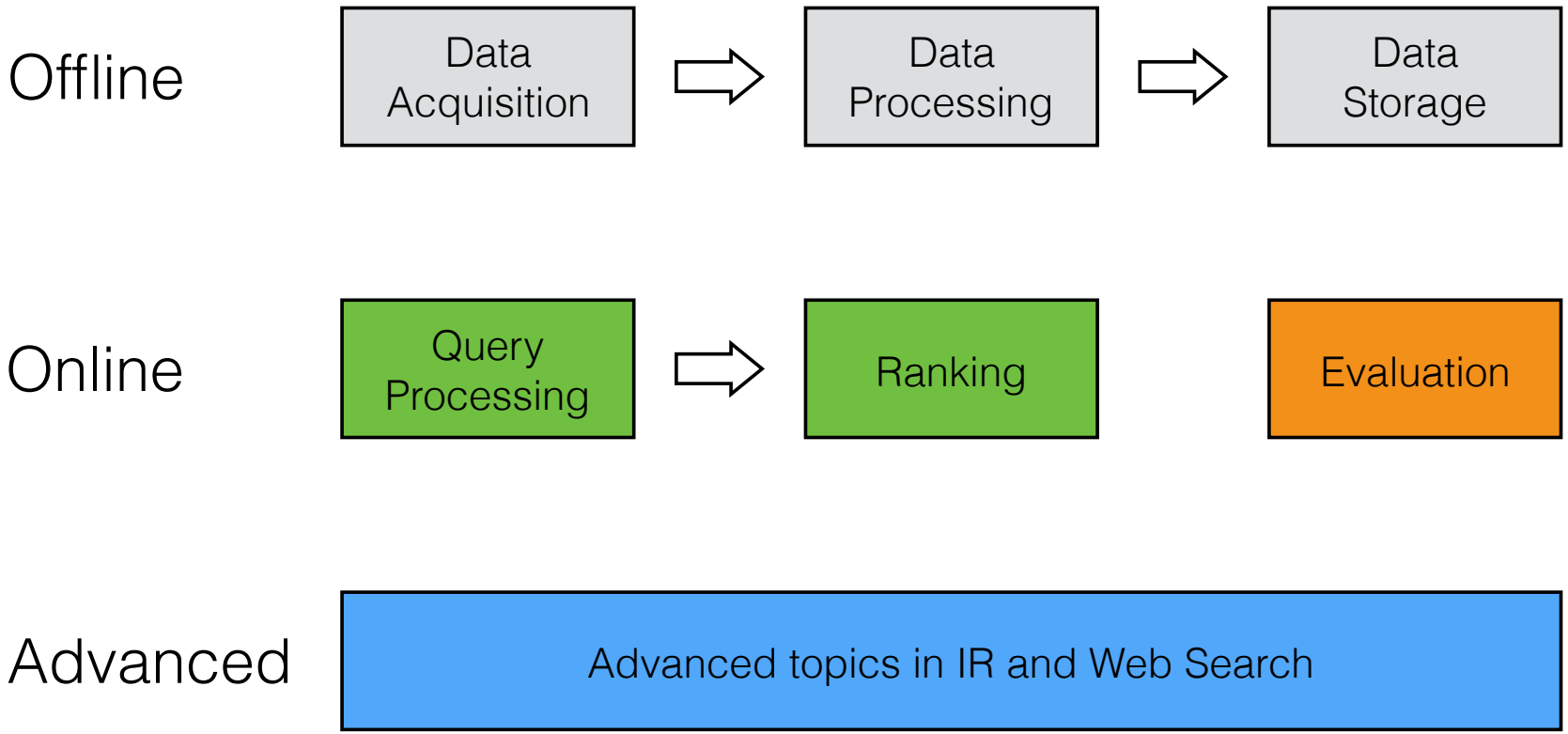
Information Retrieval

Click Models

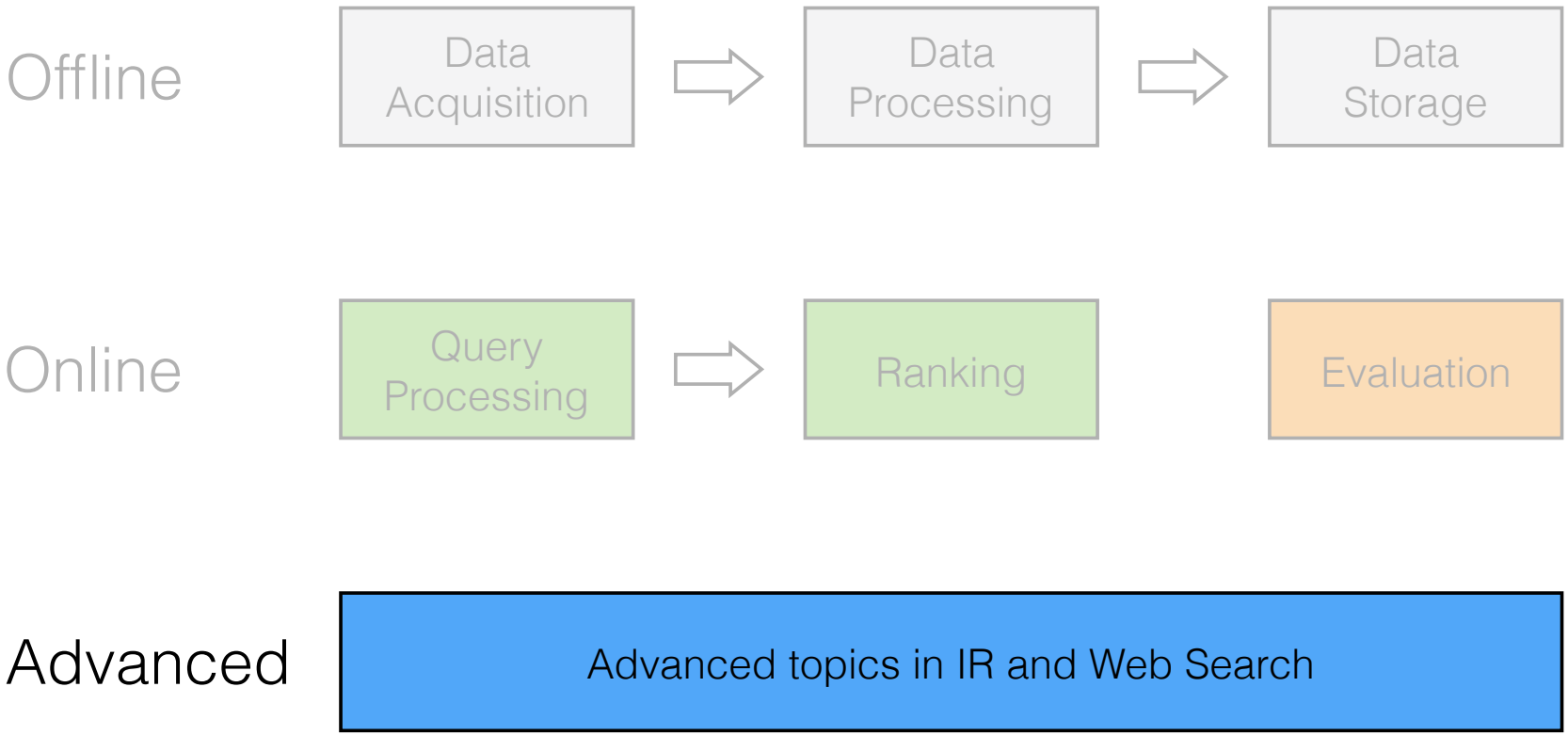
Ilya Markov
i.markov@uva.nl

University of Amsterdam

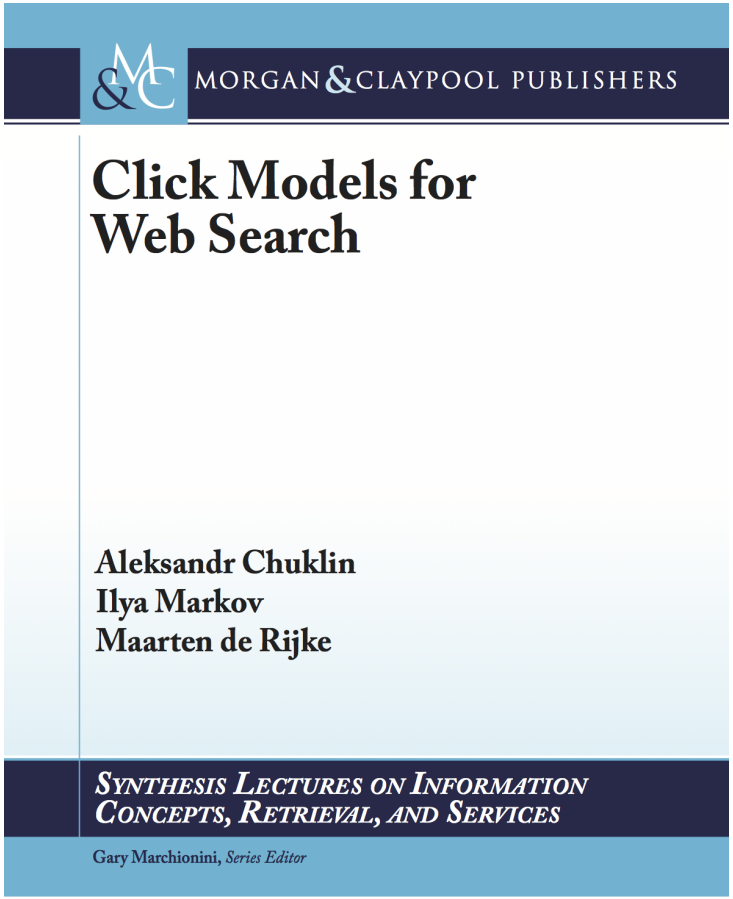
Course overview



Advanced topics in IR



The book



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

Tutorials

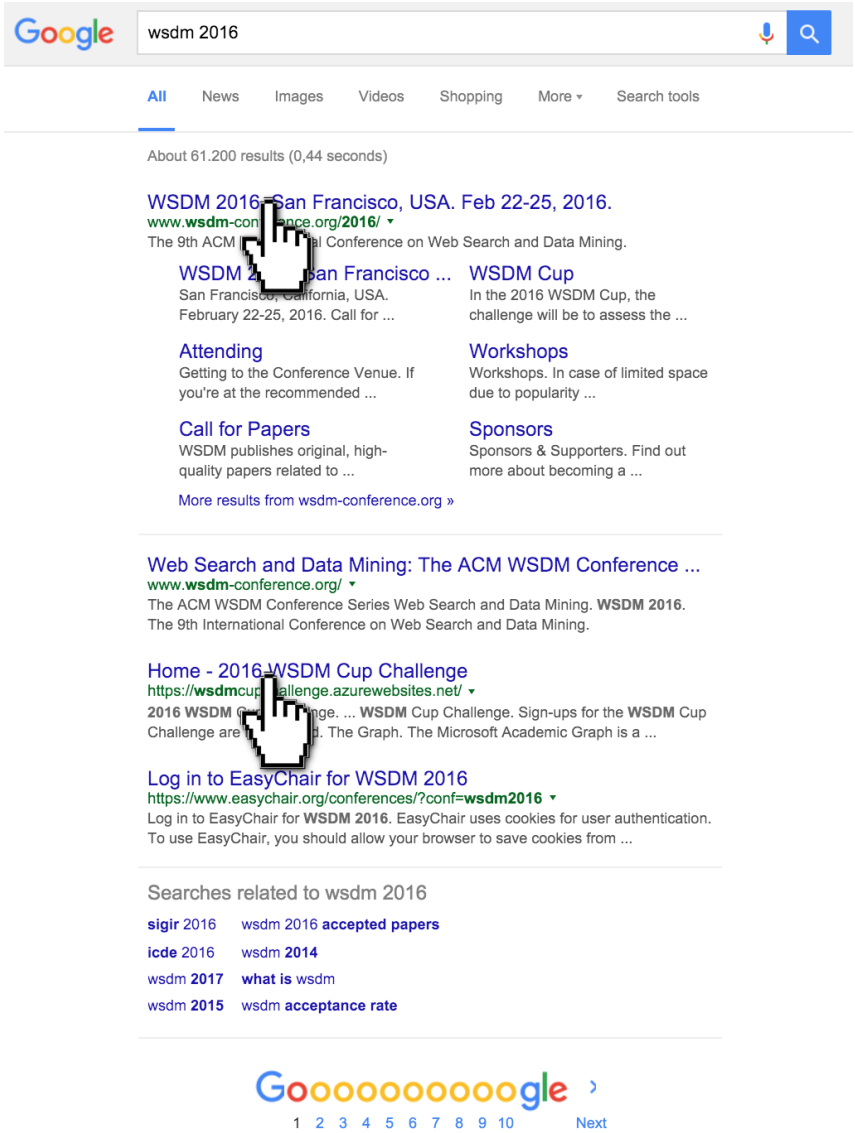
- SIGIR 2015, Santiago, Chile
- AINL-ISMW FRUCT 2015, St. Petersburg, Russia
- WSDM 2016, San Francisco, USA
- RuSSIR 2016, Saratov, Russia

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

