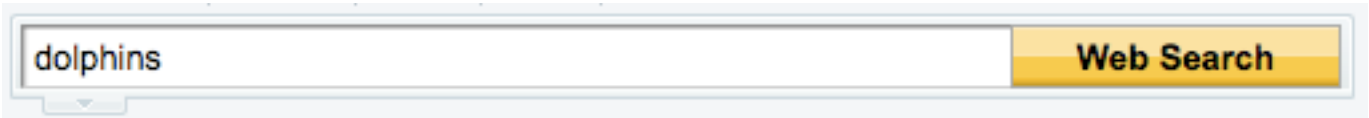
Search Result Diversity for Informational Queries

Michael Welch, Junghoo Cho, Christopher Olston
mjwelch@yahoo-inc.com, cho@cs.ucla.edu, olston@yahoo-inc.com

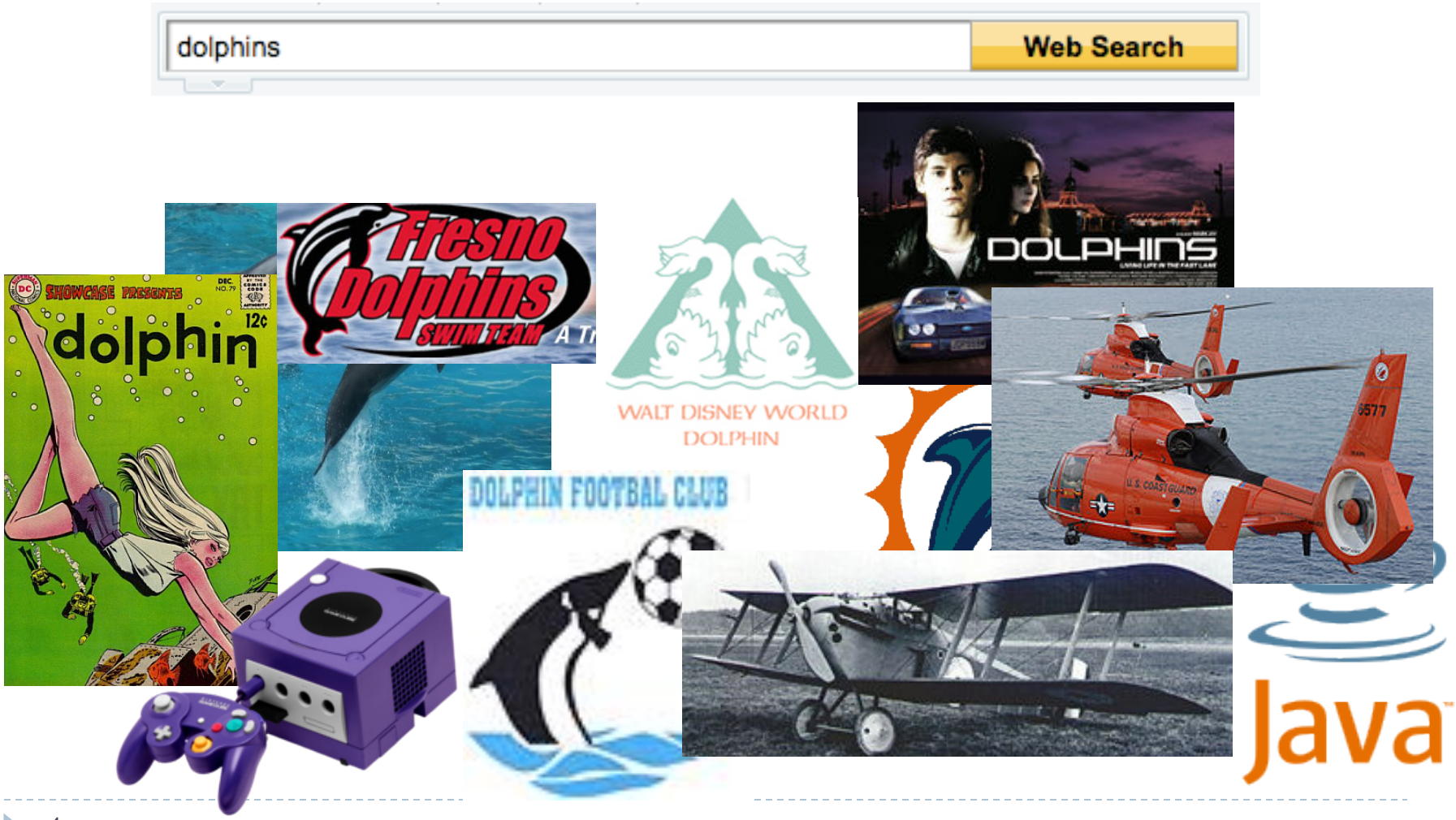
Example



Example

dolphins Web Search

Example


dolphins
Web Search
SHOWCASE PRESENTS
dolphin
12¢
Fresno Dolphins
SWIM TEAM
WALT DISNEY WORLD
DOLPHIN
DOLPHINS
DOLPHIN FOOTBAL CLUB
Java

Also try: miami **dolphins**, pictures of **dolphins**, More...

Dolphins - Image Results



More **dolphins** images

Yahoo! Shortcut - About

Dolphin - Wikipedia, the free encyclopedia

Origin of the name | Taxonomy | Evolution and anatomy | Behaviour

Dolphins are marine mammals that are closely related to whales and porpoises. There are almost forty species of **dolphin** in seventeen genera. They vary in size from 1.2 m (4 ft) and 40 kg (90 lb), up to 9.5 m (30 ft) and 10 tonnes . They are found worldwide, mostly in...

en.wikipedia.org/wiki/Dolphin - 125k - Cached

Swim with the **dolphins** at Dolphin Research Center Marathon FL, Dolphin ...

