

Information Retrieval 1

IR-User Interaction

Ilya Markov

i.markov@uva.nl

University of Amsterdam

Learning to rank

Evaluation

Document
representation
& matching

Conversational
search

Learning to rank

IR—user
interaction

Recommender
systems

IR-User Interaction

Evaluation

Document
representation
& matching

Conversational
search

Learning to rank

IR—user
interaction

Recommender
systems

User interactions

- Queries
- Interactions with a SERP (clicks, mousing, scrolling, etc.)
- Time between user actions
- Closing browser
- Interactions beyond search
- Etc.

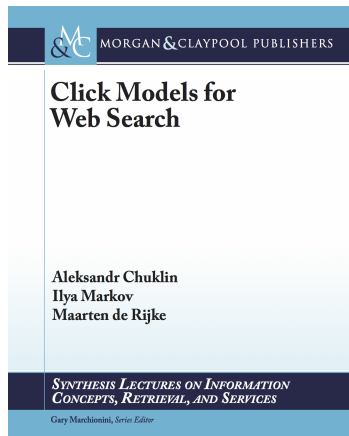
Why are user interactions important?

- Evaluate IR systems
- Improve IR systems

Models of user search interactions

- Click models
- Models of mouse hovering
- Models of time between user actions

Click models



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

Outline

- 1 Basic click models
- 2 Estimation
- 3 Applications

Basic click models

- Position-based model
- Cascade model

Position-based model

Yandex

san francisco — 62 million answers



Search

Web

San Francisco Travel

[sanfrancisco.travel](#) ▾

San Francisco is home to a little bit of everything. Whether you're a first time visitor or a long-time local. This is the place to find out about all things **San Francisco**.

Images

Video

Translate

More

San Francisco - Wikipedia, the free encyclopedia

[en.wikipedia.org](#) ▸ **San Francisco** ▾

San Francisco (/ˈsæn frənˈsikoʊ/), officially the City and County of **San Francisco**, is the cultural, commercial, and financial center of Northern California and the only consolidated city-county in California.

San Francisco travel guide - Wikitravel

[wikitravel.org](#) ▸ [en/San_Francisco](#) ▾

San Francisco is a major city in California, the centerpiece of the Bay Area, well-known for its liberal community, hilly terrain, Victorian architecture, scenic beauty, summer fog, and great ethnic and cultural diversity.

San Francisco City Guide | Hotels, Restaurants, Nightlife, Real...

[sanfrancisco.com](#) ▾

The job market may seem daunting to navigate these days, but employment and career opportunities can be found in **San Francisco's** Financial District and Silicon Valley's...

City and County of San Francisco

[sfgov.org](#) ▾

SFGov Visitors Key Services: **SF** Travel Resources. ... Table of links to **San Francisco** districts and supervisors. District. Supervisor.

$$P_{exam}(1), P_{attr}(qd_1)$$

$$P_{exam}(2), P_{attr}(qd_2)$$

$$P_{exam}(3), P_{attr}(qd_3)$$

$$P_{exam}(4), P_{attr}(qd_4)$$

$$P_{exam}(5), P_{attr}(qd_5)$$

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

Yandex

san francisco — 62 million answers



Search

Web

San Francisco Travel

[sanfrancisco.travel](#) ▾

San Francisco is home to a little bit of everything. Whether you're a first time visitor or a long-time local. This is the place to find out about all things **San Francisco**.

Images

Video

Translate

More

San Francisco - Wikipedia, the free encyclopedia

[en.wikipedia.org](#) > [San Francisco](#) ▾

San Francisco (/ˈsæn frənˈsikoʊ/), officially the City and County of **San Francisco**, is the cultural, commercial, and financial center of Northern California and the only consolidated city-county in California.

San Francisco travel guide - Wikitravel

[wikitravel.org](#) > [en/San_Francisco](#) ▾

San Francisco is a major city in California, the centerpiece of the Bay Area, well-known for its liberal community, hilly terrain, Victorian architecture, scenic beauty, summer fog, and great ethnic and cultural diversity.

