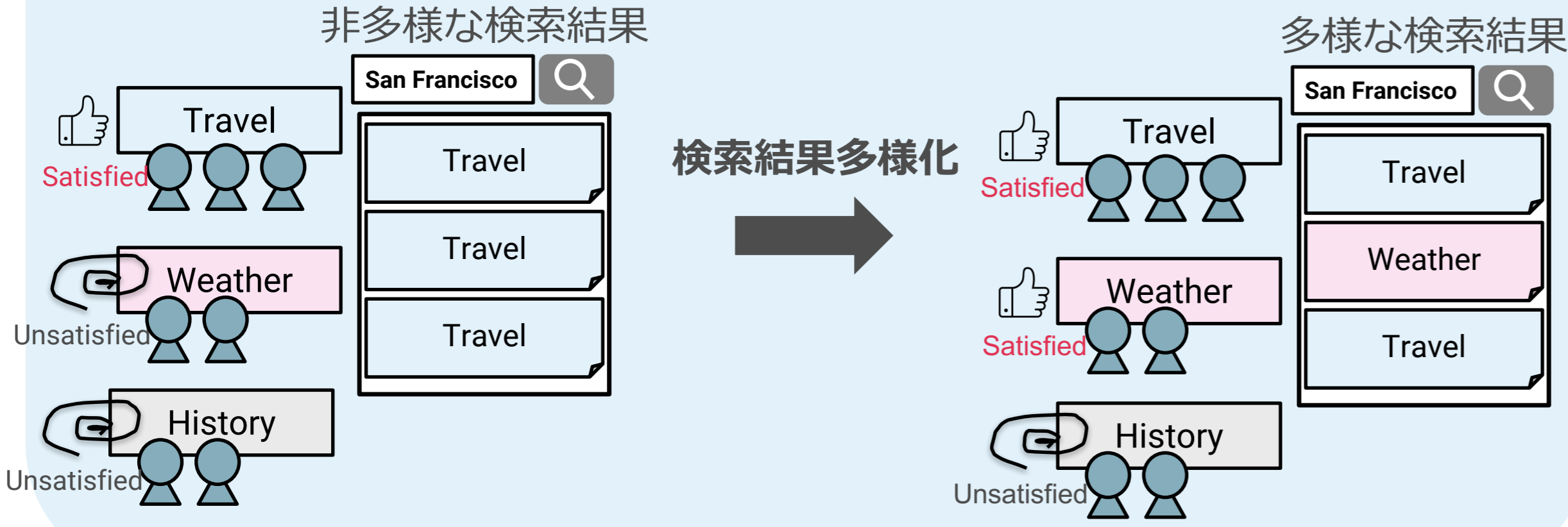


# NTCIR-9 INTENT-1 利用 曖昧なクエリに対する NTCIR-10 INTENT-2 利用 