The Effectiveness of Results Re-Ranking and Query Expansion in Cross-language Information Retrieval

Dong Zhou Vincent Wade
Centre for Next Generation Localisation, University of Dublin, Trinity College, Dublin 2, Ireland
dongzhou1979@hotmail.com Vincent.Wade@cs.tcd.ie

ABSTRACT
This paper presents the technique details and experimental results of the information retrieval system with which we participated at the NTCIR-8 ACLIA (Advanced Cross-language Information Access) IR4QA (Information Retrieval for Question Answering) task. Document corpus in Simplified Chinese (CS) and Traditional Chinese (CT) with topics in English, CS and CT were used in our experiments. We combined the query expansion and initial retrieval results re-ranking techniques as main retrieval approach. The experimental results confirmed that query expansion based on Bose-Einstein distribution and re-ranking method based on Latent Dirichlet Allocation (LDA) are able to consistently bring significant improvements over various baseline systems. Especially the approach is capable of processing mixed-multilingual text obtained by a machine translator for cross-language information retrieval (CLIR). The results obtained might provide us more insight and understanding into cross-language query expansion and document re-ranking.

Categories and Subject Descriptors
H.3.3 [**Information Systems**]: Information Storage and Retrieval - *Information Search and Retrieval;* I.2.7 [**Computing Methodologies**]: Artificial Intelligence - *Natural Language Processing.*

General Terms
Algorithms, Measurement, Performance, Experimentation, Languages.

Keywords
Cross-language Information Retrieval, Query Expansion, Document Re-Ranking, Latent Dirichlet Allocation, Machine Translation.

1. INTRODUCTION
This year's participation in the NTCIR-8 IR4QA monolingual and bilingual track [7] was motivated by a desire to test and extend a newly developed cross-language document re-ranking method, together with query expansion technologies. The merge of these two methods has proven to improve the retrieval performance in cross-language settings in European languages [10]. This paper tries to extend the previous approach from two aspects. One is to apply these techniques to two very different languages Chinese and English for cross-language information retrieval. Another one is to using mixed-languages obtained by a machine translator for cross-language document re-ranking.

Firstly we consider a typical information retrieval (IR) re-ranking problem, let $\mathbb{D} = \{d_1, d_2, \dots, d_n\}$ denote the set of documents to be retrieved. Given a query q, a set of initial results $\mathbb{D}_{init} \in \mathbb{D}$ of top documents, sorted in a decreasing order of their relevancy, are returned by a standard IR model (initial ranker). However, the initial ranker tends to be imperfect. The purpose of the re-ranking stage is to re-order a set of documents $\mathbb{D}'_{init}$ so as to improve retrieval accuracy at the very top ranks of the final results. IR systems capable of achieving such goal would be of obvious benefit to human users, and could also aid query expansion approaches, question-answering systems, and other applications that use IR engines as a core component.

In the setting of cross-language IR, this problem becomes more complicated. Cross-language information retrieval (CLIR) focuses on research in the retrieval of documents written in languages different to the language in which the query is expressed. The overwhelming majority of CLIR systems in existence today function via query translation, which is adopted in our experiments. Assume that query translation is employed, the simplest approaches to address the cross-language re-ranking problem is probably directly apply the monolingual methods on the results obtained by using the translated query. Obviously, the drawback of this approach is that, errors inherited from the translation noise will be passed to the re-ranking process, resulting in unsatisfactory performance.

In order to take the query language and the document language into consideration, external resources will be required to conduct an appropriate representation space. In this paper, we attempt a simple strategy to obtain this multilingual text. That is to translate the returned results back to the query language and enhance the re-ranking performance.

Based on above discussion, we participated in the CS-CS (Simplified Chinese), EN-CS, CT-CT (Traditional Chinese) and EN-CT subtasks of IR4QA. We applied direct re-ranking and mixed-language re-ranking in selected English-Chinese runs. Query expansion was employed across several permutations, with or without re-ranking process. The official evaluation results suggest that our method achieves very good performance. It shows that we managed to cover 95% of known relevant documents in CS runs and 96% of known relevant documents in CT as well as we ranked the top for many CS and CT runs.

