Comparing and Combining Methods for Automatic Query Expansion *

José R. Pérez-Agüera[1] and Lourdes Araujo[2]
jose.aguera@fdi.ucm.es, lurdes@lsi.uned.es

[1]Dpto. de Ingeniería del Software e Inteligencia Artificial, UCM, Madrid 28040, Spain,
[2]Dpto. Lenguajes y Sistemas Informáticos. UNED, Madrid 28040, Spain,

Abstract. Query expansion is a well known method to improve the performance of information retrieval systems. In this work we have tested different approaches to extract the candidate query terms from the top ranked documents returned by the first-pass retrieval. One of them is the cooccurrence approach, based on measures of cooccurrence of the candidate and the query terms in the retrieved documents. The other one, the probabilistic approach, is based on the probability distribution of terms in the collection and in the top ranked set. We compare the retrieval improvement achieved by expanding the query with terms obtained with different methods belonging to both approaches. Besides, we have developed a naïve combination of both kinds of method, with which we have obtained results that improve those obtained with any of them separately. This result confirms that the information provided by each approach is of a different nature and, therefore, can be used in a combined manner.

1 Introduction

Reformulation of the user queries is a common technique in information retrieval to cover the gap between the original user query and his need of information. The most used technique for query reformulation is query expansion, where the original user query is expanded with new terms extracted from different sources. Queries submitted by users are usually very short and query expansion can complete the information need of the users.

A very complete review on the classical techniques of query expansion was done by Efthimiadis [5]. The main problem of query expansion is that in some cases the expansion process worsens the query performance. Improving the robustness of query expansion has been the goal of many researchers in the last years, and most proposed approaches use external collections [17, 16, 15], such as the Web documents, to extract candidate terms for the expansion. There are other methods to extract the candidate terms from the same collection that

* Supported by project TIC2007-67581-C02-01/ and Dirección General de Universidades e Investigación de la Consejería de Educación de la Comunidad de Madrid and Universidad Complutense de Madrid (Grupo de investigación consolidado 910494)

© A. Gelbukh (Ed.)
Advances in Natural Language Processing and Applications
Research in Computing Science 33, 2008, pp. 177-188

Received 15/10/07
Accepted 07/12/07
Final Version 17/01/08

the search is performed on. Some of these methods are based on global analysis where the list of candidate terms is generated from the whole collection, but they are computationally very expensive and its effectiveness is not better than that of methods based on local analysis [11, 6, 14]. We also use the same collection that the search is performed on, but applying local query expansion, also known as pseudo-feedback or blind feedback, which does not use the global collection or external sources for the expansion. This approach was first proposed by Xu and Croft [18] and extracts the expansion terms from the documents retrieved for the original user query in a first pass retrieval.

In this work we have tested different approaches to extract the candidate terms from the top ranked documents returned by the first-pass retrieval. After the term extraction step, the query expansion process requires a further step, namely to re-compute the weights of the query terms that will be used in the search process. We have also tested different methods for this step.

There exist two main approaches to rank the terms extracted from the retrieval documents. One of them is the cooccurrence approach, based on measures of cooccurrence of the candidate and the query terms in the retrieved documents. The other one is the probabilistic approach, which is based on the differences between the probability distribution of terms in the collection and in the top ranked set. In this paper we are interested in evaluating the different techniques existing to generate the candidate term list. Our thesis is that the information obtained with the cooccurrence methods is different from the information obtained with probabilistic methods and these two kinds of information can be combined to improve the performance of the query expansion process. Accordingly, our goal has been to compare the performance of the cooccurrence approach and the probabilistic techniques and to study the way of combining them so as to improve the query expansion process. We present the results of combining different methods for the term extraction and the reweighting steps.

Two important parameters have to be adjusted for the described process. One of them is the number of documents retrieved in the first pass to be used for the term extraction. The other one is the number of candidate terms that are finally used to expand the original user query. We have performed experiments to set both of them to its optimal value in each considered method in our configuration.

The rest of the paper proceeds as follows: sections 2 and 3 describe the cooccurrence and probabilistic approaches, respectively; section 4 presents our proposal to combine both approaches; section 5 describes the different reweighting methods considered to assign new weights to the query terms after the expansion process; section 6 is devoted to show the experiments performed to evaluate the different expansion techniques separately and combined and section 7 summarizes the main conclusions of this work.

