United we fall, Divided we stand: A Study of Query Segmentation and PRF for Patent Prior Art Search

Debasis Ganguly Johannes Leveling Gareth J. F. Jones
Centre for Next Generation Localisation
School of Computing, Dublin City University
Dublin 9, Ireland
{dganguly, jleveling, gjones}@computing.dcu.ie

ABSTRACT

Previous research in patent search has shown that reducing queries by extracting a few key terms is ineffective primarily because of the vocabulary mismatch between patent applications used as queries and existing patent documents. This finding has led to the use of full patent applications as queries in patent prior art search. In addition, standard information retrieval (IR) techniques such as query expansion (QE) do not work effectively with patent queries, principally because of the presence of *noise* terms in the massive queries. In this study, we take a new approach to QE for patent search. Text segmentation is used to decompose a patent query into self-coherent sub-topic blocks. Each of these much shorted sub-topic blocks which is representative of a specific aspect or facet of the invention, is then used as a query to retrieve documents. Documents retrieved using the different resulting sub-queries or query streams are interleaved to construct a final ranked list. This technique can exploit the potential benefit of QE since the segmented queries are generally more focused and less ambiguous than the full patent query. Experiments on the CLEF-2010 IP prior-art search task show that the proposed method outperforms the retrieval effectiveness achieved when using a single full patent application text as the query, and also demonstrates the potential benefits of QE to alleviate the vocabulary mismatch problem in patent search.

Categories and Subject Descriptors

H.3.3 [**INFORMATION STORAGE AND RETRIEVAL**]: Information Search and Retrieval—*Query formulation, Relevance Feedback*; H.3.1 [**INFORMATION STORAGE AND RETRIEVAL**]: Content Analysis and Indexing—*Abstracting methods*

General Terms

Experimentation, Performance, Measurement

Keywords

Query segmentation, query expansion, pseudo relevance feedback, patent prior art search

PaIR'11, October 24, 2011, Glasgow, Scotland, UK.
.

1. INTRODUCTION

Patent prior art search involves retrieval of already filed patents which are potential candidates to constitute prior art for a patent claim, henceforth referred to as a query patent in this paper, to help patent examiners check the novelty of the claimed work. The major differences in the characteristics of patent queries with those of traditional ad-hoc search and web search queries are: a) queries being full patent claims comprise of thousands of words on average, whereas queries in ad-hoc search and web search are very short often consisting of only two or three words; b) the amount of vocabulary mismatch between a patent query and patent documents is higher due to the obscure style of writing a patent claim (patentese) as compared to that of standard text, e.g. TREC style ad-hoc collections and topics which involve news articles comprising a much simpler and non technical vocabulary.

Due to the above characteristics of patent queries, standard IR techniques are not very effective for patent prior art search because: firstly the queries are too long to be unambiguously describing a specific information need, and secondly recall is often low due to the higher degree of vocabulary mismatch. Naive methods of extraction of *key* terms to form reduced queries resembling standard IR ad-hoc search queries, increase the vocabulary mismatch further and thus lead to a degradation of recall. A long query consisting of thousands of terms on the other hand is likely to contain a lot of terms which are heavily distributed in documents not relevant to the query, thus drifting the query away from the relevant documents and hurting precision. So while on one hand it is desirable to reduce the queries to gain more specificity in the description of the information need, on the other hand there is a desire to expand queries to alleviate the vocabulary mismatch.

Previous research on patent prior art search has reported failures to improve retrieval effectiveness by query expansion (QE) over the initial retrieval (Section 2 overviews some of this work). The reason behind the failure of such a standard and successfully time-tested methodology of improving ad-hoc search can be attributed to the very different characteristics of patent queries in comparison to standard short ad-hoc search queries as mentioned earlier. QE for patent queries tends to make the massive queries more ambiguous, thus hurting retrieval effectiveness. This paper proposes a methodology to adapt the standard QE technique for patent prior art search by addressing the root cause of failure of QE. In our proposed method we decompose each patent query into self coherent sub-topics, which are less ambiguous as compared to the whole query, and hence more precise in pointing to a specific information need. These individual sub-topics segments are used as separate queries (which we call a query stream) for initial retrieval. Pseudo relevance feedback (PRF) is then applied individually on each of these retrieval streams. The final result set for the query is then obtained by merging the results from each of the streams. The underlying hypothesis behind this idea is that the individual query streams, being less ambiguous, can retrieve more documents focusing on each sub-topic of the whole patent application, and thus are potentially better candidates for QE.

