

# Comments-Oriented Query Expansion for Opinion Retrieval in Blogs

Jose M. Chenlo<sup>1</sup>, Javier Parapar<sup>2</sup>, and David E. Losada<sup>1</sup>

<sup>1</sup> Centro de Investigación en Tecnoloxías da Información (CITIUS)  
Universidad de Santiago de Compostela, Spain  
{josemanuel.gonzalez,david.losada}@usc.es

<sup>2</sup> IRLab, Computer Science Department  
University of A Coruña, Spain  
javierparapar@udc.es

**Abstract.** In recent years, Pseudo Relevance Feedback techniques have become one of the most effective query expansion approaches for document retrieval. Particularly, Relevance-Based Language Models have been applied in several domains as an effective and efficient way to enhance topic retrieval. Recently, some extensions to the original RM methods have been proposed to apply query expansion in other scenarios, such as opinion retrieval. Such approaches rely on mixture models that combine the query expansion provided by Relevance Models with opinionated terms obtained from external resources (e.g., opinion lexicons). However, these methods ignore the structural aspects of a document, which are valuable to extract topic-dependent opinion expressions. For instance, the sentiments conveyed in blogs are often located in specific parts of the blog posts and its comments. We argue here that the comments are a good guidance to find on-topic opinion terms that help to move the query towards burning aspects of the topic. We study the role of the different parts of a blog document to enhance blog opinion retrieval through query expansion. The proposed method does not require external resources or additional knowledge and our experiments show that this is a promising and simple way to make a more accurate ranking of blog posts in terms of their sentiment towards the query topic. Our approach compares well with other opinion finding methods, obtaining high precision performance without harming mean average precision.

**Key words:** Information retrieval, opinion mining, blogs, comments, relevance models, pseudo relevance feedback, query expansion

## 1 Introduction and Motivation

The blogosphere is one of the most important sources of opinion in the Internet [1]. Given a query and a collection of blogs, several methods have been proposed to retrieve opinions related to the query topic [1]. The most popular choice is to consider this task as a two-stage process that involves a topic retrieval stage (i.e., retrieve on-topic posts), and a re-ranking stage based on opinion features

[2]. The first stage usually involves ad-hoc search with popular Information Retrieval (IR) models (e.g., BM25). The second stage is a more complex task with many unresolved issues (e.g., irony, off-topic opinions, mixed polarity). Most successful approaches search for documents that are both opinionated and on-topic by considering positional information as the best guidance to find on-topic opinions [3,4]. Pseudo Relevance Feedback (PRF) combined with external opinion resources has also been proposed to support opinion finding [5]. However, most studies ignore the structural aspects of a blog post to determine opinions. This is unfortunate because sentiments often appear in specific locations of the text. For instance, in the study of blog comments presented in [6], Mishne and Glance found that comments constitute a substantial part of the blogosphere, accounting for up to 30% of the total volume of blog data.

In this paper we present a simple PRF strategy that exploits the potential opinions provided in the comments to improve opinion finding in blogs. In particular, we use one of the most robust and effective PRF techniques: Relevance-Based Language Models (RM) [7]. Several estimations for RM have been proposed in the literature, being the so-called RM3 [8] the approach that performs best [9]. In this work we present an alternative RM3 estimation for selecting expansion terms from comments. We estimate the relevance model from these highly opinionated parts of the blogs with the objective of selecting opinionated and on-topic expansion terms. We compare the performance of our comments-based approach against the standard RM3 formulation. Our experiments show that the new expansion method is promising when compared to global approaches that consider the whole document to do expansion.

## 2 Background

Nowadays, advanced search tasks need to go beyond a ranked list of relevant documents. One of these tasks is opinion retrieval [10,1], where opinions need to be integrated within the retrieval task. For instance, in the TREC Blog Track [2] the participants are asked to search for blog pages that express an opinion about a given topic. This task can be summarised as: *What do people think about X?* [2] and is often addressed in two stages. First, a ranking of documents related to the topic ( $X$ ) is obtained and, next, the initial list is re-ranked using opinion-based features. The output is a ranking of documents in decreasing order of their estimated subjectivity with respect to the query.