San Francisco City Guide | Hotels, Restaurants, Nightlife, Real...

[sanfrancisco.com](#) ▾

The job market may seem difficult to navigate these days, but employment and career opportunities can be found in **San Francisco's** Financial District and Silicon Valley's...

City and County of San Francisco

[sfgov.org](#) ▾

SFGov Visitors Key Services: **SF** Travel Resources. ... Table of links to **San Francisco** districts and supervisors. District. Supervisor.

$$\gamma_1, P_{attr}(d_1q)$$

$$\gamma_2, P_{attr}(d_2q)$$

$$\gamma_3, P_{attr}(d_3q)$$

$$\gamma_4, P_{attr}(d_4q)$$

$$\gamma_5, P_{attr}(d_5q)$$

Position-based model: attractiveness

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

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

Position-based model

The screenshot shows a Yandex search results page for the query "san francisco". The search bar indicates 62 million answers. The results are categorized by type: Web, Images, Video, Translate, and More. The first five results are annotated with a hand cursor icon and a label γ_i, α_{qd_i} .

Category	Result	Annotation
Web	San Francisco Travel sanfrancisco.travel San Francisco is home to a bit of everything. Whether you're a first time visitor or a long-time local. This is the place to find out about all things San Francisco.	γ_1, α_{qd_1}
Web	San Francisco - Wikipedia, the free encyclopedia en.wikipedia.org > San Francisco San Francisco (/sæn frənˈsiskoo/), officially the City and County of San Francisco, is the cultural, commercial, and financial center of Northern California and the only consolidated city-county in California.	γ_2, α_{qd_2}
Web	San Francisco travel guide - Wikitravel wikitravel.org > en/San Francisco San Francisco is a major city in California, the centerpiece of the Bay Area, well-known for its liberal community, hilly terrain, Victorian architecture, scenic beauty, summer fog, and great ethnic and cultural diversity.	γ_3, α_{qd_3}
Web	San Francisco City Guide Hotels, Restaurants, Nightlife, Real... sanfrancisco.com The job market may seem to navigate these days, but employment and career opportunities can be found in San Francisco's Financial District and Silicon Valley's...	γ_4, α_{qd_4}
Web	City and County of San Francisco sf.gov SFGov Visitors Key Services: SF Travel Resources. ... Table of links to San Francisco districts and supervisors. District. Supervisor.	γ_5, α_{qd_5}

Position-based model: summary

Yandex

Web **San Francisco Travel**
[sanfrancisco.travel](#)
 San Francisco is home to a lot of everything. Whether you're a first time visitor or a long...

Images

More

San Francisco (/sæn frænˈsɪskoʊ/), officially the City and County of **San Francisco**, is the cultural, commercial, and financial center of Northern California and the only consolidated city-county in California.

♥ **San Francisco City Guide | Hotels, Restaurants, Nightlife, Real...**
[sanfrancisco.com](#)
 The job market may seem to navigate these days, but employment and career opportunities can be found in San Francisco's Financial District and Silicon Valley's...

City and County of San Francisco
[sf.gov.org](#)
 SFGov Visitors Key Services: SFG Travel Resources. ... Table of links to San Francisco districts and supervisors. District. Supervisor.