2 Cooccurrence Methods

The methods based on term cooccurrence have been used since the 70's to identify some of the semantic relationships that exist among terms. In the first works of Keith Van Rijsbergen [12] we find the idea of using cooccurrence statistics to detect some kind of semantic similarity between terms and exploiting it to expand the user's queries. In fact, this idea is based on the Association Hypothesis:

> *If an index term is good at discriminating relevant from non-relevant documents then any closely associated index term is likely to be good at this.*

The main problem with the cooccurrence approach was mentioned by Peat and Willet [8] who claim that similar terms identified by cooccurrence tend to occur also very frequently in the collection and therefore these terms are not good elements to discriminate between relevant and non-relevant documents. This is true when the cooccurrence analysis is done on the whole collection but if we apply it only on the top ranked documents discrimination does occur.

For our experiments we have used the well-know Tanimoto, Dice and Cosine coefficients:

$$\mathrm{Tanimoto}(\mathrm{t_i}, \mathrm{t_j}) = \frac{\mathrm{c_{ij}}}{\mathrm{c_i} + \mathrm{c_j} - \mathrm{c_{ij}}} \tag{1}$$

$$\mathrm{Dice}(\mathrm{t_i}, \mathrm{t_j}) = \frac{2\mathrm{c_{ij}}}{\mathrm{c_i} + \mathrm{c_j}} \tag{2}$$

$$\mathrm{Cosine}(\mathrm{t_i}, \mathrm{t_j}) = \frac{\mathrm{c_{ij}}}{\sqrt{\mathrm{c_i c_j}}} \tag{3}$$

where c_i and c_j are the number of documents in which terms t_i and t_j occur, respectively, and $c_{i,j}$ is the number of documents in which t_i and t_j cooccur.

We apply these coefficients to measure the similarity between terms represented by the vectors. The result is a ranking of candidate terms where the most useful terms for expansion are at the top.

In the selection method the most likely terms are selected using the equation

$$\mathrm{rel}(\mathrm{q}, \mathrm{t_e}) = \sum_{\mathrm{t_i} \in \mathrm{q}} \mathrm{q_i} \mathrm{CC}(\mathrm{t_i}, \mathrm{t_e}) \tag{4}$$

where CC is one of the cooccurrence coefficients: Tanimoto, Dice, or Cosine. Equation 4 boosted the terms related with more terms of the original query.

The results obtained with each of these measures, presented in section 6, show that Tanimoto performs better.

3 Distribution Analysis Approaches

One of the main approaches to query expansion is based on studying the difference between the term distribution in the whole collection and in the subsets of documents that can be relevant for the query. One would expect that terms with little informative content have a similar distribution in any document of the collection. On the contrary, terms closely related to those of the original query are expected to be more frequent in the top ranked set of documents retrieved with the original query than in other subsets of the collection.

3.1 Information-Theoretic Approach

One of the most interesting approaches based on term distribution analysis has been proposed by C. Carpineto et. al. [3], and uses the concept the Kullback-Liebler Divergence [4] to compute the divergence between the probability distributions of terms in the whole collection and in the top ranked documents obtained for a first pass retrieval using the original user query. The most likely terms to expand the query are those with a high probability in the top ranked set and low probability in the whole collection. For the term t this divergence is:

$$KLD_{(PR,PC)}(t) = P_R(t)\log\frac{P_R(t)}{P_C(t)} \tag{5}$$

where $P_R(t)$ is the probability of the term t in the top ranked documents, and $P_C(t)$ is the probability of the term t in the whole collection.

3.2 Divergence from Randomness Term Weighting Model

The Divergence From Randomness (DFR) [2] term weighting model infers the informativeness of a term by the divergence between its distribution in the top-ranked documents and a random distribution. The most effective DFR term weighting model is the *Bo1 model* that uses the Bose-Einstein statistics [10, 7]:

$$w(t) = \mathrm{tf}_x \log_2(\frac{1+P_n}{P_n}) + \log(1+P_n) \tag{6}$$

where tf_x is the frequency of the query term in the x top-ranked documents and P_n is given by $\frac{F}{N}$, where F is the frequency of the query term in the collection and N is the number of documents in the collection.