クエリ推薦と検索結果提示の最適化

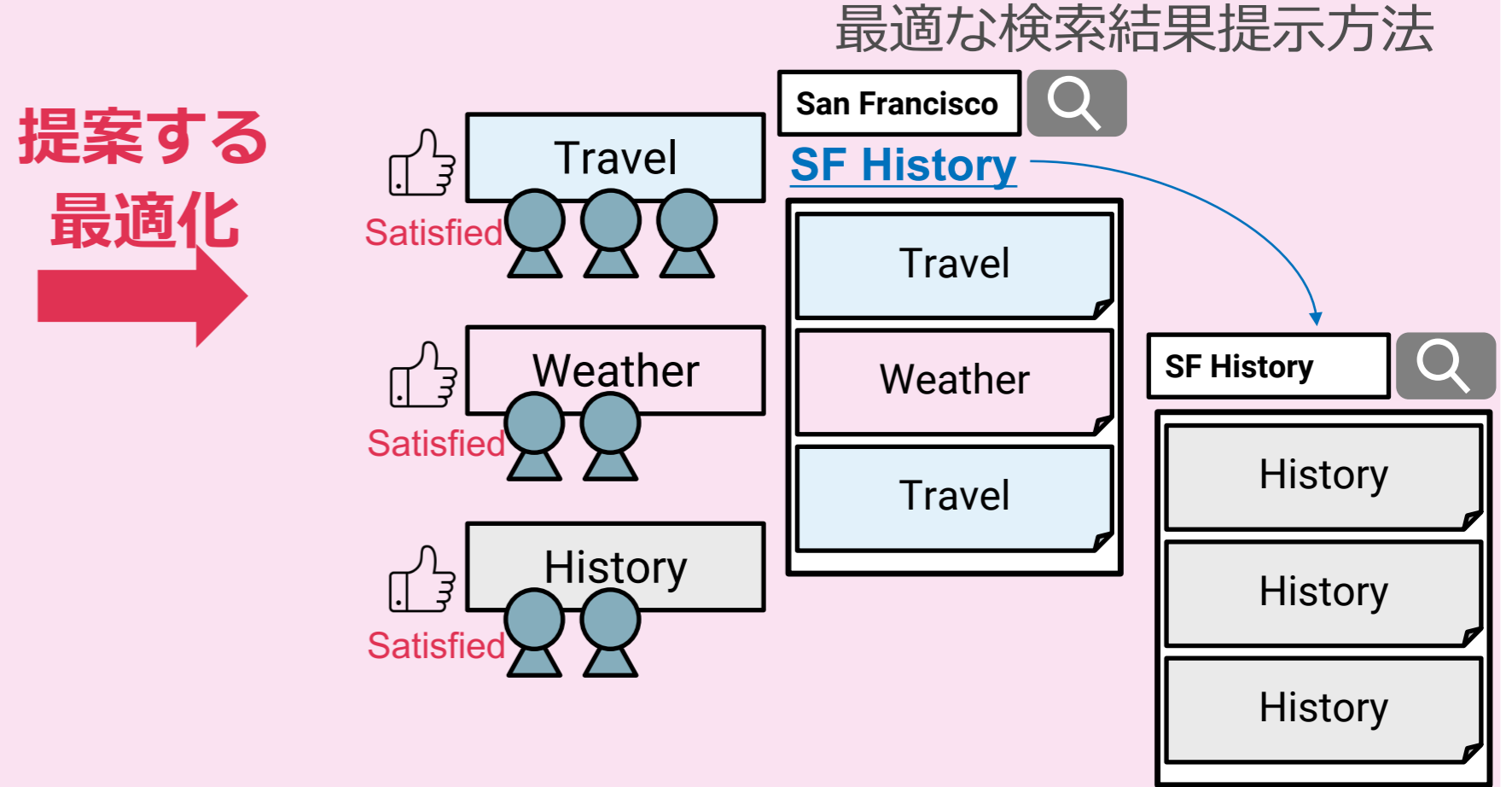
加藤 誠, 田中克己 (京都大学) <http://www.mpkato.net/>

