ICTNET at Session Track TREC 2013

Zhenhong Chen[1,2], Long Xia[1,2], Xiaoming Yu[1], Yue Liu[1], Xueqi Cheng[1]

1. Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190
2. University of Chinese Academy of Sciences, Beijing, 100190

{chenzhenhong,xialong, yuxiaoming}@software.ict.ac.cn
{liuyue,cxq}@ict.ac.cn

1 Introduction

In this paper, we describe our solutions to the Session Track at TREC 2013. There are three main differences compared to our last year's submission[2]. Firstly, we use Indri[3] to build our retrieval system. Due to computing resource limited, we only index the Category B collection. Secondly, we try to take advantage of FreeBase[4] to recognize the entities in the given query and then weight each term or phrase accordingly. Lastly, we discard the Google virtual document and page rank features from our last year's learning to rank model. The rest of this paper is organized as follows.We detail our research structure in section 2. Section 3 describes our experiments and evaluation results. Conclusions are made in the last section.

2 Our approach

Our research structure of TREC 2013 Session Track is listed in table 1.

Table 1: Methods of submitted runs

	RL1	RL2	RL3
ICTNET13SER1	Indri SpamFilter	Indri SpamFilter SVM^{rank}	Indri SpamFilter Query Expansion SVM^{rank}
ICTNET13SER2	Indri SpamFilter Freebase Weighted	Indri SpamFilter SVM^{rank} Freebase Weighted	Indri SpamFilter Query Expansion SVM^{rank} Freebase Weighted
ICTNET13SER3	Indri SpamFilter Freebase Weighted NoveltyFilter	Indri SpamFilter SVM^{rank} Freebase Weighted NoveltyFilter	Indri SpamFilter Query Expansion SVM^{rank} Freebase Weighted NoveltyFilter

2.1 Indri

We use Indri to index the Clueweb12 Category B collection. The default query model for the current query in the session is calculated as (1).

$$p(w|q) = \frac{count(w, q_m)}{\sum_{w' \in q_m} count(w', q_m)}$$

2.2 Spam Filter

We use Waterloo spam ranking score [5] to filter documents with "fusion" spam score [6] less than 60%.

2.3 Learning to Rank

We implement the SVM^{rank} [7] to learn from explicit relevance judgments on TREC 2011 Session Track. Used features are listed in table 2. Detail about each feature please referred to [2].

Table 2: Features Used in SVM^{rank}

Feature	Feature Description
QE	the score of Query Expansion Model
SessionVD	the score of Session Virtual Document Model
CAT	the score of Optimization Based on User's Attention Time Model
CosSimQT	cosine similarity between query and title
BM25QC	BM25 score between query and content

2.4 Query Expansion

Our query expansion uses the historical queries and the current query to construct the indri query. Let q_1 to q_{m-1} stand for previous queries and q_m stands for the current query, then $p(w|q)$ denotes the weight of the word w in indri query q, calculating as (2).

$$p(w|q) = \lambda \frac{count(w, q_m)}{\sum_{w' \in q_m} count(w', q_m)} + (1-\lambda) \sum_{i=1}^{m-1} \frac{count(w, q_i)}{\sum_{w' \in q_i} count(w', q_i)} \quad (2)$$

In our experiments, we set λ as 0.7.

2.5 Freebase Weighted

Given a query q, we get unigram, bigram, 3-gram and 4-gram phrases as the entity candidates. Then we use Freebase API to query with each candidate and get the first ten results. For unigram candidates, if there is a result's notable id contains "location" and "country" or "location" and "citytown", we add it to the entity set ES_1. For other candidates, if there is a result's name equal to the given candidate, we add it to the entity set ES_{234}. Finally, the query model with Freebase Weighted is estimated as follows.

$$p(w|q) = \lambda \frac{count(w, q_m) + \beta_1 * count(w, q_m - ES_1 - ES_{234}) + \beta_2 * 1_{w \in ES_1}}{\sum_{w' \in q_m} count(w', q_m)} + (1-\lambda) \sum_{i=1}^{m-1} \frac{count(w, q_i) + \beta_1 * count(w, q_i - ES_1 - ES_{234})}{\sum_{w' \in q_i} count(w', q_i)}$$

$$p(e|q) = \lambda \frac{\beta_3 * 1_{e \in q_m \cap ES_{234}}}{\sum_{w' \in q_m} count(w', q_m)} + (1-\lambda) \sum_{i=1}^{m-1} \frac{\beta_3 * 1_{e \in q_i \cap ES_{234}}}{\sum_{w' \in q_i} count(w', q_i)}$$

In our experiments, we set λ=0.7, β_1=0.2,β_2=0.5,β_3=0.1.

2.6 Novelty Filter

We filter documents that had been shown in previous queries of the session.

3 Experiments

In this section, we firstly introduce the preprocess we apply to clean the session data. Then we will discuss the evaluation results on TREC 2013 Session Track.

3.1 Session Data Preprocess

For historical queries in the same session, if the edit distance between q_i and q_{i+1} is not greater than one, then we discard the interaction with fewer clicks. If two interactions have same clicks we discard the previous. For historical interactions in the same session, if they share the same results, then we discard the interaction with fewer clicks. If two interactions have same clicks we discard the previous.

3.3Evaluation Results

Evaluation results of our submissions at TREC 2013 Session Track are showed in Table 3.The highest score for each experimental condition is indicated in bold. According to Table 3, we observe consistentlyimprove as last year's results. We obtain **42.65%** of performance increase when we compare ICTNET13SER1.RL2 with ICTNET13SER1.RL1.

Table 3: Results on 2012 Session Track, in terms of NDCG@10

	RL1	RL2	RL3
ICTNET13SER1	0.1170	0.1669	**0.1659**
ICTNET13SER2	**0.1179**	**0.1670**	0.1649
ICTNET13SER3	0.1179	0.1617	0.1608

4 Conclusions

In this paper, we detail our work at TREC 2013 Session Track. The evaluation results show that our approaches can significantly improve the effectiveness of the search results. Though, it is still far away to outperform the state-of-the art. For the future work, we will focus on using Freebase more effectively to understand user's behavior and intent in the search session.

5 Acknowledgements

We would like to thank all organizers and assessors of TREC and NIST. This work is sponsored by NSF of China Grants No. 61100083&61272536, and by 242 Program of China Grants No. 2012F124, and by the National Key Technology R&D Program (2012BAH39B04).

6 References

[1] Session Track 2012 Overview

[2] ICTNET at Session Track TREC 2012

[3] http://www.lemurproject.org/indri/

[4] http://www.freebase.com/

[5] Gordon V. Cormack, Mark D. Smucker, and Charles L. A. Clarke. Efficient and effective spam filtering and re-ranking for large web datasets. *Information Retrieval*, 14(5), 2011: 441-465.

[6] http://www.mansci.uwaterloo.ca/~msmucker/cw12spam/

[7] T. Joachims, *Training Linear SVMs in Linear Time,* Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD), 2006.