Personalized Interactive Faceted Search

Jonathan Koren[*], Yi Zhang[*], and Xue Liu[†]

[*]University of California, Santa Cruz

[†]McGill University

0:00 - 0:20

Outline

- Introduce Faceted Search
- Identify Problems with Current FS Tech
- Propose a Solution
- Novel Evaluation Methodology
- Experiments
- Conclusions

0:20 - 1:00

Faceted Search is Everywhere

1:00 – 1:40

“Even if you dont know the term faceted search, you already know what it is. It’s a popular search and navigation interfaced used in ecommerce and digital libraries.”

examples: home depot, taobao, american library of congress

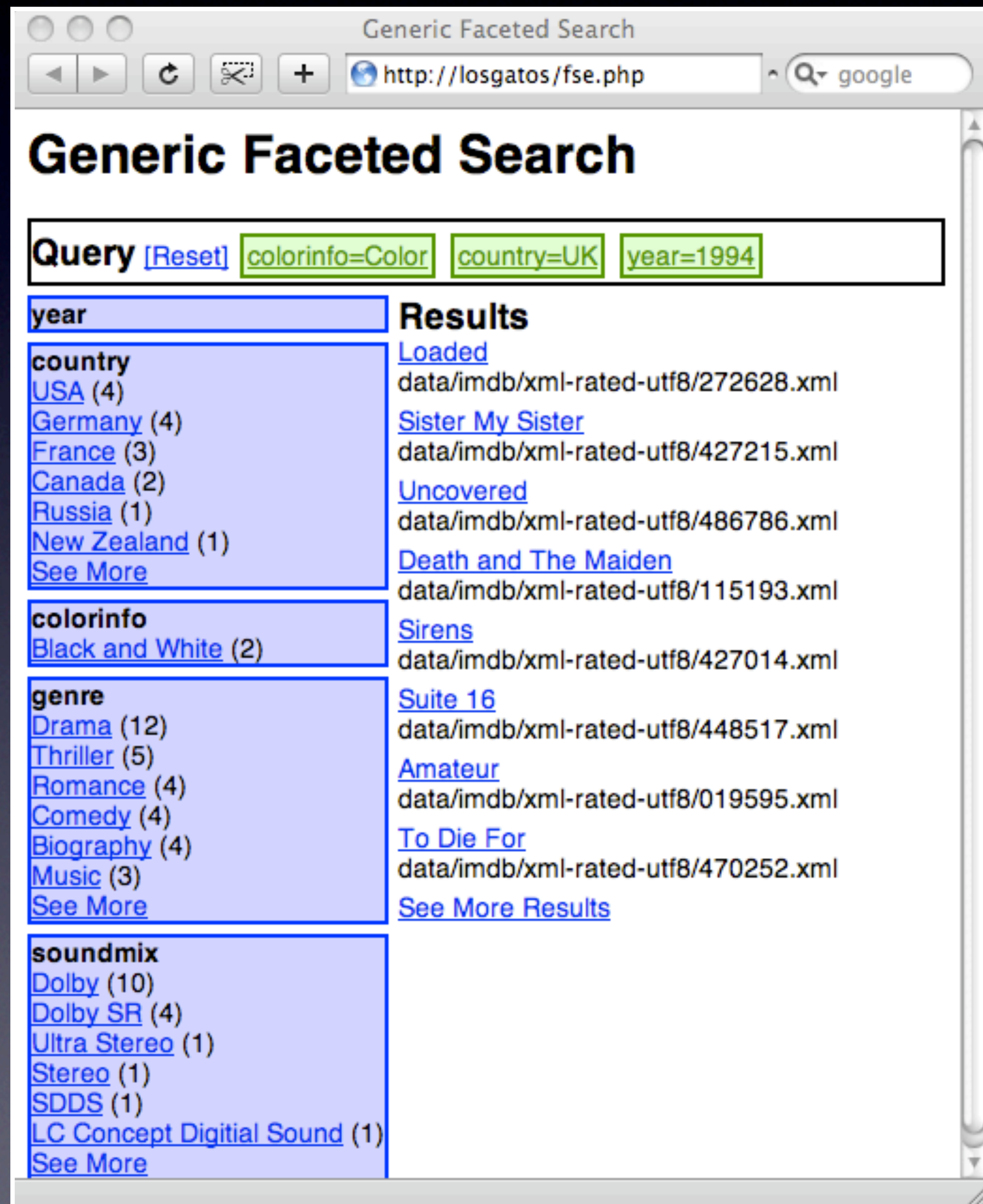
Formal Definition

- Interactive Structured Search Using Key-Value Metadata
- Parallel Hierarchies of Documents
- Point and Click Structured Query Generation

