A Survey of Query Expansion until June 2012

Yogesh Kakde
Indian Institute of Technology, Bombay

25th June 2012

Abstract

Here we present a survey of important work done on Query Expansion (QE) between the period 1970 to 2012. Queries formed by search engine users, are generally short and vague. It is very difficult to estimate the exact user need. Sometimes, the queries are well formed and clear, but the document collection does not contain the information in the same form as that of the query. Synonymous terms or related terms are present in the collection. Query expansion adds terms to the original query, which provides more information about the user need. There are various approaches to query expansion. Some of them use lexical resources or dictionaries for query expansion. Lexical methods cause more damage through topic drift. Relevance feedback approaches perform better than lexical approaches, but they depend on the performance of initial retrieval. In recent years Wikipedia is being used as an external knowledge source for QE. In this survey we cover important work done using all these approaches.

1 Query Expansion

Query expansion is an effective technique to improve the performance of information retrieval system. Studies have shown that the average query length is only around 2-3 words. Due to this, it is very difficult to understand the user intention behind the query and provide the ideal set of relevant documents. For example, if the user query is *Lalbagh*, it is not clear if the user wishes to know about the a) *Lalbagh* garden in *Bangalore* b) *Lalbagh fort* in *Bangladesh* or c) Some area called *Lalbagh* in *Mumbai*. Natural language phenomenon like polysemy and synonymy further compound this problem. Consider the query *apple prices in market*, it is not clear if the word *apple* refers to the company apple or to the fruit apple. Query Expansion techniques tries to address these issues. There are many approaches used for query expansion. All these approaches can be classified into two major classes: global methods and local methods Manning et al. [2008]. Global methods are techniques for expanding or reformulating query terms independent of the query and results returned from it, so that changes in the query wording will cause the new query to match other semantically similar terms. Only individual query terms are considered for expansion. Global methods include:

- Query expansion/reformulation with a thesaurus or WordNet
- Query expansion via automatic thesaurus generation
- Techniques like spelling correction

Local methods use documents that are retrieved using unmodified query. The basic methods are:

- Relevance Feedback
- Pseudo Relevance Feedback
- Indirect relevance feedback

1.1 Global Methods

Most common query expansion methods are the one using some form of thesaurus or knowledge resource. Each term t in a query can be expanded with synonyms and related terms from the thesaurus. Each term can be associated with a weight. Expansion terms can be given less weight than the original query terms. Methods for building a thesaurus for query expansion includes:

- Use of a controlled vocabulary that is maintained by human editors. These thesauruses contain canonical term for each concept. The subject headings of traditional library subject indexes, such as the Library of Congress Subject Headings, or the Dewey Decimal system are examples of a controlled vocabulary. Use of a controlled vocabulary is quite common for well-resourced domains. A well-known example is the Unified Medical Language System (UMLS) used with MedLine for querying the biomedical research literature.

- A manual thesaurus. Human editors build sets of synonymous names for concepts, without designating a canonical term. The UMLS metathesaurus is one example of such thesaurus. Statistics Canada maintains a thesaurus of preferred terms, synonyms, broader terms, and narrower terms for matters on which the government collects statistics, such as goods and services. This thesaurus is also bilingual English and French Manning et al. [2008].

- An automatically derived thesaurus. Thesaurus is developed automatically using co-occurring terms or grammatically related terms. Terms which co-occur quiet frequently in a corpus are more likely to be related. The other approach is to analyze the corpus for grammatical dependencies. For example, entities that grow, walk or move, are more likely to be living organism or more specific, to be humans.

- Query reformulations based on query log mining. Large amount of user interaction information is available to the web search engines. This information is stored in query logs and can be used to improve the user satisfaction of the users later.

Lot of work is done on Query Expansion using thesaurus or knowledge resources. In this following section we will review some of those work.

