Click models evaluations
(IR2 project)

Luka Stout
10616713

Finde Xumara
10690832

1 Introduction

Modeling user behavior on a search engine result page is important for understanding the users and supporting simulation experiments. As result pages become more complex, click models evolve as well in order to capture additional aspects of user behavior in response to new forms of result presentation. In recent years many models have been proposed that are aimed at predicting clicks of web search users.

In this report, we will implement and evaluate different click models using evaluation metrics. These click models include dependent click model (DCM) [3], dynamic bayesian network click model (DBN) [1], user browsing model (UBM) [2], and task-centric click model (TCM) [6]. The evaluation metrics include likelihood, perplexity, click through rate prediction, relevance prediction, computation time and ranking performance. By doing this experiment we can know the performance of each different click model and this information can be used as a benchmark of the next click model proposal.

This report is organized as follows. Section 2 describes the methods and algorithms used in our implementation. The experiments and data source information will be covered in Section 3, followed by the Section 4 for the performance analysis.

2 Methodology

We compare UBM, DCM, DBN and TCM on the Yandex dataset [4]. In this section, we briefly describe their main characteristics and differences. We have implemented TCM ourselves, the other algorithms were taken from PyClick [5].

2.1 DCM

The Dependent Click Model (DCM) was first proposed by Guo et al. in [3]. In the paper they propose a new click model which can handle multiple clicks per query by introducing a position dependent parameter $lambda_j$ to reflect the chance that the user would like to see more results after a click at position j. A graphical representation of the model is presented in Figure1

2.2 DBN

The Dynamic Bayesian Network is an extension to the traditional Cascade model proposed by Chapelle and Zhang in [6]. For a given position j, in addition to observed variable C_j indicating whether there was a click or not at this position, the following latent variable are defined to model examination, perceived relevance and actual relevance, respectedly:

- E_j: did the user $examine$ the document?
- A_j: was the user $attracted$ by the document?
- S_j: was the user $satisfied$ by the clicked document?

They introduce a variable s_u for each document u which describes the relevance of the document for this query. When the user clicks on this document, there is a certain chance that the user will be satisfied. If the user is not satisfied, he

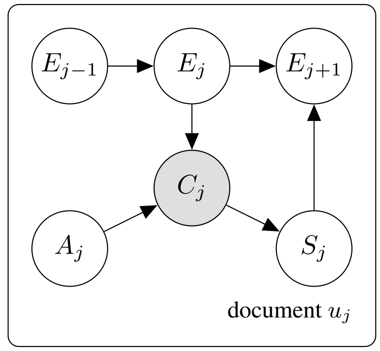

Figure 1: The graphical model of DCM.

continues to examine the next document with a probability γ and stops otherwise. The parameter γ is known as the 'perseverance'. A graphical representation of the model is presented in Figure 2.

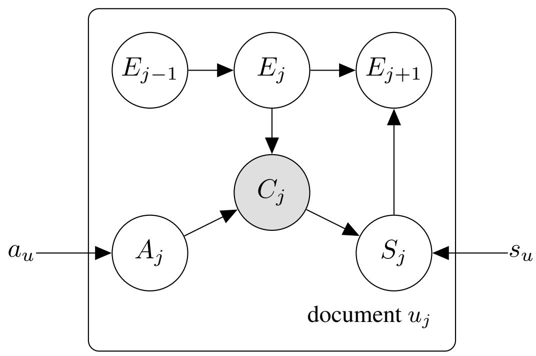

Figure 2: The graphical model of DBN.

2.3 UBM

In [2], Dupret and Piwowarski propose a new click model called the User Browsing Model (UBM). The main difference between UBM and other models is that UBM takes the distance into account from the current document u_j to the last clicked document $u_{j'}$ for determining the probability that the user continues browsing:

$$P(E_j = 1 \mid C_{j'} = 1, C_{j'+1} = 0, \ldots, C_{j-1} = 0) = \gamma_{jj'}$$

The probability that a document at rank r is examined E_j therefore depends on all possible paths the user could have taken to arrive at this document:

$$P(E_j = 1) = \sum_{i=1}^{j-1} \gamma_{ji}$$

A graphical representation of the model is presented in Figure **??**.

2.4 TCM

3 Evaluation

3.1 Evaluation Criteria

In order to compare performances of our systems we used the ...

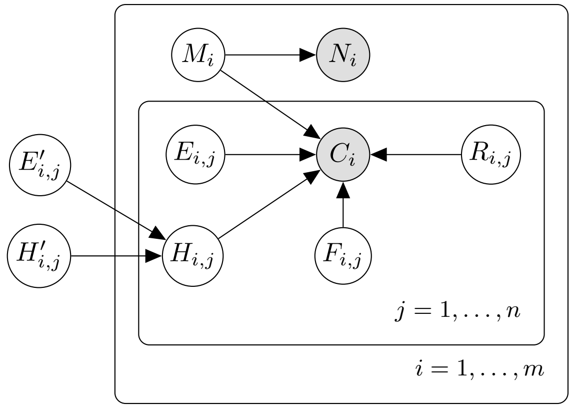

Figure 3: The graphical model of TCM.

3.2 Dataset

The dataset used are ...

Log listing 1: Example rows from XXXX

```
0 qid:18219 1:0.052893 2:1.000000 3:0.750000 4:1.000000 ...  46:0.966667
     #docid = GX004-93-7097963 inc = 0.0428115405134536 prob = 0.860366
1 qid:18219 1:0.026446 2:0.750000 3:0.750000 4:0.500000 ...  46:0.266667
     #docid = GX020-25-8391882 inc = 1 prob = 0.115043
0 qid:18219 1:0.029752 2:0.000000 3:1.000000 4:1.000000 ...  46:0.100000
     #docid = GX025-94-0531672 inc = 1 prob = 0.141903
```

3.3 Evaluation setup

The experiment was run ...

3.4 Results

In Table 1 one can see the results of the experiments. It can be seen that ...

	Log-likelihood	Perplexity	Computation Time
Click Model A	0	0	
Click Model B	0	0	

Table 1: Results

4 Analysis

After running the experiments we were able to evaluate the different algorithms based on the ...

5 Conclusions

In this paper we showed that ...

In our implementation, we did not ...

References

[1] Olivier Chapelle and Ya Zhang. A Dynamic Bayesian Network Click Model for Web Search Ranking Categories and Subject Descriptors. *Www*, pages 1–10, 2009.

[2] Georges E Dupret and Benjamin Piwowarski. A user browsing model to predict search engine click data from past observations. *Sigir '08*, pages 331–338, 2008.

[3] Fan Guo, Chao Liu, and Yi Min Ym Wang. Efficient multiple-click models in web search. *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pages 124–131, 2009.

[4] Yandex LLC. Task and datasets, n.d.

[5] Ilya Markov. Click models for web and aggregated search, 2015.

[6] Yuchen Zhang, Weizhu Chen, Dong Wang, and Qiang Yang. User-click modeling for understanding and predicting search-behavior. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11*, page 1388, 2011.