PAPER REVIEW

Presented By: Udochukwu Nweke

CS734: Introduction to Information Retrieval

Dr. Michael Nelson

November 2, 2017

Paper One

Generating query substitutions

Rosie Jones, Benjamin Rey, and Omid Madani
Yahoo! Research
3333 Empire Avenue
Burbank, CA, 91504, USA
{jonesr,benjamir,madani}@yahoo-inc.com

Wiley Greiner
Los Angeles Software Inc.
1329 Pine Street
Santa Monica, California 90405
w.greiner@lasoft.com

Date: May 23 - 26, 2006

Paper Two

SearchTogether:
An Interface for Collaborative Web Search

Meredith Ringel Morris
Microsoft Research
One Microsoft Way
Redmond, WA, 98052, USA
merrie@microsoft.com

Eric Horvitz
Microsoft Research
One Microsoft Way
Redmond, WA, 98052, USA
horvitz@microsoft.com

Date: October 07 - 10, 2007

Query Substitution

- Query substitution is a technique which improves retrieval effectiveness.
- Users' search query may be an imperfect description of their needs and Search Engines or information retrieval systems are not able to retrieve documents matching the given queries
- Query substitution attempts to identify words that are very related to specific query words rather than topic
- For example, the query "cat cancer", can be replaced with "feline cancer".

Existing Techniques of Query Substitution

- **Pseudo-relevance feedback**: This is a query expansion technique where the retrieved query is modified by adding additional terms based to the retrieved document. This has limitations in effectiveness.
- **Substituting query with related terms**: Substitution is done based on retrieved documents. This also has limitations in effectiveness since it depends on the retrieved document.
- **Deleting query terms:** Deleting a query term might lead to loss of specificity

Contributions of First Paper

- Identification of a new source of data for identifying similar queries and phrases
- Definition of a scheme for scoring query suggestions
- An algorithm for combining query and phrase suggestions which identifies highly relevant queries and phrases
- Identification of features which are predictive of highly relevant query suggestions

Problem Statement

- Given a query q_i, the objective is to generate a modified query q_j which is related to the original query. This can be written as:
- $q_i \rightarrow q_j$ where q_j is the suggestion for q_i.
- There are several ways q_j may be related to q_i, some of these ways might improve q_i and some might result in loss of originality

Classes of Relevance Suggestion

- **Precise Rewriting (Class 1)** : the rewritten query preserves the main meaning of the initial query while allowing for extremely minor variations in scope
- **Approximate Rewriting (Class 2)**: scope of the initial query is narrowed or broadened, but the relationship or category is retained.
- **Possible Rewriting (Class 3)**: rewritten form is different from user intent.
- **Clear Mismatch (Class 4)**: no relationship with original query.

Acceptable Classes of Rewriting Queries

- Based on the task, different classes of rewriting queries may be acceptable
- **Specific Rewriting (Class 1+2)**: 1s and 2s are acceptable if the goal is to have a closely related query. New result will be highly relevant
- **Broad Rewriting (Class 1+2+3)**: Classes 1, 2, and 3 are acceptable if the goal is to perform re-ranking of results retrieved with initial query. These 3 classes are also acceptable for query expansion.

Examples of Queries and Query Suggestions

Class	Score	Examples		
Precise rewriting	1	automotive insurance	$\mapsto$	automobile insurance
		corvette car	$\mapsto$	chevrolet corvette
		apple music player	$\mapsto$	apple ipod
		apple music player	$\mapsto$	ipod
		cat cancer	$\mapsto$	feline cancer
		help with math homework	$\mapsto$	math homework help
Approximate rewriting	2	apple music player	$\mapsto$	ipod shuffle
		personal computer	$\mapsto$	compaq computer
		hybrid car	$\mapsto$	toyota prius
		aeron chair	$\mapsto$	office furniture
Possible rewriting	3	onkyo speaker system	$\mapsto$	yamaha speaker system
		eye-glasses	$\mapsto$	contact lenses
		orlando bloom	$\mapsto$	johnny depp
		cow	$\mapsto$	pig
		ibm thinkpad	$\mapsto$	laptop bag
Clear mismatch	4	jaguar xj6	$\mapsto$	os x jaguar
		time magazine	$\mapsto$	time and date magazine

Table 1: Example queries and query-suggestions for each of the four classes.