A Florida nonprofit education and research facility, home to a family of Atlantic Bottlenose **Dolphins** and California Sea Lions. Offers educational programs that ...

www.dolphins.org - Cached

DOLPHINS

All **dolphins** are toothed whales belonging to the sub-order, odontocetes, of the ... In addition, although the terms **dolphins** and porpoises are often used ...

www.earthtrust.org/wlcurric/**dolphins**.html - Cached

Miami **Dolphins**

Official site of the Miami **Dolphins**. Includes schedule, news, multimedia, photos, player information, statistics, team store, tickets, and more.

www.miamidolphins.com - 1289k

MiamiDolphins.com - Official Website of the Miami **Dolphins**

www.miamidolphins.com/newsite/index.asp - Cached

Dolphins and Man.Equals?

Just how intelligent are **dolphins**? Can humans understand dolphin intelligence? ... Apparently there is something quite impressive about **Dolphins**. ...

www.littletownmart.com/dolphins - Cached

Bottlenose Dolphin - Wikipedia

Description | Taxonomy | Behavior | Intelligence

Bottlenose **dolphins**, the genus Tursiops, are the most common and well-known members of the family Delphinidae, the family of oceanic **dolphins**. Recent molecular studies show the genus contains two species, the Common...

en.wikipedia.org/wiki/Bottlenose_Dolphin - 266k - Cached

(Lack of) Diversity in Results

- **In the top 10 results from a search engine:**
  - 8 are about the mammal
  - 1 is for the NFL team (rank 5)
  - 1 is for an IMAX movie about the mammals (rank 8)
- **What about the other interpretations?**
  - Users interested in them will be dissatisfied

Motivational Questions

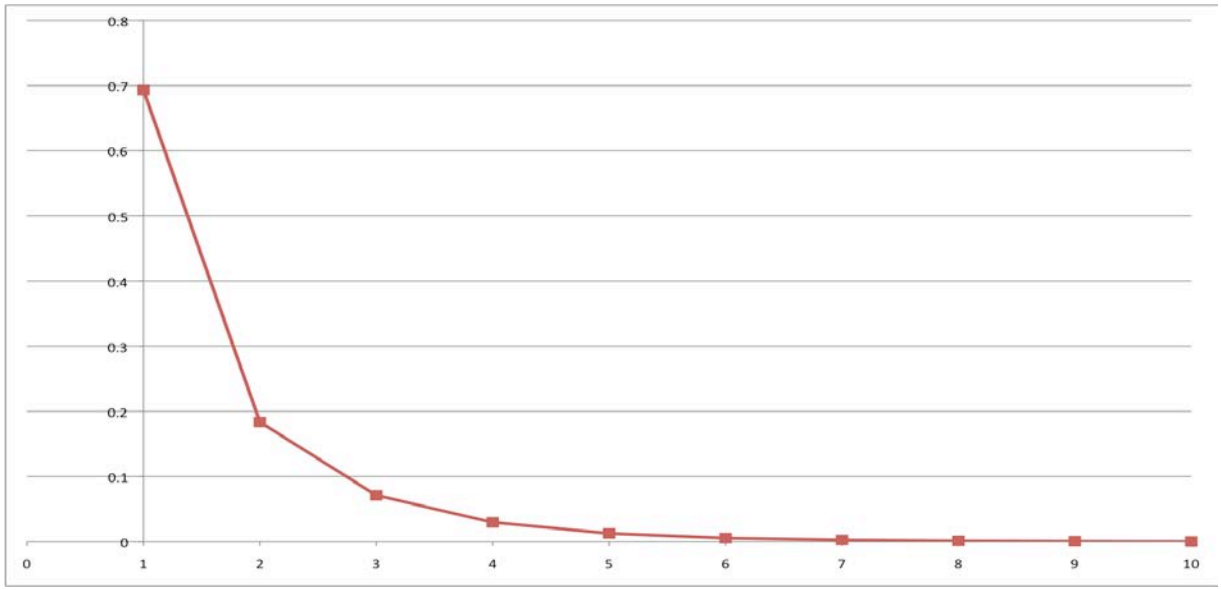
- **How many relevant results do users want?**
  - **Did we need to show 8 pages about the mammal?**
  - **Is one page enough? Two pages? Three?**
- **Are ambiguous queries really a problem?**
  - **16% of Web queries are ambiguous [Song '09]**
- **Can we better allocate the top *n* results to cover a more diverse set of subtopics?**
  - **While maintaining user satisfaction for the common subtopics**

A Quick Survey of Related Work

- **Personalized search**
  - User profiles and page taxonomies
  - [Pretschner '99, Liu '02]
- **Content based approaches**
  - Tradeoffs between relevancy, novelty, and risk
  - [Carbonell '98], [Zhai '03], [Chen '06], [Wang '09]
- **Hybrid approaches**
  - Use probabilistic measures of user intent and document classification for a set of subtopics
  - [Agrawal '09]

Is One Relevant Document Enough?

- **Most existing work assumes a *single* relevant document is sufficient**
- **Informational queries typically result in multiple clicks [Lee '05]**



Our Model for Ambiguous Queries

- User queries for topic *T* with subtopics $T_1 \ldots T_m$
- User has some number of pages *J* that they want to see for their subtopic
  - Click on *J* relevant pages if they are available
  - Clicks on fewer if less than *J* pages are relevant
- User *U* wants *J* relevant pages with *Pr(J|U)*

Our Model (cont.)