Relevance Models explicitly introduced the concept of relevance in the Language Modeling (LM) framework [7]. In RM, the original query is considered a very short sample of words obtained from a relevance model  $R$  and relevant documents are larger samples of text from the same model. From the words already seen, the relevance model is estimated. If more words from  $R$  are needed then the words with the highest estimated probability are chosen. The terms in the

vocabulary are therefore sorted according to these estimated probabilities. After doing some assumptions the RM1 method is defined as:

$$P(w|R) \propto \sum_{d \in \mathcal{C}} P(d) \cdot P(w|d) \cdot \prod_{i=1}^n P(q_i|d) \quad (1)$$

Usually,  $P(d)$  is assumed to be uniform.  $\prod_{i=1}^n P(q_i|d)$  is the query likelihood given the document model, which is traditionally computed using Dirichlet smoothing.  $P(w|d)$  accounts for the importance of the word  $w$  within the document  $d$ . The process follows four steps:

1. Initially, the documents in the collection ( $\mathcal{C}$ ) are ranked using a standard LM retrieval model (e.g., query likelihood with Dirichlet smoothing).
2. The top  $r$  documents from the initial retrieval are taken for driving the estimation of the relevance model. In the following, this pseudo relevant set will be referred to as  $RS$ .
3. The relevance model’s probabilities,  $P(w|R)$ , are calculated from the estimate presented in Eq. 1, using  $RS$  instead of  $\mathcal{C}$ .
4. The expanded query is built with the  $e$  terms with highest estimated  $P(w|R)$ .

RM3 [8] is a later extension of RM that performs better than RM1. RM3 interpolates the terms selected by RM1 with a LM computed from the original query:

$$P(w|q') = (1 - \lambda) \cdot P(w|q) + \lambda \cdot P(w|R) \quad (2)$$

Negative cross entropy with the expanded query is used to get the final ranking.

### 3 Comments-Biased Relevance Model

As we discussed in Section 1, people tend to express opinions related to the topic of the blog post when they write comments. We argue that the comments of a blog post are more densely populated by opinions than other parts of the document. Therefore, we hypothesize that terms in comments are highly opinionated and on-topic and therefore, a simple PRF technique that takes advantage of these specific words to expand the original query will be a very promising tool to improve opinion finding in blogs. We have designed an alternative RM3 estimation in which Eq. 1 is modified to promote terms that appear in the comments of the blog post:

$$P(w|R) \propto \sum_{d \in RS} P(d) \cdot P(w|d_{comm}) \cdot \prod_{i=1}^n P(q_i|d) \quad (3)$$

where  $w$  is any word appearing in the set of comments associated to documents in  $RS$  and  $P(w|d_{comm})$  is computed as the probability of  $w$  in the set of comments of document  $d$ . In this way, the comments act as proxies of the documents in terms of opinion. Observe that the estimation of the query likelihood remains at document level because the effect of topic relevance on the estimation of the relevance model is better encoded using the whole document.

Finally, both  $P(w|d_{comm})$  and  $P(q_i|d)$  are estimated using Dirichlet smoothing:

$$P(w|d_{comm}) = \frac{f_{w,d_{comm}} + \mu \cdot P(w|C_{comm})}{|d_{comm}| + \mu} \quad (4)$$

$$P(q_i|d) = \frac{f_{q_i,d} + \mu \cdot P(q_i|C)}{|d| + \mu} \quad (5)$$

where  $f_{q_i,d}$  is the number of times that the query term  $q_i$  appears in document  $d$ , and  $f_{w,d_{comm}}$  is the number of times that the word  $w$  appears in the document that is constructed by concatenating all the comments associated to  $d$  ( $d_{comm}$ ).  $|d|$  and  $|d_{comm}|$  are the number of words in  $d$  and  $d_{comm}$ , respectively.  $P(q_i|C)$  is the probability of  $q_i$  in the collection of documents  $C$  and  $P(w|C_{comm})$  is the probability of  $w$  in the collection of comments.  $\mu$  is a smoothing parameter that we have to train.

