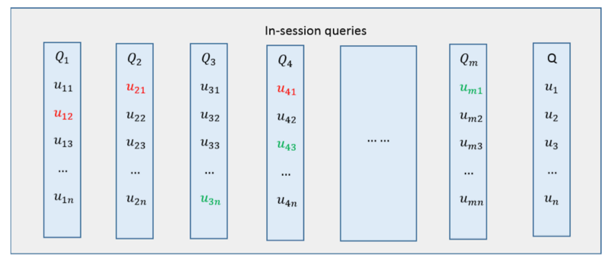
Deep Learning Powered In-Session Contextual Ranking using Clickthrough Data

Xiujun Li[1], Chenlei Guo[2], Wei Chu[2], Ye-Yi Wang[2], Jude Shavlik[1]
[1]Department of Computer Sciences, University of Wisconsin-Madison
[2]Microsoft Bing, Redmond, Washington, U.S.A.

Problem

1. $Q_1, Q_2, \dots, Q_m$ are the history queries in one session.
2. $u_{i1}, u_{i2}, \dots, u_{in}$ are the URLs in one impression page for query Q_n, it contains sat and unsat clicks.
3. Q is the current query.
4. $u_1, u_2, \dots, u_n$ are the URLs in the impression page for query Q based on the current ranking algorithm.

Goal: re-rank the $u_1, u_2, \dots, u_n$ based on some signals (i.e. topic, domain) derived from the in-session clickthrough data.



An Example

Table 1 shows two queries (2nd and 6th queries) in one session from real log data. The difference between query 2 and query 6 is query specification by adding the term of "wiki" in query 6. In query 2, the corresponding Wikipedia page was clicked as a unsatisfied click (red color in the table), and in this case it is good to demote the Wikipedia page from the first position of the 6th query since it was examined as a unsatisfied page in query 2, while, in reality, the wikia page in the second position of the query 6 is a satisfied click (green color in the table) from users' log. Thus, it is better to promote the position of wikia page in the ranking result.

Query 2: the dangerous days of daniel x	Query 6: the dangerous days of daniel x Altreda wiki
http://en.wikipedia.org/wiki/The_Dangerous_Days_of_Daniel_X	http://en.wikipedia.org/wiki/The_Dangerous_Days_of_Daniel_X
http://www.amazon.com/The-Dangerous-Days-Daniel-X/dp/0316119709	**http://fanon.wikia.com/wiki/The_Dangerous_Days_of_Daniel_X**
http://www.goodreads.com/book/show/2235597.The_Dangerous_Days_of_Daniel_X	http://www.amazon.com/The-Dangerous-Days-Daniel-X/dp/0316119709
http://www.daniel-x.co.uk/books/dangerous-days/	http://danielx.wikia.com/wiki/The_Dangerous_Days_of_Daniel_X_(novel)
http://www.freebooknotes.com/summaries-analysis/the-dangerous-days-of-daniel-x/	http://danielx.wikia.com/wiki/Daniel_X
http://www.jamespatterson.com/books_danielX.php#.VCR0vOfUe1A	http://www.goodreads.com/series/49946-daniel-x
http://jamespatterson.com/books_daniel_x.php#.VCR01efUe1A	http://en.wikipedia.org/wiki/Daniel_X:_Watch_the_Skies
http://books.google.com/books/about/The_Dangerous_Days_of_Daniel_X.html?id=2UBONTvr_BEC	http://www.jamespatterson.com/books_danielX.php

Table 1. Two queries from the current ranker

Our Approach

1. Correlate users' clicks with their satisfaction and derive a set of fine-grained features (i.e. URL-level discriminative features, Click-Based features) to explore the relations between in-session history queries, clicked URLs and current query, URLs.
2. Employ the semantic deep learning models (i.e. XCode, DSSM, CDSSM) to measure the similarity for the semantic features.
3. Incorporate these features into the ranking model, to re-rank the results.