4 Combined Query Expansion Method

The two approaches tested in this work can complement each other because they rely on different information. The performance of the cooccurrence approach is reduced by words which are not stop-words but are very frequent in the collection [8]. Those words, which represent a kind of noise, can reach a high position in the term index, thus worsening the expansion process. However, precisely because of their high probability in any set of the document collection, these words tend to have a low score in KLD or Bo1. Accordingly, combining the cooccurrence measures with others based on the informative content of the terms, such as KLD or Bo1, helps to eliminate the noisy terms, thus improving the retrieved information with the query expansion process.

Our combined model amounts to applying both, a coocurrence method and a distributional method and then obtaining the list of candidate terms by intersecting the lists provided by each method separately. Finally, the terms of the resulting list are assigned a new weight by one of the reweighting method considered.

In the combined approach the number of selected terms depends of the overlapping between the term sets proposed by both approaches. To increase the intersection area and obtain enough candidate terms in the combined list it is necessary to increase the number of selected terms for the non-combined approaches. This issue has been studied in the experiments.

5 Methods for Reweighting the Expanded Query Terms

After the list of candidate terms has been generated by one of the methods described above, the selected terms which will be added to the query must be re-weighted. Different schemas have been proposed for this task. We have compared these schemas and tested which is the most appropriate for each expansion method and for our combined query expansion method.

The classical approach to term re-weighting is the Rocchio algorithm [13]. In this work we have used Rocchio's beta formula, which requires only the β parameter, and computes the new weight *qtw* of the term in the query as:

$$\mathrm{qtw} = \frac{\mathrm{qtf}}{\mathrm{qtf_{max}}} + \beta \frac{\mathrm{w(t)}}{\mathrm{w_{max}(t)}} \tag{7}$$

where $w(t)$ is the old weight of term t, $w_{\max}(t)$ is the maximum $w(t)$ of the expanded query terms, β is a parameter, qtf is the frequency of the term t in the query and $\mathrm{qtf_{max}}$ is the maximum term frequency in the query q. In all our experiments, β is set to 0.1.

We have also tested other reweighting schemes, each of which directly comes from one of the proposed methods for the candidate term selection. These schemes use the ranking values obtained by applying the function defined through each method. Each of them can only be applied to reweight terms selected with the method it derives from. This is because these methods require data, collected during the selection process, which are specific of each of them.

For the case of the reweighting scheme derived from KLD, the new weight is directly obtained applying KLD to the candidate terms. Terms belonging to the original query maintain their value [3].

For the scheme derived from the cooccurrence method, that we called *SumCC*, the weights of the candidate terms are computed by:

$$\mathrm{qtw} = \frac{\mathrm{rel(q, t_e)}}{\sum_{\mathrm{t_i \in q}} \mathrm{q_i}} \tag{8}$$

where $\sum_{t_i \in q} q_i$ is the sum of the weights of the original terms [19].

Finally, for the reweighting scheme derived from the Bose-Einstein statistics, a normalization of Bo1 that we call *BoNorm*, we have defined a simple function based in the normalization of the values obtained by Bose-Einstein computation:

$$\mathrm{qtw} = \frac{\mathrm{Bo(t)}}{\sum_{\mathrm{t \in cl}} \mathrm{Bo(t)}} \tag{9}$$

where Bo(t) is the Bose-Einstein value for the term t, and the sum runs on all terms included in the candidate list obtained applying Bose-Einstein statistics.

6 Experiments

We have used the Vector Space Model implementation provided by Lucene[1] to build our information retrieval system. Stemming and stopword removing has been applied in indexing and expansion process. Evaluation is carried out on the Spanish EFE94 corpus, which is part of the CLEF collection [9] (approximately 215K documents of 330 average word length and 352K unique index terms) and on the 2001 Spanish topic set, with 100 topics corresponding to 2001 and 2002 years, of which we only used the title (of 3.3 average word length). Nevertheless, there is evidence in the literature [1, 3] that some of the presented methods are also valid for other languages (English, French, Italian and Spanish).