1:40 - 3:40
Keys are called “facets”
Metadata comes from multiple sources: explicit in documents, user provided (e.g. tagging), automatically extracted (e.g. ontology construction)
Traditional search box causes users to only use 2-2.5 terms per query and very rarely use advance features like “this term must be contained in this field” or “retrieve documents of only this type”
Search goes from a more of a guessing game to more of a browsing scenario

Problems

- Too Many Facets and Values
- Existing approach: Ad Hoc Value Presentation
- Proposed Solution: **Personalization** and **Collaborative** faceted search for interactive system **utility** optimization



3:40 – 5:40
Problems: Too many facets and facet values to show; current interfaces use: show all, alphabetical, most frequent
Problems: Ad hoc, Not necessarily what users search by, Different Users have different needs;
Proposed Solution: Personalization and Collaborative filtering
“In order to achieve Personalization and Collaboration <next-slide> we uses these two statistical models”

Statistical Modeling Framework

- Document Model
- User Relevance Model

5:40 - 6:00
“In order to achieve Personalization and Collaboration <this-slide> , similar to normal search, we develop two statistical models”

Document Model

- Docs are Unique Facet-Value Pairs
- Facets Come in Different Types
  - Facet-Type Suggests Statistical Model
- Docs Modeled as a Combination of Statistical Models

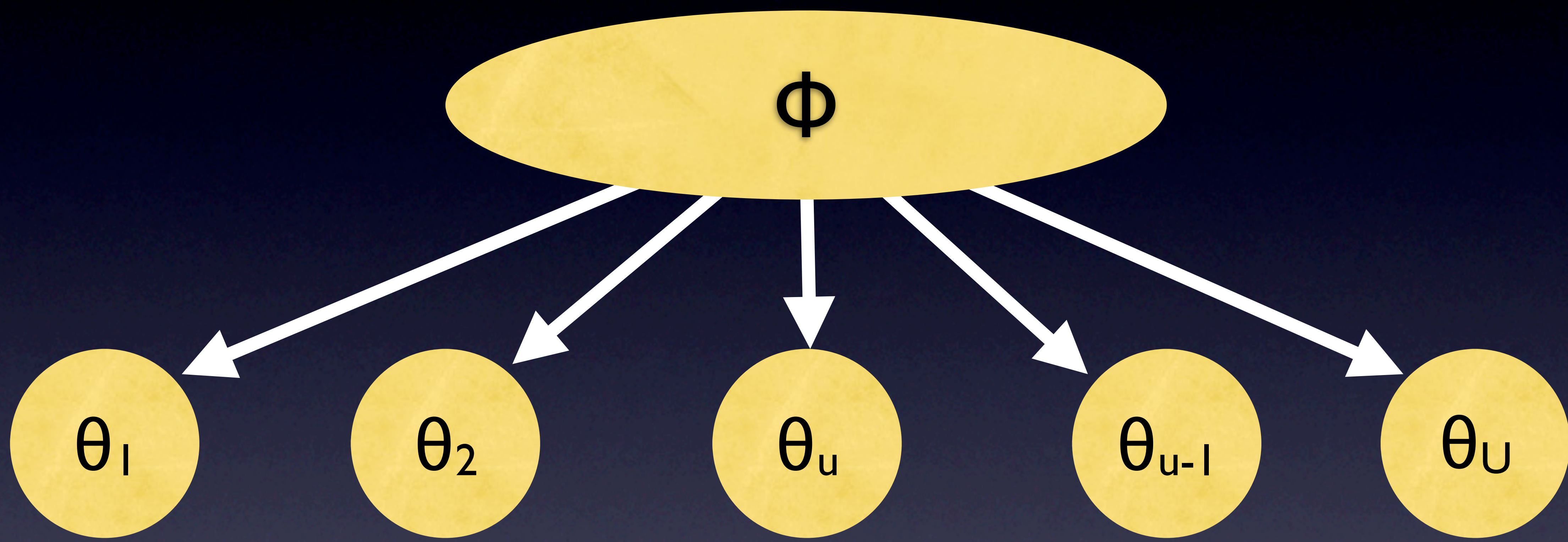
5:40 – 6:40
“In our paper, we extend the idea of language modeling framework to handle faceted documents that contain numbers, ordinal values, nominal values, free text…”
Documents may have multiple facets, and multiple values for a particular facet. (actor=Johnny Depp)
Facet types and models: nomial=bernoulli; ordinal=gaussian; freetext=multinomial