The remainder of the paper is organized as follows: Section 2 surveys related work on patent prior art search, Section 3 describes our proposed method of query segmentation in details, Section 4 describes the experiments and discusses the results followed by Section 5 which provides further per query stream analysis, and finally Section 6 concludes the paper with directions for future work.

2. RELATED WORK

The real-life working principle undertaken by patent examiners to manually formulate queries for invalidating claims involves selecting high frequency terms from the query-patent claim text. Some early work on keyword extraction to form a reduced query, modelled on this real-life methodology of patent searchers includes that by [14, 4]. A more recent work by Xue and Croft [21] advocates the use of full patent text as the query to reduce the burden on the patent examiners, and concludes with the fact that usage of the whole patent text with raw term frequencies gives the best mean average precision (MAP). Some recent work on the CLEF-IP [1] task showed that the best retrieval results are obtained when terms are used from all the fields of the query patents [18]. While one of these works uses a simple frequency threshold to filter out low-frequency terms from the query [8], others involve computing KL-divergence [11] and supervised learning, trained on features such as position of a term within a patent-query and collection based measures of *tf-idf* and Generalized Dice Coefficient [6]. Fujii [2] shows that retrieval effectiveness can be improved by merging IR methods with citation extraction. Magdy et al. [10] show that the best performing run of CLEF-IP 2010 uses citations extracted by training a Conditional Random Field, whereas the second best run uses a list of citations extracted from the patent numbers within the description field of some patent queries. They also show that a simple IR approach of using terms from all fields of a patent-query with a frequency of at least two, merged with extracted citations achieves a statistically indistinguishable performance compared to the best run which employs complex retrieval methods using two complementary indexes, one constructed by extracting terms from the patent collection and the other built by merging several terminological resources such as Wikipedia etc. As the baseline run for this paper we follow the simpler approach of the second ranked participating group.

Ad-hoc IR on news and web data has been shown to improve with respect to both MAP and average recall measures by the use of PRF, due to the fact that additional terms from pseudo-relevant documents bridge the vocabulary gap between the query and the documents relevant to it in the collection [20, 19]. However, QE is associated with the risk of additional terms contributing to a drift in the original information need with a resultant degradation in retrieval effectiveness in the feedback step [17]. Unfortunately all the existing work on PRF coupled with QE on patent prior art search report a degradation in retrieval effectiveness [5, 16, 9].

Takaki et al. [15] prescribe decomposing a patent query into sub-topics and form the final retrieval results by fusing the individual retrieval results for the decomposed queries by a weighted combined summation of similarities. Although our work at a first glance might appear similar to [15], there are some major differences which we explain as follows: a) the existing work involved segmenting query patents into sub-topics and extracting keywords from each of these sub-topics for retrieval, whereas we use the full text of each of the segments as individual sub-queries conforming to more recent findings suggesting the use of full patent text as queries [21, 18]; b) we do not distinguish between the relative importance of the individual sub-topics by specificity measures as was done in [15]; c) the existing work used a standard fusion technique of weighted COMBSUM [1], whereas we use a 1-way interleaving of the individual result-lists; d) our motivation for query segmentation is driven by an effort to adapt QE for patent prior art search whereas QE was not addressed in [15];

Magdy and Jones [7] report that MAP can be a misleading metric for patent prior art search because of its inherent characteristic of favouring precision more than recall and proposed a metric named PRES (Patent Retrieval Evaluation Score) which focuses on recall at early ranks conforming to the objective of patent prior art search. In our experiments, we report improvement of MAP, PRES and average recalls over the baseline.