Query Substitutions For Infrequent Queries

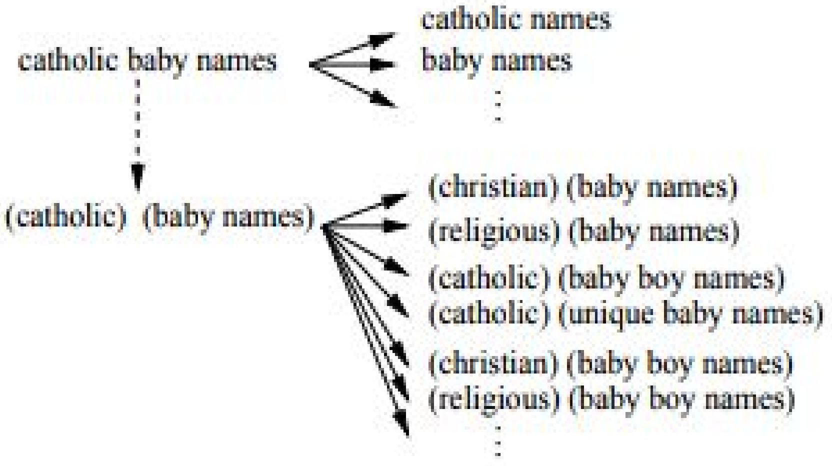

Figure 1: We generate multiple candidates Q' for query rewriting by both substituting for the whole query atomically, and by breaking it into phrases and generating substitutions for each phrase.

- Frequent queries have more related queries than infrequent queries.
- Infrequent queries can be broken down into phrases.
- Infrequent queries should not be ignored because they have higher ranking (Zipf's law)

What is a Query Pair?

- The data used for this experiment is gotten from logs of users web accesses on a daily basis. The data is unique to each user with user ID and timestamp
- A candidate reformulation is a pair of consecutive queries issued by a single user on a single day.
- Candidate reformulation will also be referred to as query pairs

$$\text{candidateQueryPairs}(user_i, day_j) = \{< q_1, q_2 >: (q_1 \neq q_2) \wedge \\ \exists t : query_t(user_i, q_1) \wedge query_{t+1}(user_i, q_2)\}$$

Identifying Significant Query Pairs and Phrase Pairs

- Hypothesis likelihood ration is used to distinguish related query and phrase pairs from candidate pairs that are unrelated.
- This metric tests the hypothesis that the probability of term q_2 is the same whether term q_1 has been seen or not

$$H_1 : P(q_2|q_1) = p = P(q_2|\neg q_1)$$

$$H_2 : P(q_2|q_1) = p_1 \neq p_2 = P(q_2|\neg q_1),$$

The likelihood score is

$$\lambda = \frac{L(H_1)}{L(H_2)}$$

The test statistic $-2 \log \lambda$ is asymptotically χ^2 distributed. Therefore we work with the log likelihood ratio score:

$$\mathrm{LLR} = -2 \log \lambda = -2 \log \frac{L(H_1)}{L(H_2)}$$

Query Substitutables

dog → dogs		9185	(pluralization)
dog → cat		5942	(both instances of 'pet')
dog → dog breeds		5567	(generalization)
dog → dog pictures		5292	(more specific)
dog → 80		2420	(random junk or noise)
dog → pets		1719	(generalization – hypernym)
dog → puppy		1553	(specification – hyponym)
dog → dog picture		1416	(more specific)
dog → animals		1363	(generalization – hypernym)
dog → pet		920	(generalization – hypernym)

Table 2: Terms and queries which can be substituted for the term or query "dog", along with likelihood ratios, based on user query rewriting sessions. The semantic relationship is shown for explanatory purposes only.