- **Probabilistic user intent in subtopics**
  - Most users interested in a single subtopic
  - User *U* interested in subtopic T_i with $Pr(T_i|U)$
- **Probabilistic document categorization**
  - Most documents belong to a single subtopic
  - Document *D* belongs to subtopic T_i with $Pr(T_i|D)$

Measuring User Satisfaction

- **How do we evaluate user satisfaction?**
  - **"Happy or not" isn't an adequate model**
  - **Measure the expected number of *hits***
  - **Hit: expected click on a relevant document**
- **Model the expected user satisfaction with a returned set of documents**
  - **Optimize document selection for that model**

Perfect Document Classification

- **Assume we know the correct subtopic for each document**

$$E(R) = \sum_{j=1}^{n} \sum_{i=1}^{m} \Pr(T_i|U) \Pr(J = j|U) \min(j, K_i)$$

- **R: a set of *n* documents**
- **User is shown K_i pages from subtopic T_i**
- **How many pages K_i should we show from each subtopic T_i?**

Choosing Optimal K_i Values

- **Selecting *n* documents from *m* topics:** $\binom{n+m-1}{n}$
- **Lemma (proof given in paper)**
  - **Label subtopics $T_1 \ldots T_m$ such that $Pr(T_1|U) \geq Pr(T_2|U) \geq \ldots Pr(T_m|U)$**
  - **Optimal solution has property $K_1 \geq K_2 \geq \ldots K_m$**
- **Can use this property to create ordering of documents in a greedy fashion**

KnownClassification Algorithm

- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$

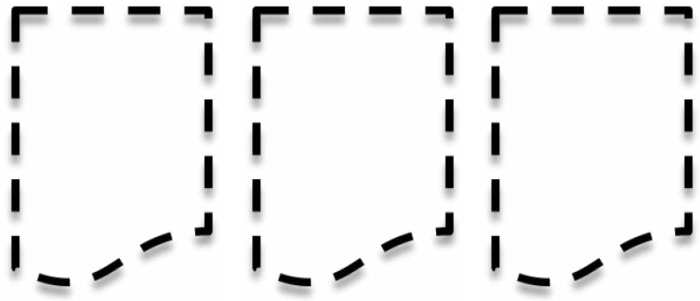
KnownClassification Algorithm

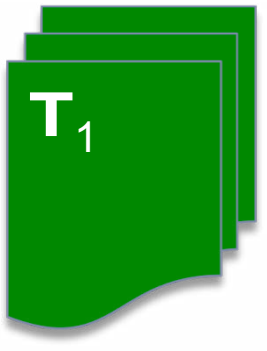
- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$
- $K_1 = 0$, $K_2 = 0$

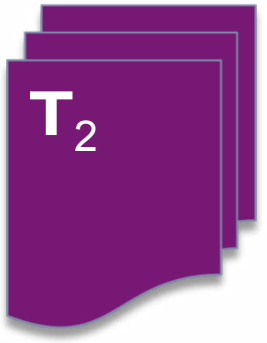
$$\Delta E(T_1) = \sum_{j=1}^{n} \Pr(T_1 \mid U)\Pr(J = j \mid U)\min(j, K_1) = 0.7 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.7$$

$$\Delta E(T_2) = \sum_{j=1}^{n} \Pr(T_2 \mid U)\Pr(J = j \mid U)\min(j, K_2) = 0.3 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.3$$

R =







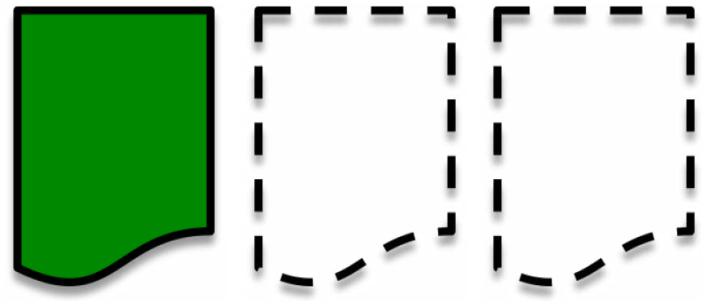
KnownClassification Algorithm

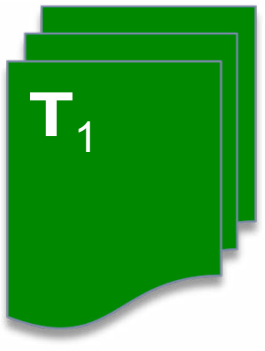
- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$
- $K_1 = 1$, $K_2 = 0$

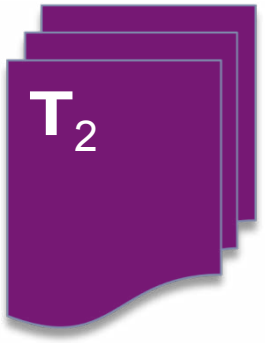
$$\Delta E(T_1) = \sum_{j=1}^{n} \Pr(T_1 \mid U)\Pr(J = j \mid U)\min(j, K_1) = 0.7 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.7$$

$$\Delta E(T_2) = \sum_{j=1}^{n} \Pr(T_2 \mid U)\Pr(J = j \mid U)\min(j, K_2) = 0.3 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.3$$