3. QUERY REDUCTION

3.1 Motivation

The important observations to be made from existing work as discussed in Section 2 are that patent prior art search achieves best retrieval performance when: i) information from all fields of the query patents are used; ii) unit frequency terms (UFT), i.e. terms which occur only once in the patent query are eliminated; and iii) no PRF is applied. The last point in fact provides the motivation to analyze the ulterior reasons for the failure of QE and to devise a method leading to better PRF performance for the patent prior art search task; this forms the basis of our work described in this paper.

Expository texts such as a patent application comprise of a multitude of self-coherent densely discussed sub-topics [3]. Each of these sub-topics in a patent query typically expresses a particular aspect of the claimed invention, and the prior art search task requires all existing patent documents to be retrieved for each sub-topic. Using a full patent claim as a query is therefore associated with a risk of under specifying each sub-topic precisely, leading to an ambiguous expression of the information need. Also in such a scenario, expanding the query further contributes to a degradation in the specificity of information need hurting retrieval effectiveness further.

To alleviate the above issues, we propose to use each of the sub-topics or segments of a whole patent as queries to produce individual query streams to be given as inputs to a retrieval system, and then to merge the retrieval results from each of the individual streams to construct the final ranked list for the whole original query. Using each sub-topic as a query stream should enable a retrieval model to retrieve related documents from the collection in a more precise way and should also allow the PRF algorithm to work on a more focused set of pseudo-relevant documents.

3.2 Segmented Retrieval Algorithm

The general outline of our proposed method of retrieval using query segments or streams is as follows:

1. Segment each patent query Q into the constituent fields: title (Q_t), abstract (Q_a), description (Q_d), and claim (Q_c).
2. Segment each Q_d into $\eta(Q_d)$ segments Q_d^1, $Q_d^2 \ldots Q_d^{\eta(Q_d)}$ by TextTiling [3].
3. Remove UFTs from each query stream.

[1] http://www.ir-facility.org/clef-ip

4. Run retrieval on each query stream: Q_t, Q_a, Q_c and $Q_d^1 \ldots Q_d^{\eta(Q_d)}$. Let the list of documents retrieved for a stream q be $L(q)$.
5. Interleave one document from each $L(q)$, eliminating duplicates while interleaving, in a round-robin manner to construct the initial retrieval ranked list for the whole query Q.
6. Perform QE using R pseudo-relevant documents from each initial ranked list $L(q)$ and add T terms to each query stream q. Call the expanded query stream q'.
7. Perform retrieval to obtain feedback ranked lists $L(q')$ on each expanded query stream q', and build up the feedback retrieval result for the original query Q in the exact same manner as Step 5.

We explain the rationale behind each step as follows. In step 2, we segment only the description field because the description field is the longest among all patent fields, comprising multiple subtopics or aspects of the claimed invention, thus making it a suitable candidate for text segmentation. In step 5 we use interleaving as opposed to the more standard fusion techniques such as COMBSUM etc. [1]. This is because COMBSUM is particularly useful for merging results retrieved by different retrieval algorithms executed against the same query, but in our case it is the queries which are different and not the retrieval algorithm. More precisely speaking, every query stream is a sub-topic or one specific aspect of the whole information need and we expect that the relevant set should comprise of documents from each of these query streams. This is what we do by the one-way interleaving or choosing documents in a round-robin manner from the ranked lists retrieved against each query stream. Thus, in the merged result set we end up with documents from each sub-topic.

4. EXPERIMENTS

4.1 Description and Parameter Settings

To evaluate our approach, we use the patent document corpus of CLEF-IP 2010, which consists of 2.68 million patents from the European Patent Office (EPO). We restrict our retrieval experiments only to the English subset which constitutes 68% of the collection. The topic set comprises of 1348 topics which are patent applications having *title*, *abstract*, *claims* and *description* fields. For testing our query segmentation approach, we restrict our investigation on a total of 50 topics, these being the top 50 in the list of query names ordered lexicographically. In addition to using a standard list of stopwords[2], we also removed formulae, numeric references, chemical symbols and the patent jargons such as *method*, *system*, *device* etc. Porter stemmer [13] was used to stem the words. Language Modeling (LM), involving Jelineck-Mercer smoothing implemented in SMART,[3] was used for retrieval. As a value of the smoothing parameter, we use $\lambda = 0.6$ for all our retrieval experiments.