1.1.1 Manually Developed Lexical Resources

Manually developed lexical resources like WordNet, UMLS metathesaurus, *etc.* are commonly used for query expansion. In Voorhees [1994a], author used WordNet for query expansion. Synonyms terms were added to the query. He observed that this approach makes a little difference in retrieval effectiveness if the original query is well formed. Query which are not well formed can be improved significantly. WordNet is also used by Smeaton et al. [1995] along with POS tagging for query expansion. Each query term was expanded independently and equally. For the expansions, parents of query terms (P) were given weights, as were grandparents, children (C) and grandchildren (GC). For example, the original query term *prison* can be expanded to

prison 1.0, penal institution 0.5, institution 0.25, camp 0.5, work camp 0.25, prison camp 0.25, prison farm 0.25, prison camp 0.25, internment camp 0.25, prisoner of war camp 0.25, POW camp 0.25, college 0.5, house of correction 0.5, hold 0.5, keep 0.5, jail 0.5, jailhouse 0.5, gaol 0.5, lockup 0.5, lock up 0.5, cooler 0.5, pokey 0.5

The interesting thing about their method is they ignored the original query terms after expansion. As a results of this precision and recall dropped, but they were able to retrieve documents which didn't contain any of the query terms but are relevant to the query.

1.1.2 Automatically Derived Thesaurus

There are many approaches which use corpus or lexical resources to automatically develop a thesaurus. Most of these methods are used in domain specific search engines or applications. In Gong et al. [2005] author used WordNet and TSN (Term Semantic Network) developed using word co-occurrence in corpus. As seen in the earlier section, lexical resources alone cannot significantly improve the performance. Here, the author used TSN as a filter and supplement for WordNet. They observed that when WordNet and TSN are used together gives better results than any of them used alone. Qiu and Frei [1993] constructed a term-vs-term similarity matrix based on how the terms of the collection are indexed. A probabilistic method is used to estimate the probability of a term similar to a given query in the vector space model. Even when adding hundreds of terms into queries, this approach showed that the similarity thesaurus can improve performance significantly. Adding so many terms may not, however, be efficient for large information retrieval systems. It should be noted that their improved results on the NLP collection is still lower than the baselines used in this paper. Jing and Croft [1994] used a Phrase finder program that automatically constructs a thesaurus using text analysis and text feature recognition. Phrase finder considers co-occurrence between phrases and terms as associations. They reported 30% increase in performance.

1.1.3 Query Expansion Based on Query Log Mining

With the increase in usage of Web search engines, it is easy to collect and use user query logs. Cui et al. [2002] developed a system that extracts probabilistic correlations between query terms and documents terms using query logs. These correlations are then used to find good expansion terms for new queries. Cui et al. [2003] assumed that the documents visited by users are relevant to the query. They maintained a list of all the documents visited for a particular query. Probability of document being visited when a particular query word is present in a query is calculated to find the relevance of the document. Larger query logs improves the retrieval accuracy of the system. One type of query expansion is to identify the phrases from the query. de Lima and Pedersen [1999] used Web query log and some hand written context-free-grammar to estimate the parameters of a probabilistic context-free grammar. Using this grammar they extract phrases from the query. They observed $12 - 65\%$ improvement on different datasets. Yin et al. [2009] considered query log as bipartite graph that connects the query nodes to the URL nodes by *click* edges. Given a query node q and a URL node u, there will be an edge (q, u) if u is among the clicked answers for query q. The weight of the edge (q, u) is computed by the click frequency C_{qu}. Random walk probability is used to find n most similar queries to q and combine them to form an expanded query. They have recorded more than 10% improvement over the baseline.

2 Relevance Feedback

Relevance Feedback is one specific type of query expansion technique.

In RF, the user query is used to retrieve the initial list of ranked documents. User judge the relevance of the results. These judgments are used by RF to refine the query and a new ranked list is produced using the refined query. The above process takes place iteratively. The actual algorithm employed to refine the query based on relevance feedback depends on the underlying retrieval model used.

2.1 Types of Relevance Feedback

Based on the way in which feedback information is gathered, RF is categorized into:

- Explicit Feedback: User gives information about relevance of a document.
- Implicit Feedback: Implicit feedback is inferred from user behavior, such as noting which documents they do and do not select for viewing, the duration of time spent viewing a document, or page browsing or scrolling actions For example, the similar pages features in *Google*.
- Pseudo Relevance Feedback: In PRF top k initial retrieved documents are considered as relevant and are used to refine the query. The refined query is then used to retrieve the documents.

In following sections, we review the literature related to RF and PRF. We discuss the actual techniques proposed for query refinement in RF. Later, we also discuss the techniques proposed to improve PRF. Many algorithms discussed in the context of explicit RF like Rocchio, Probabilistic RF are also relevant in the context of PRF where the top k is taken as the relevant feedback.

