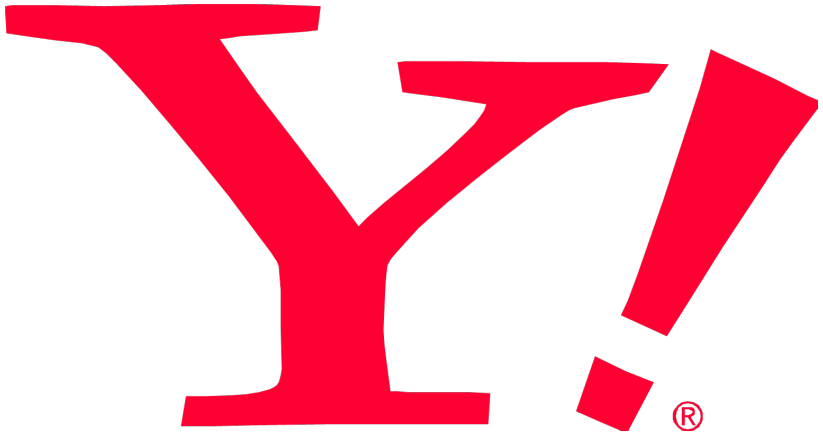


Filling Context-Ad Vocabulary Gaps with Click Logs

2014/9/2

Yahoo Japan Corporation
Yukihiro Tagami, Toru Hotta, Yusuke Tanaka,
Shingo Ono, Koji Tsukamoto, Akira Tajima

First of all…

- Our paper pdf in the proceedings seems to be broken...
- Please refer to the followings and get the pdf
  - http://dl.acm.org/citation.cfm?id=2623334
  - http://research-lab.yahoo.co.jp/ml/20140701_81.html

Summary

- For contextual advertising, we propose a method of translating contextual information into the textual features of ads by using past click data
  - In this study, we focus on increasing the click-through rate(CTR)
- We applied our approach to a real ad serving system and achieved an improvement over the existing production system

Background

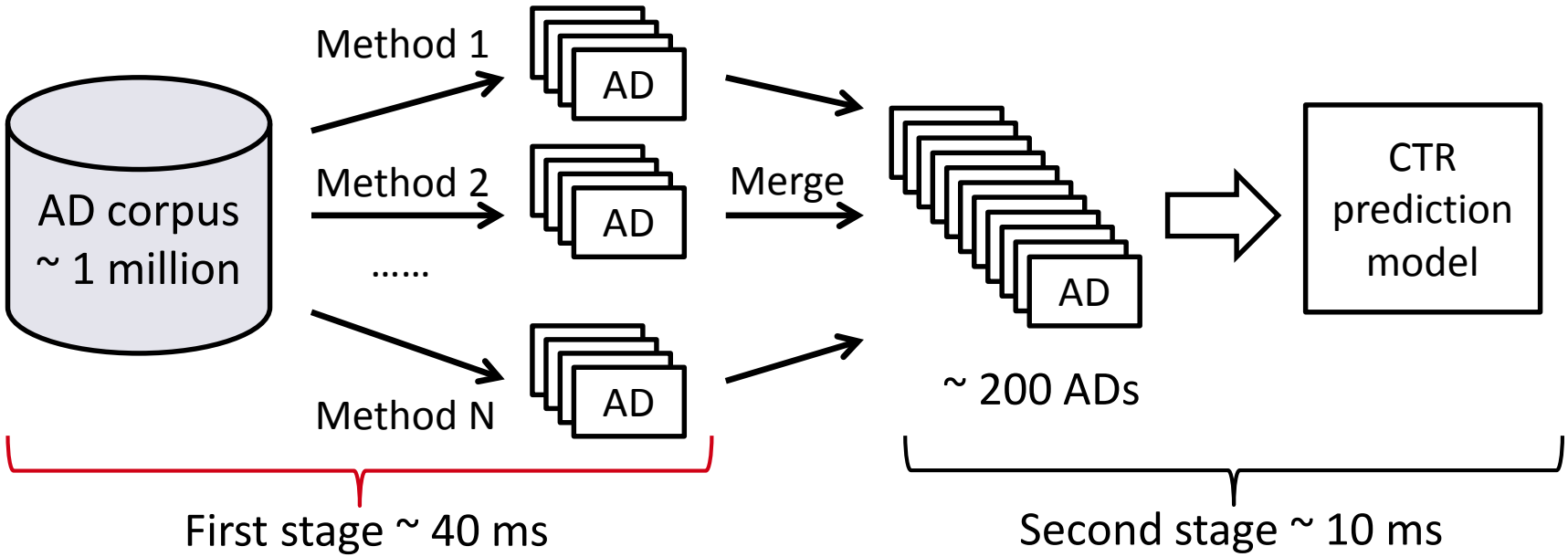
- Contextual advertising for Japanese market
  - Developed by Yahoo! JAPAN
  - Annual revenue: 30 billion JPY (300 million USD)