Step 1. Fine-grained Features

Features	Level	Description	Category
Semantic Features	URL-URL	The similarity between the URLs of previous queries and URLs of current query	sat
			unsat
	Query-URL	The similarity between the previous queries and URLs of current query	sat
			unsat
	Weighted URL-URL	The weighted similarity between the URLs of previous queries and URLs of current query	sat
			unsat
Click-Based Features	URL Level	The domain and history click statistics of current URL in the previous queries	history click
			domain click
	Impression Level	The sat/unsat statistics of in-session history queries	clicked URL
			clicked count

- **Semantic features** measure the similarity between previous queries, clicked URLs to the current query, URLs from a semantic level. All of them are URL-level and can distinguish one URL from another.
- **Click-Based features** encode the user interaction behaviors and content preferences in web search, e.g. unsatisfied/satisfied click count, domain unsatisfied/satisfied click count etc.

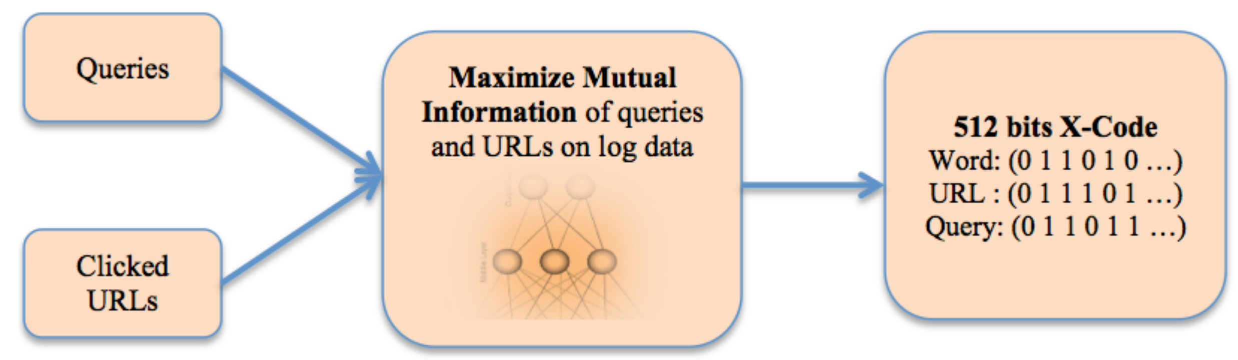
Acknowledgements

1. The first author would thank the support and funding from Microsoft Research.
2. JS was supported by DARPA Grant FA8750-13-2-0039.