## 4 Experiments

In our experiments we used the well-known BLOGS06 test collection [11]. We considered the TREC 2006, TREC 2007, and TREC 2008 blog track’s benchmarks, all of which have the BLOGS06 as the reference collection. One of the core tasks in these tracks is the opinion finding task, i.e., given a query topic, systems have to return a ranking of subjective blog posts related to the query. As usual in TREC, each query topic contains three different fields (title, description, and narrative). We only used the title field, which is short and the best representation of real user web’s queries [2]. Documents were pre-processed and segmented into posts and comments following the heuristic method proposed in [12]. We also removed 733 common words from documents and queries.

Documents were judged by TREC assessors in two different aspects: i) Topic relevance: a post can be relevant, not relevant, or not judged, ii) Opinion: whether or not the on-topic documents contain explicit expression of opinion or sentiment about the topic. In this paper we are interested in this second level of judgements, focusing our attention on retrieving documents that express an explicit opinion about the query (regardless of the polarity of the opinion).

### 4.1 Baselines

In TREC 2008, to promote the study of the performance of opinion-finding methods against uniform retrieval rankings, a set of five topic-relevance retrieval runs was provided. These standard baselines use a variety of retrieval approaches, and have varying retrieval effectiveness<sup>3</sup>.

It is a standard practice to use these baselines as initial input for the opinion retrieval stage. We followed this evaluation design and applied the proposed

<sup>3</sup> Baselines were selected from the runs submitted to TREC Blog Retrieval Task 2008

RM estimation to re-rank the baselines. The measures adopted to evaluate the opinion retrieval effectiveness are mean average precision (MAP), Precision at 10 (P@10), and the Reliability of Improvement (RI) [13], which is a commonly used robustness measure for PRF methods:

$$RI(q) = \frac{n_+ - n_-}{|q|} \quad (6)$$

where  $q$  is the set of queries tested,  $n_+$  is the number of improved queries,  $n_-$  the number of degraded queries and  $|q|$  is the total number of queries in  $q$ . Observe that the gold-standard is obtained from the documents that were assessed as subjective with respect to the query topic.

## 4.2 Query Formulation

We used the Indri retrieval platform for both indexing and retrieval<sup>4</sup>. In order to apply our RM estimation under this framework, Equation 2 is implemented in the Indri’s query language as follows:

$$\begin{aligned} &\#weight(\lambda \#combine(q_1 \cdots q_{|n|}) \\ &(1 - \lambda) \#weight(P(t_1|R) \cdot t_1 \cdots P(t_e|R) \cdot t_e)) \end{aligned} \quad (7)$$

where  $q_1 \cdots q_{|n|}$  are the original query terms,  $t_1 \cdots t_e$  are the  $e$  terms with highest probability according to Equation 3, and  $\lambda$  is a free parameter to control the trade-off between the original query and the expanded terms. We selected Dirichlet [14] as the smoothing technique for our experiments.

## 4.3 Training and Testing

We trained our methods with the 100 topics provided by TREC 2006 and TREC 2007 blog track (optimising MAP) and then we used the 50 TREC 2008 topics as the testing query set. The parameters trained were the following: the smoothing parameter of Dirichlet  $\mu$  ( $\mu \in \{10, 100, 1000, 2000, 3000, 4000, 5000, 6000\}$ ), the number of documents in the pseudo relevant set  $r = |RS|$ , ( $r \in \{5, 10, 25, 50, 75, 100\}$ ), the number of terms selected for expansion  $e$  ( $e \in \{5, 10, 25, 50, 75, 100\}$ ) and the interpolation weight  $\lambda$  ( $\lambda \in \{0, .1, .2, .3, .4, .5, .6, .7, .8, .9, 1\}$ ). The parameters were tuned (independently for each baseline) for both the classical  $RM3$  estimated from the whole documents (post and comments) and for our proposal (labelled as  $RM3_C$ ) following an exhaustive exploration process (grid search).