We have used different measures to evaluate each method. Each of them provides a different estimation of the precision of the retrieved documents, which is the main parameter to optimize when doing query expansion, since recall is always improved by the query expansion process. The measures considered have been MAP[2] *(Mean Average Precision)*, GMAP[3], Precision@X[4], R-Precision[5].

First of all we have tested the different cooccurrence methods described above. Table 1 shows the results obtained for the different measures considered in this work. We can observe that Tanimoto provides the best results for all the measures, except for P@10, but in this case the difference with the result of Dice, which is the best, is very small. According to the results we have selected the Tanimoto similarity function as coocurrence method for the rest of the work.

Table 1. Comparing different cooccurrence methods. The Baseline row corresponds to the results of the query without expansion. P@5 stands for precision after the first five documents retrieved, P@10 after the first ten, and R-PREC stands for R-precision. Best results appear in boldface.

	MAP	GMAP	R-PREC	P@5	P@10
Baseline	0.4006	0.1941	0.4044	0.5340	0.4670
Cosine	0.4698	0.2375	0.4530	0.6020	0.5510
Tanimoto	**0.4831**	**0.2464**	**0.4623**	**0.6060**	0.5520
Dice	0.4772	0.2447	0.4583	0.6020	**0.5530**

[1] http://lucene.apache.org

[2] the average of the precision value (percent of retrieved documents that are relevant) obtained for the top set documents existing after each relevant document is retrieved.

[3] a variant of MAP, that uses a geometric mean rather than an arithmetic mean to average individual topic results.

[4] precision after X documents (whether relevant or non-relevant) have been retrieved.

[5] measures precision after R documents have been retrieved, where R is the total number of relevant documents for a query.

6.1 Selecting the Reweighting Method

The next set of experiments have had the goal of determining the most appropriate reweighting method for each candidate term selection method. Table 2 shows the results of different reweighting methods (Rocchio and SumCC) applied after selecting the candidate terms by the cooccurrence method. We can observe that the results are quite similar for both reweighting methods, though Rocchio is slightly better.

Table 2. Comparing different reweighting methods for cooccurrence. *CooRocchio* corresponds to using cooccurrence as selection terms method and Rocchio as reweighting method. *CooSumCC* corresponds to using cooccurrence as selection terms method and *SumCC* as reweighting method. Best results appear in boldface.

	MAP	GMAP	R-PREC	P@5	P@10
Baseline	0.4006	0.1941	0.4044	0.5340	0.4670
CooRocchio	**0.4831**	**0.2464**	0.4623	0.6060	**0.5520**
CooSumCC	0.4798	0.2386	**0.4628**	**0.6080**	0.5490

Table 3. Comparing different reweighting methods for KLD. *KLDRocchio* corresponds to using KLD as selection terms method and Rocchio as reweighting method. *KLDkld* corresponds to using KLD as selection terms method and *kld* as reweighting method. Best results appear in boldface.

	MAP	GMAP	R-PREC	P@5	P@10
Baseline	0.4006	0.1941	0.4044	0.5340	0.4670
KLDRocchio	0.4788	0.2370	0.4450	0.5960	0.5480
KLDkld	**0.4801**	**0.2376**	**0.4526**	**0.6080**	**0.5510**

Table 3 shows the results of different reweighting methods (Rocchio and kld) applied after selecting the candidate terms with KLD. The best results are obtained using kld as reweighting method.

Table 4 shows the results of different reweighting methods (Rocchio and BoNorm) applied after selecting the candidate terms with Bo1. In this case, the best results are obtained using BoNorm as reweighting method.

The results of this section show that the best reweighting method after selecting terms by cooccurrence is Rocchio, while for the distributional methods in the term selection process, the best reweighting is obtained with the method derived from themselves, though Rocchio also provides results very close to the best one.

Table 4. Comparing different reweighting methods for Bo1 *BoRocchio* corresponds to using Bo1 as selection terms method and Rocchio as reweighting method. *BoBoNorm* corresponds to using Bo1 as selection terms method and *BoNorm* as reweighting method. Best results appear in boldface.

