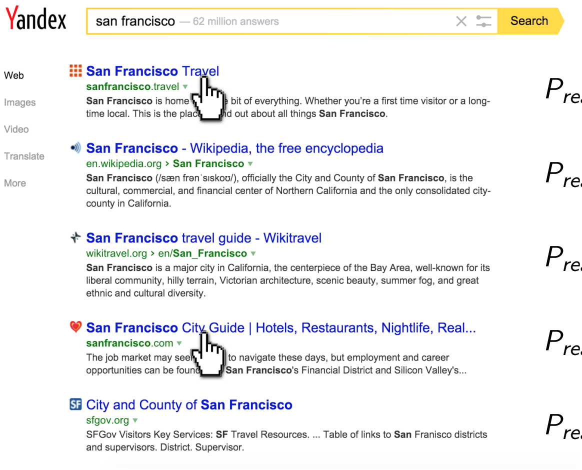
$$P(C_u = 1) = \rho_{uq}$$



Outline

2 Basic click models
- Random click model
- Position-based model
- Cascade model
- Click probabilities
- Evaluation
- Parameter estimation

Position-based model



$P_{read}(1)$, $P_{click}(u_1 q)$

$P_{read}(2)$, $P_{click}(u_2 q)$

$P_{read}(3)$, $P_{click}(u_3 q)$

$P_{read}(4)$, $P_{click}(u_4 q)$

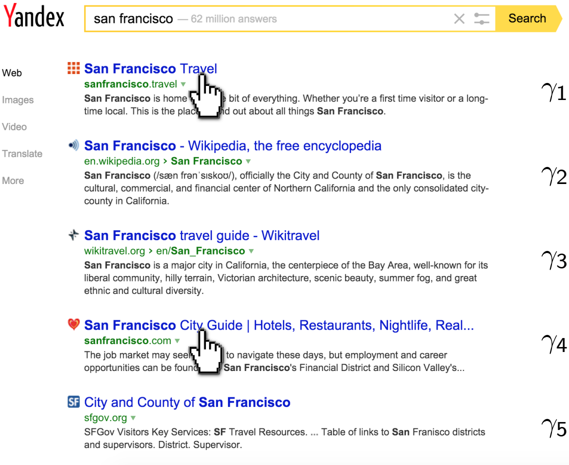
$P_{read}(5)$, $P_{click}(u_5 q)$

Position-based model: examination

- Terminology
  - Examination = reading a **snippet**
  - E_r – binary random variable denoting examination of a snippet at rank r
- Position-based model (PBM)
  - Examination depends on rank

$$P(E_r = 1) = \gamma_r$$

Position-based model



$\gamma_1 \; , \; P_{click}(u_1 q)$

$\gamma_2 \; , \; P_{click}(u_2 q)$

$\gamma_3 \; , \; P_{click}(u_3 q)$

$\gamma_4 \; , \; P_{click}(u_4 q)$

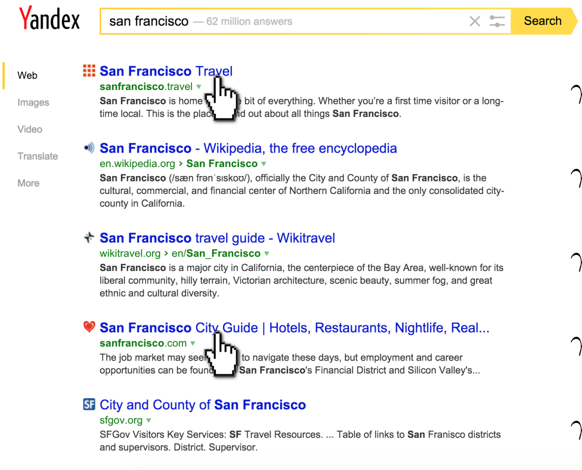
$\gamma_5 \; , \; P_{click}(u_5 q)$

Position-based model: attractiveness

- Terminology
  - Attractiveness = a user wants to click on a document after examining (reading) its snippet
  - A_u – binary random variable showing whether document u is attractive to a user, given query q
- Position-based model (PBM)
  - Attractiveness depends on a query-document pair