Web検索クエリはしばしば曖昧・不明瞭。  
そのため、ユーザの多様な検索意図にもれなく対応できるよう  
検索結果は多様化されている必要がある



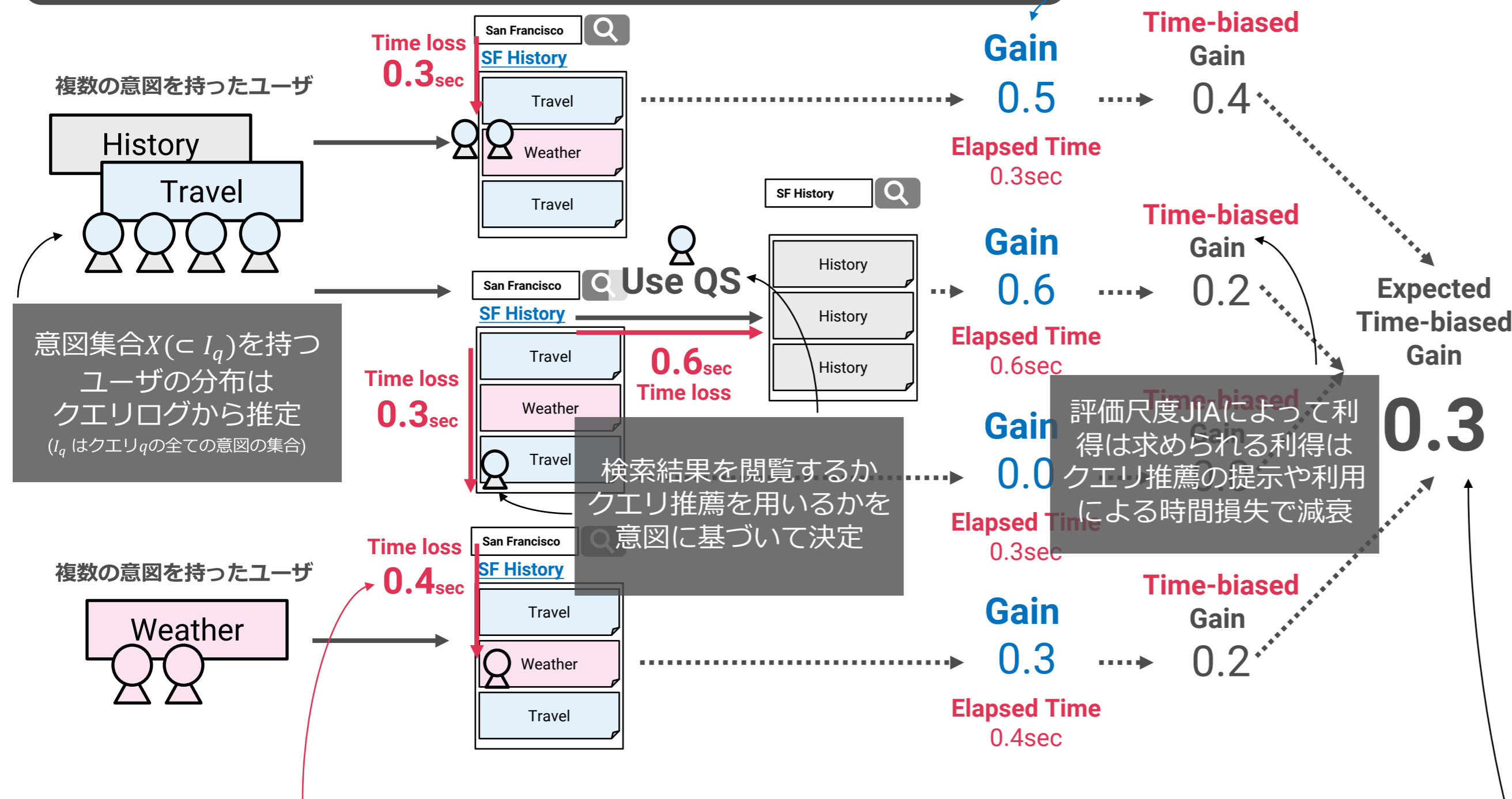
詳細はこちら:

Makoto P. Kato, Katsumi Tanaka (2016) To Suggest, or Not to Suggest for Queries with Diverse Intents: Optimizing Search Result Presentation, WSDM2016, pp. 133-142.



## 3つの新しいアイデアに基づき 最適な検索結果提示方法 (文書ランキング+クエリ推薦) を提案

### アイデア 1: クエリ推薦を利用するユーザの確率的モデル



### アイデア 3: 同時意図を考慮した評価指標

Agrawalらの意図を考慮した評価指標

仮定 ユーザは1つの意図を持つ

$$IA(d) = \sum_{i \in I_q} P(i|q) M_i(d)$$

e.g. DCG or ERR

- $I_q$ : クエリ $q$ の意図集合
- $P(i|q)$ : ユーザがクエリ $q$ の意図 $i$ を持つ確率
- $M_i$ : 意図 $i$ を考慮したある評価指標の値
- $d$ : 文書列