	MAP	GMAP	R-PREC	P@5	P@10
Baseline	0.4006	0.1941	0.4044	0.5340	0.4670
BoRocchio	0.4765	0.2381	0.4450	0.5880	0.5450
BoBoNorm	**0.4778**	**0.2388**	**0.4470**	**0.5960**	**0.5470**

6.2 Parameter Study

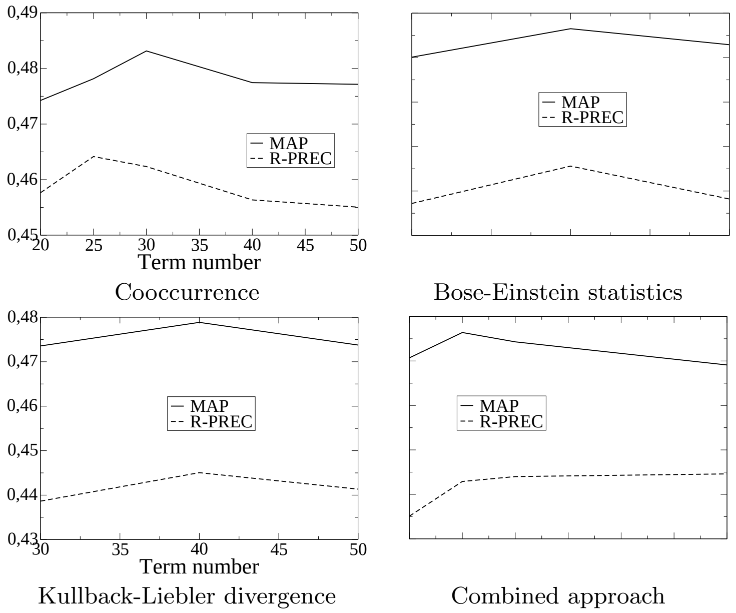
We have studied two parameters that are fundamental in query expansion, the number of candidate terms to expand the query and the number of documents from the top ranked set used to extract the candidate terms. The optimal value of these parameters can be different for each method, and thus we have studied them for each case. The reweighting used for each method has been the one that provides de best results, and Rocchio for the combined approach.

Figure 1 shows, for the different expansion methods considered, the MAP and R-PREC measures with different numbers of candidate terms to expand the original query. We can observe that the results of both measures, MAP and R-PREC, indicate similar values, and that this value is different for each considered method: around 25 terms for the cooccurrence method, 40 terms for Bose-Einstein statistics and Kullback-Liebler divergence and 75 terms for our combined approach. The combined approach requires a larger number of selected terms from each basic approach in order to have enough expansion terms in the intersection list.

Figure 2 shows, for the different expansion methods considered, the MAP and R-PREC measures with different numbers of documents used to extract the set of candidate query terms. We can observe that in all cases the best value is around 10 documents.

6.3 Comparing and Combining Both Approaches

The next step of our experiments has been comparing the overall retrieval performance of the different expansion method considered, including our combined approach. The reweighting used for each method has been the one that provides de best results, and Rocchio for the combined approach. Table 5 shows MAP and GMAP measures, while table 6 shows R-precision, precision after 5 documents retrieved (P@5) and after 10 documents (P@10). We can observe that for nearly every measure (except for P@5) the best results are obtained by the combination of the Bo1 model with cooccurrence. The next best result is provided by the other combination considered, KLD with cooccurrence. These results prove that the information provided by methods belonging to different approaches, cooccurrence and distributional analysis, is different and thus its combination improves the results obtained by any of them separately.