$$P(A_{uq} = 1) = \alpha_{uq}$$

Position-based model



$\gamma_1\ ,\ \alpha_{u_1 q}$

$\gamma_2\ ,\ \alpha_{u_2 q}$

$\gamma_3\ ,\ \alpha_{u_3 q}$

$\gamma_4\ ,\ \alpha_{u_4 q}$

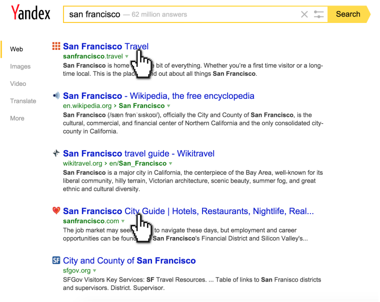
$\gamma_5\ ,\ \alpha_{u_5 q}$

Position-based model: summary

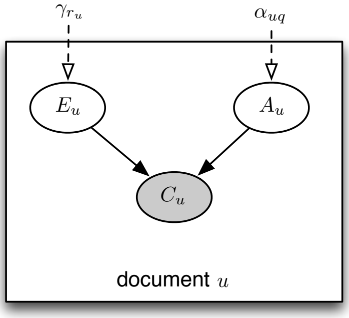
$$P(E_{r_u} = 1) = \gamma_{r_u}$$

$$P(A_u = 1) = \alpha_{uq}$$

$$P(C_u = 1) = P(E_{r_u} = 1) \cdot P(A_u = 1)$$



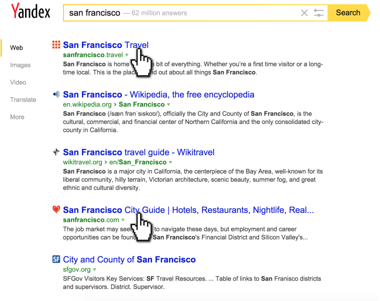
Position-based model: probabilistic graphical model



Position-based model: exercises

$$P(E_{r_u} = 1) = \gamma_{r_u}$$
$$P(A_u = 1) = \alpha_{uq}$$
$$P(C_u = 1) = P(E_{r_u} = 1) \cdot P(A_u = 1)$$

$$E_{r_u} = 0 \Rightarrow C_u = 0$$
$$A_u = 0 \Rightarrow C_u = 0$$
$$E_{r_u} = 1 \Rightarrow (C_u = 1 \iff A_u = 1)$$
$$A_u = 1 \Rightarrow (C_u = 1 \iff E_{r_u} = 1)$$



Outline

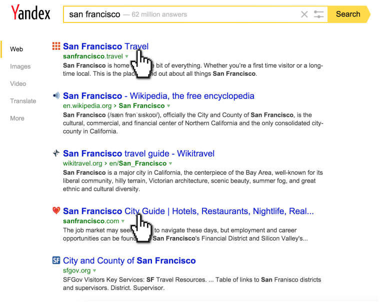
2. Basic click models
   - Random click model
   - Position-based model
   - Cascade model
   - Click probabilities
   - Evaluation
   - Parameter estimation