2.1.1 Rocchio Algorithm

Rocchio's algorithm is a classic algorithm for incorporating relevance feedback into the Vector Space Model. It was introduced by Salton's SMART system in 1970 Salton [1971]. The main intuition behind this approach is that the modified query vector is made to move towards the centroid of the relevant documents and away from the centroid of the irrelevant documents. Given a query vector $\overrightarrow{q}$ and partial information about relevant and irrelevant document sets D_r and D_{nr}, the algorithm modifies the query vector as:

$$\overrightarrow{q_m} = \alpha \cdot \overrightarrow{q_0} + \beta \cdot \frac{1}{|D_r|} \sum_{\overrightarrow{d_i} \in D_r} \overrightarrow{d_i} - \gamma \cdot \frac{1}{|D_{nr}|} \sum_{\overrightarrow{d_j} \in D_{nr}} \overrightarrow{d_j} \tag{2.1}$$

Here, α, β, γ control the relative importance given to the initial query vector $\overrightarrow{q_0}$, set of relevant documents $|D_r|$ and the set of irrelevant documents $|D_{nr}|$.

Any negative term weights are ignored and set to 0. Positive feedback was usually found to be more useful than negative feedback and hence the weights are usually set as $\alpha = 1$, $\beta = 0.75$ and $\gamma = 0.15$.

2.1.2 Pseudo-Relevance Feedback (PRF)

As mentioned in section 2.1, PRF methods assume top k documents to be relevant and the query is refined using the feedback documents. The basis behind PRF is that documents which are similar to the user's initial query will lead us to more relevant terms which when augmented with the query will lead to an improvement in performance. Croft and Harper Croft and Harper [1979] first suggested this technique for estimating the probabilities within the probabilistic model. However, they also highlighted one fundamental problem - *query drift*. Query drift is caused as a result of adding terms which have no association with the topic of relevance of the query. This happens only when there are only a few or no relevant documents in the top k feedback documents. Due to this sensitivity to the quality of top k documents, PRF only improves the performance of queries which have good or reasonable initial retrieval performance. It can't be used to improve the performance of bad or failed queries which do not retrieve anything relevant initially. Hence, it is not robust to the quality of the initial retrieval.

Several approaches have been proposed to improve the robustness of PRF. The main strategies used could be categorized as follows:

- To refine the feedback document set so that instead of using all the top k documents, we could choose a subset of it which is likely to be "highly relevant" Mitra et al. [1998], Sakai et al. [2005].
- Instead of using all the terms obtained through feedback for query refinement, only use a subset of important terms to avoid introducing query drift Cao et al. [2008].
- Dynamically decide when to apply PRF instead of using it for all queries Amati et al. [2004], Cronen-Townsend et al. [2004].
- Varying the importance of each feedback document Tao and Zhai [2006].
- Using a large external collection like Wikipedia or the web as a source of expansion terms besides those obtained through PRF. The intuition

behind this approach is that if the query does not have many relevant documents in the collection then any improvements in the modeling of PRF is bound to perform poorly due to query drift Voorhees [2006], Xu et al. [2009a].

- Recently, Lv and Zhai [2010] proposed a *positional relevance model* where the terms in the document which are nearer to the query terms are assigned more weight.

Several approaches have been proposed which use different types of lexically and semantically related terms during query expansion. Voorhees et al. Voorhees [1994b] use Wordnet for query expansion and report negative results. Random walk models Collins-Thompson and Callan [2005], Lafferty and Zhai [2001] have been used to learn a rich set of term level associations by combining evidence from various kinds of information sources mentioned so far like WordNet, co-occurrence relationships, web, morphological variants *etc.* Metzler and Croft [2007] propose a feature based approach called latent concept expansion to model term dependencies.

2.1.3 PRF in Language Modeling Framework

The Language Modeling (LM) Framework for IR offers a principled approach to model PRF. In the LM approach, the document and query are modeled using multinomial distribution over words called *document language model* $P(w|D)$ and *query language model* $P(w|\Theta_Q)$ respectively. For a given query, the document language models are ranked based on their proximity to the query language model, measured using KL-Divergence.

$$\begin{aligned} Rank(D,Q) &= KL(\Theta_Q||D) \\ &= \sum_w P(w|\Theta_Q) \cdot log\frac{P(w|\Theta_Q)}{P(w|D)} \end{aligned}$$

Since the query length is short, it is difficult to estimate the query language model accurately using the query alone. In PRF, the top k documents obtained through the initial ranking algorithm are assumed to be relevant and used as feedback for improving the estimation of Θ_Q. The feedback documents contain a mix of both relevant and noisy terms. The actual relevant terms modeled using the feedback language model Θ_F is inferred from D_F based on a Generative Mixture Model Zhai and Lafferty [2001] formulation.