Example on a Web page of trip for Hawaii (ハワイ)

ハワイ絶品グルメ
travel.yahoo.co.jp
ヘルシー志向のオーガニックグルメ♪自家農園で育てた自然の恵みを堪能
ハワイの最旬旅行情報
travel.yahoo.co.jp
アロハスピリッツを体感する旅へ！パワースポットでリラックス体験
ハワイをもっと知ろう！
travel.yahoo.co.jp
ハワイ最大の博物館で歴史・文化を学ぼう／ビショップ博物館（オアフ島）
ハワイのおすすめホテル
travel.yahoo.co.jp
お気に入りが見つかったら早速予約！ホテル指定のツアーもおすすめ
フラに触れハワイを知る
travel.yahoo.co.jp
優雅な踊りに込められた神への祈りやハワイの歴史を体験しよう

Two-stage approach

- Two-stage approach in our ad serving system
  1. Ads are retrieved by multiple methods using inverted index
  2. The ads are merged and passed to click-through rate(CTR) prediction model

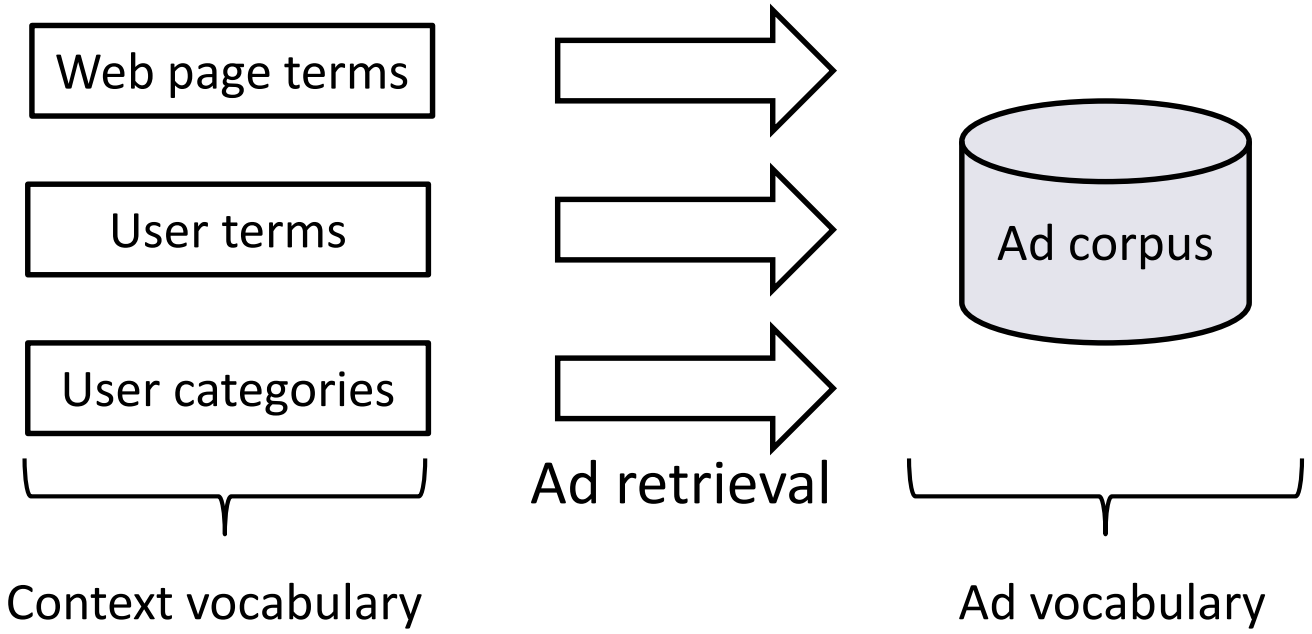


Objective of selecting ads

- Ad serving system selects ads that are relevant to the page content and/or user information
  - Advertisers are primarily interested in targeting relevant users
  - Publishers, which is a Web page owner, are concerned with keeping the user experience pleasant

Typical approach (1/2)

- To measure the word overlap between two types of information
  - In practice, these vocabularies are different
    → context-ad vocabulary gaps

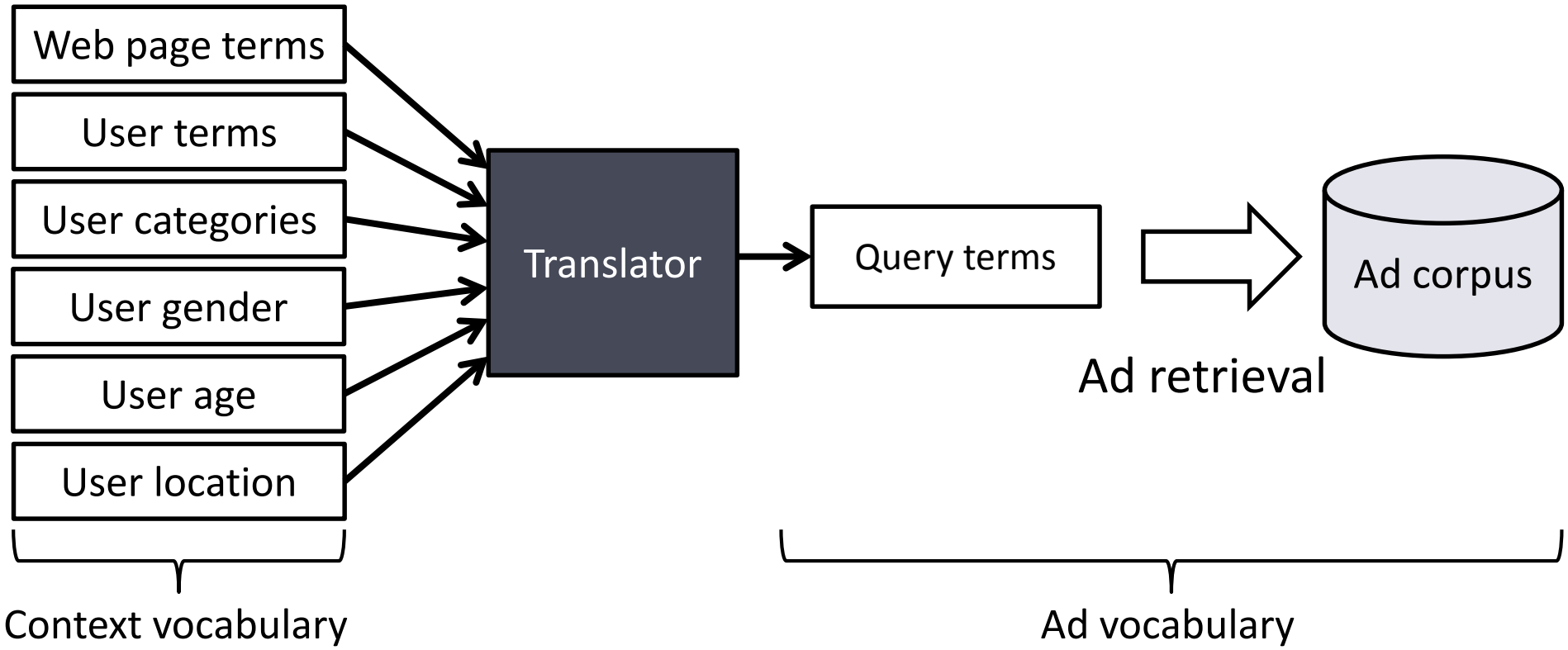


Typical approach (2/2)