Step 2. Semantic Features Calculation

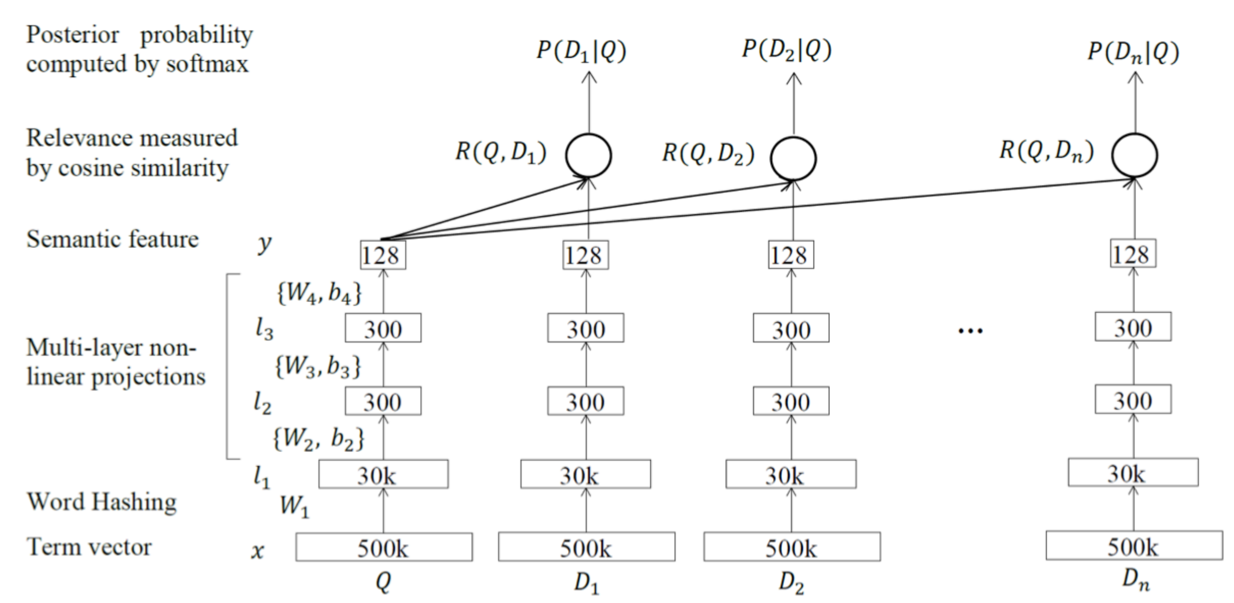
Model 1 XCode

- Map every word, query, URL into an intent space.
- Each URL, query and word has a unique 512-bit XCode.
- Measure the semantic similarity in this intent space.



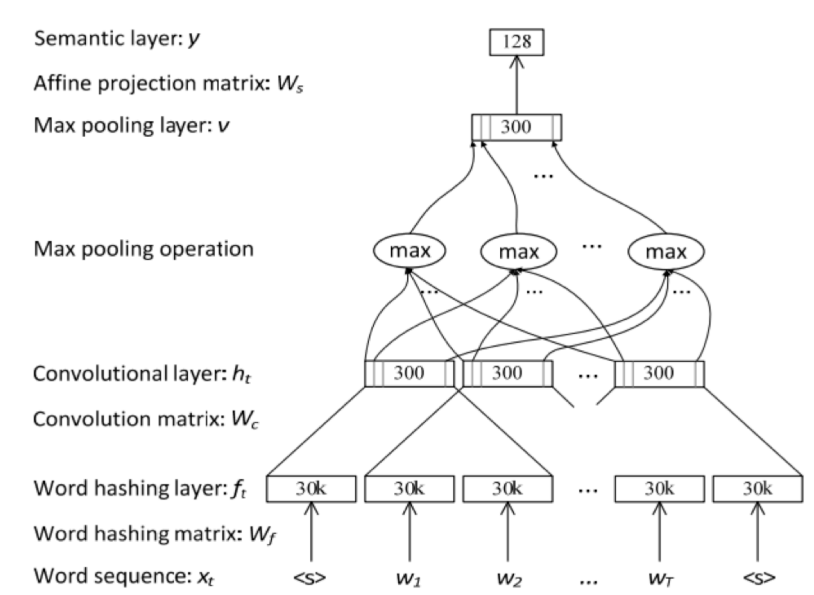
Model 2 DSSM

- **Representation**: use deep neural networks to extract abstract semantic representations.
- **Learning**: maximize the similarity between the query and the documents.
- **Scalability**: use sub-word unit (i.e. letter-ngram) as raw input feature to handle a unlimited vocabulary (*word hashing*).



Model 3 Convolutional DSSM

- **Difference** between CDSSM and DSSM
  1. Sliding window: capture the contextual feature for a word.
  2. Convolutional layer: project each word within a context window to a local contextual feature vector.
  3. Max pooling layer: extract the most salient local features to form a fixed-length global contextual feature vector.