Mixture Model for Estimating Feedback Model: Let $D_F = \{d_1, d_2, \ldots, d_k\}$ be the top k documents retrieved using the initial ranking algorithm. Zhai and Lafferty [2001] model the feedback document set D_F as a mixture of two distributions: (a) the *feedback language model* and (b) the *collection model* $P(w|C)$. Assuming a fixed mixture proportion λ in the feedback document set, the feedback language model is inferred using the EM Algorithm Dempster et al. [1977]. The EM update equations are given by:

$$t^{(n)}(w) = \frac{\lambda \cdot p^{(n)}(w|D_F)}{\lambda \cdot p^{(n)}(w|D_F) + (1-\lambda) \cdot P(w|C))} \tag{2.2}$$

$$p^{(n+1)}(w|D_F) = \frac{\sum_{d \in D_F} c(w;d) \cdot t^{(n)}(w)}{\sum_{w \in W} \sum_{d \in D_F} c(w;d) \cdot t^{(n)}(w)} \tag{2.3}$$

In the EM algorithm, the feedback model is iteratively refined by accumulating probability mass on most *distinguishing* terms which are more frequent in the feedback document set and less frequent across the entire collection. Let Θ_F be the final converged feedback model. Later, in order to keep the query focus, Θ_F is interpolated with the initial query model Θ_Q to obtain the final query model Θ_{Final}.

$$\Theta_{Final} = (1-\alpha) \cdot \Theta_Q + \alpha \cdot \Theta_F \tag{2.4}$$

Θ_{Final} is used to re-rank the corpus using the KL-Divergence ranking function to obtain the final ranked list of documents.

2.1.4 Multilingual PRF

Many of the above methods suffer from following drawbacks

- Terms having co-occurrence relationship with the query are considered but the terms which are semantically and lexically related to the query are not explicitly considered.
- Result of PRF is dependent on quality of initially retrieved top k documents and thus is not robust. Result is biased towards these initial documents. Chinnakotla [2010] suggests the use of one language to

help retrieval of another language. This approach is called Multilingual PRF (MultiPRF). In Multilingual PRF, query in language L_1 is translated into another language L_2 . PRF is performed on a collection in L_1 as well as in L_2 . The feedback model for query in L_2 is translated back from L_2 to L_1 . This feedback model is combined with the feedback model for query in L_1 . Final list of relevant documents can be obtained by using this feedback model(refined query). MultiPRF gives better performance than PRF as it uses the PRF in two collections of different languages, it is less likely to drift from the original topic. As it is using assisting language, both co-occurrence based terms as well as lexically and semantically related terms are taken into consideration. For languages with poor coverage of the collection, this approach can be very useful

3 Query Expansion Using Wikipedia

Wikipedia is the biggest encyclopedia available freely on the web. Though developed by people around the globe, Wikipedia content is well structured and correct. Being develop by people is an advantage, its growing rapidly and contains wide variety of topics. These features make Wikipedia a good knowledge source for query expansion. Recently, many approaches are being developed which use Wikipedia for query expansion. In this section we will see some of those approaches.

Li et al. [2007] proposed query expansion using Wikipedia by utilizing the category assignments of its articles. The base query is run against a Wikipedia collection and each category is assigned a weight proportional to the number of top-ranked articles assigned to it. Articles are then re-ranked based on the sum of the weights of the categories to which each belongs. The method shows improvement over PRF in measures favoring weak queries.

A thesaurus-based query expansion method using Wikipedia was proposed by Milne et al. [2007a]. The thesaurus is derived with criteria such that topics relevant to the document collection are included. They propose to extract significant topics from a query by checking consecutive sequences of words in the query against the thesaurus. However, query dependent knowledge is not taken into consideration by the thesaurus. Xu et al.