The rest of the paper is organized as follows. Section 2 describes the methodology used in our experiments together with technical details. System description and framework are briefly summarized in section 3. In Section 4 we report on the official results of experiments performed over NTCIR test collections in simplified and traditional Chinese. Finally, Section 5 concludes the paper and speculates on future work.

2. Methodology

2.1 Monolingual Retrieval

Firstly we describe the monolingual retrieval models used in the experiment. All of the retrieval runs ran queries (translated or original) against the document collections using the Terrier toolkit [5] and the following retrieval functions:

- DLH13

$$Score(q,d) = \sum_{t \in q} tf_q(t) \frac{tf_d(t) log \frac{tf_d(t) avg|d| |\mathbb{D}|}{|d| tf_{\mathbb{D}}(t)} + 0.5 log(2\pi tf_d(t)\left(1 - \frac{tf_d(t)}{|d|}\right))}{tf_d(t) + 0.5}$$

This model is a parameter-free weighting model. Even if the user specifies a parameter value, it will not affect the results.

- BB2

$$Score(q,d) = \sum_{t \in q} tf_q(t) \frac{tf_{\mathbb{D}}(t) + 1}{df(t)(ntf_d(t) + 1)} tf_q(t) \, (-log(|\mathbb{D}| - 1) + \emptyset(|\mathbb{D}| + tf_{\mathbb{D}}(t) - 1, |\mathbb{D}| + tf_{\mathbb{D}}(t) - ntf_d(t) - 2) - \emptyset(tf_d(t), tf_d(t) - ntf_d(t))$$

$$ntf_d(t) = tf_d(t) log(1 + \frac{avg|d|}{|d|})$$

$$\emptyset(n,m) = m + 0.5 log \frac{n}{m} + (n - m) log n$$

- BM25

$$Score(q,d) = \sum_{t \in q} w_t \frac{(k_1 + 1) tf_d(t)}{(K + tf_d(t))} \frac{(k_3 + 1) tf_d(t)}{k_3 + tf_q(t)}$$

$$w_t = log \frac{|\mathbb{D}| - df(t) + 0.5}{df(t) + 0.5}$$

$$K = k_1((1 - b) + b \frac{|d|}{avg|d|})$$

Where w_t is the Robertson/Sparck Jones weight of t, $k1$, b, and $k3$ are parameters (set to 1.2, 0.75, and 8 respectively). $tf_q(t)$, $tf_d(t)$ and $tf_{\mathbb{D}}(t)$ stand for term frequency in a query, a document and a document collection, respectively. $df(t)$ is document frequency of a given term. $|d|$ and $|\mathbb{D}|$ represent document length and number of documents and $avg|d|$ stands for average document length.

2.2 Query Expansion

The query expansion usually refers to the technique that uses blind relevance feedback to expand a query with new query terms, and reweigh the query terms, by taking into account a pseudo relevance set (usually top-ranked documents) [6]. The mechanism is based on Divergence from Randomness (DFR) theory. Like the language model approach of Ponte and Croft, a nonparametric model is derived as a combination of different probability distributions. The DFR paradigm [1] is a generalization of Harter's 2-poisson indexing model [2]. In the DFR approach, a query term is weighted by how different its term distribution in the document *d* is, compared to the whole collection. The more divergence of the within document term frequency from its frequency within the collection, the more the information carried by term *t* in document *d*. Currently, Terrier deploys the Bo1 (Bose-Einstein 1), Bo2 (Bose-Einstein 2) and KL (Kullback-Leibler) term weighting models. The DFR term weighting models follow a parameter-free approach in default. From an in-house experiment over NTCIR-7 IR4QA test collections[1], Bo1 model seems to perform best in average and is defined as:

$$inf E_q(t) = -log\left(\frac{1}{1 + \lambda_{E_q}}\right) - tf_{E_q} log\left(\frac{\lambda_{E_q}}{1 + \lambda_{E_q}}\right)$$

$$\lambda_{E_q} = \frac{tf_{E_q}}{N}$$

Where E_q denotes the elite set of the query, the set of topmost documents satisfying the query q according to some weighting function. It was adopted as main query expansion approach. In addition, we choose 20 terms added to the source query from top 5 documents returned with or without re-ranking process.