同時意図を考慮した評価指標

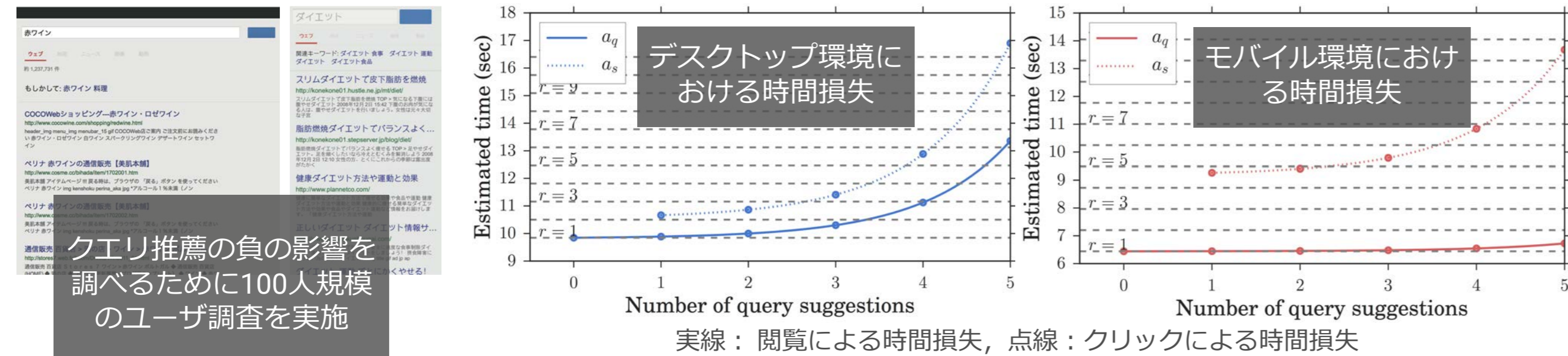
仮定 ユーザは複数の意図を持つ

$$JIA(d) = \sum_{X \in 2^{I_q}} P(X|q) M_X(d)$$

- $P(X|q)$ : ユーザがクエリ $q$ の意図集合 $X$ を持つ確率
- もし $M_X$ が $M_i$ の(重み付き)和で表されるならJIAを効率的に計算できることを証明
- 時間バイアスを評価指標に含めることを提案

### アイデア 2: クエリ推薦提示・利用による時間損失の計測

クエリ推薦を提示することで  
それらを読覧・クリックすることによる時間損失が発生



- 多くのクエリ推薦が提示されると多くの時間損失が発生
- モバイル環境では閲覧による損失は小、クリックによる損失は大

## 検索結果提示の最適化

時間バイアス利得の期待値を  
最大化するクエリ推薦を貪欲に追加

クエリ推薦が1つも提示されない場合より  
期待値が高くなることを保証

## 実験

NTCIR-9 INTENT-1 および  
NTCIR-10 INTENT-2 で構築された  
テストコレクションを利用  
各クエリに意図が付与されClueWeb09-JAの文書  
に意図ごとの適合性判定がされている