- A high value for likelihood ratio indicates that there is a relationship between terms q1 and q2.
- Query pairs and phrase pairs above a threshold for the LLR score are referred to as query substitutables

Assessing The Quality of Substitutables

Random Ranking Algorithm

- Set the maximum number of whole query alternatives m to 10
- Segment the query, and assign number of alternatives per phrase, k, as a function of the number of constituent phrases n:

```
if (n > 5)
      k = 5
else
      k = 10
```

- Whole query and phrase suggestion LLR score should be 50 at least.
- The likelihood of selection for substitution is a phrase over the whole query

Substitution Type and Log-likelihood Ratio Score

LLRNumSubt using the following intuitions

- First try whole-query suggestions ranked by LLR score
- Then try suggestions which change a single phrase, ranked by the phrase substitution LLR score
- Then try suggestions which change two phrases, ordered by the LLR scores of the two phrases

Query Substitution with Machine Learning

- Evaluation is based on binary classification tasks (Acceptable Classes of Rewriting)
  - broad (classes 1+2+3) for which the negative class will be rewritings labeled 4
  - specific (classes 1+2) for which the negative class will be rewritings labeled 3+4.
- Labeled Data: Training Data is a sample of top-ranked suggestion for each 1000 queries
- The top-ranked suggestion was generated using LLRNumSubs rankikng scheme

Linear Regression Classifier

For this type of classifier, labels {1, 2, 3, 4} were used and standard linear regression was performed

The simplest best fit was obtained with the following features:

- **Word distance**: prefer suggestions with more words in common with the initial query
- **Normalized edit distance**: prefer suggestions with more letters in common with the initial query
- **Number of substitutions**: prefer whole query suggestions over phrase suggestions, prefer fewer phrases changed

Precision of Top-Ranked Suggestion For Each Ranking Scheme

Ranking Scheme	{1+2}	{1 + 2 + 3}
random	55%	-
LLR	6%	87.5%
numSubstEdit	74%	87.5%

Precision is the fraction of retrieved instances among relevant instances.

Table 6: Precision of top-ranked suggestion for each ranking scheme for precise (1+2) and broad (1+2+3) rewriting.

Paper Two

SearchTogether: An Interface for Collaborative Web Search

Meredith Ringel Morris
Microsoft Research
One Microsoft Way
Redmond, WA, 98052, USA
merrie@microsoft.com

Eric Horvitz
Microsoft Research
One Microsoft Way
Redmond, WA, 98052, USA
horvitz@microsoft.com

Collaborative Web Search

- Web search is considered a solitary activity: browsers and search engine homepages are designed to support single-user scenarios

- Users often desire to collaborate on search tasks like planning a trip, researching medical condition, working on a common project , etc.

- Collaboration in web search supports synchronous remote collaboration

Passive Collaborative Search Systems

- In passive collaboration, search is focused on using data generated by different number of users interactions with a system to automatically refine the search behavior.
- Systems use query log to generate query substitutions

Multi-User Search

https://hci.stanford.edu/publications/2006/teamsearch.pdf

Figure 1. A four-person group uses TeamSearch at a DiamondTouch table to find photos from a metadata-tagged repository.

- Prior systems designed to allow search collaboration had some limitations
- For example, TeamSearch where a group of co-located people search through a database for images through visual query language

SearchTogether

- SearchTogether is an interface for collaborative search that allows group of any number to collaboratively search the site from any location

- SearchTogether encourages direct collaboration among friends, family, and colleagues in information seeking and review

- SearchTogether enables either synchronous or asynchronous remote collaboration

SearchTogether Design Goals

- **Awareness-Features**
- Awareness of other group member's activities has the tendency of reducing undesired duplication of effort.
- SearchTogether promotes learning of search techniques by going through the awareness of other group members' search strategies
- SearchTogether reduces the wasted effort of query substitution

Division Of Labor-Features

- SearchTogether provides several mechanisms for encouraging division of labor among collaborators
    - Explicitly spliting keywords, search engines, or sub-task and assigning them to individual group member avoids unnecessary duplication of effort.
- Recommendation mechanism: group member recommends a page to a specific user to read
- Instant Messaging: group members discuss current task and coordinate their efforts

Persistence-Feature

- Storing shared messages promotes asynchronous search sessions
- Search summary: Pages which received positive ratings are showed with title, URL, and comments
- SearchTogether allows users to store search history

SearchTogether Query Interface

Figure 1. The SearchTogether client. (a) integrating messaging, (b) query awareness, (c) current results, (d) recommendation queue, (e)(f)(g) search buttons, (h) page-specific metadata, (i) toolbar, (j) browser

Relationship Between Paper One And Paper Two

- In paper one, query modification is based on users log sessions where the data used to modify the initial query is generated from user query log sessions.
- Paper two contradicts paper one because searchTogether uses active collaborative approach where query modification is made by group members.