2.3 Document Re-Ranking

We adopted a document re-ranking method based on Latent Dirichlet Allocation (LDA) [11] which exploits implicit structure of the documents with respect to original queries. Rather than relying on graph-based techniques to identify the internal structure, the approach tries to directly model the latent structure of "topics" or "concepts" in the initial retrieval set. Then we can compute the distance between queries and initial retrieval results based on latent semantic information inferred. To prevent the problem of topic drift, the generative probability of a document is summed over all topics induced. By combining the initial retrieval scores calculated by language models, we are able to gather important information for re-ranking purposes. The intuition behind this method is the hidden structural information among the documents: *similar documents are likely to have the same hidden information with respect to a query*. In other words, if a group of documents are talking about the same topic which shares a strong similarity with a query, in our method they will get allocated similar ranking as they are more likely to be relevant to the query. In addition, the refined ranking scores should be relevant to the initial ranking scores, which, in our method, are combined together with the re-ranking score either using a linear fashion or multiplication process.

Our method is based on LDA. The basic generative process of LDA closely resembles PLSA [3]. LDA extends PLSA method by defining a complete generative model of text. The topic mixture is drawn from a conjugate Dirichlet prior that remains the same for all documents. The process of generating a document corpus is as follows:

1) Pick a multinomial distribution $\vec{\varphi}_z$ for each topic k from a Dirichlet distribution with hyperparameter $\vec{\beta}$.

2) For each document d, pick a multinomial distribution $\vec{\theta}_d$, from a Dirichlet distribution with hyperparameter $\vec{\alpha}$.

[1] http://research.nii.ac.jp/ntcir/ntcir-ws7/ws-en.html

3) For each word token w in document d, pick a topic $z \in \{1 \dots k\}$ from the multinomial distribution $\vec{\theta}_d$.

4) Pick word w from the multinomial distribution $\vec{\varphi}_z$.

Thus, the likelihood of generating a corpus is:

$$p(d_1, \dots, d_n|\vec{\alpha}, \vec{\beta}) = \iint \prod_{d=1}^{n} p(\vec{\theta}_d|\vec{\alpha}) \cdot \prod_{z=1}^{k} p(\vec{\varphi}_z|\vec{\beta}) \cdot \prod_{i=1}^{N_d} \sum_{z_i=1}^{k} p(z_i|\vec{\theta}_d) p(w_i|z, \vec{\varphi}_z)\, d\vec{\theta}_d d\vec{\varphi}_z$$

Unlike PLSA model, LDA possesses fully consistent generative semantics by treating the topic mixture distribution as a k-parameter hidden random variable. LDA offers a new and interesting framework to model a set of documents. The documents and new text sequences (for example, queries) could be easily connected by "mapping" them to the topics in the corpus. In the next subsection we will introduce how to achieve this goal and apply it to document re-ranking.

In the re-ranking setting, we estimate that the probability of a document d generates w, using a mixture model LDA. It uses a convex combination of a set of component distributions to model observations. In this model, a word w is generated from a convex combination of some hidden topics z:

$$LDA_d(w) = \sum_{z=1}^{k} p(w|z)p(z|d)$$

where each mixture model $p(w|z)$ is a multinomial distribution over terms that correspond to one of the latent topics z. This could be generated to give a distribution on a sequence of text:

$$LDA_d(w_1 w_2 \dots w_n) \stackrel{\text{def}}{=} \prod_{j=1}^{n} LDA_d(w_j)$$

Then the distance between a query and a document based on this model can be obtained. The method we propose here adopts the KL divergence between the query terms and document terms to compute a Re-Rank score RS_{LDA}^{KL1}:

$$RS_{LDA}^{KL1} = -D(MLE_q(\cdot)||LDA_d(\cdot))$$

This method also has the property of length-normalization to ameliorate long document bias problems [4]. The KL divergence is defined as:

$$D(p)||(q) = \sum_x p(x) log \frac{p(x)}{q(x)}$$

Motivated by the significant improvement obtained by [8] and [9], we formulate our method through a linear combination of the re-ranking scores based on initial ranker and the latent document re-ranker, shown as follow:

$$RS1 = (1 - \lambda) \cdot OS + \lambda \cdot RS_{LDA}^{KL1}$$

where OS denotes original scores returned by the initial ranker and λ is a parameter that can be tuned with $\lambda = 0$ meaning no re-ranking is performed.

