A Click Model for Time-sensitive Queries

Seung Eun Lee
Daum Communications
Daum Space.1 2181 Yeongpyeong-dong
Jeju-si, Jeju-do, South Korea
seunglee@daumcorp.com

Dongug Kim
Daum Communications
Daum Space.1 2181 Yeongpyeong-dong
Jeju-si, Jeju-do, South Korea
abia@daumcorp.com

ABSTRACT

User behavior on search results pages provides a clue about the query intent and the relevance of documents. To incorporate this information into search rankings, a variety of click modeling techniques have been proposed so far and now they are widely used in commercial search engines. For time-sensitive queries, however, applying click models can degrade the search relevance because the best document in the past may not be the current best answer. To address this problem, it is required to detect a time point, a *turning point*, where the search intent for a given query changes and to reflect it in click models. In this work, we devised a method to detect the turning point of a query from its search volume history. The proposed click model is designed to take only user behavior observed after the turning points. We applied our model in a commercial search engine and evaluated its relevance.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms

Algorithms, Experimentation

Keywords

Click modeling, Time-sensitive queries

1. INTRODUCTION

Click behavior on a search engine results page is an implicit feedback from users about search quality and document relevance. There have been many click models proposed to analyze this behavior and they are now widely adopted as ranking factors in commercial search engines. The underlying assumption here is that search results which satisfied users in the past are still good ones. However, the relevance of documents is not fixed and it can change over time. For instance, people who submitted a query about a weekly television series would be searching for contents about the latest episode. Hence, the most satisfactory document a week ago may be not the best any more.

WWW 2013 Companion, May 13–17, 2013, Rio de Janeiro, Brazil.
ACM 978-1-4503-2038-2/13/05.

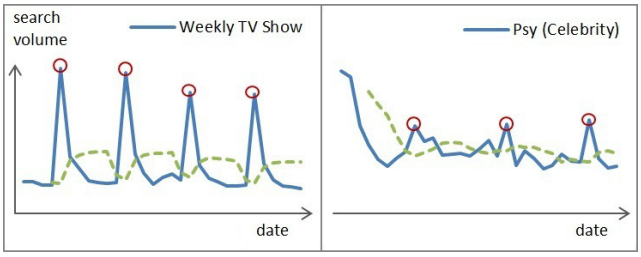

Figure 1: Search volume trends of queries. "Infinite Challenge", a popular Korean weekly TV show (Left) "Psy", a popular Korean singer (Right)

To address this problem, previous works proposed to give a weight according to the time click occurred [5] or to emphasize the real-time click feedback [4]. Unfortunately, there remain cases where these approaches do not work well. For example, Figure 1 shows the trend of search volume of a query indicating a weekly TV show broadcasted every Saturday. (Left) The popularity dramatically increased on every Saturday, peaked while and short after the show aired, and stayed low during weekdays. For a user who issued the query on Thursday, both past click behavior occurred 1 day ago and 3 days ago are equally important and there is no reason to give higher weights on more recent clicks. In addition, every time-sensitive query has its own temporal aspect. Periodically recurring queries have their cycles (ex. monthly magazines and yearly holidays) and breaking news queries have a pattern of getting popular in a short period. Dong et al [2] identified breaking news queries based on the temporal difference of query likelihood. Kulkarni et al [3] studied the relationship between the temporal popularity of a query and it search intent change. Ignoring these various temporal aspects of queries and applying a static weighting scheme in click models can degrade its performance.

To improve the timeliness of click analysis, we need to detect a moment where the search intent of a given query changes. We will call this the *turning point* of a query. The turning point of a TV show would be the time when the latest episode was on air. In the case of celebrities, the turning point is the moment they made news (ex. a singer released a new album.) In this work, we made two assumptions.

1. Every time-sensitive query has its own turning point, and the relevance of documents for the query can change since it.

2. At the turning point of a query, its search popularity will increase rapidly.

In this paper, we propose a click model based on a query-dependent turning point.

2. A BURST-AWARE CLICK MODEL

Our model consists of two modules. A click model is used to compute the relevance scores of documents. A turning point determines which data to use in the model.

2.1 Modeling User Click

There are click models proposed for predicting the document satisfaction from user behavior on search engine result pages ([1], [6].) In those researches, several measures are extracted for each (query-document) pair and the final score is computed from them. With reference to the studies, we extracted 4 features for each document in response to a query.