Position-based model

$$\color{red}{P(E_{r_u} = 1) = \gamma_{r_u}}$$

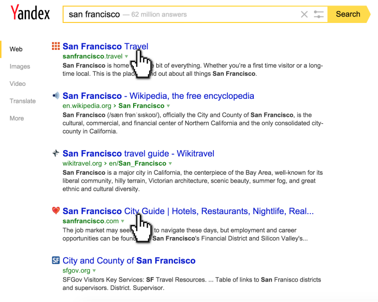
$$P(A_u = 1) = \alpha_{uq}$$

$$P(C_u = 1) = P(E_{r_u} = 1) \cdot P(A_u = 1)$$



Cascade model

1. Start from the first document
2. Examine documents one by one
3. If click, then stop
4. Otherwise, continue



Cascade model

$$E_r = 1 \text{ and } A_{u_r} = 1 \Leftrightarrow C_r = 1$$

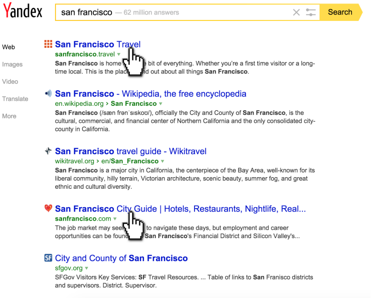
$$P(A_{u_r} = 1) = \alpha_{u_r q}$$

$$\underbrace{P(E_1 = 1)}_{\text{start from first}} = 1$$

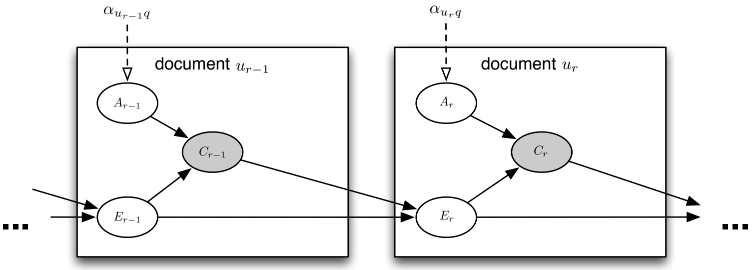
$$\underbrace{P(E_r = 1 \mid E_{r-1} = 0)}_{\text{examine one by one}} = 0$$

$$\underbrace{P(E_r = 1 \mid C_{r-1} = 1)}_{\text{if click, then stop}} = 0$$

$$\underbrace{P(E_r = 1 \mid E_{r-1} = 1, C_{r-1} = 0)}_{\text{otherwise, continue}} = 1$$



Cascade model: probabilistic graphical model



Basic click models summary

- CTR models
  - \+ count clicks (simple and fast)
  - – do not distinguish examination and attractiveness
- Position-based model (PBM) $\longrightarrow$ **User browsing model**
  - \+ examination and attractiveness
  - – examination of a document at rank r does not depend on examinations and clicks above r
- Cascade model (CM) $\longrightarrow$ **Dynamic Bayesian network**
  - \+ cascade dependency of examination at r on examinations and clicks above r
  - – only one click is allowed

Outline

2 Basic click models
- Random click model
- Position-based model
- Cascade model
- Click probabilities
- Evaluation
- Parameter estimation

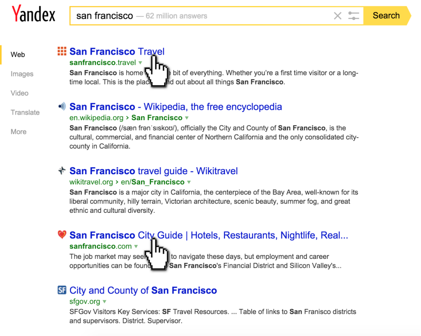
Click probabilities

- Full probability – probability that a user clicks on a document at rank r

$$P(C_r = 1)$$

- Conditional probability – probability that a user clicks on a document at rank r given previous clicks

$$P(C_r = 1 \mid C_1, \ldots, C_{r-1})$$



Click probabilities

- Full probability

$$P(C_{r+1} = 1) =$$

$$\alpha_{u_{r+1}q}\epsilon_r \cdot \left(\begin{array}{l} P(E_{r+1} = 1 \mid E_r = 1, C_r = 1) \cdot P(C_r = 1 \mid E_r = 1) \\ + \\ P(E_{r+1} = 1 \mid E_r = 1, C_r = 0) \cdot P(C_r = 0 \mid E_r = 1) \end{array} \right)$$

- Conditional probability

$$P(C_{r+1} = 1 \mid C_1, \ldots, C_r)$$

$$= \alpha_{u_{r+1}q} \cdot \left(\begin{array}{l} P(E_{r+1} = 1 \mid E_r = 1, C_r = 1) \cdot c_r^{(s)} \\ + \\ P(E_{r+1} = 1 \mid E_r = 1, C_r = 0) \cdot \dfrac{\epsilon_r(1 - \alpha_{u_r q})}{1 - \alpha_{u_r q}\epsilon_r} \cdot (1 - c_r^{(s)}) \end{array} \right)$$