- To use semantic taxonomy or hidden classes
  - It is necessary to expand the ad retrieval system or build new index to handle it
  - In the operation, the following tasks are periodically required
    - A review of the number and hierarchical structure of categories
    - A re-creation of clusters
  - In practice, these are not always easy to perform

Proposed approach

- Our approach translates contextual information into the textual space of ads
- The translation table is learned with click logs



A score proportional to CTR

$$score(\boldsymbol{q}, \boldsymbol{a}) = \boldsymbol{q}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{a} + \boldsymbol{w}_{basic}^{\mathrm{T}} \boldsymbol{x}_{basic}$$

$$= \sum_{i=1}^{D_q} \sum_{j=1}^{D_a} w_{ij} q_i a_j + \boldsymbol{w}_{basic}^{\mathrm{T}} \boldsymbol{x}_{basic}$$

$\boldsymbol{q} \in \mathbb{R}^{D_q}$: contextual feature vector

$\boldsymbol{a} \in \mathbb{R}^{D_q}$: ad feature vector

$\boldsymbol{x}_{basic}$: features such as ad's display position

} features

$\boldsymbol{w}_{basic}$: weight vector corresponding to $\boldsymbol{x}_{basic}$

$\boldsymbol{W}$: translation matrix

} parameters

Feature selection

- Feature selection using past click statistics
  - To learn the translation matrix efficiently
  - To make the transformed vector sparse
- The score is represented linear form

$$score(\boldsymbol{q}, \boldsymbol{a}) = \sum_{(i,j) \in P} w_{ij} q_i a_j + \boldsymbol{w}_{basic}^{\mathrm{T}} \boldsymbol{x}_{basic}$$

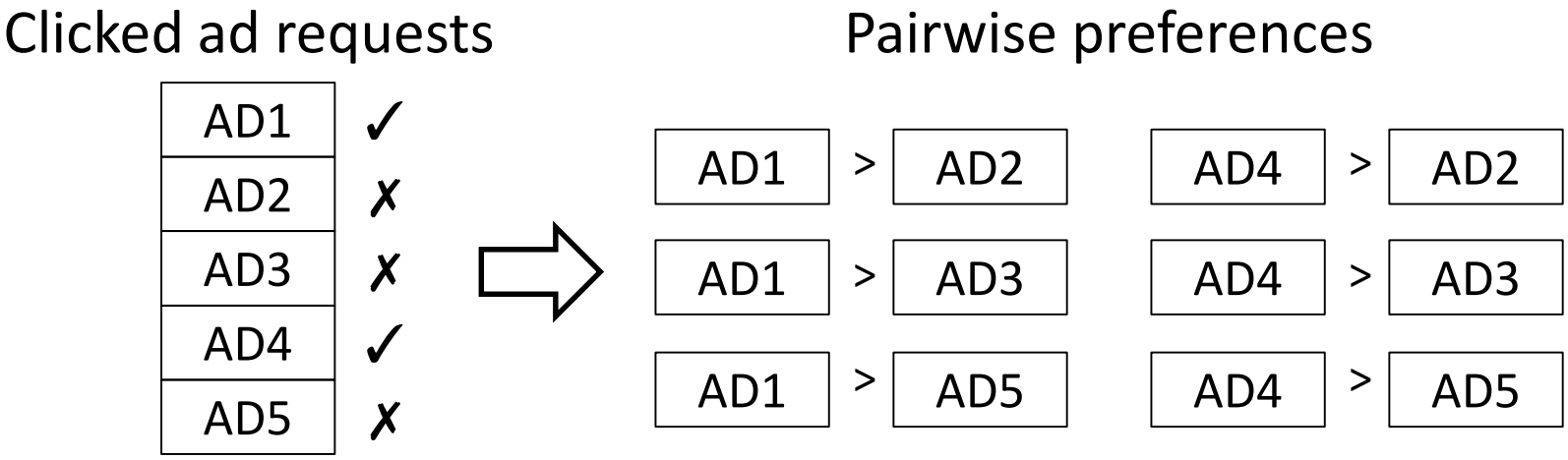
$$= \boldsymbol{w}_{match}^{\mathrm{T}} \boldsymbol{x}_{match} + \boldsymbol{w}_{basic}^{\mathrm{T}} \boldsymbol{x}_{basic}$$