R =





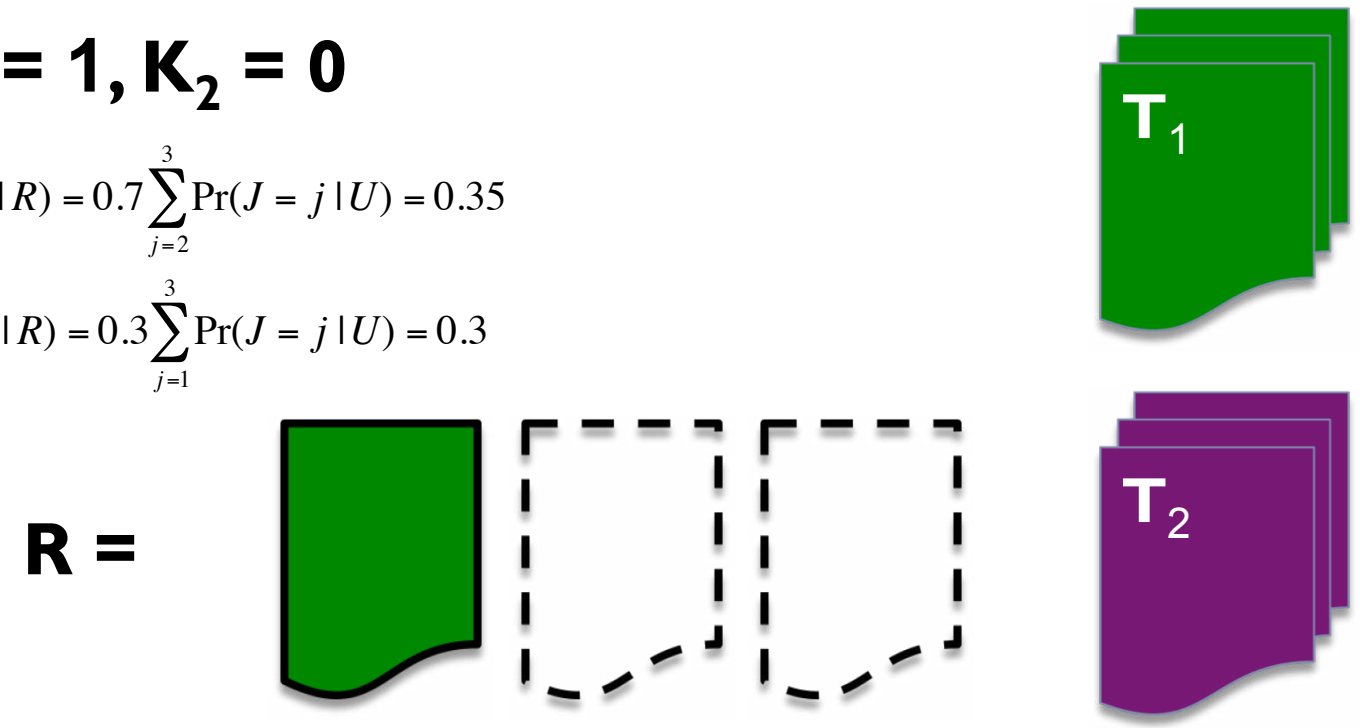


KnownClassification Algorithm

- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$
- $K_1 = 1, K_2 = 0$

$$\Delta E(T_1 \mid R) = 0.7 \sum_{j=2}^{3} \Pr(J = j \mid U) = 0.35$$

$$\Delta E(T_2 \mid R) = 0.3 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.3$$

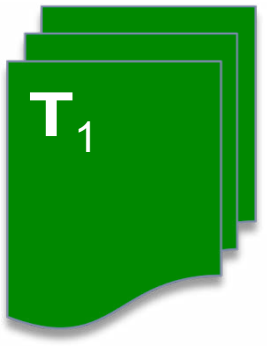


KnownClassification Algorithm

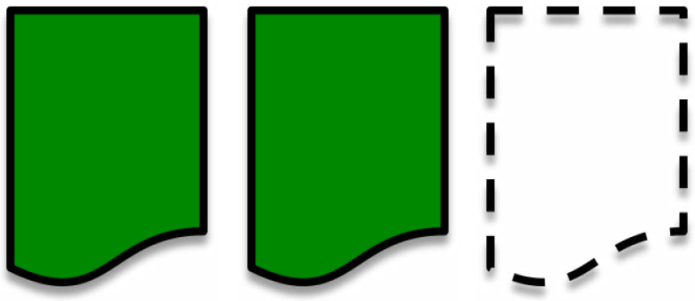
- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$
- $K_1 = 2, K_2 = 0$

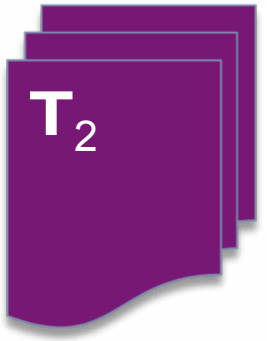
$$\Delta E(T_1 \mid R) = 0.7 \sum_{j=2}^{3} \Pr(J = j \mid U) = 0.35$$

$$\Delta E(T_2 \mid R) = 0.3 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.3$$



R =



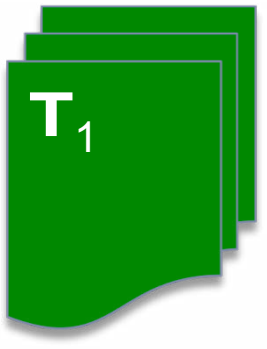


KnownClassification Algorithm

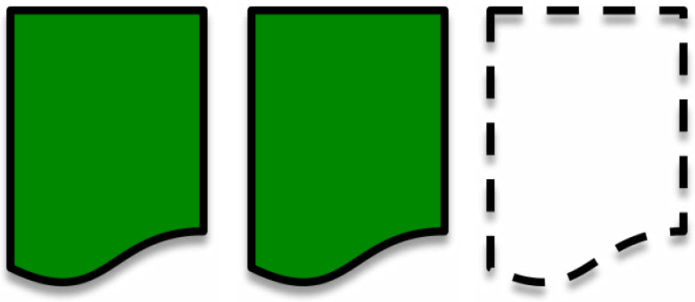
- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$
- $K_1 = 2, K_2 = 0$