## 4.4 Results

Table 1 and Table 2 report the experimental results. Each run was evaluated in terms of its ability to retrieve subjective documents higher up in the ranking. The

<sup>4</sup> <http://www.lemurproject.org/indri.php>

**Table 1.** Opinion finding MAP results for the TREC 2008 dataset. The symbols  $\blacktriangle(\blacktriangledown)$  and  $\triangle(\nabla)$  indicate a significant improvement(decrease) over the original baselines and the *RM3* method respectively.

Baseline	orig.	<i>RM3</i>		<i>RM3<sub>C</sub></i>	
	MAP	MAP	RI	MAP	RI
baseline1	.3239	<u>.3750<math>\blacktriangle</math></u> (+16%)	.60	<u>.3653<math>\blacktriangle</math></u> (+13%)	.56
baseline2	.2639	<u>.3117<math>\blacktriangle</math></u> (+18%)	.36	<u>.3244<math>\blacktriangle</math></u> (+23%)	<u>.52</u>
baseline3	.3564	.3739 (+5%)	.08	<u>.3753</u> (+5%)	<u>.12</u>
baseline4	<u>.3822</u>	.3652 (-4%)	<u>-.04</u>	.3688 (-4%)	-.08
baseline5	.2988	<u>.3383<math>\blacktriangle</math></u> (+13%)	.44	<u>.3385<math>\blacktriangle</math></u> (+13%)	<u>.48</u>
average	.3251	<u>.3528<math>\blacktriangle</math></u> (+8%)	.29	<u>.3545<math>\blacktriangle</math></u> (+9%)	<u>.32</u>

**Table 2.** Opinion finding P@10 results for the TREC 2008 dataset. The symbols  $\blacktriangle(\blacktriangledown)$  and  $\triangle(\nabla)$  indicate a significant improvement(decrease) over the original baselines and the *RM3* method respectively.

Baseline	orig.	<i>RM3</i>		<i>RM3<sub>C</sub></i>	
	P@10	P@10	RI	P@10	RI
baseline1	.5800	.6140 (+6%)	.18	<u>.6360</u> (+10%)	<u>.20</u>
baseline2	.5500	.5560 (+1%)	.04	<u>.6340<math>\blacktriangle\Delta</math></u> (+15%)	<u>.18</u>
baseline3	.5540	.5800 (+5%)	-.02	<u>.6460<math>\blacktriangle\Delta</math></u> (+17%)	<u>.30</u>
baseline4	.6160	.6140 (-0%)	-.04	<u>.6560</u> (+6%)	<u>.18</u>
baseline5	.5300	.5940 (+12%)	.18	<u>.6660<math>\blacktriangle\Delta</math></u> (+26%)	<u>.54</u>
average	.5660	.5916 (+5%)	.07	<u>.6476<math>\blacktriangle\Delta</math></u> (+14%)	<u>.28</u>

best value for each baseline and performance measure is underlined. Statistical significance was estimated using the Wilcoxon test at the 95% level. The symbols  $\blacktriangle$  and  $\blacktriangledown$  indicate a significant improvement or decrease over the original baselines and the symbols  $\triangle$  and  $\nabla$  indicate a significant improvement (resp. decrease) with respect to the standard *RM3* method.

**Opinion retrieval performance.** Both *RM3* and *RM3<sub>C</sub>* outperform the original baselines but *RM3<sub>C</sub>* performs the best. In terms of *MAP*, *RM3* is able to achieve improvements that are similar to those found with *RM3<sub>C</sub>*. However, in terms of *P@10*, *RM3<sub>C</sub>* shows significant improvements with respect to the baselines and with respect to *RM3*. Furthermore, *RM3<sub>C</sub>* shows higher values of *RI*. This indicates that the improvements obtained using queries expanded with terms from comments are more consistent than those obtained with terms from the whole document. These results also highlight the importance of comments to enhance precision without harming recall (*MAP* is roughly the same with either RM methods). This suggests that subjective words estimated from comments

**Table 3.** Average opinion finding MAP performance over the 5 different baselines re-ranked by TREC 2008 systems against the results achieved by  $RM3_C$  on top of those systems. The symbols  $\blacktriangle$ ( $\blacktriangledown$ ) indicate a significant (resp. decrease) improvement over the TREC systems. TREC systems that were able to outperform the original 5 topic-retrieval baselines are in bold.