Outline

2. Basic click models
   - Random click model
   - Position-based model
   - Cascade model
   - Click probabilities
   - Evaluation
   - Parameter estimation

Evaluation

Click model's output	Evaluation
Full click probabilities	Perplexity
Conditional click probabilities	Log-likelihood

Perplexity

Perplexity measures how well a click model estimates full click probabilities (i.e., when clicks are not observed).

$$p_r(M) = 2^{-\frac{1}{|\mathcal{S}|}\sum_{s \in \mathcal{S}} \left(\log_2 \overbrace{P_M(C_r^{(s)} = c_r^{(s)})}^{\text{full click probability}} \right)}$$

$$p_r(M) \in [1..2]$$

Likelihood

Likelihood measures how well a click model estimates conditional click probabilities given observed clicks.

$$\mathcal{LL}(M) = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \log P_M \left(C_1 = c_1^{(s)}, \ldots, C_n = c_n^{(s)} \right)$$

$$= \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \sum_{r=1}^{n} \log \underbrace{P_M \left(C_r = c_r^{(s)} \mid \mathbf{C}_{<r} = \mathbf{c}_{<r}^{(s)} \right)}_{\text{conditional click probability}}$$

$$\mathcal{LL}(M) \in [-\infty..0]$$

Outline

2 Basic click models
- Random click model
- Position-based model
- Cascade model
- Click probabilities
- Evaluation
- Parameter estimation

Parameter estimation

- Maximum likelihood estimation
- Expectation-maximization

MLE for random click model

$$P(C_u = 1) = \rho$$

$$\mathcal{L} = \underbrace{\prod_{s \in \mathcal{S}} \prod_{u \in s} \rho^{c_u^{(s)}} (1 - \rho)^{1 - c_u^{(s)}}}_{\text{likelihood of Bernoulli random variable}}$$

$$\mathcal{LL} = \sum_{s \in \mathcal{S}} \sum_{u \in s} \left(c_u^{(s)} \log(\rho) + (1 - c_u^{(s)}) \log(1 - \rho) \right)$$

$$\rho = \frac{\sum_{s \in \mathcal{S}} \sum_{u \in s} c_u^{(s)}}{\sum_{s \in \mathcal{S}} |s|} = \frac{\#\text{ clicks}}{\#\text{ shown docs}}$$

Expectation maximization

1. Set parameters to some initial values
2. Repeat until convergence
    - E-step: derive the expectation of the likelihood function
    - M-step: maximize this expectation

Expectation maximization

$$Q(\theta_c) = \sum_{s \in \mathcal{S}} \mathbb{E}_{\mathbf{X} | \mathbf{C}^{(s)}, \mathbf{\Psi}} \left[\log P\left(\mathbf{X}, \mathbf{C}^{(s)} \mid \mathbf{\Psi}\right) \right]$$

$$= \sum_{s \in \mathcal{S}} \mathbb{E}_{\mathbf{X} | \mathbf{C}^{(s)}, \mathbf{\Psi}} \left[\sum_{c_i \in s} \left(\mathcal{I}\left(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p}\right) \log(\theta_c) + \mathcal{I}\left(X_{c_i}^{(s)} = 0, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p}\right) \log(1 - \theta_c) \right) + \mathcal{Z} \right]$$

$$= \sum_{s \in \mathcal{S}} \sum_{c_i \in s} \left(P\left(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right) \log(\theta_c) + P\left(X_{c_i}^{(s)} = 0, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right) \log(1 - \theta_c) \right) + \mathcal{Z}$$

$$ESS(x) = \sum_{s \in \mathcal{S}} \sum_{c_i \in s} P\left(X_{c_i}^{(s)} = x, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right)$$