- Both DSSM and CDSSM provides a way to encode how well contextual information is matched at the semantic level.
  1. Map queries/titles/brokenURLs to a common low dimensional semantic space using DNN
  2. Leverage the contextual information, not depend only on context terms

CDSSM / DSSM Double Model Training		
Training Data	**Source Model**	**Target Model**
<Query, Title>	Query Model	Title Model
<Query, BrokenURL>	Query Model	BrokenURL Model

- Model Semantics Comparison

Model	Distance	Range	Semantics
XCode	Hamming Distance	[0, 512]	0 is the most similar, 512 is the least similar
DSSM	Cosine Distance	[-1, 1]	1 is the most similar, -1 is the least similar
CDSSM	Cosine Distance	[-1, 1]	1 is the most similar, -1 is the least similar

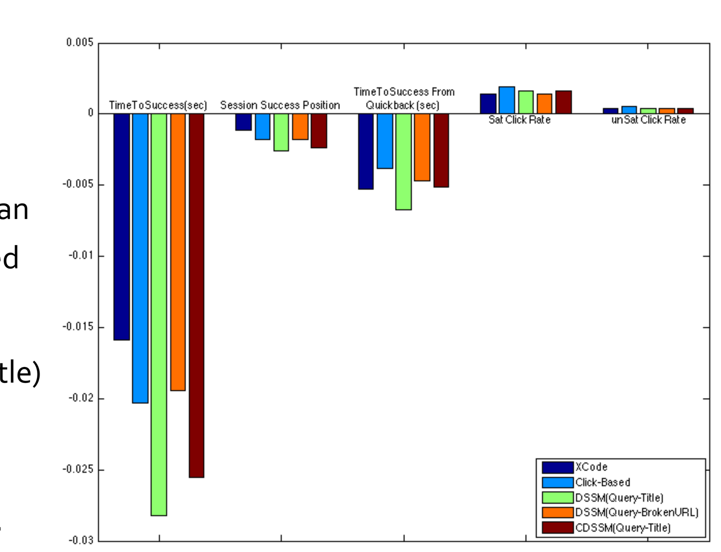
Click-Based Features

Level	Category	Semantics	Features
URL Level	History Click	How many sat/unsat clicks for the current URL in the previous queries in the same session	clickcount_url_total clickcount_url_sat clickcount_url_unsat
	Domain Click	How many sat/unsat clicks for the domain of current URL in the previous queries in the same session	domain_click_total domain_click_sat domain_click_unsat
Impression Level	Clicked URL	Some statistics information from in-session history queries	sat_click_urls unsat_click_urls
	Clicked Count		sat_click_count unsat_click_count

Step 3. Performance Evaluation

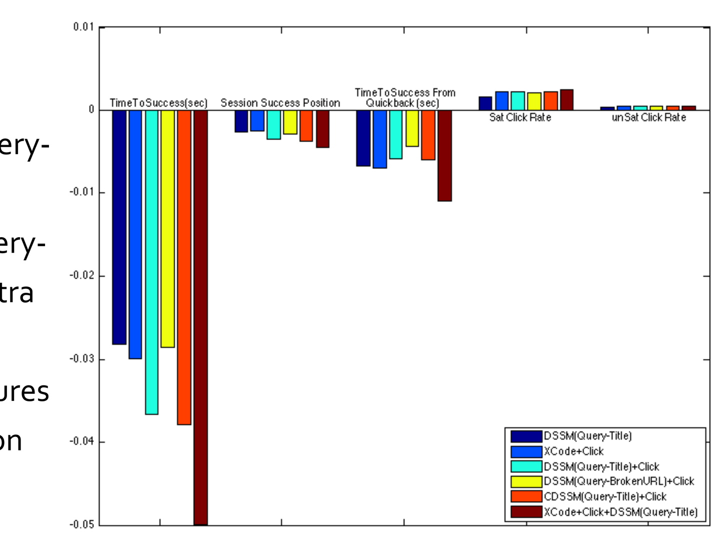
Individual Features

- Conduct large-scale analysis on real clickthrough log data, 15 days raw data, around 2TB.
- Measure the performance using the mean average precision (MAP) of the re-ranked lists.
- For Semantic features, DSSM (Query-Title) is the best group.
- The Click-Based features is better than XCode features, and easy to be flighted.