Examination depends on rank of doc d

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

Attractiveness depends on a query-document pair
Attractiveness is true relevance

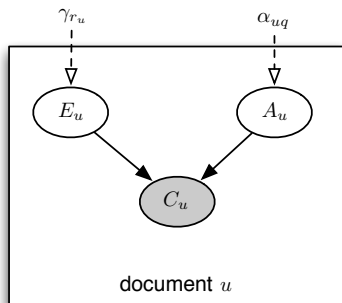
$$P(A_{qd} = 1) = \alpha_{qd}$$

$P(C_d = 1) = P(E_{rd} = 1) \cdot P(A_{qd} = 1)$

click is Biased

attractiveness is Unbiased

Position-based model: probabilistic graphical model



Position-based model

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

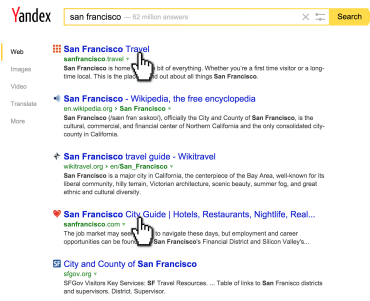
$$P(A_{qd} = 1) = \alpha_{qd}$$

$$P(C_d = 1) = P(E_{rd} = 1) \cdot P(A_{qd} = 1)$$

The screenshot shows a Yandex search engine interface. The search bar contains the text "san francisco" and indicates "62 million answers". Below the search bar, there are several search results categorized by type (Web, Images, Video, Translate, More). The first result is "San Francisco Travel" from sanfrancisco.travel, which includes a brief description of the city. The second result is "San Francisco - Wikipedia, the free encyclopedia" from en.wikipedia.org. The third result is "San Francisco travel guide - Wikitravel" from wikitravel.org. The fourth result is "San Francisco City Guide | Hotels, Restaurants, Nightlife, Real..." from sanfrancisco.com. The fifth result is "City and County of San Francisco" from sf.gov.org.

Cascade model

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



Cascade model

we don't have para which depends on ranked position

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

$$P(A_{d_r} = 1) = \alpha_{qd_r}$$

$$P(E_1 = 1) = 1$$

start from first

$$P(E_r = 1 \mid E_{r-1} = 0) = 0$$

examine one by one

$$P(E_r = 1 \mid C_{r-1} = 1) = 0$$

if click, then stop

$$P(E_r = 1 \mid E_{r-1} = 1, C_{r-1} = 0) = 1$$

otherwise, continue

Yandex san francisco — 62 million answers Search

Web **San Francisco Travel**
www.sanfrancisco.travel

Images
 Video
 Translate
 More

if we have not examined doc 3, then we will not examine doc 4

San Francisco travel guide - Wikitravel
wikitravel.org/en/San_Francisco
 San Francisco is a major city in California, the centerpiece of the Bay Area, well-known for its liberal community, hilly terrain, Victorian architecture, scenic beauty, summer fog, and great ethnic and cultural diversity.

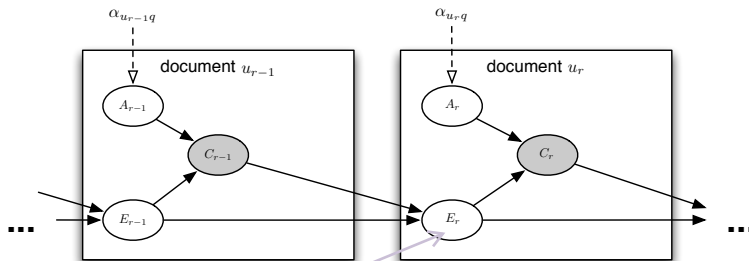
San Francisco City Guide | Hotels, Restaurants, Nightlife, Real...
sanfrancisco.com
 The job market may seem to navigate these days, but employment and career opportunities can be found in San Francisco's Financial District and Silicon Valley's...

City and County of San Francisco
sf.gov
 SF Gov Visitors Key Services: SF Travel Resources. ... Table of links to San Francisco districts and supervisors. District Supervisor

if we click doc 3, we don't exam doc 4

if we examined doc 3, and did not click doc 3, then we must examine doc 4

Cascade model: probabilistic graphical model



user will examine doc 3 only when:
the user has examined doc 2
and did not click doc 2

Basic click models summary

has bias

bias is removed
it is unbiased

- Position-based model (PBM)
 - + examination and attractiveness
 - examination of a document at rank r does not depend on examinations and clicks above r
- Cascade model (CM)
 - + cascade dependency of examination at r on examinations and clicks above r

比较：

position based model: click happen in the previous time step has NO influence on whether the user will click the rest of snippets. they are indep. it only depends on the rank of snippet. PBM has 2 para: attractiveness, rank.