Outline

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
```

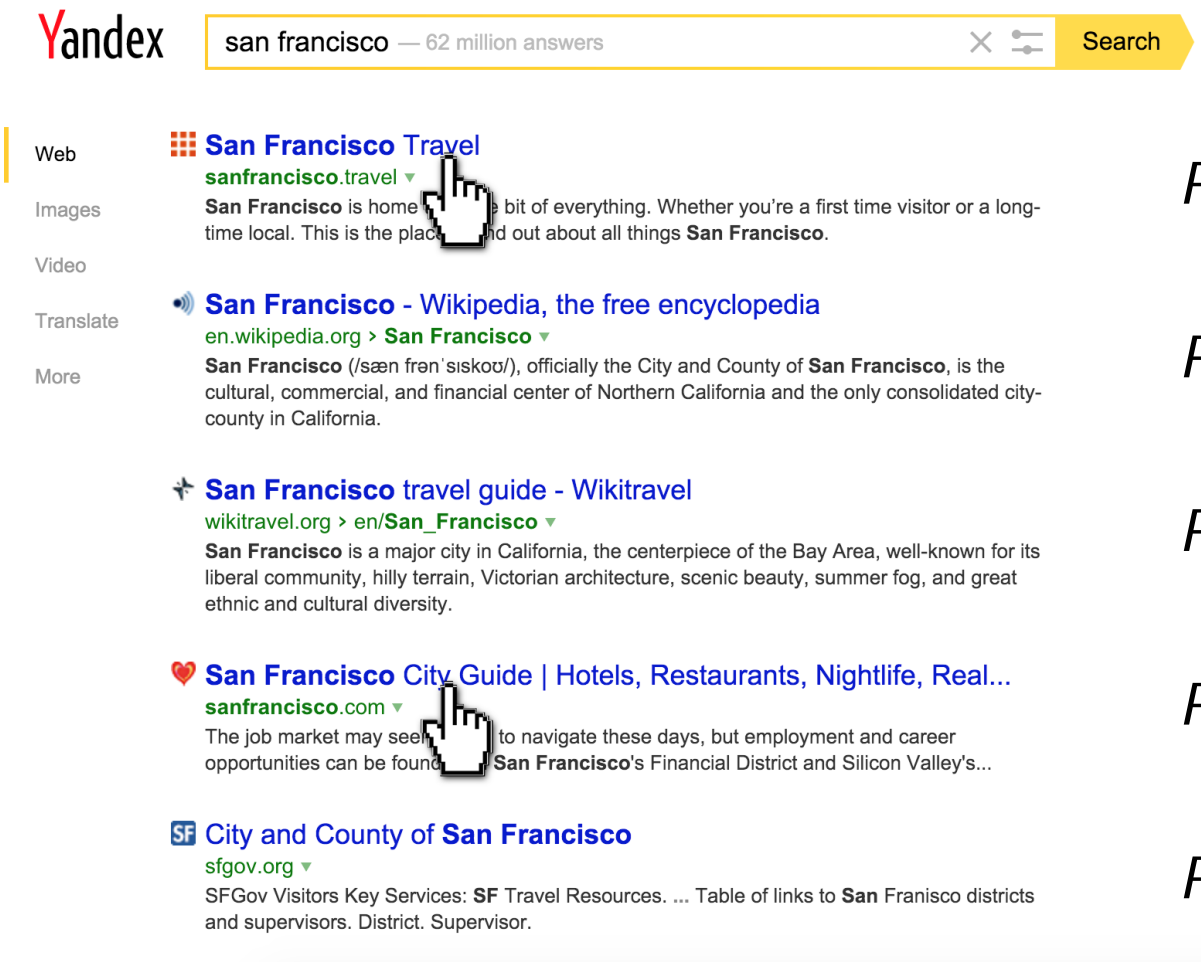
Yandex Relevance Prediction Challenge

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

Outline

Outline

Random click model

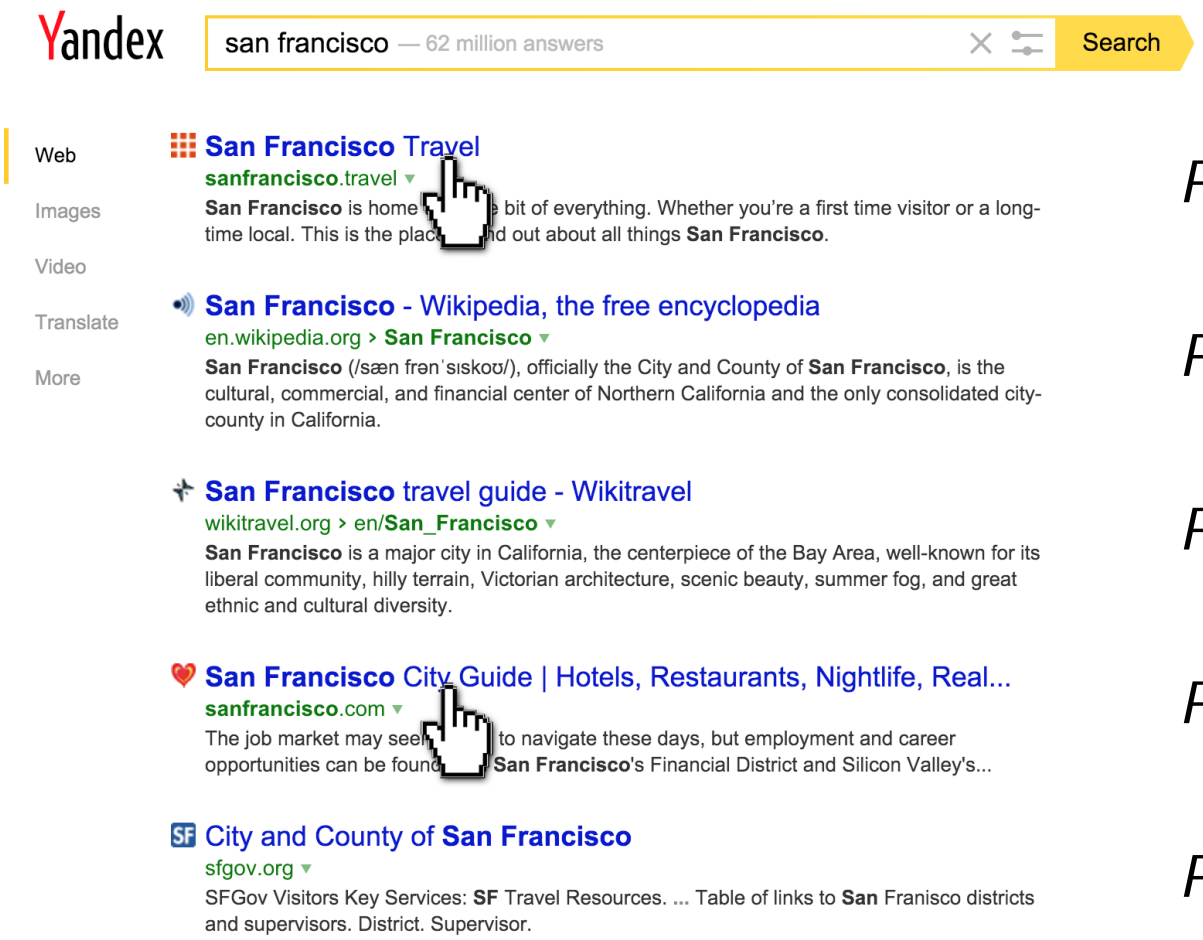


Random click model

- 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):

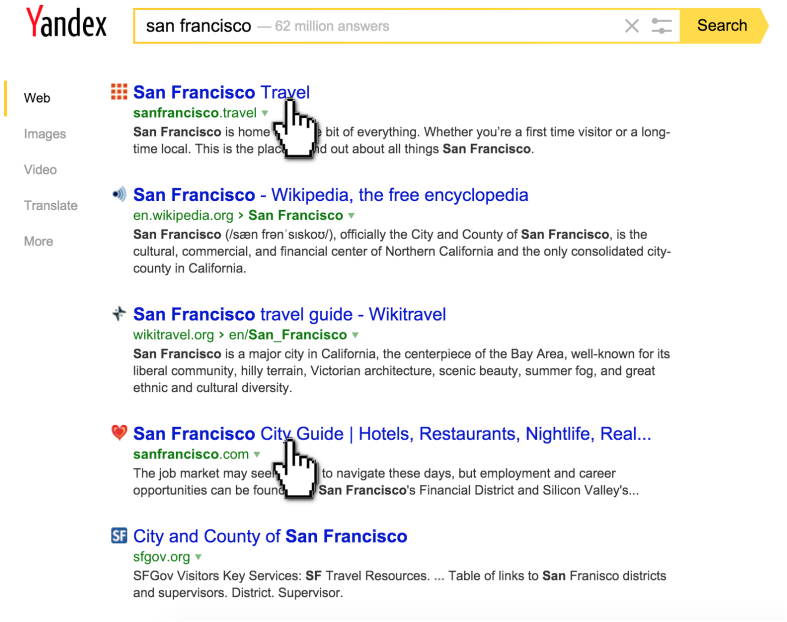
$$P(C_u = 1) = \rho$$

Rank-based CTR:

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

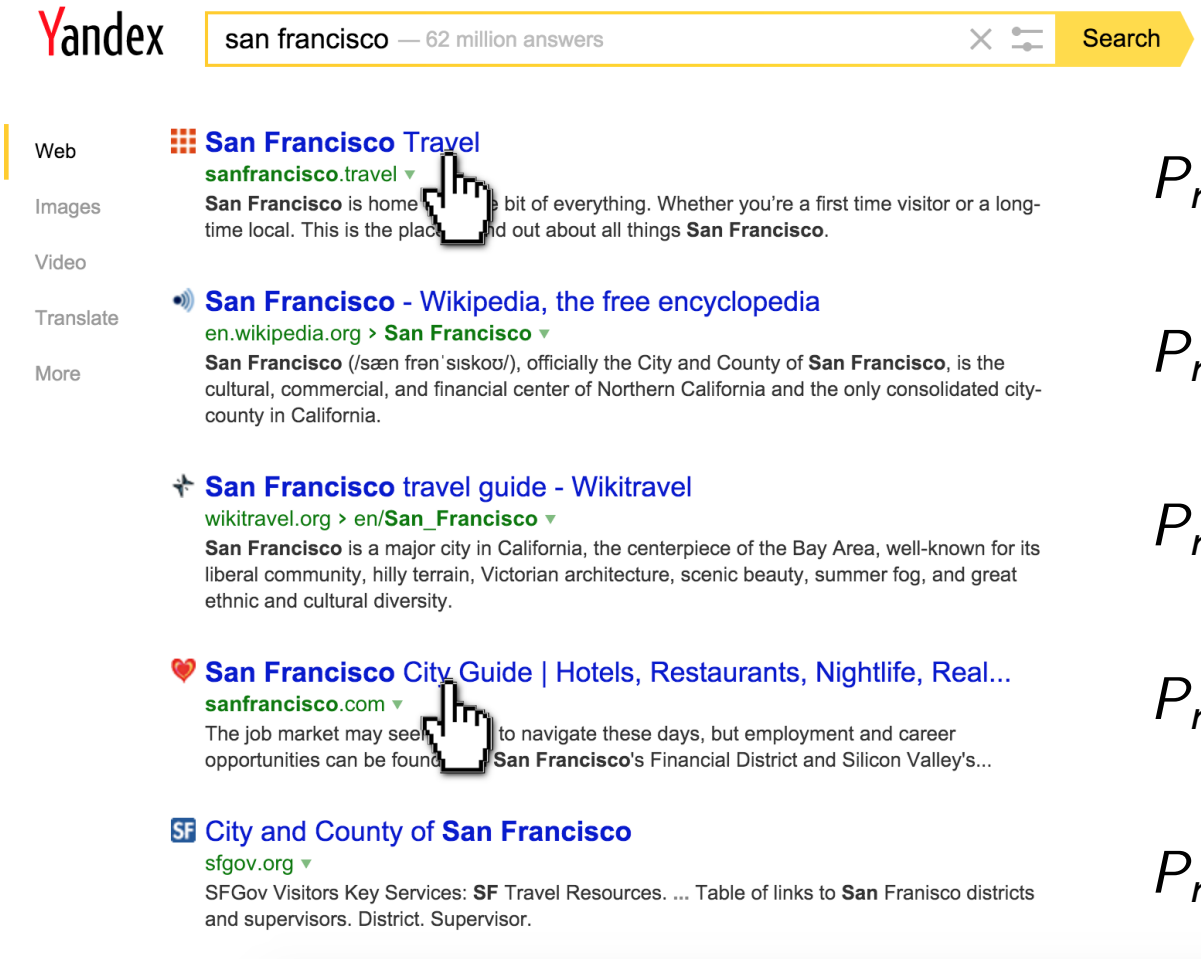