The objective of our experiments is two-fold: firstly to demonstrate that decomposing a query into segments and retrieving with the individual streams can perform better than retrieving with the whole query; and secondly to show that QE can perform better on the individual streams. For the former case, our baseline is a reproduction of the methodology of the second best performing run of CLEF-IP 2010, which is statistically indistinguishable than the best run [10]. We call this baseline `WHOLE`. For the later case, we have two baselines, the first being PRF on the retrieval run `WHOLE`, which we call `WHOLE_PRF`, to show the degree of effectiveness of PRF when whole patents are used as queries. To test the effectiveness of PRF on the query streams, it is also compared against the initial retrieval obtained from the streams, which forms the second baseline for our QE run. As a QE technique, we use the LM score based QE as proposed by Ponte and Croft [12] on the whole query and on the respective query streams.

4.2 Query Segmentation Results

In this section, we report the results of executing the algorithm described in Section 3.2 without the QE step.

Table 1: Segmented vs. whole query retrieval.

Run Name	Parameters		Evaluation metric		
	Segmented	Fusion method	PRES	MAP	Recall
`WHOLE`	No	N/A	0.4413	0.0899	0.5310
`SEG_COMBSUM`	Yes	COMBSUM	0.1545	0.0308	0.1759
`SEG_RR`	Yes	Round-robin	**0.4949**	**0.0947**	**0.5982**

It can be seen from Table 1 that the method of retrieving by separate query streams works well in conjunction with the 1-way interleaving of documents returned for each query stream. By comparison combination of documents by the standard fusion technique produces very poor results. The reason is due to the fact that the standard fusion techniques have been devised to merge retrieval results obtained for the same query by different retrieval techniques. But in our case we obtain the query streams by applying TextTiling to the full query description, which draws boundaries at sharp valleys of plotted cosine similarities between consecutive blocks of sentences. Thus the query streams, comprising of the textual contents of the output of TextTiling, are minimally similar to each other. The documents retrieved for each of the individual streams are mostly expected to be non overlapping, and hence not conducive to be fused by the standard technique of COMBSUM.

4.3 PRF Results

In this section we report the post feedback results both on whole queries and segmented queries. Table 2 summarizes the results, in this we include the whole and the segmented runs from Table 1 for the sake of continuity. Both the segmented runs reported in this table use 1-way interleaving which is not shown as a separate parameter. The columns R and T denote the number of pseudo relevant documents, and terms added for QE respectively. After a set of initial experiments we found that the feedback performs best with the setting of $(R, T) = (10, 10)$ and hence the reported results in this table use the same.

Table 2: Pseudo Relevance Feedback on segmented retrieval.

Run Name	Parameters				Evaluation metric		
	Segmented	PRF	R	T	PRES	MAP	Recall
`WHOLE`	No	No	-	-	0.4413	0.0899	0.5310
`WHOLE_PRF`	Yes	Yes	10	10	0.4415	0.0889	0.5333
`SEG`	Yes	No	-	-	0.4949	0.0947	0.5982
`SEG_PRF`	Yes	Yes	10	10	**0.5033**	**0.1025**	**0.6166**

The table shows that the relative gains from QE are higher if it is performed on each of the streams separately, and the results then merged, as is evident from comparing the results of `SEG_PRF` and `WHOLE_PRF`. `WHOLE_PRF` results in almost negligible gains in PRES and average recall, and a very slight decrease of MAP. A possible

[2] `http://members.unine.ch/jacques.savoy/clef/`
[3] `ftp://ftp.cs.cornell.edu/pub/smart`

reason for this very small change in the results can be due to the fact that the queries are already very large and that an additional 10 terms cannot produce a pronounced change in the retrieval effect. Whereas for the segmented case, since the queries are much shorter, an additional 10 terms can play a pivotal role in changing retrieval results. This can be verified from the fact that SEG_PRF retrieves relevant documents, as can be seen from the 3.1% relative increase in average recall compared to the run SEG.