TREC Run	orig.	$TREC\ run+RM3_C$		
	MAP	MAP		RI
<b>uicop1bl1r</b>	.3614	.3524	(-2%)	-.18
<b>B1PsgOpinAZN</b>	.3565	.3558	(-2%)	.10
<b>uogOP1PrintL</b>	.3412	.3510	(+3%)	.10
<b>NOpMM107</b>	.3273	.3532 $\blacktriangle$	(+8%)	.38
UWnb1Op	.3215	.3538 $\blacktriangle$	(+10%)	.33
FIUBL1DFR	.2938	.3520 $\blacktriangle$	(+20%)	.61
UniNEopLRb1	.2118	.2121	(+0%)	.18
uams08b1pr	.1378	.3347 $\blacktriangle$	(+43%)	.93

**Table 4.** Average opinion finding P@10 performance over the 5 different baselines re-ranked by TREC 2008 systems against the results achieved by  $RM3_C$  on top of those systems. The symbols  $\blacktriangle$ ( $\blacktriangledown$ ) indicate a significant (resp. decrease) improvement over the TREC systems. TREC systems that were able to outperform the original 5 topic-retrieval baselines are in bold.

TREC Run	orig.	$TREC\ run+RM3_C$		
	P@10	P@10		RI
<b>uicop1bl1r</b>	.6020	.6264	(+4%)	.14
<b>B1PsgOpinAZN</b>	.6204	.6512 $\blacktriangle$	(+5%)	.30
<b>uogOP1PrintL</b>	.5964	.6320 $\blacktriangle$	(+6%)	.25
<b>NOpMM107</b>	.5744	.6432 $\blacktriangle$	(+12%)	.37
UWnb1Op	.6068	.6500	(+7%)	.25
FIUBL1DFR	.4804	.6392 $\blacktriangle$	(+33%)	.76
UniNEopLRb1	.6156	.6464	(+5%)	.29
uams08b1pr	.1284	.6100 $\blacktriangle$	(+375%)	1.0

lead to a more accurate query-dependent opinion vocabulary. Furthermore, the independence of our method of any external lexicon is important because, in many domains and languages, there is a lack of good opinion resources.

**Comparison against TREC systems.** Our technique does not use any specific opinion lexicon. It simply re-ranks documents based on a comments-oriented query expansion method that works from an initial ranked set of documents. This brings us the opportunity to apply our methods on top of effective opinion finding methods. To test this combination we considered the systems proposed by teams participating in the last TREC blog opinion retrieval task (TREC2008) [2]. Observe that this subjective task was quite challenging: half of TREC systems failed to retrieve more subjective documents than the baselines [2]. In Ta-

ble 3 and Table 4 we report the mean performance (over the five baselines) of the TREC systems against the average performance achieved by applying  $RM3_C$  on top of those systems' runs. Observe that our methods and these TREC systems were evaluated under the same testing conditions (i.e., re-ranking performance against the 5 topic-retrieval baselines). The systems in bold were the only ones able to show improvements with respect to the original five retrieval baselines in terms of MAP. We can observe that our  $RM3_C$  approach is often able to improve the performance of these methods, showing usually significant improvements in terms of  $P@10$ , as well as good  $RI$  scores. This demonstrates that our method is able to improve strong subjective rankings. Table 3 and Table 4 also show that our expansion approach is robust because  $RM3_C$  is able to outperform all types of opinion retrieval systems regardless of their original performance. Observe also that the average  $P@10$  of our method in Table 2 (.6476) is clearly higher than the  $P@10$  obtained by any TREC participant.