User Relevance Model

$$\theta_u = \{\mathrm{P}(rel \mid u), \mathrm{P}(x_k \mid rel, u), \mathrm{P}(x_k \mid non, u)\}$$

6:40 – 8:40
documents are either relevant or nonrelevant to a user
x_k = is a particular facet value pair
u = user
rel/non x_k in a relevant document or nonrelevant doc
The distributions for P(x_k|rel u) and P(x_k|non u) are the same as the document generative model.

User Collaboration



- Φ is the Conjugate Prior to θ_u
- Φ Fills in Gaps in Individual User Models

8:40 – 12:00
Assume users have somewhat similar preferences. Shared prior allows users to tell other users about useful parts of the search space
Prior helps new users achieve good performance quickly
Shared prior is also a tuple, with each field being the conjugate prior on the corresponding field in the user models
Learn user models. Prior estimated from user models
“With this models we can build an adaptive interface, but how do we evaluate this interface? <slide>”

Interface Evaluation

- User Studies are Expensive
- New Complementary Approach
  - Expected User Interface Utility
  - Simulated Interaction with Pseudousers

12:00 - 13:00
Doesn’t replace user studies, but rather complements them
Simulated interaction has been used in other fields such as evaluating speech systems

User Interface Utility

- Identify Types of Actions
- Assign Costs to Actions
- Reward for Relevant Docs Retrieved
- Calculate Utility for Entire Search Session

13:00 - 15:00
Action types: Select/Deselect FVP, View More FVPs, Mark Doc as Rel/Nonrel, View more docs
Cost of each action dependent on user effort (physical + cognitive) to carry out the action

Expected User Interface Utility

$$E[\mathrm{U}] = \sum_{u \in \mathcal{U}} \sum_{D \in \mathcal{D}} E[\mathrm{U}(u, D)] \mathrm{P}(D \mid u) \mathrm{P}(u)$$

$$E[\mathrm{U}(u, D)] = \sum_{t=0} \sum_{a \in \mathcal{A}_t} \mathrm{R}(q_{t+1}, a, q_t) \mathrm{P}(q_{t+1} \mid a, q_t, u)$$

$$\mathrm{P}(a \mid q_t, u, D) \mathrm{P}(q_t \mid q_{t-1}, u, D)$$

15:00 - 16:00

Assumptions

1. Users Need to Satisfy a Need with a Set of Documents
2. Users Can Recognize Relevant Documents and Facet-Value Pairs
3. Users Continue to Perform Actions Until Their Need is Met

16:00 – 16:30

Pseudousers

- Stochastic Users
- First-Match Users
- Myopic Users
- Optimal Users

16:30 - 17:00
Can think of these as either heuristics for human behavior or as bots that interact with the system being evaluated
The next slides very briefly illustrate the differences between these pseudousers
User does not issue complicated queries to retrieve relevant documents
User can recognize relevant documents and facets in the documents if they are shown