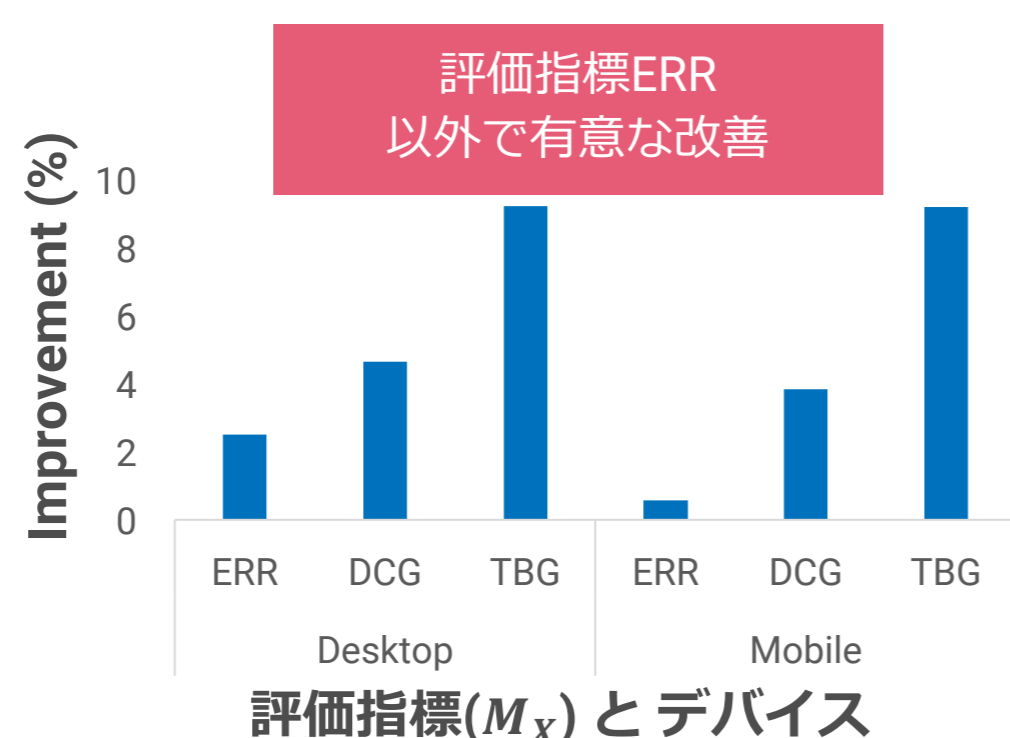
Baseline: クエリ推薦を含めない検索結果最適化

RQ1での設定:  
正解の意図確率と意図ごとの適合性を最適化に  
利用

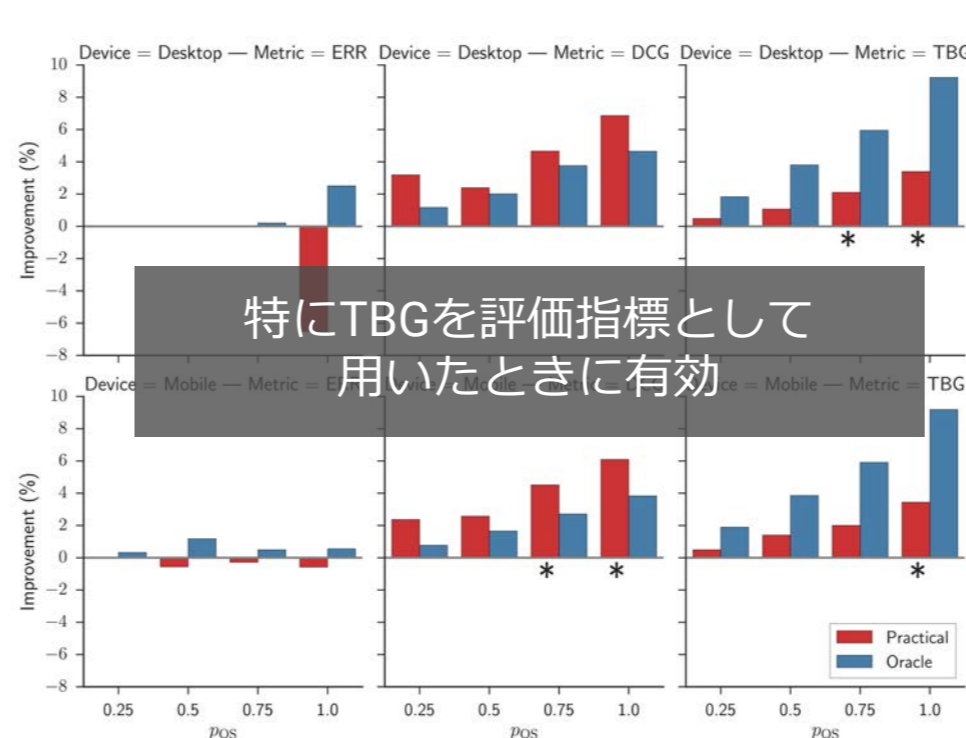
RQ2での設定:  
意図確率と適合性を推定  
パラメータ:

$p_{QS}$ : ユーザが自信の持つ意図に適合するクエリ  
推薦を利用する確率

RQ1. どのシナリオで有意な改善が可能か?



RQ2. 実用的な設定でも改善がされるか?



RQ3. どのようなクエリが恩恵を受けるか?