$$\frac{\partial Q(\theta_c)}{\partial \theta_c} = \sum_{s \in \mathcal{S}} \sum_{c_i \in s} \left(\frac{P\left(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right)}{\theta_c} - \frac{P\left(X_{c_i}^{(s)} = 0, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right)}{1 - \theta_c} \right) = 0$$

$$\theta_c^{(t+1)} = \frac{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} P\left(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right)}{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} \sum_{x=0}^{x=1} P\left(X_{c_i}^{(s)} = x, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right)}$$

$$= \frac{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} P\left(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right)}{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} P\left(\mathcal{P}(X_{c_i}^{(s)}) = \mathbf{p} \mid \mathbf{C}^{(s)}, \mathbf{\Psi}\right)} = \frac{ESS^{(t)}(1)}{ESS^{(t)}(1) + ESS^{(t)}(0)}$$

Click models summary so far

- Basic click models
  - CTR models
  - Position-based model
  - Cascade model
- Click probabilities
  - Full click probabilities
  - Conditional click probabilities
- Evaluation
  - Perplexity
  - Log-likelihood
- Parameter estimation
  - Maximum likelihood estimation
  - Expectation-maximization

What do click models give us?

General:

- Understanding of user behavior

Specific:

- Conditional click probabilities
- Full click probabilities
- Attractiveness and satisfactoriness for query-document pairs

Applications

Click model's output	Application
Understanding of user behavior	User interaction analysis
Conditional click probabilities	User simulation
Full click probabilities	Model-based metrics
Parameter values	Ranking

Outline

User interaction analysis

- Random click model (global CTR): $\rho = 0.122$
- Rank-based CTR:
  $\rho_1 = 0.429, \rho_2 = 0.190, \rho_3 = 0.136, \ldots, \rho_{10} = 0.048$
- Position-based model:
  $\gamma_1 = 0.998, \gamma_2 = 0.939, \gamma_3 = 0.759, \ldots, \gamma_{10} = 0.260$
- Dynamic Bayesian network model: $\gamma = 0.9997$

Click models are trained on the first 10K sessions of the WSCD 2012 dataset.

Simulating users

Algorithm Simulating user clicks

Input: click model M, query session s

Output: vector of simulated clicks $(c_1, \ldots, c_n)$

1: **for** $r \leftarrow 1$ to $|s|$ **do**

2: $P_r \leftarrow \underbrace{P_M(C_r = 1 \mid C_1 = c_1, \ldots, C_{r-1} = c_{r-1})}_{\text{conditional click probability}}$

3: Generate c_r from $Bernoulli(P_r)$

4: **end for**

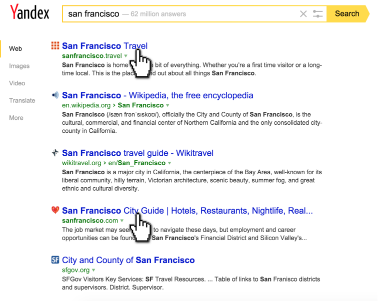
Model-based metrics

- Utility-based metrics

$$uMetric = \sum_{r=1}^{n} P(C_r = 1) \cdot U_r$$

- Effort-based metrics

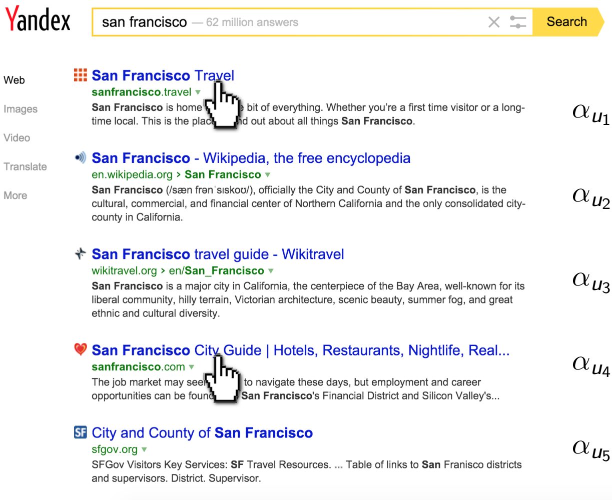
$$eMetric = \sum_{r=1}^{n} P(S_r = 1) \cdot F_r$$