Stochastic Users

- Picks Relevant FVP at Random

A	Nonrelevant	(14 matches)
B	Relevant	(17 matches)
C	Relevant	(11 matches)
D	Nonrelevant	(12 matches)
E	Nonrelevant	(12 matches)
F	Relevant	(15 matches)
G	Relevant	(13 matches)
H	Nonelevant	(4 matches)
I	Relevant	(13 matches)
J	Nonrelevant	(16 matches)

17:00 - 17:30

First-Match Users

- Scans list for Relevant FVPs from Top to Bottom, Picking the First

A Nonrelevant	(14 matches)
B Relevant	(17 matches)
C Relevant	(11 matches)
D Nonrelevant	(12 matches)
E Nonrelevant	(12 matches)
F Relevant	(15 matches)
G Relevant	(13 matches)
H Nonelevant	(4 matches)
I Relevant	(13 matches)
J Nonrelevant	(16 matches)

17:30 - 18:00

Myopic Users

- Picks Relevant FVP that is Contained in the Least Number of Documents

A Nonrelevant	(14 matches)
B Relevant	(17 matches)
C Relevant	(11 matches)
D Nonrelevant	(12 matches)
E Nonrelevant	(12 matches)
F Relevant	(15 matches)
G Relevant	(13 matches)
H Nonelevant	(4 matches)
I Relevant	(13 matches)
J Nonrelevant	(16 matches)

18:00 - 18:30

Optimal Users

- Examines the Complete Interface
- Executes the Action that Maximizes the Utility

A	Nonrelevant	(14 matches)
B	Relevant	(17 matches)
C	Relevant	(11 matches)
D	Nonrelevant	(12 matches)
E	Nonrelevant	(12 matches)
F	Relevant	(15 matches)
G	Relevant	(13 matches)
H	Nonelevant	(4 matches)
I	Relevant	(13 matches)
J	Nonrelevant	(16 matches)

18:30 – 19:00
“In this case, choosing I causes the target document to appear on the first page of the matching documents list”

Evaluation Review

- Each Pseudouser Logs into the Search Interface
- Pseudouser Interacts with Interface to Retrieve a Set of Documents.
- Interface Receives a Score for the Session.
- Expected Utility = Average Score for all Sessions

19:00 – 19:30

Personalization Experiments

- Facet-Value Pair Suggestion
  - Most Frequent
  - Most Probable (Collaborative)
  - Most Probable (Personalized)
  - Mutual Information
- Start Page Personalization
  - Empty Page
  - Collaborative Page
  - Personalized page