- CTR: click through rate of the document
- dwell-time: an average dwell time on the document
- short-rate: a rate that users left the document in short time. A high rate implies low quality of the document
- last-rate: a rate that users stopped clicking after clicking the document

We asked evaluation experts to rate documents for queries and then used it as training data for our regression model.

2.2 Detecting a Turning Point

Our click model takes only user behavior observed after the turning point. The turning point of a query is detected based on its search popularity burst. We used a simple moving window method. The turning point is defined as the last date when the search volume is 1.5 times larger than the average volume of past 5 days. In Figure 1, the dashed line is a moving average with window size 5 and the red circles are displayed on detected burst dates.

3. EVALUATION

The main idea of this work is to control the time span of click analysis for time-sensitive queries. To evaluate the proposed model (*d-burst*) which uses only click behavior after the turning point, we compared it with three baseline models that take behavior of the fixed intervals: 1, 7, and 30 day(s). (*d-1*, *d-7*, and *d-30*, respectively) Another baseline is a model taking behavior before the turning point (*d-old*.)

We conducted an evaluation on a commercial search engine. Most of major search engines already have their own strategies to provide fresh search results for temporal queries. This experiment is designed to measure and compare the performance of click models on this default ranking scheme.

3.1 Experiments

We collected queries whose turning points are between December 31, 2012 and January 29, 2013. Sample queries are randomly chosen so that their turning points are evenly distributed among dates. We used click models to compute document scores for these queries. 4316 documents for 474 queries were obtained and then evaluated by experts in 5-point scales on January 30, 2013. Evaluators were noticed that timeliness of documents should be reflected in rating.

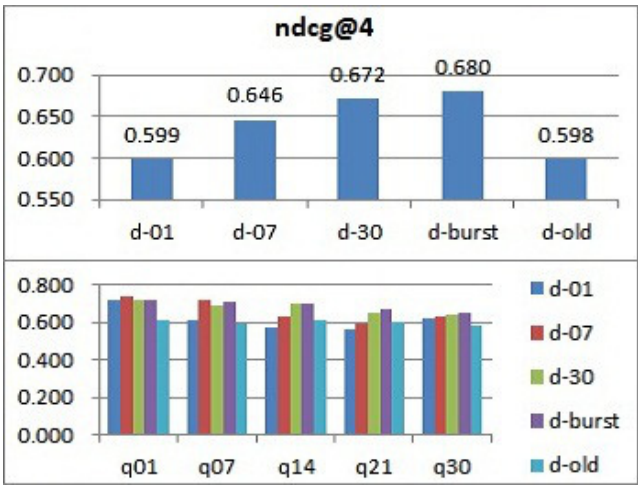

Figure 2: Evaluation results. q01 is a query group whose turning points are 1 day old. q07 is one whose turning points are between 2 and 7 days old.

We used NDCG@4 as an evaluation metric. For some queries, baseline models could not generate scores of any documents due to the lack of user behavior data. In this case, we ordered documents by random permutation.

3.2 Results

The evaluation result is summarized in Figure 2. In terms of average NDCG, the d-burst model showed the best result. We divided queries into five groups according to their turning points and compared the performance of each model. Our d-burst performed consistently well over all ranges of turning points. We found that d-7 is the best for queries having turning points in recent 7 days while d-burst and d-30 are better for queries with older ones. This result supports our hypothesis that every time-sensitive query has its own turning point and this query-specific temporal aspect needs to be incorporated into click models for finding timely relevant documents. Based on the results of this work, we have a plan to develop more sophisticated query turning point detection algorithms.

4. REFERENCES

[1] O. Chapelle and Y. Zhang. A dynamic bayesian network click model for web search ranking. WWW '09.

[2] A. Dong, Y. Chang, Z. Zheng, G. Mishne, J. Bai, R. Zhang, K. Buchner, C. Liao, and F. Diaz. Towards recency ranking in web search. WSDM '10

[3] A. Kulkarni, J. Teevan, K. M. Svore, and S. T. Dumais. Understanding temporal query dynamics. WSDM '11

[4] T. Moon, L. Li, W. Chu, C. Liao, Z. Zheng, and Y. Chang. Online learning for recency search ranking using real-time user feedback. CIKM '10

[5] Y. Inagaki, N. Sadagopan, G. Dupret, C. Liao, A., Y. Chang, and Z. Zheng. Session based click features for recency ranking. AAAI '10

[6] F. Zhong, D. Wang, G. Wang, W. Chen, Y. Zhang, Z. Chen, and H. Wang. Incorporating post-click behaviors into a click model. SIGIR '10