Expected reciprocal rank

$$
\begin{aligned}
ERR &= \sum_r \frac{1}{r} \cdot P(S_r = 1) \\
&= \sum_r \frac{1}{r} \cdot R_{u_r q} \cdot \prod_{i=1}^{r-1} \left(\gamma \cdot (1 - R_{u_i q})\right)
\end{aligned}
$$

Features for ranking



$\alpha_{u_1 q}$

$\alpha_{u_2 q}$

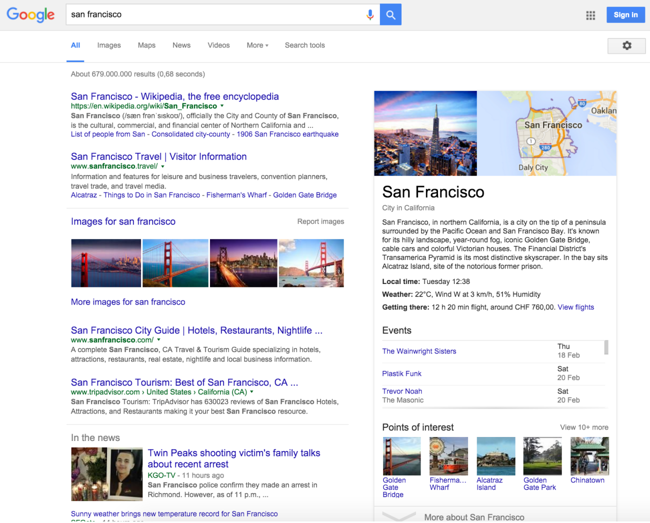
$\alpha_{u_3 q}$

$\alpha_{u_4 q}$

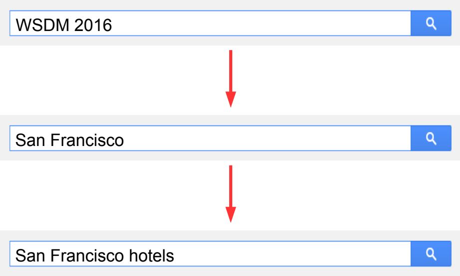
$\alpha_{u_5 q}$

Outline

Aggregated SERPs



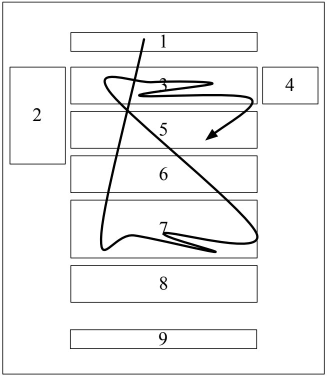
Search tasks



Using features

ID	Feature description
1	Term frequency (TF) of body
2	TF of anchor
3	TF of title
4	TF of URL
5	TF of whole document
6	Inverse document frequency (IDF) of body
7	IDF of anchor
8	IDF of title
9	IDF of URL
10	IDF of whole document
11	TF*IDF of body
12	TF*IDF of anchor
13	TF*IDF of title
14	TF*IDF of URL
15	TF*IDF of whole document
16	Document length (DL) of body
17	DL of anchor
18	DL of title
19	DL of URL
20	DL of whole document
21	BM25 of body
22	BM25 of anchor
23	BM25 of title
24	BM25 of URL
25	BM25 of whole document

Beyond clicks



Picture taken from F. Diaz, R.W. White, G. Buscher, and D. Liebling. Robust models of mouse movement on dynamic web search results pages. In *CIKM*, 2013. ACM Press

Outline

Incorporating clicks, attention and satisfaction

Incorporating clicks, attention and satisfaction into a search engine result page evaluation model