Then we apply this method into the cross-language re-ranking by concatenating texts from different languages into several dual-language documents and a single dual-language query (in this case, by translating using a machine translator). An LDA analysis of these texts results in a multilingual semantic space in which terms from both languages are presented. Hence force the re-ranking process can be carried out by directly model the latent structure of multilingual "topics" or "concepts" in this enriched initial retrieval set. The similarity of "contexts" in which the terms appear is guaranteed to capture the inter-relationship between texts in different languages. To speed processing, only first 50 documents were re-ranked. We will report the effectiveness and shortage of this choice in the experimental section.

3. SYSTEM DESCRIPTION

Our final experimental system works by linking the two methods discussed above so that the output of the re-ranking process constitutes the input to the query expansion algorithm. A summary of the process chain is as follows:

(1) First, the target query is translated using a machine translator, a process which creates a translated query set. Google[2] translate was chosen for this purpose here.

(2) Next, a multilingual corpus with documents written in both query and document languages is built. Re-ranking is performed by applying the LDA based method on this multilingual space (with the translated and the original query).

(3) Next, the most significant terms are extracted from the top documents produced by step 2.

(4) Finally, the augmented query is passed to the normal IR engine involving three retrieval functions.

4. EXPERIMENTAL SETUP AND RESULTS

Our participation in NTCIR-8 is aimed at evaluating the following thematically related questions:

(1) Do the query expansion and document re-ranking techniques alone provide improvements over simpler techniques without relevance feedback and re-ranking?

(2) How the different retrieval models behaviour in the settings of English-Chinese information retrieval?

(3) Is the combination of these techniques (in the form of a hybrid retrieval method) effective when used in the context of a CLIR experiment?

4.1 Pre-processing

The most strenuous part of the text preprocessing stage involved character encoding and segmentation issues. All of the Chinese documents in our test collection were encoded using BIG5 and GB encodings. We first convert the encoding of all of the documents in the test collection to Unicode UTF-8. This was accomplished using an encoding converter written in the Java programming language. Following conversion, all of the Chinese documents were processed using a segmentation tool[3] and the English queries were sent to Google for translation. Finally, the test collection was indexed using the Terrier toolkit.

[2] http://www.google.com/translate

[3] http://www.mandarintools.com/

Table 1. Overview of NTCIR-8 Submitted Runs

Run Name	Model	Re-Ranking	Query Expansion	Fields
KDEG-CS-CS-01-T	BB2	N	Y	Title
KDEG-CS-CS-02-DN	DLH13	Y	Y	Title+Description
KDEG-CS-CS-03-T	BB2	N	N	Title
KDEG-CS-CS-04-T	BM25	N	N	Title
KDEG-CS-CS-05-T	BM25	Y	N	Title
KDEG-EN-CS-01-T	BB2	N	Y	Title
KDEG-EN-CS-02-DN	DLH13	Y	Y	Title+Description
KDEG-EN-CS-03-T	BB2	N	N	Title
KDEG-EN-CS-04-T	BM25	N	N	Title
KDEG-EN-CS-05-T	BM25	Y	N	Title
KDEG-CT-CT-01-T	BB2	N	Y	Title
KDEG-CT-CT-02-DN	DLH13	Y	Y	Title+Description
KDEG-CT-CT-03-T	BB2	N	N	Title
KDEG-CT-CT-04-T	BM25	N	N	Title
KDEG-CT-CT-05-T	BM25	Y	Y	Title
KDEG-EN-CT-01-T	BB2	N	Y	Title
KDEG-EN-CT-02-DN	DLH13	Y	Y	Title+Description
KDEG-EN-CT-03-T	BB2	N	N	Title
KDEG-EN-CT-04-T	BM25	N	N	Title
KDEG-EN-CT-05-T	BM25	Y	Y	Title