## 5 Related Work

Relevance Feedback and Query Expansion techniques have been considered as an efficient, effective and natural way to enhance the effectiveness of retrieval systems [15]. RF methods use the information provided by relevant documents from an initial retrieval to rewrite and improve the quality of the original query [16]. However, in many scenarios, the applicability of RF is limited because of the lack of relevance judgements. In order to deal with the absence of judgements, Pseudo Relevance Feedback strategies were proposed [17,18]. These methods do not need explicit relevance judgements because they assume that some of the documents retrieved by an IR system are relevant to the original query. How to select the pseudo-relevance documents and also how to use them to improve the original query varies from one PRF method to another.

Relevance Models have emerged as one of the most effective and efficient PRF approaches. As a result of this, different estimations have been proposed [9] and applied in all sorts of IR problems. In particular, for the opinion retrieval task, Huang and Croft [5] proposed a RM estimation based on a mixture with external opinion resources. This approach showed satisfactory results. However, the information provided by the documents' structure to search for opinions is often ignored. This is unfortunate because the comments supply valuable information, as demonstrated in ad-hoc IR retrieval tasks [19,20], summarisation [21] and snippet generation problems [22].

Several blog opinion retrieval methods have been proposed in the literature. The most successful studies in this subject are those focused on finding documents that are both opinionated and on-topic [4,3,23]. To meet this aim, some papers consider term positional information to find opinionated information related to the query. Santos et al. [4] applied a novel opinion mining approach that takes into account the proximity of query terms to subjective sentences in a document. Gerani et al. [3,23] proposed proximity-based opinion propagation methods to calculate the opinion density at the position of each query term in



a document. These two studies led to improvements over state of the art baselines for blog opinion retrieval. The main concern for applying these methods is their computational cost. For example, in [3,23], it is necessary to apply a kernel function at each opinion term to propagate their sentiment scores to every query term in the document. Furthermore, these methods are dependent on external opinion resources. These resources might not be available for a particular domain or language. We designed here a simple PRF method that, focusing the query expansion process on comments, performs well without the need of any external information. Moreover, as we explained in section 4, our proposal is complementary to other opinion finding techniques.

## 6 Conclusions and Future Work

In this paper we have proposed a RM estimation focused on the comments of the blog posts to support opinion finding. Under this framework, the original query is expanded with salient words supplied by a relevance model constructed from the comments of the blog posts. The proposed method significantly outperforms the classical *RM3* estimation for an opinion finding task. We provided experimental evidence showing that the comments are very useful to move the query towards opinionated words. This novel expansion approach is particularly consistent as a high precision mechanism.

One of the characteristics of our approach is that we apply an homogeneous treatment for all types of queries. However, in some cases this could be not desirable. In this respect, we would like to study methods to dynamically adapting our expansion techniques depending on the quality of the initial query [13]. In the near future, we also want to study the effect of spam comments on our expansion approach.

**Acknowledgments:** This work was funded by *Secretaría de Estado de Investigación, Desarrollo e Innovación* from the Spanish Government under project TIN2012-33867.

## References

1. Santos, R.L.T., Macdonald, C., McCreadie, R., Ounis, I., Soboroff, I.: Information retrieval on the blogosphere. *Found. Trends Inf. Retr.* **6**(1) (January 2012) 1–125
2. Ounis, I., Macdonald, C., Soboroff, I.: Overview of the TREC 2008 blog track. In: *Proc. TREC 2008, the 17th Text Retrieval Conference*, Gaithersburg, United States, NIST (2008)
3. Gerani, S., Carman, M.J., Crestani, F.: Proximity-based opinion retrieval. In: *Proc. 33rd international ACM SIGIR conference on Research and development in information retrieval. SIGIR '10*, New York, NY, USA, ACM (2010) 403–410
4. Santos, R.L.T., He, B., Macdonald, C., Ounis, I.: Integrating proximity to subjective sentences for blog opinion retrieval. In: *Proc. 31th European Conference on IR Research on Advances in Information Retrieval. ECIR '09*, Berlin, Heidelberg, Springer-Verlag (2009) 325–336