Aleksandr Chuklin, Maarten de Rijke

Proceedings of CIKM 2016, Indianapolis, USA

A neural click model for web search

A neural click model for web search

Alexey Borisov, Ilya Markov, Maarten de Rijke, Pavel Serdyukov

Proceedings of WWW 2016, Montreal, Canada

A context-aware time model for web search

A context-aware time model for web search

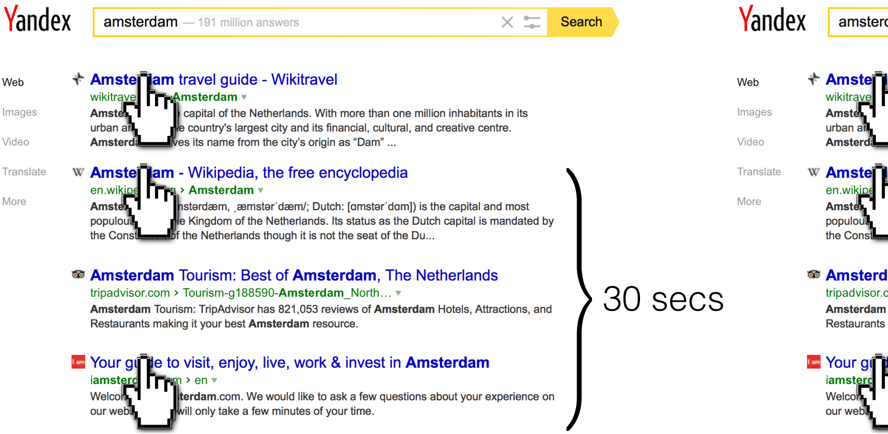
Alexey Borisov, Ilya Markov, Maarten de Rijke, Pavel Serdyukov

Proceedings of SIGIR 2016, Pisa, Italy
best student paper award

Time between user actions

- **Time between clicks**
- Time to first click
- Time to last click
- Time between queries

Time between clicks



there is *context bias*

Modeling time

- Average

$$Time(\text{“}Amsterdam\text{”}, \text{“}wikipedia.org\text{”}) = \frac{120 + 60 + 30}{3}$$

- Probability distribution

$$Time(\text{“}Amsterdam\text{”}, \text{“}wikipedia.org\text{”}) \sim Gamma(\boldsymbol{k}, \boldsymbol{\theta})$$

where $(\boldsymbol{k}, \boldsymbol{\theta})$ are estimated from 120, 60, 30

context bias is not modeled

Context-aware time modeling

$$Time(\text{"Amsterdam"}, \text{"wikipedia.org"}, context_1) \sim Gamma(\boldsymbol{k_1}, \boldsymbol{\theta_1})$$

$$Time(\text{"Amsterdam"}, \text{"wikipedia.org"}, context_2) \sim Gamma(\boldsymbol{k_2}, \boldsymbol{\theta_2})$$

$$Time(\underbrace{\text{"Amsterdam"}, \text{"wikipedia.org"}}_{\text{user action}}, \underbrace{context_3}_{\text{context}}) \sim Gamma(\boldsymbol{k_3}, \boldsymbol{\theta_3})$$

Context-aware time modeling

$$
\begin{aligned}
Time(action, context) \sim Gamma(&\\
&\boldsymbol{a}_k(ctx) \cdot \boldsymbol{k}(act) + \boldsymbol{b}_k(ctx),\\
&\boldsymbol{a}_\theta(ctx) \cdot \boldsymbol{\theta}(act) + \boldsymbol{b}_\theta(ctx))
\end{aligned}
$$

Parameter estimation

$$
\begin{aligned}
Time(action, context) \sim Gamma(&\\
& \boldsymbol{a}_k(ctx) \cdot \boldsymbol{k}(act) + \boldsymbol{b}_k(ctx),\\
& \boldsymbol{a}_\theta(ctx) \cdot \boldsymbol{\theta}(act) + \boldsymbol{b}_\theta(ctx))
\end{aligned}
$$

1. Fix context-independent parameters
2. Optimize context-dependent parameters using *neural networks*
3. Fix context-dependent parameters
4. Optimize context-independent using *gradient descent*
5. Repeat until convergence

Parameter estimation

- We do not know the form of context-dependent parameters $\Longrightarrow$ neural networks
- We know the form of context-independent parameters (Gamma distribution) $\Longrightarrow$ direct optimization