4.2 Description of Submitted Runs

In order to investigate the effectiveness of the techniques and to study the effect of combining them within a single process, we submitted a set of runs, which is summarized in Table 1. So we have 8 runs (2 CS-CS, 2 EN-CS, 2 CT-CT and 2 EN-CT) to test the effectiveness of the query expansion, 4 runs (2 CS-CS and 2 EN-CS) to examine the effectiveness of results re-ranking and 4 runs (2 CT-CT and 2 EN-CT) to compare the combined methodology with the simpler one and 4 runs (using DLH13 models) just to show the effectiveness of overall systems.

4.3 Experimental Results

The first set of results involves several runs to test the query expansion techniques. As shown in Table 2[4], the improvements of this technique were observed across all runs under different measurements, namely Mean AP, Mean Q and Mean NDCG. Generally the results suggest that there are greater improvements in cross-language runs than in monolingual runs (19.41% to 13.45% in CS runs and 26.06% to 15.61% in CT runs in terms of Mean AP). This at least shows that with the BB2 model, the query expansion technique is quite effective. This phenomenon is re-confirmed later with the BM25 model.

[4] Please note that IR4QA organizers released a bug-fix version of results, which were adopted here.

Table 2. Effectiveness of Query Expansion AFTER bug fix

Run Name	Mean AP	Mean Q	Mean NDCG
KDEG-CS-CS-03-T	0.3941	0.4336	0.6424
KDEG-CS-CS-01-T	**0.4471**	**0.4865**	**0.6759**
KDEG-EN-CS-03-T	0.2829	0.3231	0.5342
KDEG-EN-CS-01-T	**0.3378**	**0.3766**	**0.5778**
KDEG-CT-CT-03-T	0.4155	0.4539	0.6728
KDEG-CT-CT-01-T	**0.4818**	**0.5227**	**0.714**
KDEG-EN-CT-03-T	0.2817	0.3087	0.4957
KDEG-EN-CT-01-T	**0.3551**	**0.3822**	**0.5567**

We now draw our attention to examine the effectiveness of re-ranking and the combination of the two methods using BM25 retrieval model. The results in Table 3 show that there are only minor improvements when the re-ranking technique employed alone (CS runs). This is possibly due to several reasons. Firstly, as we discussed in section 2, only first 50 documents were re-ranked due to the efficiency issue. Clearly this strategy did not work quite well. In the future we will definitely try to re-rank all the documents returned, especially in terms of the Mean AP measurement. But the good part of this experiment is that improvements are always observed in the measurements that give more credit to the first set of retrieval results, such as Mean NDCG. The second reason is that the segmentation tool used in CS produced quite different results from those in CT runs. This maybe another reason of why this method works less well in CS runs.

However, the combination of the methods worked excellent. It provided much more improvements over the runs that only employed the query expansion method. Again the cross-language runs worked better than monolingual runs and the most improvement is up to 42.28%. We also depicted per-topic analysis of selected runs (shown in Figure. 1.). The results showed that not only for overall performance, the improved runs were managed to gain higher performance for each individual topic that was evaluated.

The runs using DLH13 model that are just to show the effectiveness of the overall system showed quite good performance. We managed to rank the top on the CS-CS, CT-CT and EN-CT runs and rank quite high in EN-CS runs.

Finally, we take a look at the cross-language performance in our submitted runs. We managed to achieve 66.25% to 85.72% monolingual performance under Mean AP, 66.54% to 86.55% monolingual performance under Mean Q and 71.88% to 90.28% monolingual performance under Mean NDCG. This shows that the cross-language performance is quite acceptable by using our method.

5. CONCLUSION

In Summary, our participation in NTCIR-8 is shown to be quite successful in the settings of monolingual and cross-language information retrieval. Particularly the re-ranking method is shown to have the potential to further improve the performance.