Combined Features

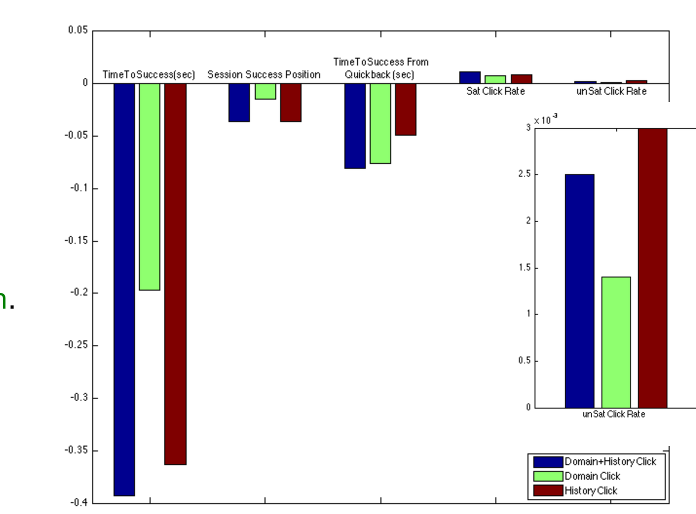
- Combined features is better than any individual group of features.
- For two groups of features, CDSSM (Query-Title) + Click-Based is the best.
- The Combination of XCode + DSSM(Query-Title) + Click-Based features can gain extra improvement.
- Semantic features and Click-Based features can provide different context information to the web search task.



Click-Based Features

Conduct a depth analysis for the domain clicks features and history clicks features:

- Domain Click features can effectively decrease the unsat clicks.
- History Click features can improve the TimeToSuccess, SessionSuccessPosition.



Revisit the Example

Since the Wikipedia domain has 5 unsatisfied clicks in the query 2, the new ranker will demote the Wikipedia page, and promotes the wikia page in the query 6. And this wikia page is a satisfied click at the query 6, perfectly meeting the user's needs.

Query 6: the dangerous days of daniel x Altreda wiki	Query 6: the dangerous days of daniel x Altreda wiki
http://en.wikipedia.org/wiki/The_Dangerous_Days_of_Daniel_X	**http://fanon.wikia.com/wiki/The_Dangerous_Days_of_Daniel_X**
http://fanon.wikia.com/wiki/The_Dangerous_Days_of_Daniel_X	**http://en.wikipedia.org/wiki/The_Dangerous_Days_of_Daniel_X**
http://www.amazon.com/The-Dangerous-Days-Daniel-X/dp/0316119709	http://danielx.wikia.com/wiki/The_Dangerous_Days_of_Daniel_X_(novel)
http://danielx.wikia.com/wiki/The_Dangerous_Days_of_Daniel_X_(novel)	http://www.goodreads.com/series/49946-daniel-x
http://danielx.wikia.com/wiki/Daniel_X	http://danielx.wikia.com/wiki/Daniel_X
http://www.goodreads.com/series/49946-daniel-x	http://www.amazon.com/The-Dangerous-Days-Daniel-X/dp/0316119709
http://en.wikipedia.org/wiki/Daniel_X:_Watch_the_Skies	http://www.jamespatterson.com/books_danielX.php
http://www.jamespatterson.com/books_danielX.php	http://en.wikipedia.org/wiki/Daniel_X:_Watch_the_Skies

Table 2. The new ranking results, the left is based on baseline ranker, the right is from the new ranker.

Discussion

- In the Individual features, it shows DSSM model is slight better than CDSSM, while we expect CDSSM would provide better performance. Some potential reasons follow:
  1. DSSM was trained a very large dataset, while CDSSM can only be trained on a small dataset with a up-bounded size (i.e. 1M queries); the size of training dataset might affect the prediction ability of the model.
  2. Due to time constraint, not all parameters of CDSSM in the experiment were optimally tuned.
- All the experiments were conducted offline over real clicked log data, we are going to verify the gain in the online settings.