Query-document CTR:

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



Outline

Position-based model



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

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

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

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

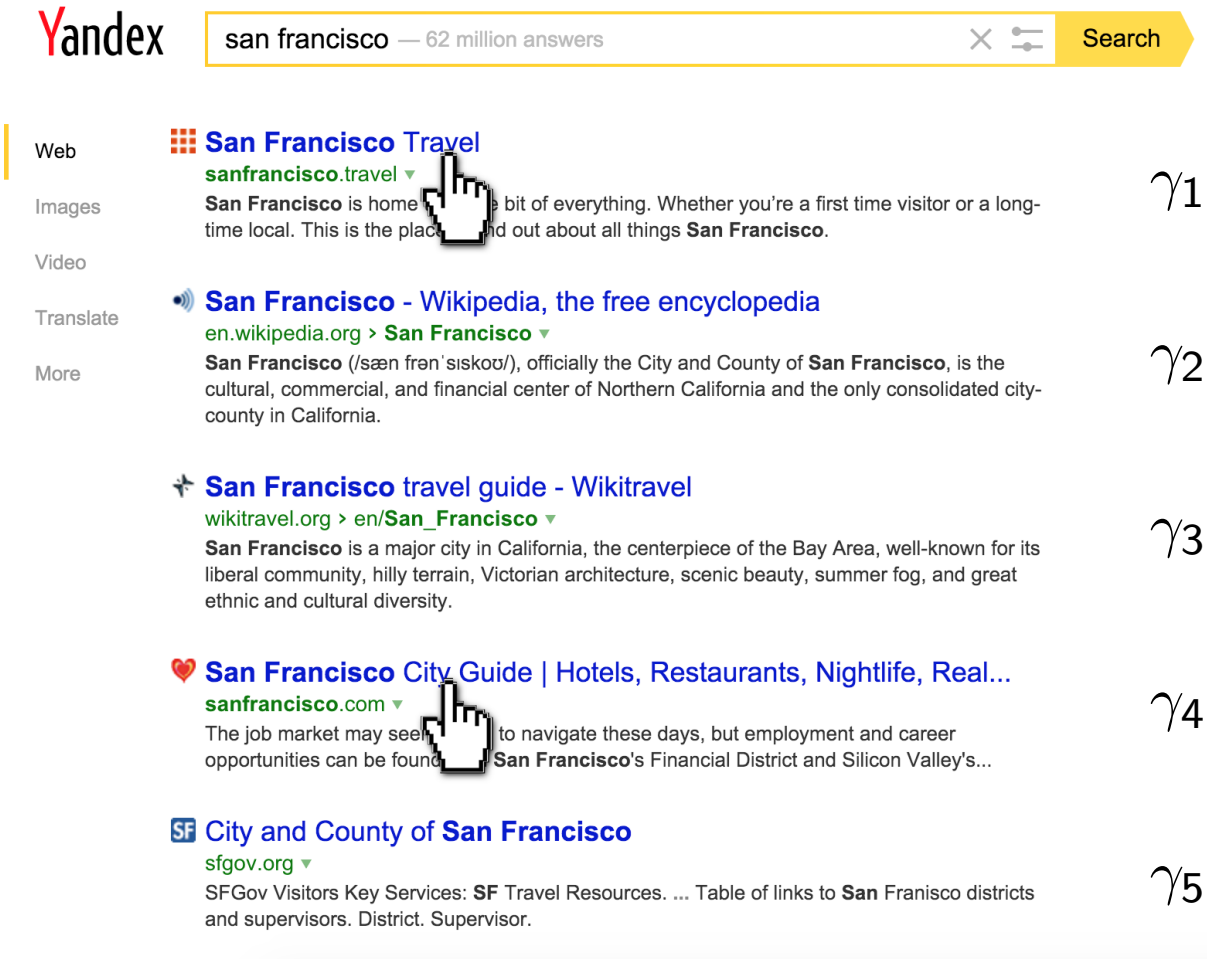
$P_{read}(5)$, $P_{click}(u_5 q)$

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



γ_1 , $P_{click}(u_1 q)$

γ_2 , $P_{click}(u_2 q)$

γ_3 , $P_{click}(u_3 q)$