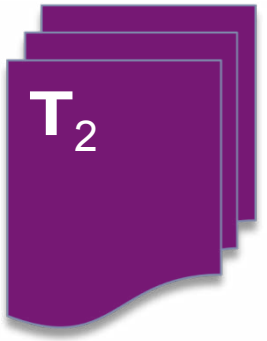
$$\Delta E(T_1 \mid R) = 0.7 \sum_{j=3}^{3} \Pr(J = j \mid U) = 0.07$$

$$\Delta E(T_2 \mid R) = 0.3 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.3$$



R =



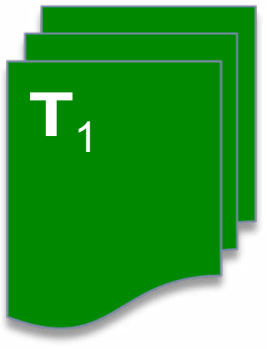


KnownClassification Algorithm

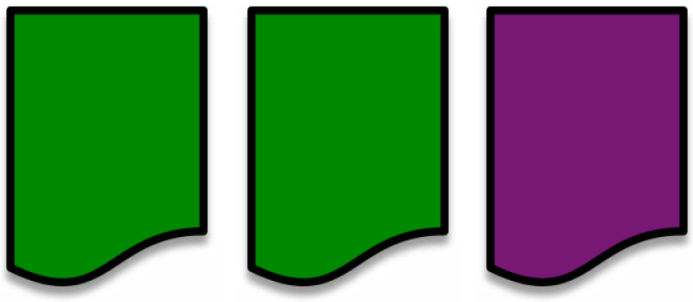
- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$
- $K_1 = 2, K_2 = 1$

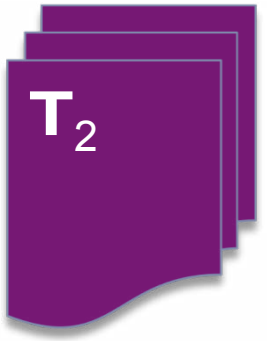
$$\Delta E(T_1 \mid R) = 0.7 \sum_{j=3}^{3} \Pr(J = j \mid U) = 0.07$$

$$\Delta E(T_2 \mid R) = 0.3 \sum_{j=1}^{3} \Pr(J = j \mid U) = 0.3$$



R =



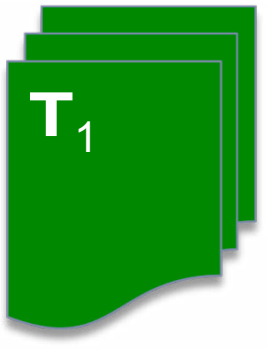


KnownClassification Algorithm

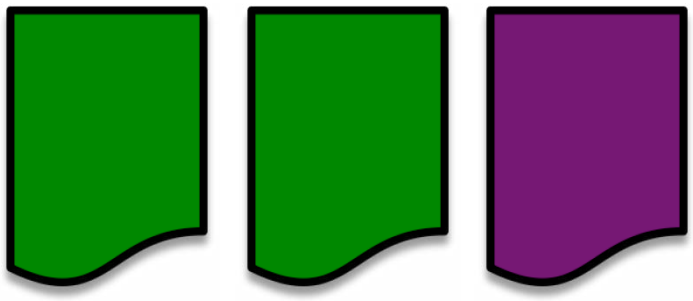
- $Pr(T_1|U) = 0.7$ and $Pr(T_2|U) = 0.3$
- $Pr(J=1|U) = 0.5$, $Pr(J=2|U) = 0.4$, $Pr(J=3|U) = 0.1$
- $n = 3$
- $K_1 = 2, K_2 = 1$

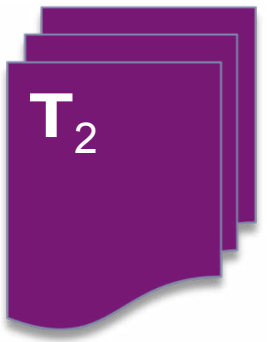
$$\Delta E(T_1 \mid R) = 0.7 \sum_{j=3}^{3} \Pr(J = j \mid U) = 0.07$$

$$\Delta E(T_2 \mid R) = 0.3 \sum_{j=2}^{3} \Pr(J = j \mid U) = 0.15$$



R =





Diversity-IQ Algorithm

- **Given all three probability distributions, we define the expected hits as:**

$$E(R) = \sum_{j=1}^{n} \sum_{i=1}^{m} \Pr(T_i|U) \Pr(J = j|U) \sum_{k=1}^{n} \Pr(K_i = k|R) \min(j, k)$$

- **Algorithm follows a similar greedy approach**
- **K_i values are now probabilistic**
  - **ΔE computation is now O(|R| • n • m) = O(n^2)**