cascade: clicks happen above(排序靠前的snippet) influences whether user will click in the following snippet (排序靠后的snippets). they are dependent. PBM has 1 para: attractiveness. Cascade does not have rank para r

Outline

- 1 Basic click models
- 2 Estimation
- 3 Applications

Parameter estimation

- Maximum likelihood estimation
- Expectation-maximization

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

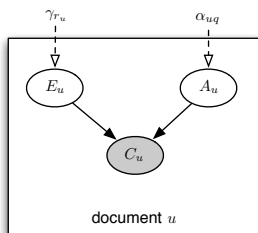
$$\begin{aligned}
 Q(\theta_c) &= \sum_{s \in \mathcal{S}} \mathbb{E}_{X|C^{(s)}, \Psi} \left[\log P(X, C^{(s)} | \Psi) \right] \\
 &= \sum_{s \in \mathcal{S}} \mathbb{E}_{X|C^{(s)}, \Psi} \left[\sum_{c_i \in s} \left(\mathcal{I}(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = p) \log(\theta_c) + \mathcal{I}(X_{c_i}^{(s)}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

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

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

$$\begin{aligned}
 \theta_c^{(t+1)} &= \frac{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} P(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = p | C^{(s)}, \Psi)}{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} \sum_{x=0}^1 P(X_{c_i}^{(s)} = x, \mathcal{P}(X_{c_i}^{(s)}) = p | C^{(s)}, \Psi)} \\
 &= \frac{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} P(X_{c_i}^{(s)} = 1, \mathcal{P}(X_{c_i}^{(s)}) = p | C^{(s)}, \Psi)}{\sum_{s \in \mathcal{S}} \sum_{c_i \in s} P(\mathcal{P}(X_{c_i}^{(s)}) = p | C^{(s)}, \Psi)} = \frac{ESS^{(t)}(1)}{ESS^{(t)}(1) + ESS^{(t)}(0)}
 \end{aligned}$$

EM for Position-Based Model



$$P(A_d = 1) = \alpha_{qd}$$

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

EM update rules for PBM: attractiveness

$$\alpha_{qd}^{(t+1)} = \frac{1}{|\mathcal{S}_{qd}|} \sum_{s \in \mathcal{S}_{qd}} \left(c_d^{(s)} + (1 - c_d^{(s)}) \frac{(1 - \gamma_r^{(t)}) \alpha_{qd}^{(t)}}{1 - \gamma_r^{(t)} \alpha_{qd}^{(t)}} \right)$$

- t – iteration
- \mathcal{S}_{qd} – search sessions initiated by query q and containing document u
- $c_d^{(s)}$ – observed click on document u in search session s

EM update rules for PBM: examination

$$\gamma_r^{(t+1)} = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \left(c_d^{(s)} + (1 - c_d^{(s)}) \frac{\gamma_r^{(t)} (1 - \alpha_{qd}^{(t)})}{1 - \gamma_r^{(t)} \alpha_{qd}^{(t)}} \right)$$

Outline

- 1 Basic click models
- 2 Estimation
- 3 Applications**

Click probabilities

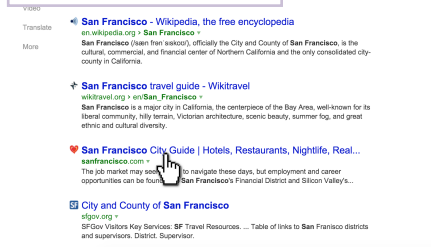
- Full probability – probability that a user clicks on a document at rank r

$$P(C_r = 1)$$

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

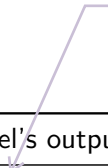
$$P(C_r = 1 \mid C_1, \dots, C_{r-1})$$

no matter what doc is located at position r , what is the prob that this position is clicked



Applications of click models

RBP and ERR belongs to this.
they are based on user based models



Click model's output	Application
Full click probabilities	Model-based metrics
Conditional click probabilities	User simulation
Parameter values	Ranking