γ_4 , $P_{click}(u_4 q)$

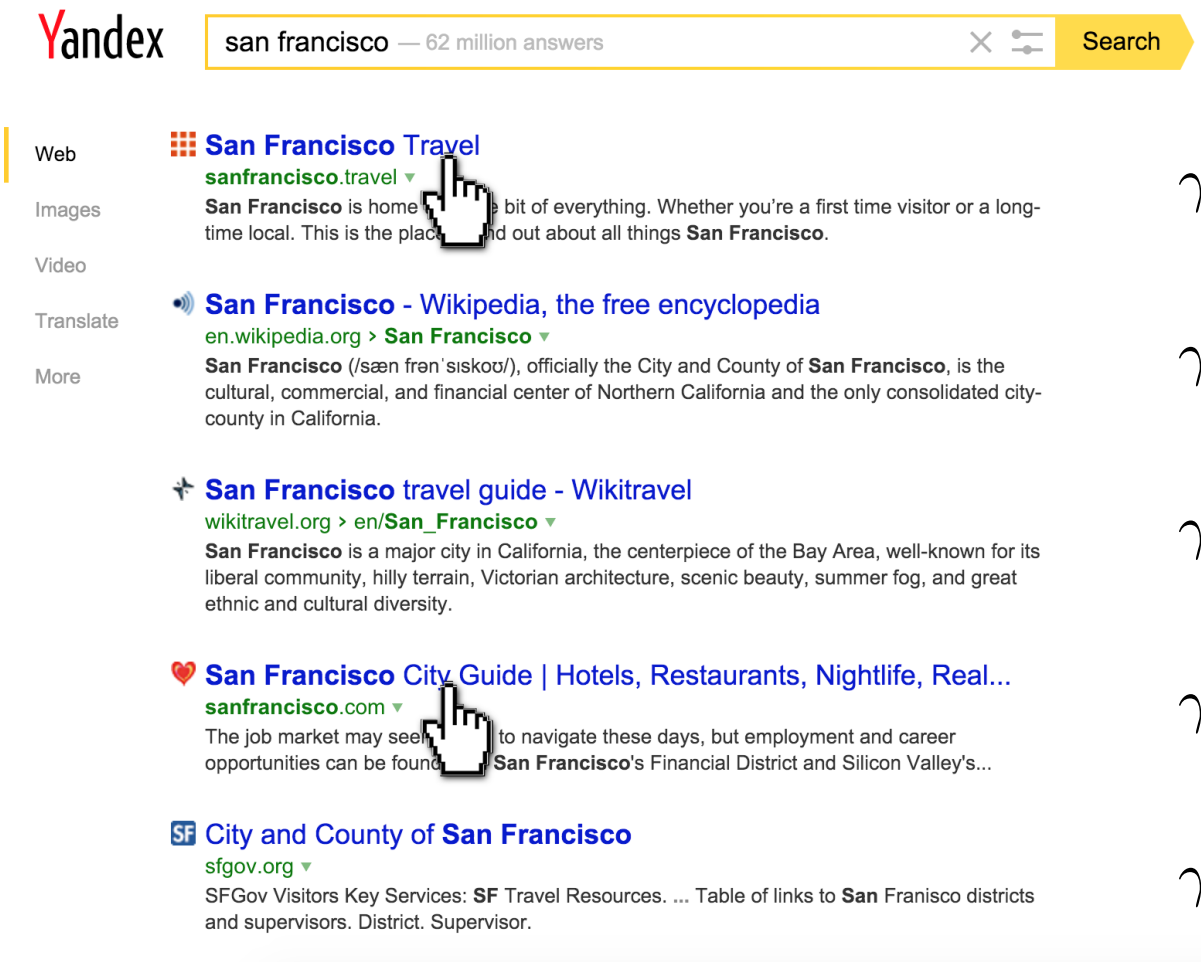
γ_5 , $P_{click}(u_5 q)$

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



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

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

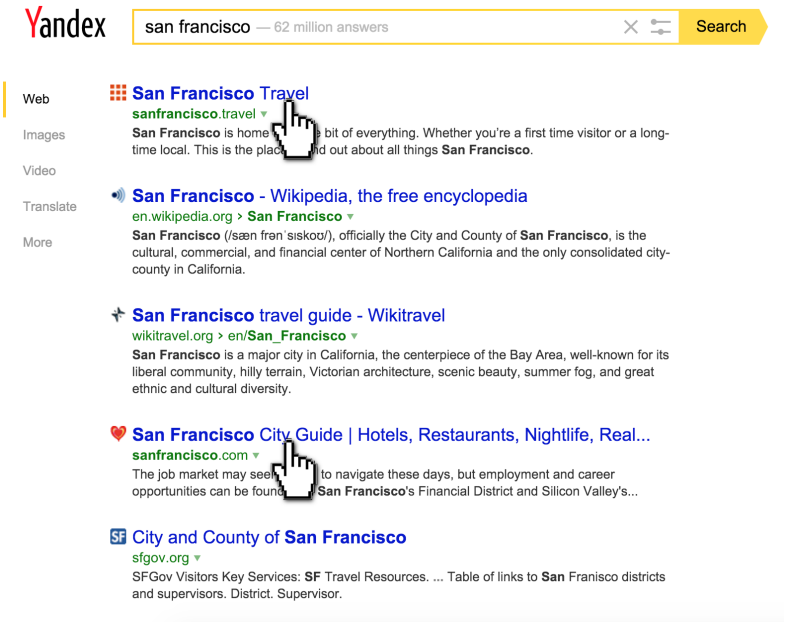
$\gamma_3 \ , \ \alpha_{u_3 q}$

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

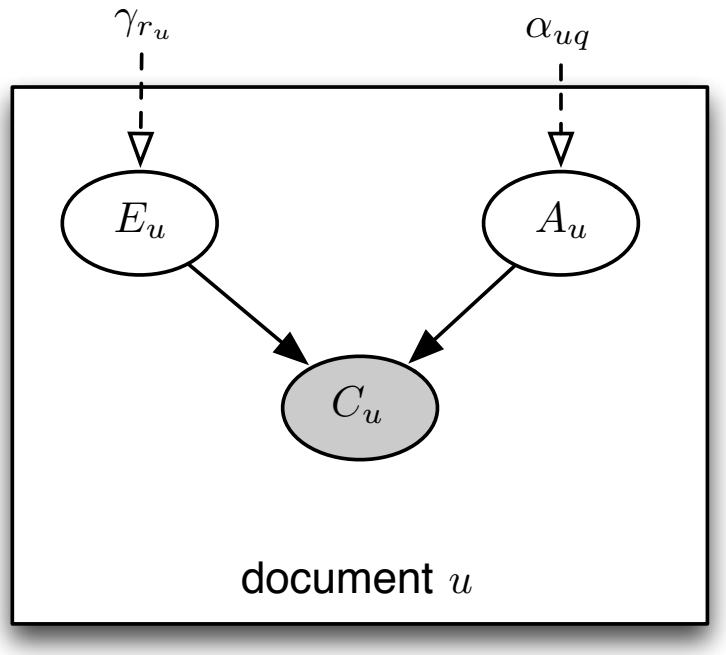
$\gamma_5 \ , \ \alpha_{u_5 q}$

Position-based model: summary

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