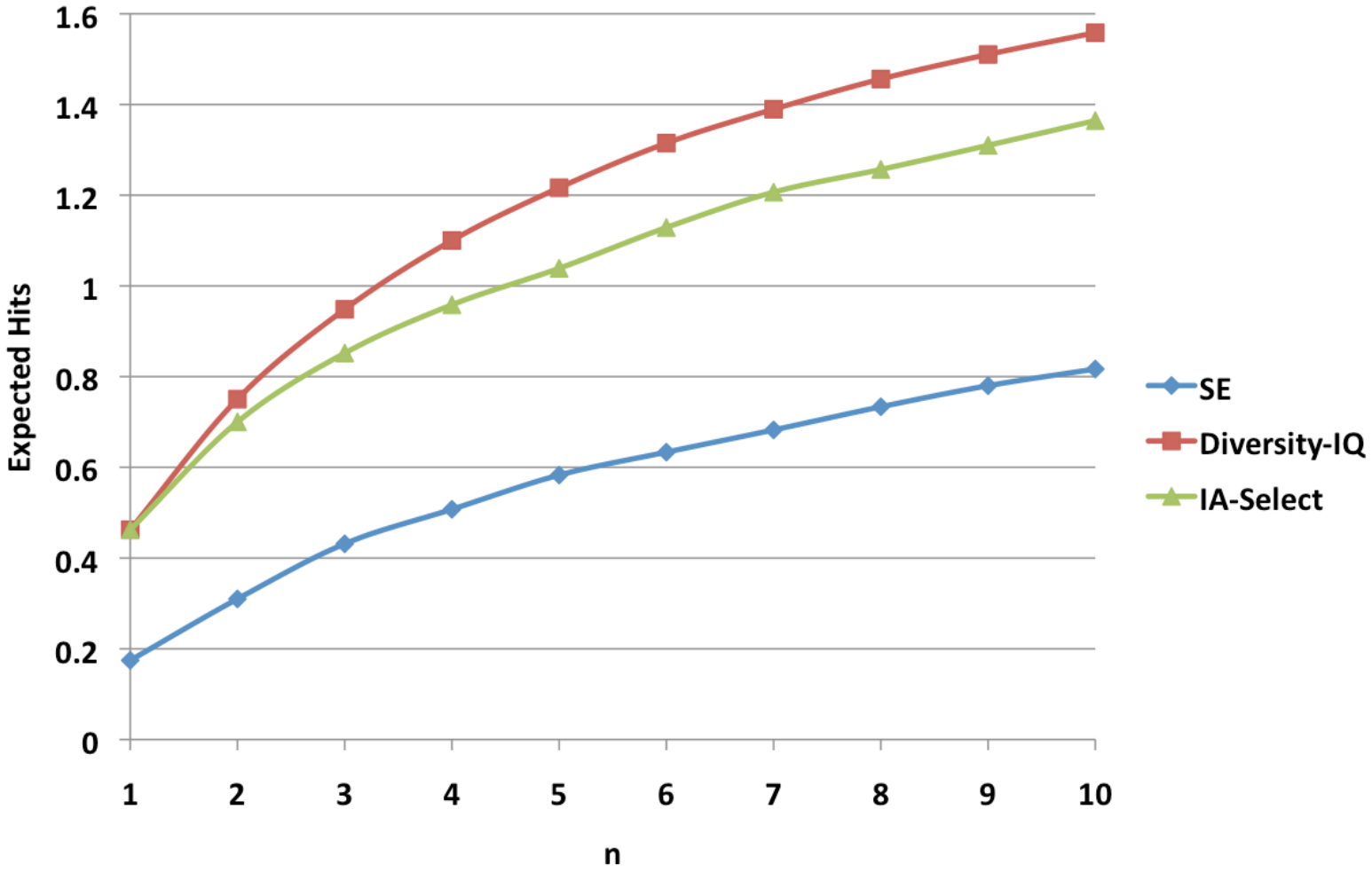
Evaluating *Diversity-IQ*

- **Generated set of 50 ambiguous test queries from a search query log**
- **Extracted subtopic categories from Wikipedia**
  - **Issued each subtopic title as query to search engine and merged top 200 results to form document set**
- **Compared with two other ranking strategies**
  - **Original search engine ranking**
  - **Ranking generated by *IA-Select* [Agrawal '09]**