5. ANALYSIS OF RETRIEVAL STREAMS

In this section we report and analyze the per stream retrieval performance for our proposed method. The first sub-section talks about the ranks of the relevant documents retrieved in each stream, followed by a subsection on the relative gains in feedback for each retrieval stream.

5.1 Per Stream Ranks of Relevant Documents

Let us assume that we need to retrieve N documents for every original patent query and let τ be the average number of query streams over the set of whole patent queries. Thus, the expected number of documents we pick up from each list to construct the final retrieved set for the whole query, is $c = N/\tau$. The potential worst case of the segmented retrieval algorithm can arise when there is no overlap in the retrieved sets of documents and all the relevant documents have been retrieved at ranks beyond c.

For the CLEF-IP task, $N = 1000$ and from the output of TextTiling on the query set we find that $\tau = 17.66$, i.e. on average we end up decomposing every whole patent query into almost 18 streams. The expected position in the ranked lists we need to visit, starting from their tops, is thus $1000/17.66 \approx 57$. Thus our proposed algorithm can work well if all the streams retrieve a good number of relevant documents within the top 57 positions, and hence it will be interesting to see the number of relevant documents retrieved within the cut-off rank of 57.

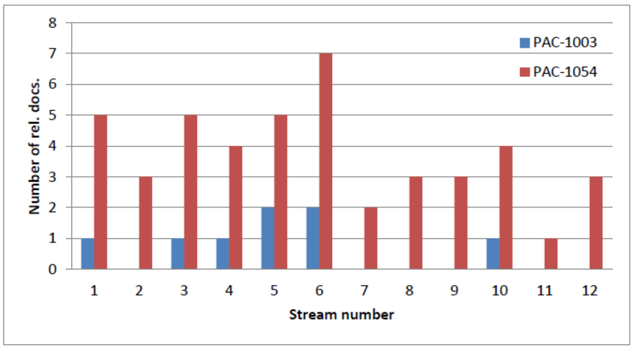

Figure 1: Per stream analysis of the best (PAC-1054) and the worst (PAC-1003) performing query.

Figure 1 shows the number of relevant documents retrieved within a cut-off value of 57 for two query instances. PAC-1054 is the best performing query in terms of gain in PRES achieved by SEG relative to WHOLE, whereas PAC-1003 is the query which suffers from maximum relative loss of PRES. The reason why query PAC-1054 is able to achieve good performance can be seen from the fact that the individual streams retrieve many relevant documents within the average rank cut-off.

5.2 Per Stream PRF Performance

The best performing query in terms of relative PRES gain (from SEG to SEG_PRF) is the query named PAC-1038 having a 59.9% increase in PRES. The best performing query, involving PRF on whole queries, is the query named PAC-1036 with a relative gain (from WHOLE to WHOLE_PRF) in PRES of only 1.48%.

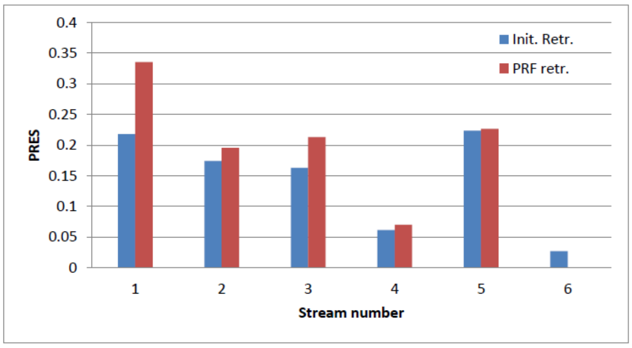

Figure 2: Feedback effect on each query stream for the best performing query PAC-1038.

Figure 2 shows that all the query streams (except the one numbered 6) register an increase in PRES. The small increases for each separate query stream contribute to the overall increase of 59.9% increase in PRES.