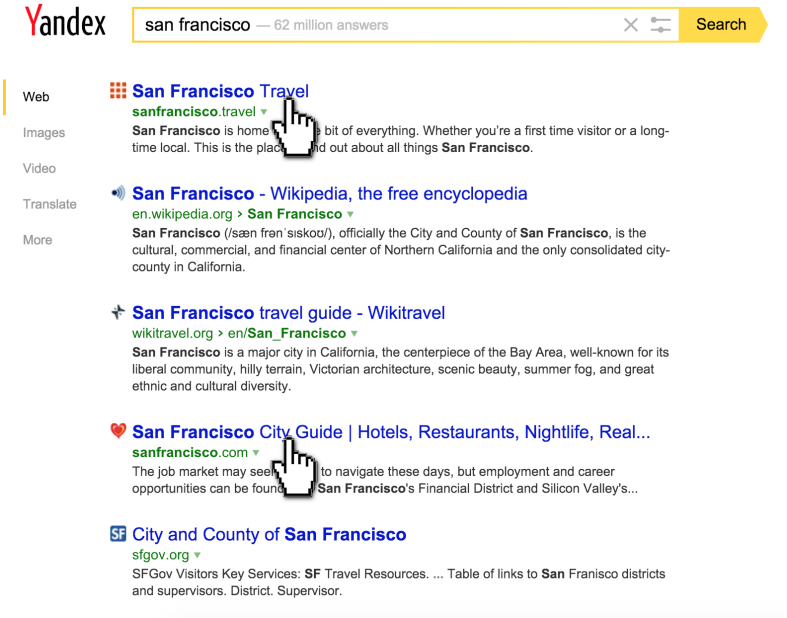


Position-based model: probabilistic graphical model



Position-based model: exercises

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

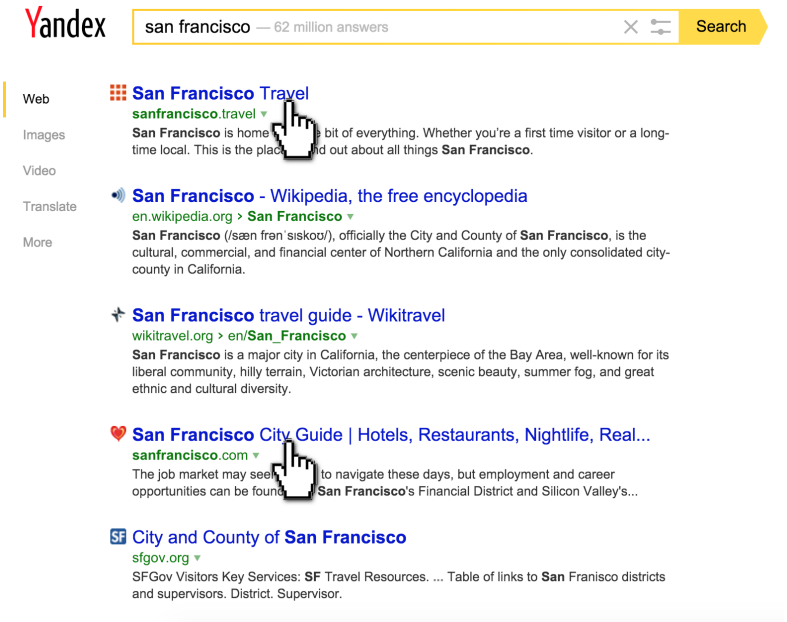


$$E_{r_u} = 0 \Rightarrow C_u = 0$$
$$A_u = 0 \Rightarrow C_u = 0$$
$$E_{r_u} = 1 \Rightarrow (C_u = 1 \iff A_u = 1)$$
$$A_u = 1 \Rightarrow (C_u = 1 \iff E_{r_u} = 1)$$

Outline

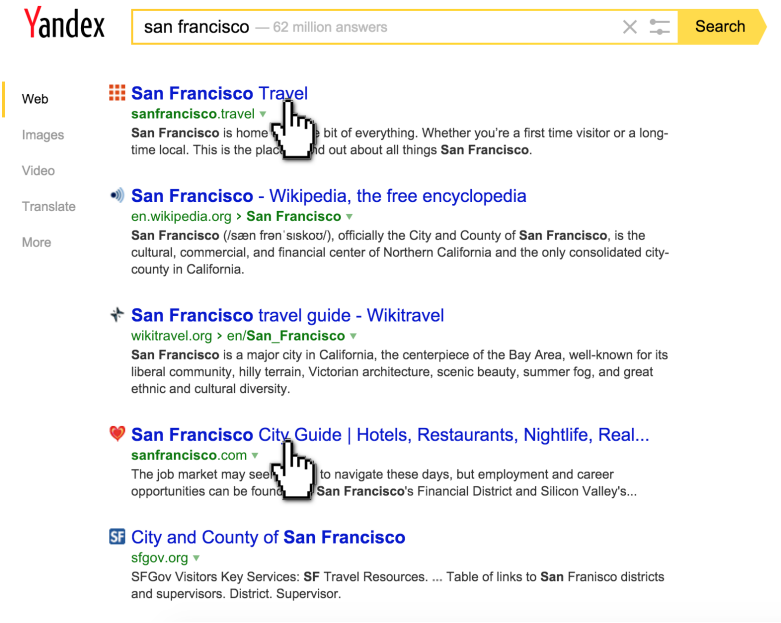
Position-based model

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



