(Not Too) Personalized Learning to Rank for Contextual Suggestion

Andrew Yates[1], Dave DeBoer[1], Grace Hui Yang[1], Nazli Goharian[1], Steve Kunath[2], Ophir Frieder[1]

Speaker: Grace Hui Yang

[1]Department of Computer Science
[2]Department of Linguistics
Georgetown University
Nov 8, 2012@TREC

1

Introduction

- TREC 2012 Context Suggestion track operates on the Open Web and aims to provide visitors suggestions (for entertainment) based on time, location, and personal interests.
- Problem Formulation:
  - Given a person P's ratings (+1, 0, -1) for 50 example suggestions in City A (Toronto in this case), provide the best 50 ranked suggestions in City B for P.

2

Assumptions

- In City B, Person P will be interested in the similar types of things/suggestions as in City A
  - Recognize type of the suggestions
    - E.g., Fresh on bloor -> restaurant, vegetarian restaurant
    - Types of suggestions are context-independent
  - Find the same suggestion types in City B
- Can we just submit queries to Google/Bing/Yelp?

3

An Issue

- *Many suggestions are local stores that seem 'not well-known' and 'not attractive' to visitors*

1. Rancho Ventavo
2. Courtyard Oxnard Ventura
3. Little Book Store
4. Café Naakio
5. The Kitchen
6. Cabo Seafood Grill & Cantina
7. Cafe Amri
8. ARC Thrift Store
9. Tomas Café
10. Peet's Coffee & Tea

Profile 23, Context 19 (Oxnard, CA, Fall weekend morning)

4

Assumptions

- In City B, Person P will be interested in things that City B is famous for.
  - E.g., visiting historical buildings in DC while you don't usually visit them in Pittsburgh;
  - E.g., visiting Falling Water in Pittsburgh and visiting Empire State Building in NYC
  - Create a city profile for each city

5

Assumptions

- In City B, Person P will be interested in things that most people are interested in.
  - E.g., People are interested in food, shopping, museums, tours
  - E.g., visiting restaurants more frequently than spa
  - Create a general profile that most people like

6

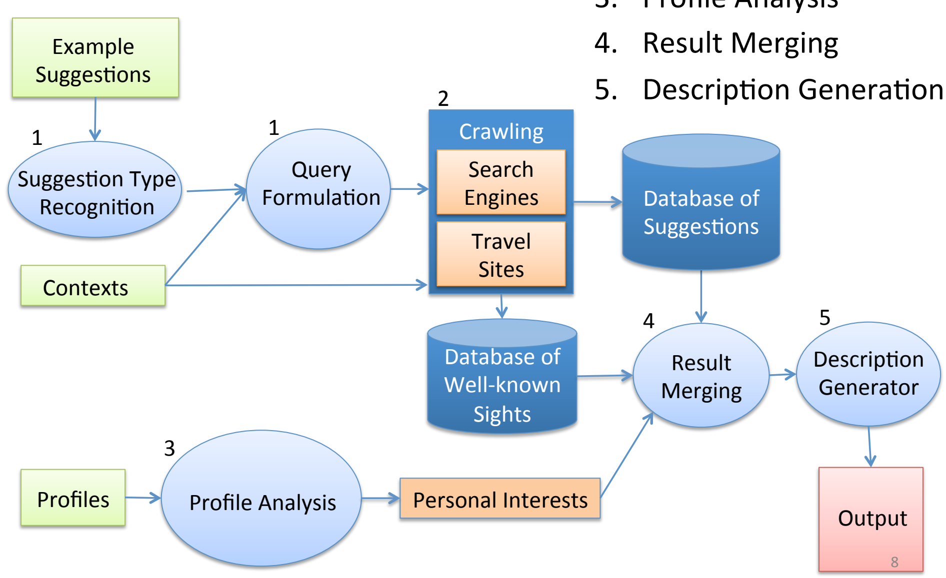
'Not too personalized' contextual suggestion engine

- Our Approach:
  - Merging and re-ranking contextual suggestions crawled from the Open Web.
  - Balancing among a person's profile, a city's profile and general population's profile

7

System Components

1. Query Formulation
2. Crawling
3. Profile Analysis
4. Result Merging
5. Description Generation



8

Suggestion Type Recognition

Example Suggestion:
<title>Fresh on Bloor</title>
<description>Our vegan menu boasts an array of exotic starters, multi-layered salads, filling wraps, high protein burgers and our signature Fresh bowls.</description>
<url>http://www.freshrestaurants.ca</url>