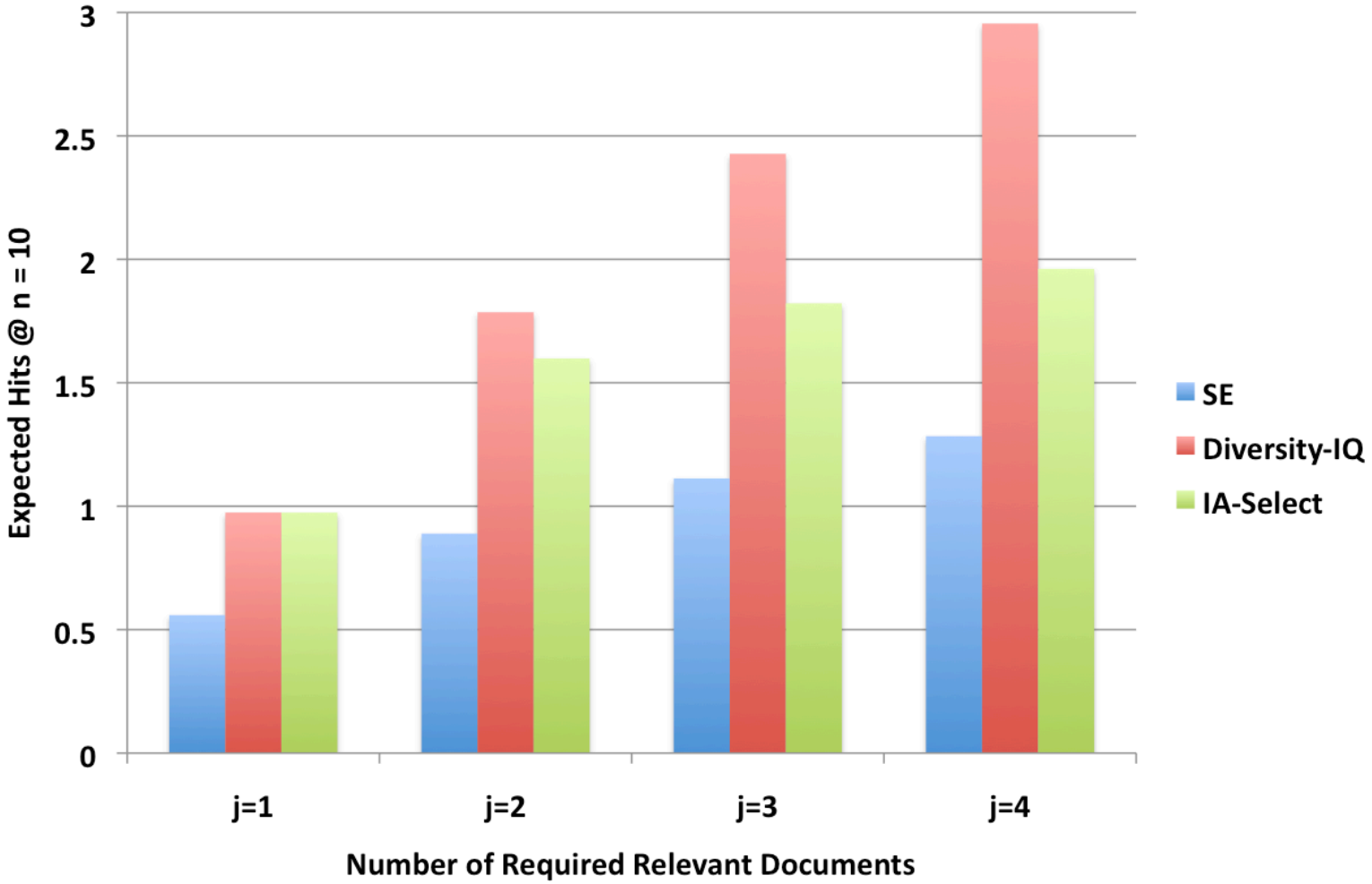
Probability Distributions for Evaluations

- **Page requirements $Pr(J|U)$**
  - **Geometric series $Pr(J=j|U) = 2^{-j}$**
    - **Click log underestimates (e.g. contains navigational)**
- **User intent $Pr(T_i|U)$**
  - **Mechanical Turk survey**
- **Document classification $Pr(T_i|D)$**
  - **Latent Dirichlet Allocation**
    - **Used resulting Θ document-topic distribution**