In order to see the feedback effects per query (or per query stream for the segmented retrieval), we categorize every query stream into buckets of initial retrieval metric ranges. This way of categorizing the queries allows us to look at the performance over a group of queries having an initial retrieval measure of very poor $(0 - 0.2)$, poor $(0.2 - 0.4)$, average $(0.4 - 0.6)$, good $(0.6 - 0.8)$ or excellent $(0.8 - 1.0)$. For example, if the initial retrieval PRES for 5 queries are 0.15, 0.23, 0.25, 0.68 and 0.52, we place the first query in bucket-1, the next two in bucket-2, the next one in bucket-4 and the last one in bucket-3. We categorize for the other metrics average precision and recall in an identical manner. Table 3 shows the relative gains in the three metrics averaged over the groups for the runs WHOLE_PRF and SEG_PRF to compare unsegmented feedback against segmented one.

From table 3 we can see that WHOLE_PRF results in a very slight increase of PRES in each query group, whereas the method SEG_PRF yields a considerable increase in percentage gain of PRES, MAP and Recall for the stream group $[0, 0.2)$. The next group also registers a good increase of PRES. This shows that feedback in this case is improving the retrieval effectiveness of query streams for which the initial retrieval results are poor.

Thus we see that although expansion of the queries as a whole produces negligible changes in average precision in each query group, the changes are non-negligible when we decompose the full patent queries into much smaller segments. This observation verifies our hypothesis that QE can be successfully applied to patent search if the queries are decomposed into shorter and unambiguous segments.

6. CONCLUSION

This paper presented a technique for applying QE in patent prior art search by decomposing each patent query into self coherent blocks of text, retrieving against each query stream thus constructed, and merging the results by a simple one way interleaving of documents from the individual lists assuming that each query stream represents a separate aspect of the information need, and that documents collected from all the individual sub-topics in a round-robin manner would satisfy each facet of the information need. We evalu-

Table 3: PRF on whole vs. segmented queries

Run name	Interval range	PRES # queries	PRES # improved	PRES % improved	MAP # queries	MAP # improved	MAP % improved	Recall # queries	Recall # improved	Recall % improved
WHOLE_PRF	[0.0,0.2)	13	3	+0.18	41	20	-1.54	12	0	+0.00
	[0.2,0.4)	8	6	**+0.25**	8	4	-0.95	5	0	+0.00
	[0.4,0.6)	11	5	+0.03	0	0	+0.00	8	0	+0.00
	[0.6,0.8)	11	4	-0.01	1	0	+0.00	10	1	**+1.68**
	[0.8,1.0]	7	6	+0.07	0	0	+0.00	15	0	+0.00
SEG_PRF	[0.0,0.2)	472	235	**+14.04**	775	433	**+157.77**	357	81	**+25.09**
	[0.2,0.4)	328	213	+3.55	9	2	-1.71	391	79	+3.33
	[0.4,0.6)	54	46	+1.33	0	0	+0.00	103	7	+0.31
	[0.6,0.8)	35	27	+0.23	0	0	+0.00	38	0	+0.00
	[0.8,1.0]	0	0	+0.00	105	47	-17.78	0	0	+0.00

ated our approach on a subset of 50 queries from the CLEF-IP 2010 patent query set. The results show that query segmentation alone can result in a 12.14% increase in PRES and segmentation coupled with QE results in a 14.05% increment of PRES. Although the improvements are statistically not significant under the Wilcoxon test with 95% confidence level, the percentage gains themselves are non negligible. Our proposed method thus helps to unleash the power of feedback which has not been found to be possible in previous research on this topic.

We do have some loop holes in our assumption that all the query streams are very dissimilar to each other because: a) segmentation is done only for the description field and a segment of the description may in fact be very similar to the abstract or to the claim query streams; b) a particular aspect of an invention claim may return in the description text as a sub-topic in which case we have near duplicate pairs of query streams. A possible solution to the first problem is to use a combined merging style interleaving for the description streams and COMBSUM for the rest. The second problem may be solved by putting the description text segments into separate clusters and use the cluster centroids as query streams. Attempting the above two possible approaches to further improve retrieval effectiveness for patent search will form the basis of our future work.