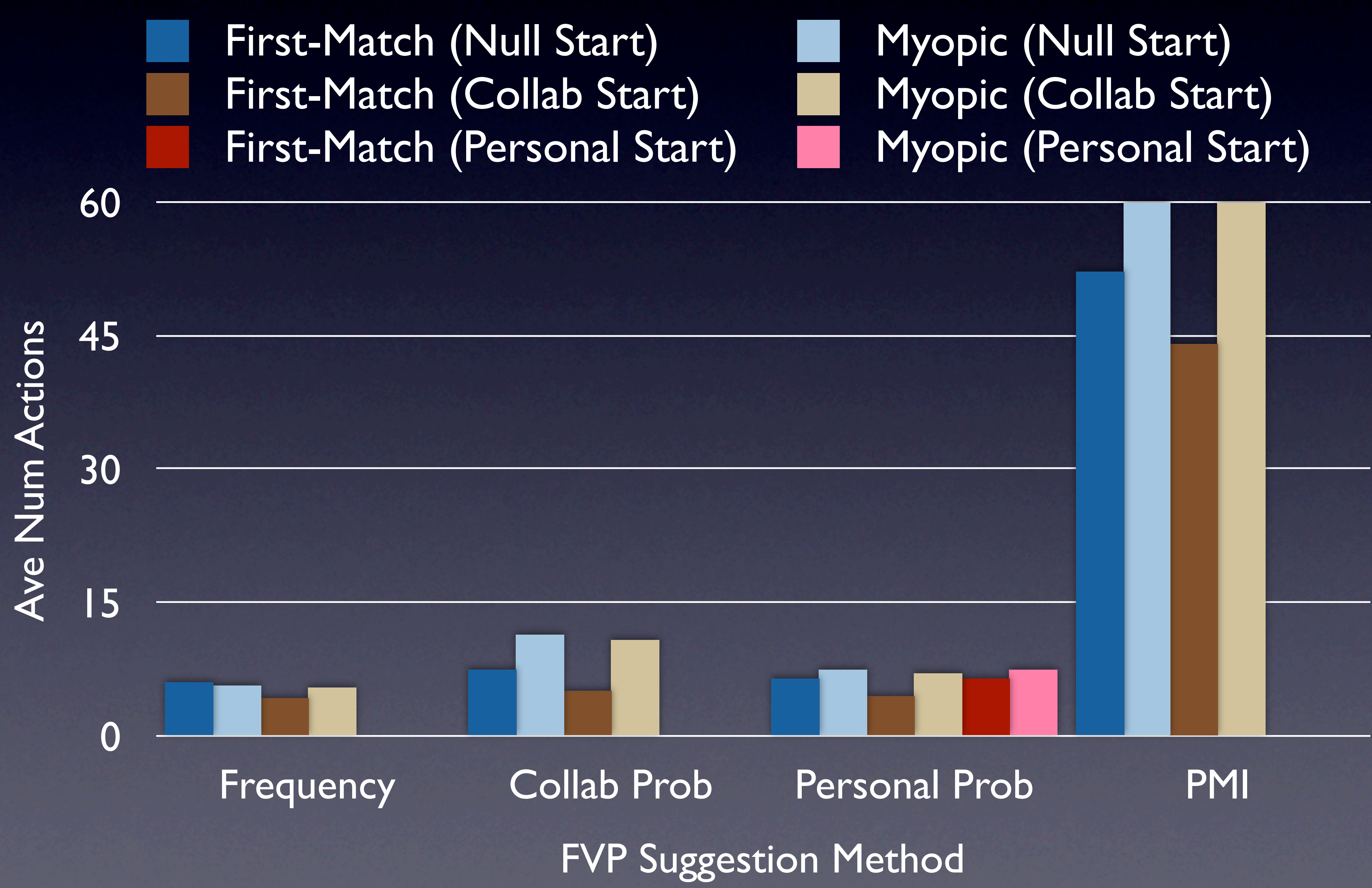
19:30 - 20:00
used naive-bayes to learn user models
Looked at first-match and myopic pseudousers
Pseudousers searched for a single relevant document, and had perfect knowledge of the doc

Document Corpora

- 8000 Documents from IMDB
  - 19 Facets and 367k Facet-Value Pairs
- 5000 Users Each from Netflix and MovieLens
  - 633k Ratings for Netflix
  - 742k Ratings for Movielens