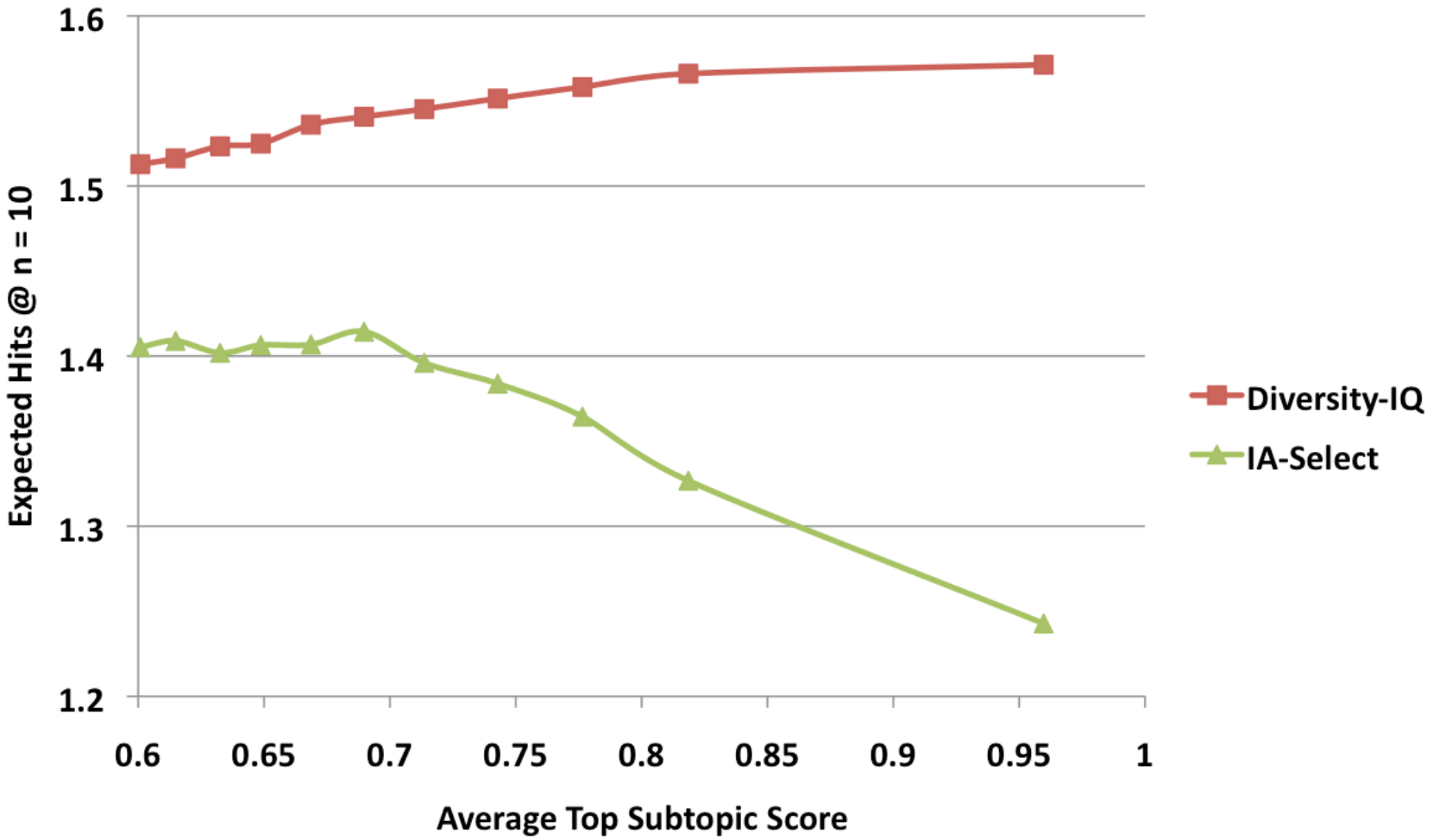
Expected Hits



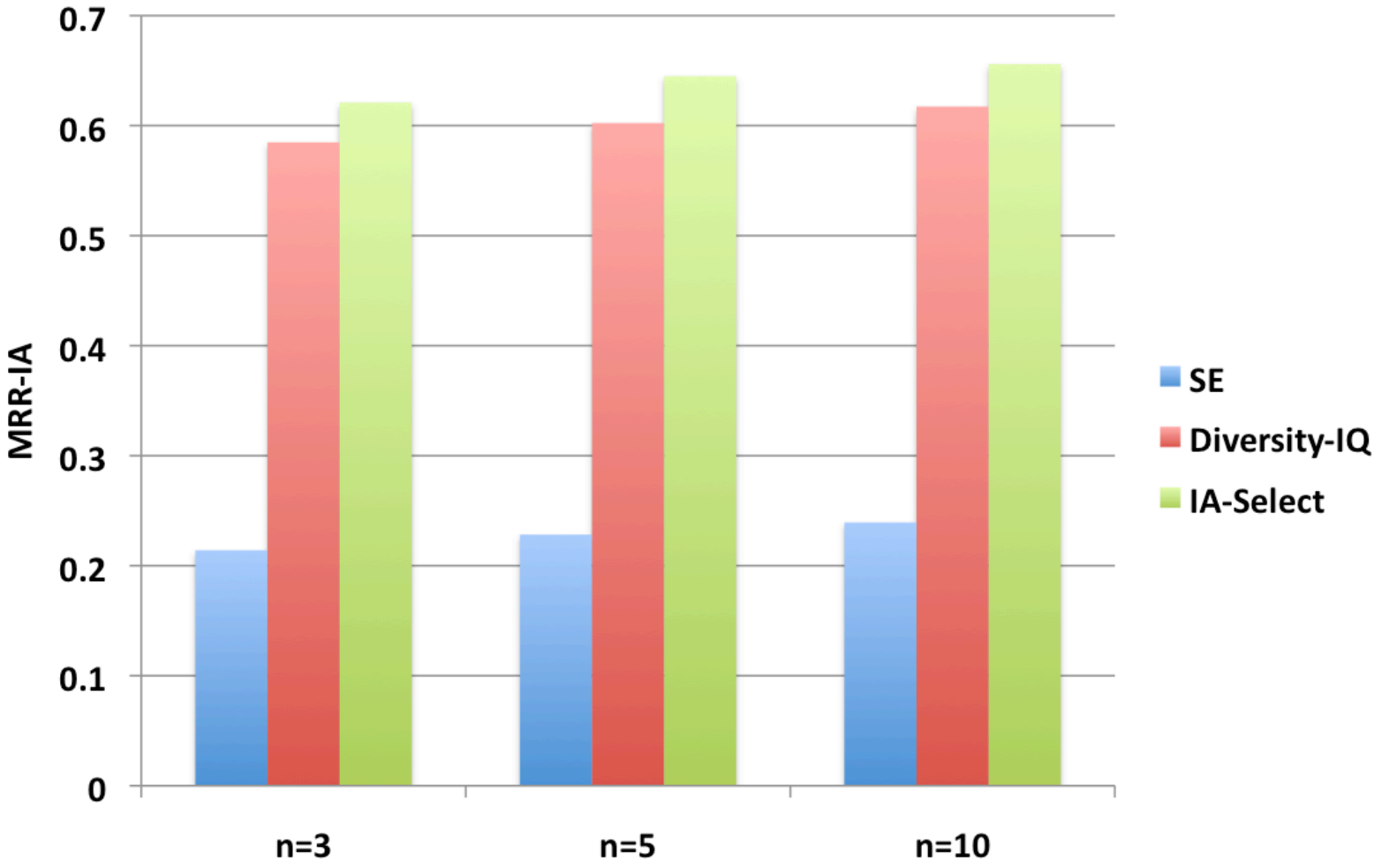
Expected Hits (varying *Pr(J|U)*)



Expected Hits (varying $Pr(T_i|D)$)



Intent-Aware Mean Reciprocal Rank



Evaluation Highlights

- ***Diversity-IQ* improves expected hits**
  - **Relative performance increases as users are expected to require additional relevant documents**
- **Improved user experience for informational queries**
- **Still outperform baseline search engine on "single document" metrics**

Summary

- Presented algorithm for diversifying search results for ambiguous queries
- Our model accounts for the unique requirements of informational queries
  - One relevant document may not be enough
- Up to 50% improvement over modern algorithms in these cases