The future work will be centered upon how to improve the cross-language re-ranking and the overall system. By using the technique described in this paper, the extracted topics or concepts can be easily dominated by the words from one language if two parallel texts are not in equal length, even one is the direct translation of another. In addition, simply combining texts into a single text unit may not be the best way to think the problem. The correspondence relationship among multilingual topics will not be fully exploited. There are plenty of rooms to explore.

6. ACKNOWLEDGMENTS

The authors would like to thank NTCIR organizers to provide the test collections and the relevance judgments. This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (www.cngl.ie) at University of Dublin, Trinity College.

7. REFERENCES

[1] Amati, G. and Van Rijsbergen, C. J. Probabilistic models of Information Retrieval based on measuring divergence from randomness. *ACM Transactions on Information Systems*, 20, 4, 2002, 357-389.

Table 3. Effectiveness of Re-Ranking and Combined Method AFTER bug fix

Run Name	Mean AP	Improvements	Mean Q	Improvements	Mean NDCG	Improvements
KDEG-CS-CS-04-T	0.3603	NA	0.3979	NA	0.606	NA
KDEG-CS-CS-05-T	**0.3603**	**0.00%**	**0.398**	**0.03%**	**0.6063**	**0.05%**
KDEG-EN-CS-04-T	0.2707	NA	0.3072	NA	0.5138	NA
KDEG-EN-CS-05-T	**0.2709**	**0.07%**	**0.3073**	**0.03%**	**0.5142**	**0.08%**
KDEG-CT-CT-04-T	0.3713	NA	0.41	NA	0.6369	NA
KDEG-CT-CT-05-T	**0.4844**	**30.46%**	**0.5242**	**27.85%**	**0.7129**	**11.93%**
KDEG-EN-CT-04-T	0.246	NA	0.2728	NA	0.4578	NA
KDEG-EN-CT-05-T	**0.35**	**42.28%**	**0.3765**	**38.01%**	**0.5414**	**18.26%**

[2] Harter, S. P. A probabilistic approach to automatic keyword indexing. *Journal of the American Society for Information Science*, 261975, 197-206.

[3] Hofmann, T. Probabilistic latent semantic indexing. In *Proceedings of the Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval,* Berkeley, California, United States, 1999. ACM.

[4] Kurland, O. and Lee, L. PageRank without hyperlinks: structural re-ranking using links induced by language models. In *Proceedings of the Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval,* Salvador, Brazil, 2005. ACM.

[5] Ounis, I., Lioma, C., Macdonald, C. and Plachouras, V. Research Directions in Terrier: a Search Engine for Advanced Retrieval on the Web. *Novatica/UPGRADE Special Issue on Next Generation Web Search*, 8, 1, 2007, 49-56.

[6] Rocchio, J. Relevance Feedback in Information Retrieval. *The SMART Retrieval System*1971, 313-323.

[7] Sakai, T., Shima, H., Kando, N., Song, R., Lin, C.-J., Mitamura, T., Sugimoto, M. and Lee, C.-W. Overview of NTCIR-8 ACLIA IR4QA. In *Proceedings of the NTCIR-8 workshop,* Tokyo, Japan, 15-18 June, 2010.

[8] Wei, X. and Croft, W. B. LDA-based document models for ad-hoc retrieval. In *Proceedings of the Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval,* Seattle, Washington, USA, 2006. ACM.

[9] Zhang, B., Li, H., Liu, Y., Ji, L., Xi, W., Fan, W., Chen, Z. and Ma, W.-Y. Improving web search results using affinity graph. In *Proceedings of the Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval,* Salvador, Brazil, 2005. ACM.

[10] Zhou, D. and Wade, V. Language Modeling and Document Re-Ranking: Trinity Experiments at TEL@CLEF-2009. In *Proceedings of the CLEF 2009: Workshop on Cross-Language Information Retrieval and Evaluation,* Corfu, Greece, 2009.

[11] Zhou, D. and Wade, V. Latent Document Re-Ranking. In *Proceedings of the Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing,* Singapore, 2009. ACL.