[2009b] explored Wikipedia to classify the query into entity queries, ambiguous queries, and broader queries. Pseudo relevant documents are generated in two ways according to the query type: 1) using top ranked articles from Wikipedia retrieved in response to the query, and 2) using Wikipedia entity pages corresponding to queries. In selecting expansion terms, term distributions and structures of Wikipedia pages are taken into account. Koru Milne et al. [2007b] a search interface which offers a domain-independent knowledge-based information retrieval is developed using Wikipedia structure for query expansion. They used Wikipedia articles as building blocks for the thesaurus, and its skeleton structure of hyper-links to determine which blocks are needed and how they should fit together. They observed that, due to the large size of Wikipedia corpus irrelevant results are also retrieved. Its essential to identify correct concept relevant to a particular document collection. Kaptein and Kamps [2009] used Wikipedia link and category information to expand the query. Category information is used by calculating distances between document categories and target categories. Observation was that category information has more value than link information.

4 Summary

In general query expansion methods improve the performance of a search engine. Use of thesaurus or WordNet gives some good expansion terms, but it also cause topic drift. These approaches gives better results if domain specific thesaurus or dictionaries are used. Domain specific resources reduce the topic drift, but these resources can only be used for domain specific search or when we know the domain of a query. Resources take a lot of time to build and are expensive. Automatically developed resources are cheap and faster, but are inaccurate. Recent increase in use of search engines, opened a new resource in terms of query logs. Query logs provide gives a good idea about the user need and help us to serve the user better.

Local methods gives better results than global methods. PRF is widely used for query expansion. PRF assumes initially retrieved top k documents to be relevant. Since the top k relevant documents are considered for expansion, the expansion terms are relevant and cause less topic drift. PRF performance decreases when the assumption fails. PRF performs better if the initial systems performance is average or above average. In case of resource scarce languages, initial retrieval is poor and that causes PRF to perform poorly. Multilingual PRF uses an assisting language for query expansion. The Language in which the initial performance is good, is choose as an assisting language and expansion terms are obtained from that language. In recent years, Wikipedia is being explored as an external resource for query expansion. Wikipedia is the largest encyclopedia freely available online. Wikipedia category structure and document structure is used for expansion.

References

Amati, G., Carpineto, C., and Romano, G. (2004). Query difficulty, robustness, and selective application of query expansion. In *ECIR*, pages 127–137.

Cao, G., Nie, J.-Y., Gao, J., and Robertson, S. (2008). Selecting good expansion terms for pseudo-relevance feedback. In *SIGIR '08: Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 243–250, New York, NY, USA. ACM.

Chinnakotla, M. K. (2010). *Information Retrieval in Multilingual Resource-Constrained Settings*. PhD thesis, Indian Institute of Technology, Mumbai.

Collins-Thompson, K. and Callan, J. (2005). Query expansion using random walk models. In *CIKM '05: Proceedings of the 14th ACM international conference on Information and knowledge management*, pages 704–711, New York, NY, USA. ACM.

Croft, W. B. and Harper, D. J. (1979). Using probabilistic models of document retrieval without relevance information. *Journal of Documentation*, 35(4):285–295.

Cronen-Townsend, S., Zhou, Y., and Croft, B. W. (2004). A framework for selective query expansion. In *CIKM '04: Proceedings of the 2004 ACM CIKM International Conference on Information and Knowledge Management*, pages 236–237. ACM.

Cui, H., Wen, J.-R., Nie, J.-Y., and Ma, W.-Y. (2002). Probabilistic query expansion using query logs. In *Proceedings of the 11th international conference on World Wide Web*, WWW '02, pages 325–332, New York, NY, USA. ACM.

Cui, H., Wen, J.-R., Nie, J.-Y., and Ma, W.-Y. (2003). Query expansion by mining user logs. *Knowledge and Data Engineering, IEEE Transactions on*, 15(4):829–839.

de Lima, E. F. and Pedersen, J. O. (1999). Phrase recognition and expansion for short, precision-biased queries based on a query log. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '99, pages 145–152, New York, NY, USA. ACM.

Dempster, A., Laird, N., and Rubin, D. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, 39:1–38.

Gong, Z., Cheang, C., and Leong Hou, U. (2005). Web query expansion by wordnet. In Andersen, K., Debenham, J., and Wagner, R., editors, *Database and Expert Systems Applications*, volume 3588 of *Lecture Notes in Computer Science*, pages 166–175. Springer Berlin / Heidelberg. 10.1007/11546924_17.

Jing, Y. and Croft, W. B. (1994). An association thesaurus for information retrieval. In *In RIAO 94 Conference Proceedings*, pages 146–160.