Cascade model

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



Cascade model

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

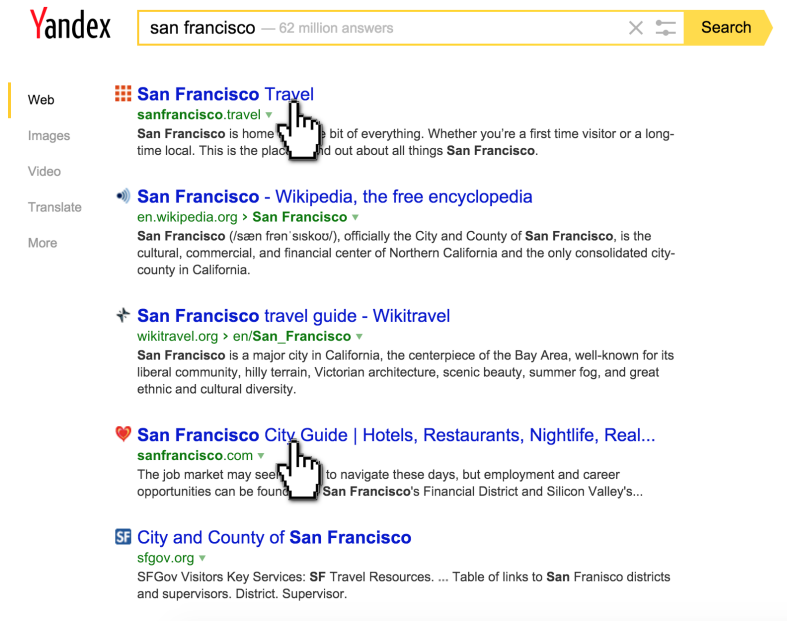
$$P(A_{u_r} = 1) = \alpha_{u_r q}$$

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

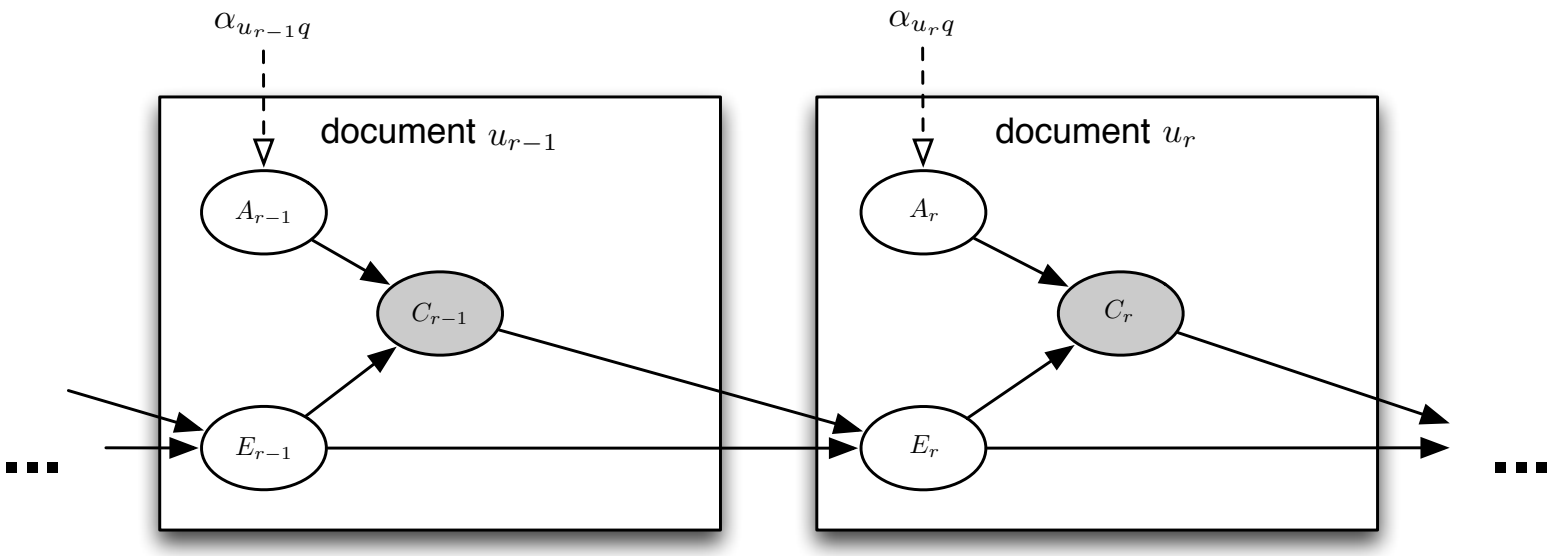
$$\underbrace{P(E_r = 1 \mid E_{r-1} = 0)}_{\text{examine one by one}} = 0$$

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

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



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