Context

General	
Is query (Q-action)	(0: no, 1: yes)
Is click (C-action)	(0: no, 1: yes)
$\log(1 +$ observed time since previous action$)$	(0: undefined)
$\log(1 +$ average time since previous action$)$	(0: undefined)
Q-action	
Is new search session	(0: no, 1: yes)
Number of terms in issued query	(0: undefined)
BM25 (issued query, previous query)	(0: undefined)
BM25 (previous query, issued query)	(0: undefined)
C-action	
Is click on the 1st position	(0: no, 1: yes)
...	...
Is click on the 10^{th} position	(0: no, 1: yes)

Dataset

3 months of log data from Yandex search engine

Time between actions	Max time	# Observations
Time-to-first-click	1 min	30,747,733
Time-between-clicks	5 min	6,317,834
Time-to-last-click	5 min	30,446,973
Time-from-abandoned-query	1 min	11,523,351

Evaluation tasks

Task1. Predict time between clicks

- Log-likelihood
- Root mean squared error (MSE)

Task2. Rank results based on time between clicks

- nDCG@$\{1, 3, 5, 10\}$

Task 1. Predicting time

Time model	Distribution	Log-likelihood	RMSE
Basic	exponential	−4.9219	60.73
	gamma	−4.9105	60.76
	Weibull	−4.9077	60.76
Context-aware	exponential	−4.8787	58.93
	gamma	−4.8556	58.98
	Weibull	−4.8504	58.94

Task 2. Ranking results

		NDCG			
Time model	Distribution	@1	@3	@5	@10
Average	—	0.651	0.693	0.728	0.812
Context-aware	exponential	0.668	0.710	0.743	0.820
	gamma	0.675	0.715	0.748	0.822
	Weibull	0.671	0.709	0.745	0.821

Other times

- Time to first click
- Time to last click
- Time between queries

Summary

- Removed context bias from time between actions
- Predicted user search interactions better (**Task 1**)
- Used the context-independent component for better document ranking (**Task 2**)

Outline

Future research

- Keep on adding new variables – not a good idea
- Parameter estimation
  - Efficiency
  - Online learning
- Other interactions and environments
  - Interactions beyond clicks
  - Devices beyond desktop computers

Future research

Model's output	Evaluation	Application
Conditional click probs	Log-likelihood	User simulation
Full click probs	Perplexity	Model-based metrics
Parameter values	Ranking evaluation	Ranking

- Why use intermediate evaluation?
  - Evaluate applications, not models
- Why maximize log-likelihood?
  - Optimize models for specific applications

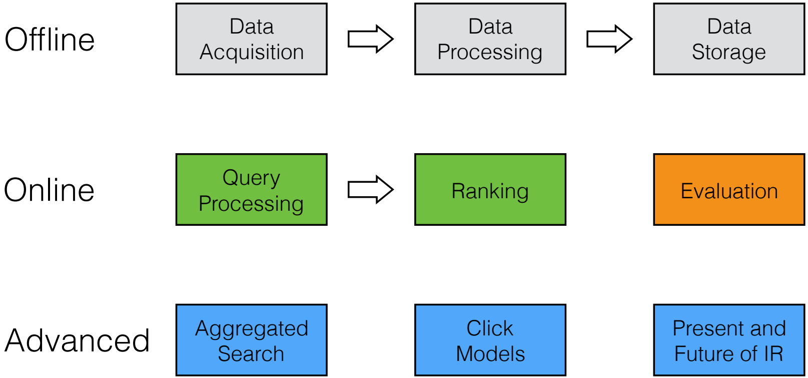
Outline

Outline

Materials

- Aleksandr Chuklin, Ilya Markov, Maarten and de Rijke
  Click Models for Web Search
  Morgan & Claypool, 2015

Thank you!


Offline
Data Acquisition
Data Processing
Data Storage
Online
Query Processing
Ranking
Evaluation
Advanced
Aggregated Search
Click Models
Present and Future of IR