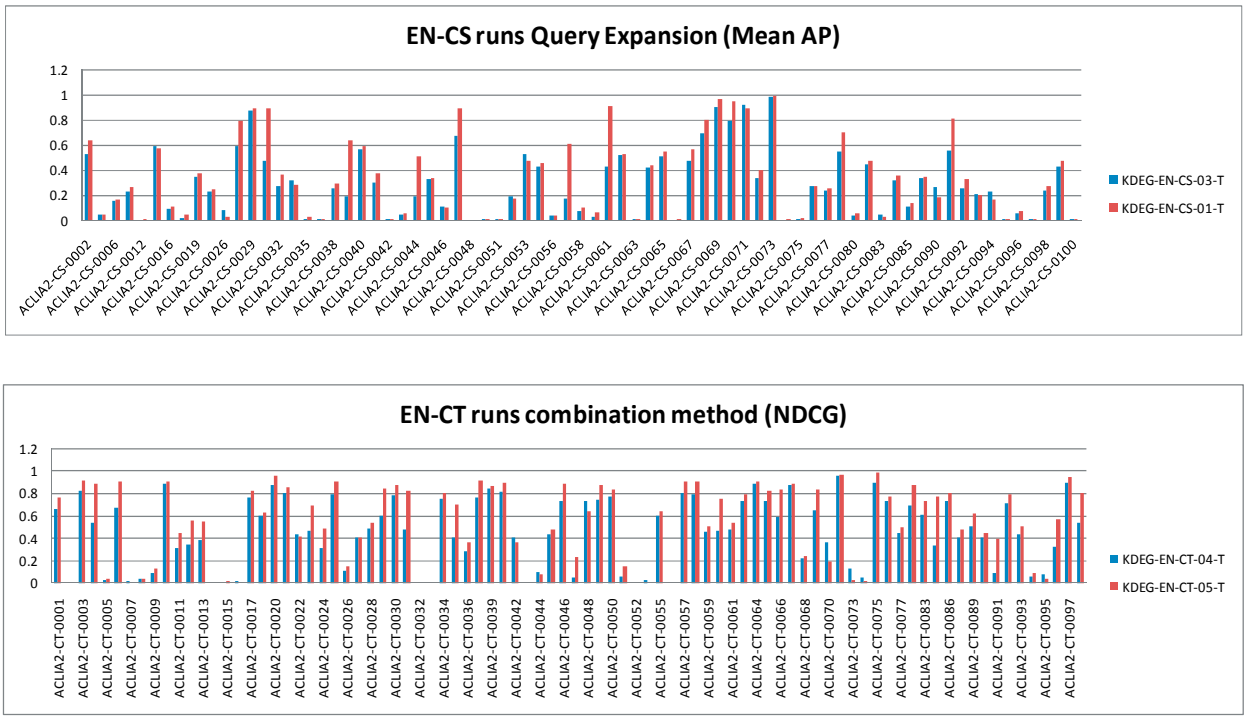

Figure 1. Per-topic analysis of selected runs

Table 4. Cross-language Performance of the Submitted Runs AFTER bug fix

RunName	Mean AP	% mono	Mean Q	% mono	Mean NDCG	% mono
KDEG-EN-CS-01-T	0.3378	75.55%	0.3766	77.41%	0.5778	85.49%
KDEG-EN-CS-02-DN	0.3847	85.72%	0.422	86.55%	0.6147	90.28%
KDEG-EN-CS-03-T	0.2829	71.78%	0.3231	74.52%	0.5342	83.16%
KDEG-EN-CS-04-T	0.2707	75.13%	0.3072	77.21%	0.5138	84.79%
KDEG-EN-CS-05-T	0.2709	75.19%	0.3073	77.21%	0.5142	84.81%
KDEG-EN-CT-01-T	0.3551	73.70%	0.3822	73.12%	0.5567	77.97%
KDEG-EN-CT-02-DN	0.3723	75.98%	0.4006	76.12%	0.5689	79.29%
KDEG-EN-CT-03-T	0.2817	67.80%	0.3087	68.01%	0.4957	73.68%
KDEG-EN-CT-04-T	0.246	66.25%	0.2728	66.54%	0.4578	71.88%
KDEG-EN-CT-05-T	0.35	72.25%	0.3765	71.82%	0.5414	75.94%