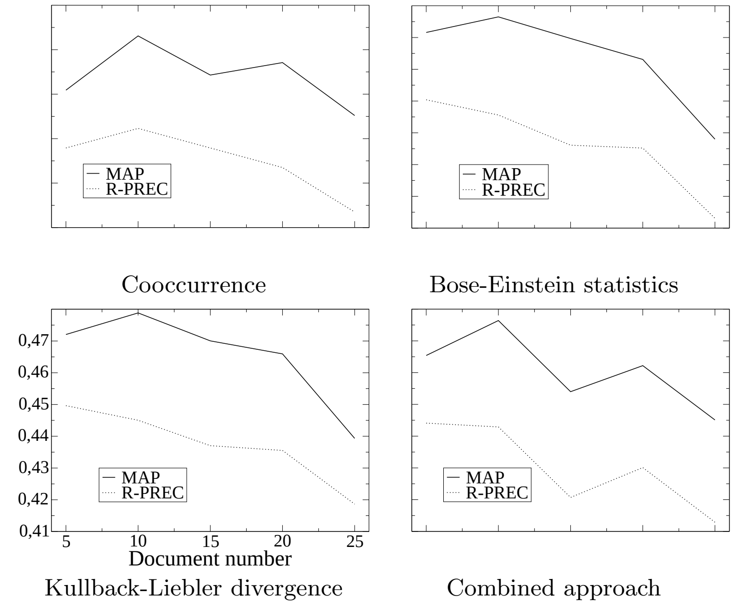
Cooccurrence

Bose-Einstein statistics

Kullback-Liebler divergence

Combined approach

Fig. 1. Study of the best number of candidate terms to expand the original query with the different considered methods. R-PREC stands for R-Precision.



Cooccurrence

Bose-Einstein statistics

Kullback-Liebler divergence

Combined approach

Fig. 2. Study of the best number of documents used to extract the set of candidate query terms. R-PREC stands for R-Precision.

Table 5. Comparing MAP and GMAP for different methods considered for query expansion.

	MAP	GMAP
Baseline	0.4006(-)	0.1941(-)
KLD	0.4801(+16.55%)	0.2376(+18.30%)
Bo	0.4778(+16.15%)	0.2388(+18.71%)
Cooccurrence	0.4831(+17.07%)	0.2464(+21.22%)
BoCo	**0.4964(+19.29%)**	**0.2570(+24.47%)**
KLDCo	0.4944(+18.97%)	0.2483(+21.82%)

Table 6. Comparing R-Precision (R-PREC), precision after 5 documents retrieved (P@5) and after 10 documents retrieved (P@10) for different methods considered for query expansion.

	R-PREC	P@5	P@10
Baseline	0.4044(-)	0.5340(-)	0.4670(-)
KLD	0.4526(+10.64%)	0.6080(+12.17%)	0.5510(+15.24%)
Bo	0.4470(+9.53%)	0.5960(+10.40%)	0.5470(+14.62%)
Cooccurrence	0.4623(+12.5%)	0.6060(+11.88%)	0.5520(+15.39%)
BoCo	**0.4629(+12.63%)**	0.6220(+14.14%)	**0.5630(+17.05%)**
KLDCo	0.4597(+12.02%)	**0.6240(+14.42%)**	0.5600(+16.60%)

6.4 Analysis of the Results

We have analyzed the results for some specific queries of our test set. Table 8 compares the MAP measure obtained for cooccurrence, Bo1 and the combination for the test set queries shown in table 7. We can observe that the best result in each case is provided by a different method, thus showing that these methods provided different information. We can also observe that the combined model not always provides the best result. This suggests to investigate other combination schemes.

Table 7. Queries used to study the performances of each method in particular cases.

C041	Pesticidas en alimentos para bebés
C049	Caída de las exportaciones de coches en Japón
C053	Genes y enfermedades
C055	Iniciativa Suiza para los Alpes
C058	Eutanasia
C122	Industria norteamericana del automóvil

These results are compatible with the previous observations of improvement with the combined approaches because we can observe that BoCo always im-

Table 8. Results of the MAP measure for the queries 41, 49, 53, 55, 58, 122. BoCo stands for the model combining Bo1 and cooccurrence. The best value appears in boldface.

Measure	41	49	53	55	58	122
Baseline	0.62	0.1273	0.3098	0.2334	0.8417	0.0760
Cooccurrence	0.9428	0.2775	**0.4901**	0.6447	0.8960	0.0588
Bo1	0.9270	**0.3594**	0.4005	0.5613	**0.9329**	0.1130
BoCo	**0.9630**	0.3072	0.4724	**0.6588**	0.9223	**0.1252**

proves some of the non-combined methods. This is an indication of the higher robustness of the combined approach. Nevertheless, it will be interesting to analysis the kind of queries more appropriate for each approach.

7 Conclusions and Future Works

We have presented a study of two different approaches, cooccurrence and distributional analysis, for query expansion. For each approach we have considered several models. Results have shown that the query expansion methodology that we apply is very robust and improves the retrieval results of the original query with all tested approaches.