5. Huang, X., Croft, B.: A unified relevance model for opinion retrieval. In: Proc. of the 18th ACM conference on Information and knowledge management. CIKM '09, New York, NY, USA, ACM (2009) 947–956
6. Mishne, G., Glance, N.: Leave a reply: An analysis of weblog comments. In: In Third annual workshop on the Weblogging ecosystem. (2006)
7. Lavrenko, V., Croft, W.B.: Relevance based language models. In: Proc. of the 24th annual international ACM SIGIR conference on Research and development in information retrieval. SIGIR'01, New York, NY, USA, ACM (2001) 120–127
8. Abdul-jaleel, N., Allan, J., Croft, W.B., Diaz, O., Larkey, L., Li, X., Smucker, M.D., Wade, C.: UMass at TREC 2004: Novelty and HARD. In: Proc. of TREC-13. NIST Special Publication, National Institute for Science and Technology (2004)
9. Lv, Y., Zhai, C.: A comparative study of methods for estimating query language models with pseudo feedback. In: Proc. of the 18th ACM conf. on Information and knowledge management. CIKM '09, New York, NY, USA, ACM (2009) 1895–1898
10. Pang, B., Lee, L.: Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* **2**(1-2) (2007) 1–135
11. Macdonald, C., Ounis, I.: The TREC Blogs 2006 collection: Creating and analysing a blog test collection. Technical Report TR-2006-224, Department of Computing Science, University of Glasgow (2006)
12. Parapar, J., López-castro, J., Barreiro, A.: Blog posts and comments extraction and impact on retrieval effectiveness. In: 1st Spanish Conference on Information Retrieval, CERI'12, Madrid (2010) 5–16
13. Sakai, T., Manabe, T., Koyama, M.: Flexible pseudo-relevance feedback via selective sampling. *ACM Transactions on Asian Language Information Processing (TALIP)* **4**(2) (2005) 111–135
14. Zhai, C., Lafferty, J.: A study of smoothing methods for language models applied to information retrieval. *ACM Trans. Inf. Syst.* **22**(2) (2004) 179–214
15. Rocchio, J.: Relevance feedback in information retrieval. In Salton, G., ed.: *The SMART Retrieval System: Experiments in Automatic Document Processing*. Prentice Hall, Inc. (1971) 313–323
16. Croft, B., Harper, D.J.: Using Probabilistic Models of Document Retrieval without Relevance Information. *Journal of Documentation* **35** (1979) 285–295
17. Ruthven, I., Lalmas, M.: A survey on the use of relevance feedback for information access systems. *Knowl. Eng. Rev.* **18**(2) (2003) 95–145
18. Lu, X.A., Ayoub, M., Dong, J.: Ad Hoc Experiments Using EUREKA. In: Proc. of TREC-5. NIST Special Publication, National Institute for Science and Technology (1996) 229–240
19. Weerkamp, W., de Rijke, M.: Credibility improves topical blog post retrieval. In: Proc. of ACL-08: HLT, Columbus, Ohio, Association for Computational Linguistics, Association for Computational Linguistics (2008) 923931
20. Mishne, G.: Using blog properties to improve retrieval. In: *International Conference on Weblogs and Social Media 2007*, Retrieved February 29, 2008. (2007)
21. Hu, M., Sun, A., Lim, E.P.: Comments-oriented blog summarization by sentence extraction. In: Proc. of the sixteenth ACM conference on information and knowledge management. CIKM '07, New York, NY, USA, ACM (2007) 901–904
22. Parapar, J., López-Castro, J., Barreiro, A.: Blog snippets: a comments-biased approach. In: Proc. of the 33rd international ACM SIGIR conference on Research and development in information retrieval. SIGIR '10, New York, NY, USA, ACM (2010) 711–712
23. Gerani, S., Keikha, M., Crestani, F.: Aggregating multiple opinion evidence in proximity-based opinion retrieval. In: *SIGIR*. (2011) 1199–1200