$$= \boldsymbol{w}^{\mathrm{T}} \boldsymbol{x}$$

$$P = \{(i,j) \mid m_{ij} > T\}, \quad m_{ij} = \frac{ctr(q_i, a_j)}{\max(ctr(q_i), ctr(a_j))}$$

Pairwise preferences

- In our setting, some ads are displayed on a page at the same time
- User clicks the ad because it is preferable in the listing



Parameter estimation

- Defining a pairwise loss function
  - RankSVM (squared hinge loss)

$$L(\boldsymbol{w}) = \sum_{r \in R^+} \sum_{i:y_i^{(r)}=1} \sum_{j:y_j^{(r)}=0} \max(0, 1 - \boldsymbol{w}^{\mathrm{T}}(\boldsymbol{x}_i^{(r)} - \boldsymbol{x}_j^{(r)}))^2$$

R^+: a set of ad requests which includes clicked ad

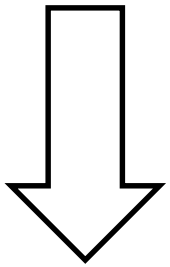
- We add a regularization term to avoid overfitting and seek the parameters

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \frac{1}{2}\|\boldsymbol{w}\|_2^2 + C \cdot L(\boldsymbol{w})$$

Ad retrieval

- With the learned matrix, the contextual feature vector is transformed into the input term vector of the ad retrieval system for each ad request

$$\boldsymbol{q}_{input} = \boldsymbol{q}^{\mathrm{T}} \hat{\boldsymbol{W}}$$



Top-*k* similarity search using inverted index (WAND algorithm)

$$(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N)$$

Online A/B testing result

- Relative gain over existing method
  - Click-through rate(CTR)
    - Colored value represents “p-value < 0.05” on chi-squared test (website A, B, D, E, H)
  - Cost per click(CPC)
  - Revenue per request(RPR)

Metrics	Website							
	A	B	C	D	E	F	G	H
CTR	-3.67%	+4.60%	+0.48%	+2.82%	+2.47%	+1.42%	+3.27%	+4.02%
CPC	+3.63%	-2.00%	+1.62%	+1.31%	-1.01%	+7.51%	-2.42%	-2.94%
RPR	-0.18%	+2.51%	+2.10%	+4.17%	+1.44%	+9.04%	+0.77%	+0.97%

Example of translation table

- Example of Web site B's table for user terms
  - "User terms" are extracted from user behavioral events

User term	Translated term	Weight
iPhone	iPhone	0.2114
	ケース(case)	0.1534
	iPad	0.0868
プリウス(Toyota Prius)	プリウス(Toyota Prius)	0.2600
	燃費(mileage)	0.0732
	HV(Hybrid Vehicle)	0.0607
温泉(hot spring)	温泉(hot spring)	0.1730
	旅館(Japanese inn)	0.1272
	露天風呂(outdoor hot spring)	0.0809

Summary

- For contextual advertising, we propose a method of translating contextual information into the textual features of ads by using past click data
  - In this study, we focus on increasing the click-through rate(CTR)
- We applied our approach to a real ad serving system and achieved an improvement over the existing production system

Questions?

Please speak clearly and slowly.

Yukihiro Tagami
yutagami@yahoo-corp.jp