Acknowledgments

This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (CNGL) project.

7. REFERENCES

[1] E. A. Fox and J. A. Shaw. Combination of multiple searches. In D. K. Harman, editor, *The Second Text REtrieval Conference (TREC-2)*, pages 243–252. National Institute for Standards and Technology. NIST Special Publication 500-215, 1994.

[2] A. Fujii. Enhancing patent retrieval by citation analysis. In *SIGIR*, pages 793–794. ACM, 2007.

[3] M. Hearst. TextTiling: Segmenting text into multi-paragraph subtopic passages. *CL*, 23(1):33–64, 1997.

[4] H. Itoh, H. Mano, and Y. Ogawa. Term distillation in patent retrieval. In *Proceedings of the ACL-2003 workshop on Patent corpus processing - Volume 20*, pages 41–45, Stroudsburg, PA, USA, 2003.

[5] K. Kishida. Experiment on pseudo relevance feedback method using taylor formula at NTCIR-3 patent retrieval task. In *NTCIR-3*, 2003.

[6] P. Lopez and L. Romary. Experiments with citation mining and key-term extraction for prior art search. In *CLEF-2010 (Notebook Papers/LABs/Workshops)*, 2010.

[7] W. Magdy and G. J. Jones. PRES: a score metric for evaluating recall-oriented information retrieval applications. In *SIGIR*, pages 611–618, 2010.

[8] W. Magdy and G. J. F. Jones. Applying the KISS principle for the CLEF-IP 2010 prior art candidate patent search task. In *CLEF-2010 (Notebook Papers/LABs/Workshops)*, 2010.

[9] W. Magdy, J. Leveling, and G. J. F. Jones. Exploring structured documents and query formulation techniques for patent retrieval. In *10th Workshop of the Cross-Language Evaluation Forum, CLEF 2009*, pages 410–417, 2010.

[10] W. Magdy, P. Lopez, and G. J. F. Jones. Simple vs. sophisticated approaches for patent prior-art search. In *ECIR*, pages 725–728, 2011.

[11] J. Pérez-Iglesias, Á. Rodrigo, and V. Fresno. Using BM25F and KLD for patent retrieval. In *CLEF-2010 (Notebook Papers/LABs/Workshops)*, 2010.

[12] J. M. Ponte. *A language modeling approach to information retrieval*. PhD thesis, University of Massachusetts, 1998.

[13] M. F. Porter. An algorithm for suffix stripping. *Program*, 14(3):130–137, 1980.

[14] T. Takaki. Query terms extraction from patent document for invalidity search. In *NTCIR-5*, 2005.

[15] T. Takaki, A. Fujii, and T. Ishikawa. Associative document retrieval by query subtopic analysis and its application to invalidity patent search. In *Proceedings of CIKM*, CIKM '04, pages 399–405, 2004.

[16] H. Takuechi, N. Uramoto, and K. Takeda. Experiments on patent retrieval at NTCIR-5 workshop. In *NTCIR-5*, 2005.

[17] E. L. Terra and R. Warren. Poison pills: harmful relevant documents in feedback. In *Proceedings of CIKM*, pages 319–320. ACM, 2005.

[18] M. Z. Wanagiri and M. Adriani. Prior art retrieval using various patent document fields contents. In *CLEF-2010 (Notebook Papers/LABs/Workshops)*, 2010.

[19] R. H. Warren and T. Liu. A review of relevance feedback experiments at the 2003 Reliable Information Access (RIA) workshop. In *SIGIR 2004*, pages 570–571. ACM, 2004.

[20] J. Xu and W. B. Croft. Improving the effectiveness of informational retrieval with Local Context Analysis. *ACM Transactions on information systems*, 18:79–112, 2000.

[21] X. Xue and W. B. Croft. Transforming patents into prior-art queries. In *SIGIR*, pages 808–809, 2009.