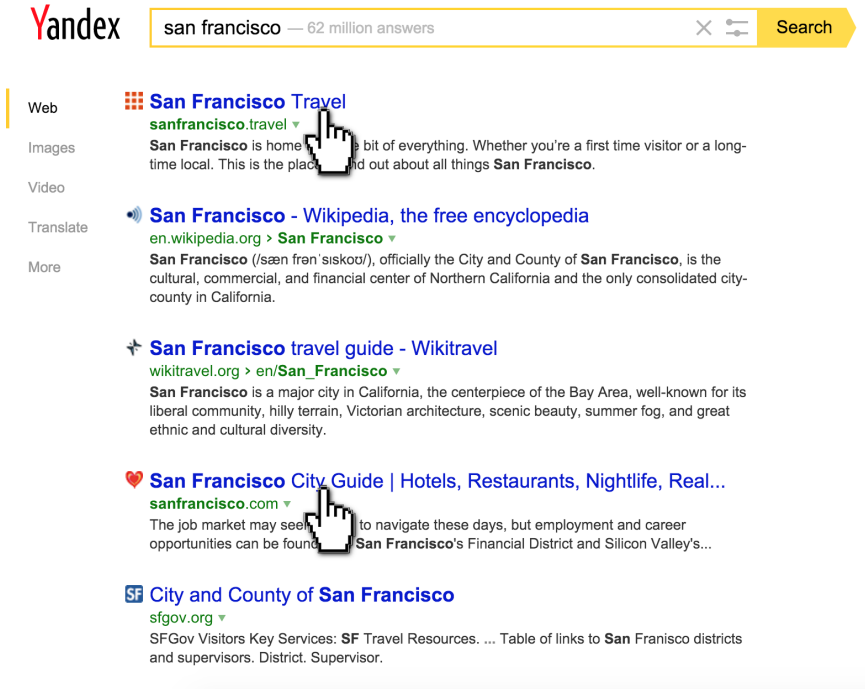
Click probabilities

- 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

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

- Conditional probability

$$P(C_{r+1}=1 \mid C_1, \ldots, C_r)$$
$$= \alpha_{u_{r+1}q} \cdot \begin{pmatrix} P(E_{r+1}=1 \mid E_r=1, C_r=1) \cdot c_r^{(s)} \\ + \\ P(E_{r+1}=1 \mid E_r=1, C_r=0) \cdot \dfrac{\epsilon_r(1-\alpha_{u_rq})}{1-\alpha_{u_rq}\epsilon_r} \cdot (1-c_r^{(s)}) \end{pmatrix}$$

Outline

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

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

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

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

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

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

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

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

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

Outline

Outline

Materials

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

Advanced topics in IR