Kaptein, R. and Kamps, J. (2009). Advances in focused retrieval. chapter Finding Entities in Wikipedia Using Links and Categories, pages 273–279. Springer-Verlag, Berlin, Heidelberg.

Lafferty, J. and Zhai, C. (2001). Document language models, query models, and risk minimization for information retrieval. In *SIGIR '01: Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 111–119, New York, NY, USA. ACM.

Li, Y., Luk, W. P. R., Ho, K. S. E., and Chung, F. L. K. (2007). Improving weak ad-hoc queries using wikipedia as external corpus. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '07, pages 797–798, New York, NY, USA. ACM.

Lv, Y. and Zhai, C. (2010). Positional relevance model for pseudo-relevance feedback. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '10, pages 579–586, New York, NY, USA. ACM.

Manning, C. D., Raghavan, P., and Schutze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Metzler, D. and Croft, W. B. (2007). Latent concept expansion using markov random fields. In *SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 311–318, New York, NY, USA. ACM.

Milne, D. N., Witten, I. H., and Nichols, D. M. (2007a). A knowledge-based search engine powered by wikipedia. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, CIKM '07, pages 445–454, New York, NY, USA. ACM.

Milne, D. N., Witten, I. H., and Nichols, D. M. (2007b). A knowledge-based search engine powered by wikipedia. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, CIKM '07, pages 445–454, New York, NY, USA. ACM.

Mitra, M., Singhal, A., and Buckley, C. (1998). Improving automatic query expansion. In *SIGIR '98: Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 206–214, New York, NY, USA. ACM.

Qiu, Y. and Frei, H.-P. (1993). Concept based query expansion. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '93, pages 160–169, New York, NY, USA. ACM.

Sakai, T., Manabe, T., and Koyama, M. (2005). Flexible pseudo-relevance feedback via selective sampling. *ACM Transactions on Asian Language Information Processing (TALIP)*, 4(2):111–135.

Salton, G. (1971). *The SMART Retrieval System: Experiments in Automatic Document Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Smeaton, A. F., Kelledy, F., and O'Donnell, R. (1995). Thresholding posting lists, query expansions with wordnet and pos tagging of spanish. In *The Fourth Text REtrieval Conference (TREC-4*, pages 373–390.

Tao, T. and Zhai, C. (2006). Regularized estimation of mixture models for robust pseudo-relevance feedback. In *SIGIR '06: Proceedings of the 29th Annual International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 162–169, New York, NY, USA. ACM Press.

Voorhees, E. (2006). Overview of the trec 2005 robust retrieval track. In *E. M. Voorhees and L. P. Buckland, editors, The Fourteenth Text REtrieval Conference, TREC 2005*, Gaithersburg, MD. NIST.

Voorhees, E. M. (1994a). Query expansion using lexical-semantic relations. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '94, pages 61–69, New York, NY, USA. Springer-Verlag New York, Inc.

Voorhees, E. M. (1994b). Query Expansion using Lexical-Semantic Relations. In *SIGIR '94: Proceedings of the 17th Annual International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 61–69, New York, NY, USA. Springer-Verlag New York, Inc.

Xu, Y., Jones, G. J., and Wang, B. (2009a). Query dependent pseudo-relevance feedback based on wikipedia. In *SIGIR '09: Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 59–66, New York, NY, USA. ACM.

Xu, Y., Jones, G. J., and Wang, B. (2009b). Query dependent pseudo-relevance feedback based on wikipedia. In *Proceedings of the 32nd inter-*

national ACM SIGIR conference on Research and development in information retrieval, SIGIR '09, pages 59–66, New York, NY, USA. ACM.

Yin, Z., Shokouhi, M., and Craswell, N. (2009). Query expansion using external evidence. In Boughanem, M., Berrut, C., Mothe, J., and Soule-Dupuy, C., editors, *Advances in Information Retrieval*, volume 5478 of *Lecture Notes in Computer Science*, pages 362–374. Springer Berlin / Heidelberg. 10.1007/978-3-642-00958-7_33.

Zhai, C. and Lafferty, J. (2001). Model-based Feedback in the Language Modeling approach to Information Retrieval. In *CIKM '01: Proceedings of the 10th International Conference on Information and Knowledge Management*, pages 403–410, New York, NY, USA. ACM Press.