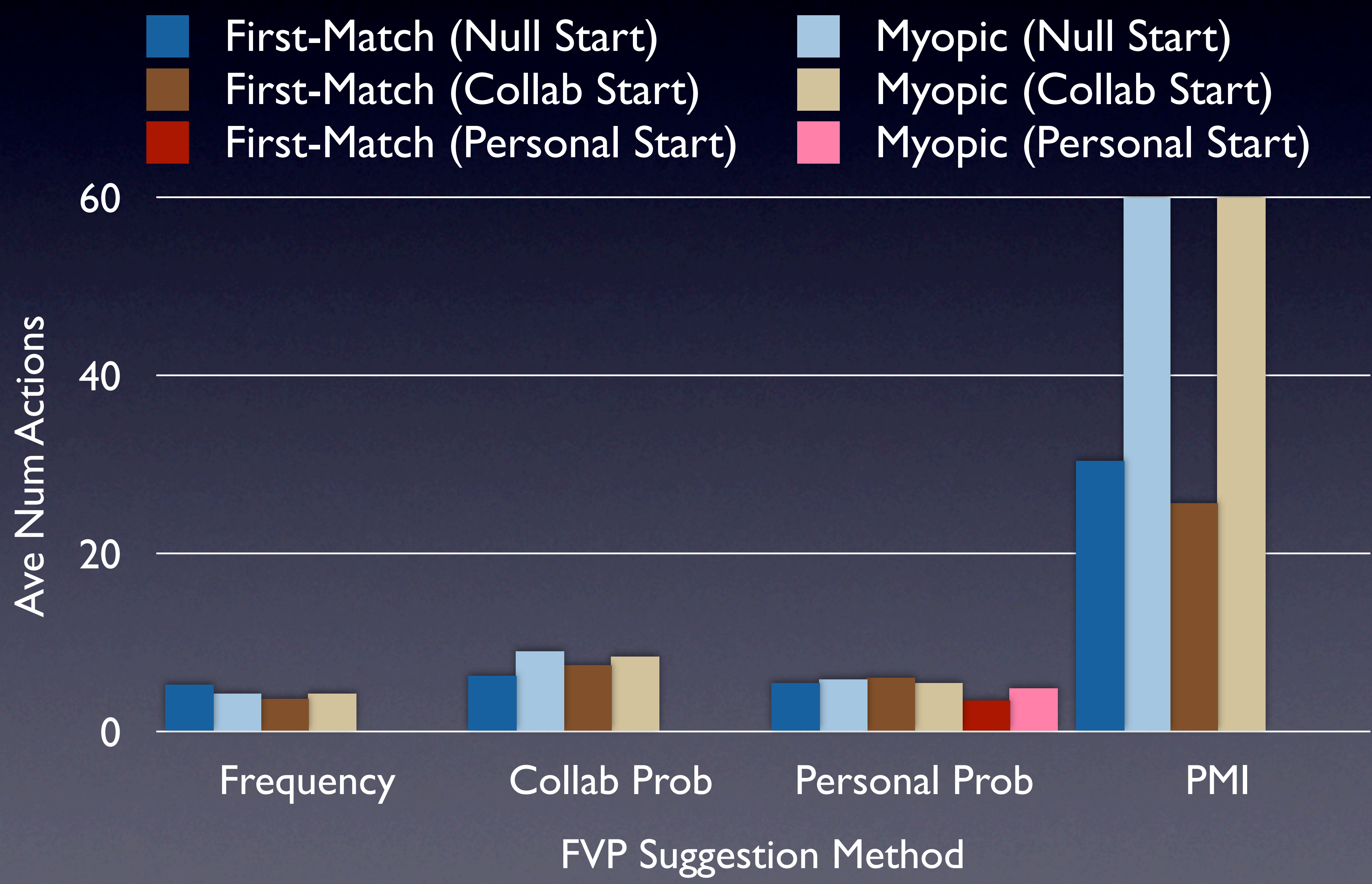
20:00 - 21:00
Used IMDB corpus with explicit real-user ratings from the MovieLens and Netflix corpora
8000 documents, with each document being relevant to at least one user

Results
(Netflix)



21:00 – 22:30
Shorter is better; uniform cost for all actions
(200 myopic null, 178 for myopic collab)
PMI had problems with infrequent but highly correlated FVPs
First match users tended to find documents faster than myopic users. (Good FVPs would disappear after reranking for myopic. Myopic always picked the overfitted FVP?)
Nothing consistently outperformed frequency as a suggestion method. Document corpus biased towards relevant docs. The personalization method was weak.
Landing page personalization helped users. (Never hurt)

Results (MovieLens)



23:30 – 24:00
Similar results as before. (myopic null = 121; myopic collab = 110)

Conclusions

- Many Facets and Values are a Problem
  - Personalized Interfaces Can Help
- Proposed Statistical Modeling Framework for Faceted-Search
- Proposed Inexpensive Repeatable Evaluation Technique for Faceted-Search Interfaces
- Personalized Start Pages are Helpful

24:00 - 25:00
Need better personalization method

fin

25:00

Example: Two Myopic Users Search for "The 'Burbs"

User: 302	User: 1329
certificate=PG	certificate=PG
soundmix=Dolby	soundmix=Dolby
genre=Comedy	genre=Comedy
	country=USA
	language=English
	colorinfo=Color
	year=1989
productiondesigner=SpencerJamesH	productiondesigner=SpencerJamesH

Myopic users, personalized probability FVP suggestion
With myopic users the order of the suggested FVPs is irrelevant