Model-based metrics

- Utility-based metrics

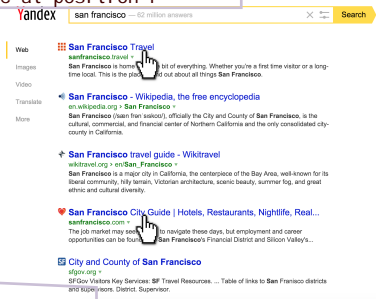
$$uMetric = \sum_{r=1}^n P(C_r = 1) \cdot U_r$$

U: utility=happiness of user, after click the doc at position r

- Effort-based metrics

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

efforts that user make to get satisfaction from the doc which rank at r position



Expected reciprocal rank

satisfaction = 1 or 0
= 1 when user is
satisfied with the doc,
after click it

r : rank of a doc

$$RR = \frac{1}{r}, \text{ where } S_r = 1$$

$$ERR = \sum_r \frac{1}{r} \cdot P(S_r = 1)$$

sum over all doc that
user is satisfied after
click them

Dynamic Bayesian network model (DBN)

$$P(A_r = 1) = \alpha_{qd_r}$$

$$P(E_1 = 1) = 1$$

$$P(E_r = 1 \mid S_{r-1} = 1) = 0$$

$$P(E_r = 1 \mid S_{r-1} = 0) = \gamma$$

$$P(S_r = 1 \mid C_r = 0) = 0$$

$$P(S_r = 1 \mid C_r = 1) = \sigma_{qd_r}$$

$$P(S_r = 1) = ?$$

DBN has 2
satisfaction para:
alpha: satisfaction
after reading the
snippet

sigma: satisfaction
after click a snippet

Video

Translate

More

San Francisco - Wikipedia, the free encyclopedia

[en.wikipedia.org](https://en.wikipedia.org/wiki/San_Francisco) > [San Francisco](#) >

San Francisco (/sænˈfrænsikoʊ/), officially the City and County of **San Francisco**, is the cultural, commercial, and financial center of Northern California and the only consolidated city-county in California.

San Francisco travel guide - Wikitravel

[wikitravel.org](https://wikitravel.org/en/San_Francisco) > [en:San_Francisco](#) >

San Francisco is a major city in California, the centerpiece of the Bay Area, well-known for its liberal community, hilly terrain, Victorian architecture, scenic beauty, summer fog, and great ethnic and cultural diversity.

San Francisco City Guide | Hotels, Restaurants, Nightlife, Real...

sanfrancisco.com >

The job market may seem to navigate these days, but employment and career opportunities can be found in San Francisco's Financial District and Silicon Valley's...

City and County of San Francisco

sf.gov >

SFGov Visitors Key Services: SFG Travel Resources... Table of links to San Francisco districts and supervisors. District Supervisor.

Search

other you're a first time visitor or a long-time San Francisco.

DBN: Satisfaction

$$\begin{aligned}P(S_r = 1) &= P(S_r = 1 \mid C_r = 1) \cdot P(C_r = 1) \\&= \sigma_{qd_r} \cdot P(C_r = 1) \\&= \sigma_{qd_r} \cdot \alpha_{qd_r} \cdot P(E_r = 1) \\&= \sigma_{qd_r} \cdot \alpha_{qd_r} \cdot \prod_{i=1}^{r-1} (\gamma \cdot (1 - \sigma_{qd_i} \cdot \alpha_{qd_i})) \\&= R_{qd_r} \cdot \prod_{i=1}^{r-1} (\gamma \cdot (1 - R_{qd_i}))\end{aligned}$$

Expected reciprocal rank

$$\begin{aligned} ERR &= \sum_r \frac{1}{r} \cdot P(S_r = 1) \\ &= \sum_r \frac{1}{r} \cdot R_{qd_r} \cdot \prod_{i=1}^{r-1} (\gamma \cdot (1 - R_{qd_i})) \end{aligned}$$

Summary

- Interactions
 - Examples of interactions
 - Applications: evaluate and improve IR
- Interaction models
 - Basic click models: PBM, CM
 - Applications: ERR

Materials

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