- Generating Context-independent Queries from Example Suggestions
  - E.g. vegetarian restaurant
- Head nouns in title: “Toronto Zoo”->”Zoo” (~30% accuracy)
- High frequency terms in descriptions/documents
  - “Hockey Hall of Fame”->”game”
- Mapping title to Yelp categories (~60% accuracy)
  - Missing entries in Yelp;
  - Unwanted category names in Yelp (e.g. “Getaways”, “Landmark and Historical Buildings”, “Sites”)
- Mapping title/description to a two-level ontology (>95% accuracy)
  - An ontology is handcrafted based on Yelp.
    - 14 top categories, 70 second level categories.
  - For each category, create a *representative document* by submitting this category name to Google and concatenating snippets and Wikipedia pages.
  - Mapping a suggestion with representative documents by BM25

9

Query Formulation

- Suggestion types are used as the context-independent queries
  - E.g., restaurant, walking tour, spa, performing arts
- Each context-independent query is paired with a city to form a *context-dependent query*
  - E.g. *restaurant Pittsburgh, spa New York City, walking tour San Francisco*

10

Crawling

- Context-dependent queries are sent to 5 online search engines:
  - Google, Google Places, Bing, Yelp and Yellow Pages.
- From each search engine, crawl the top 50 results and store metadata in a relational database.
  - Title, url, city, state, zip, address, telephone number, snippets, ratings (if any), reviews (if any), hours of operation
- Filtering Noise
  - 3rd party pages, "under construction" and "coming soon", duplicates
- Filling up Missing Values
  - *Performing arts* operate during evenings, Mon-Sun.
  - *Everything else* operates during morning and afternoon, Mon-Sun.

11

Profile Analysis

- General Profiles
  - Aim to capture relative importance among different categories of interests for general population
  - Each category is weighed by the number of suggestions of that category in Toronto examples
    - E.g. Performing arts (7) and Restaurants (5) are much more popular than Spas (1)

12

Profile Analysis

- City Profiles
  - Famous sights of a city
  - Crawl from 3 travel web sites:
    - aviewoncities.com, Wikitravel, and Wikipedia
  - Sights are ranked by their ranks in aviewoncities and the number of travel web sites it appears
  - E.g. NYC

1. Statue of Liberty
2. Empire State Building
3. Central Park
4. Brooklyn Bridge
5. Times Square

6. Chrysler Building
7. Grand Central Terminal
8. Rockefeller Center
9. Metropolitan Museum of Art
10. Fifth Avenue

13

Personal Profile

- User-Category Association Matrix
- Each cell contains the counts of positive ratings (+1) that a user judges for that category

	Performing Arts (7)	Bar (6)	Restaurant (5)	Spa (1)	…	Mean for this person	Std for this person
Profile 1	2	0	3	1		1.4	1.1
Profile 2	2	1	2	1		1.6	0.8
Profile 3	1	0	0	1		0.8	0.7
…							
Mean for this category	3	2.6	2.2	0.65		**1.85**	

14

Personal Profiles

- Detecting major, less major, minor, and negative interests for a user
  - Major interests i for profile p,
    - Score(p,i) greater than both row and col mean
  - Less major interests i for profile p,
    - Score(p,i) greater than row mean only
  - Minor interests i for profile p,
    - Score(p,i) greater than col mean only
  - Negative interests,
    - Score(p,i) smaller than both row and col mean
- Learn & Assign different weights for different types of interests

15

Personal Profiles

- Top level vs. Specific Interests
  - E.g. Restaurants vs. Sushi Bar, Game vs. Hockey
- Macro vs. Micro Interests
  - Macro: *electoral college votes*
    - Lean towards rare categories
  - Micro: *popular votes*
    - Lean towards popular categories
- Initial (Description) vs. Final (Webpage) judgments
- ***Many variations***
  - ***2^3=8 combinations, plus different weights***

16

Result Merging

- Two level Learning to rank by SVMRank
  1. Rank suggestions within each category
  2. Rank and merge results from different category
- Creation of Training Data:
  - Simulate ranked results from profiles
  - Form (q,d) pairs by (example type, example) pairs
  - Rankings follow the examples' popularity
    - ratio of number of init/final 1's to number of init/final -1's in the profiles.