The analysis of the results indicates that the statistical information exploited in each considered approach is different, and this suggests combining them to improve the results.

We have carried out experiments to measure the improvement of each method separately, and the combination of them. Results have shown that a simple combination of the different query expansion approaches is more efficient than the use of any of them separately. This result confirms our thesis that the information exploited by each approach is different and it is worthwhile to investigate more sophisticated ways of performing this combination, what we plan to do in future works.

References

1. Gianni Amati, Claudio Carpineto, and Giovanni Romano. Comparing weighting models for monolingual information retrieval. In *CLEF*, pages 310–318, 2003.
2. Gianni Amati and Cornelis Joost Van Rijsbergen. Probabilistic models of information retrieval based on measuring the divergence from randomness. *ACM Trans. Inf. Syst.*, 20(4):357–389, 2002.
3. Claudio Carpineto, Renato de Mori, Giovanni Romano, and Brigitte Bigi. An information-theoretic approach to automatic query expansion. *ACM Trans. Inf. Syst.*, 19(1):1–27, 2001.
4. Thomas M. Cover and Joy A. Thomas. *Elements of information theory.* Wiley-Interscience, New York, NY, USA, 1991.

5. E. N. Efthimiadis. Query expansion. *Annual Review of Information Systems and Technology*, 31:121–187, 1996.
6. Y. Jing and W. Bruce Croft. An association thesaurus for information retrieval. In *Proceedings of RIAO-94, 4th International Conference "Recherche d'Information Assistee par Ordinateur"*, pages 146–160, New York, US, 1994.
7. C. Macdonald, B. He, V. Plachouras, and I. Ounis. University of Glasgow at TREC 2005: Experiments in Terabyte and Enterprise Tracks with Terrier. In *Proceeddings of the 14th Text REtrieval Conference (TREC 2005)*, 2005.
8. Helen J. Peat and Peter Willett. The limitations of term co-occurrence data for query expansion in document retrieval systems. *JASIS*, 42(5):378–383, 1991.
9. Carol Peters and Martin Braschler. European research letter: Cross-language system evaluation: The clef campaigns. *JASIST*, 52(12):1067–1072, 2001.
10. V. Plachouras, B. He, and I. Ounis. University of Glasgow at TREC2004: Experiments in Web, Robust and Terabyte tracks with Terrier. In *Proceeddings of the 13th Text REtrieval Conference (TREC 2004)*, 2004.
11. Yonggang Qiu and Hans-Peter Frei. Concept based query expansion. In *SIGIR*, pages 160–169, 1993.
12. C. J. Van Rijsbergen. A theoretical basis for the use of cooccurrence data in information retrieval. *Journal of Documentation*, (33):106–119, 1977.
13. J. J. Rocchio. Relevance feedback in information retrieval. In G Salton, editor, *The SMART retrieval system*, pages 313–323. Prentice Hall, 1971.
14. Hinrich Schütze and Jan O. Pedersen. A cooccurrence-based thesaurus and two applications to information retrieval. *Inf. Process. Manage.*, 33(3):307–318, 1997.
15. Ellen M. Voorhees. Overview of the trec 2003 robust retrieval track. In *TREC*, pages 69–77, 2003.
16. Ellen M. Voorhees. The trec robust retrieval track. *SIGIR Forum*, 39(1):11–20, 2005.
17. Ellen M. Voorhees. The trec 2005 robust track. *SIGIR Forum*, 40(1):41–48, 2006.
18. Jinxi Xu and W. Bruce Croft. Improving the effectiveness of information retrieval with local context analysis. *ACM Trans. Inf. Syst.*, 18(1):79–112, 2000.
19. Angel Zazo, Carlos G. Figuerola, and José Luis Alonso Berrocal. REINA at CLEF 2006 robust task: Local query expansion using term windows for robust retrieval. In A. Nardi, C. Peters, and J.L. Vicedo, editors, *ABSTRACTS CLEF 2006 Workshop, 20-22 September, Alicante, Spain. Results of the CLEF 2006 Cross-Language System Evaluation Campaign*, 2006.