17

Rank Suggestions within Each Category

- Features about Rank, Reviews, Ratings
  - average rating across all search engines;
  - mean reciprocal rank across all search engines;
- Features about Query-Suggestion Relevancy
  - average percentage of query terms appearing in title/snippet across all search engines;
  - sum of the IDF of query terms found in reviews/snippets;
- Features about URL
  - number of slashes in the URL;
  - length of the URL in characters;
- Features about Search engines
  - a boolean indicating whether the suggestion was found by a search engine
- *Complete set of features in the notebook paper*

18

Merge Categories

- To include a variety of results from categories that the user liked
- 10 results at a time
- Among each 10, the results are merged from categories that interest user *p*
  - Categories ranked by their *Score(p,i)*
  - $Number\ of\ results\ in\ each\ 10\ from\ category \propto \frac{score(p,i)}{\sum_j\ score(p,j)}$

19

Results for Profile 23, Context 19 (Oxnard, CA, Fall weekend morning)

- *1. Rancho Ventavo*
- *2. Courtyard Oxnard Ventura*
- *3: Murphy Auto Museum*
- *4. Carnegie Art Museum*
- *5. The Kitchen*
- *6. Cabo Seafood Grill & Cantina*
- *7. Cafe Amri*
- *8. ARC Thrift Store*
- *9. Tomas Café*
- *10. Peet's Coffee & Tea*

20

Handling Contexts

- Location: by submitting context-dependent queries
- Time: Fill up missing values & Database queries
- Season:

Seasons	*Action*	*Categories*
winter	ignore	parks, gardens, cultural districts, farmers market, Landmarks, tours, zoos
spring, summer, and fall	boost	cafes, restaurants, bars

21

Natural Language Description Generation

- Use SVMRank to choose each result's best description from the favorable reviews on Yelp and Google Places.
  - Training data: manually ranked descriptions from *context-independent queries + Pittsburgh*
- Add a beginning sentence to descriptions using rules such as:
  - Title_of_the_suggestion is a [must-go | great | gorgeous | top | brilliant | cool | famous | wonderful ...] type_of_the_suggestion
  - Title_of_the_suggestion is an [amazing | excellent | attractive ...] type_of_the_suggestion
  - *More in the notebook paper*

22

Results for Profile 17, Context 12 (San Diego, Winter weekend afternoon)

- ***1. USS Midway Museum***: USS Midway Museum is an amazing museum. One of my favorite places to take friends from out of state. I'm always so impressed by its sheer mass. I highly recommend taking the tour because I know we've all wondered what the inside of an aircraft carrier looks like. Here's your chance to explore and be wowed by the countless rooms, planes on deck, mess hall, control tower and weapons hold. Don't miss out on the opportunity to see it up close and personal.
- ***2. Birch Aquarium***: Birch Aquarium is a fabulous museum. I honestly thought Birch was mediocre due to the size. Nothing really stood out for me besides view of the ocean on the outer area. It's very educational, pretty good visit for those who have never really been to any aquariums and also good place to visit for school field trips. It's definitely worth a visit if you're an aquarium/fish lover.

23

Submitted Runs

- GUInit: using initial judgments to train and rank
- GUFinal: using final judgments to train and rank
- Both runs:
  - Micro interests
  - Top categories only
  - Ratio of Major, Less Major, Minor, Negative personal interests
    - 2, 0.7, 0.3, 0

24

Experiment Results

Run	P@5 WGT	P@5 GT	P@5 G	P@5 T	P@5 W	P@5 D
	0.3235	0.6027	0.8930	0.6156	0.4599	0.3605
guinit	0.2920	0.6635	0.8802	0.6997	0.4451	0.5019
gufinal	0.2710	0.6689	0.8852	0.7031	0.4241	0.5191
	0.2481	0.4950	0.7565	0.5794	0.3500	0.2852
	0.2475	0.5464	0.9036	0.5510	0.4198	0.5160
	0.2333	0.6032	0.8148	0.6147	0.3889	0.3815
	0.2210	0.5442	0.7939	0.6210	0.3500	0.3173
	0.2185	0.5649	0.9034	0.5839	0.4049	0.4710

What works:
- **Almost everything.**

Nice surprises:
- **Assumptions based on suggestions types for time/season**
- **Description generation**

What needs to be improved:
- **Webpage (W)**

25

What happened in Webpages?

- High variance in personalized judgments ?
  – 34*50 sets of results are only evaluated for ~40 sets, and for only top 5 results per set.
  – If more result sets are judged and more results in 5+ are judged?
- We did not touch too much on content of a webpage
  – Most ranking features are from the metadata, or title, snippets
  – Not full text
- What if we handle the full text better?
  – Parsing HTML
  – Multimedia
  – Menu

26

Conclusions

- Merge suggestions crawled from online search engines by *Learning to rank*
  - *Mostly by meta data features, not full text*
- As a service designed for visitors, we argue that contextual suggestion is a personalized service but *not too personalized*
- Balanced suggestions by
  - including city profiles (famous sights) and general profiles
  - using micro level interests
  - using top level interests (more general interests)

27

Thank You

Grace Hui Yang

Department of Computer Science

Department of Linguistics

Georgetown University

Contact: huiyang@cs.georgetown.edu

28