Comparing Semantic Associations in Sentences and Paragraphs for Opinion Detection in Blogs

Malik Muhammad Saad Missen
Université de Toulouse
IRIT UMR 5505 CNRS
Toulouse, France
+ 33 5 61 55 67 65
Missen.Malik@irit.fr

Mohand Boughanem
Université de Toulouse
IRIT UMR 5505 CNRS
Toulouse, France
+ 33 5 61 55 67 65
Bougha@irit.fr

Guillaume Cabanac
Université de Toulouse
IRIT UMR 5505 CNRS
Toulouse, France
+ 33 5 61 55 67 65
Cabanac@irit.fr

ABSTRACT

Opinion Detection is one of the most interesting and challenging work in the field of Information Retrieval. Lot of research work already exists in this area with some distinctive work. A review of the reveals that researchers have been working on different levels of granularity like documents, passages, sentences and words for the task of opinion detection. In this work we revise our previous approach that combines document level heuristics with a semantic similarity based method. We evaluate this semantic similarity approach on a huge data collection using three different setups involving both sentences and passages and then compare the performance of our approach with these different setups. For evaluation purposes, we are using TREC Blog 2006 collection *(148 GB)* with 50 topics of TREC Blog 2006 over baseline obtained through Terrier Information System Platform. Results show that our approach improves the baseline opinion MAP by 28.89%, 30.13% and 32.26% using setup one, two and three respectively.

Categories and Subject Descriptors

H.3.3 **[Information Storage and Retrieval]:** Information Search and Retrieval - *Query formulation, Relevance feedback, Search process.*

General Terms

Experimentation, Verification

Keywords

Opinion Detection, Blogs, Semantic Relatedness, Passages, Sentences

1. INTRODUCTION

The field of Opinion Detection is becoming more popular as with popularity of the blogging phenomenon. Opinion Detection is the ability of recognizing and classifying opinionated text within the documents [1]. This ability is desirable for various tasks, including filtering advertisements, separating the arguments in online debate or discussions, ranking web documents cited as authorities on contentious topics, etc.

Describing in other words, Opinion Detection is the task of distinguishing between Factual and Opinionated Information. This task can be performed on different levels of granularity, i.e. on word level, sentence level or on document level. As a conclusion of this task a given word, sentence or document can be declared as of opinionated nature (or subjective) or of factual nature (objective). Text with opinionated nature can further be analyzed for having negative or positive polarity of opinion and this subtask is called *Opinion Polarity Detection*.

The task of Opinion Detection becomes more difficult and challenging when it is to be performed for Blogs. A blog (a contraction of the term Web log) is a Web site, usually maintained by an individual with regular entries of commentary, descriptions of events, or other material such as graphics or video. Entries are commonly displayed in reverse-chronological order. “Blog” can also be used as a verb, meaning to maintain or add content to a blog [2]. In blogs people (bloggers) are used to express their sentiments to public or private issues like Cigarette Smoking, Elections etc. [3]. Topic can vary from social to political issues. Therefore, blogs are the prominent source of opinions on the Web. There are several commercial blog search engines[1] that allow users to find out the opinions and thoughts of other people, who share their thoughts in blogs.

Looking at the interest of researchers in the task of Opinion Detection, *TREC (Text Retrieval Conference)* introduced a Blog track in 2006 known as TREC Blog Track with the release of blog data collection. The data collection is *148GB* in size with release of 50 topics each year from 2006 to 2008. TREC has also provided the query relevance assessments (*qrels*) for this collection in which different categories of blog documents (like relevant, non-relevant, opinionated, positive-opinionated, negative-opinionated, mixed) have been identified and each document has been marked to belong to a class of documents. This blog track is very important for evaluating a system’s performance on the blog data collection. Different groups with their systems have participated in TREC Blog 2006, TREC Blog 2007 and TREC Blog 2008. Generally two kinds of approaches

MEDES 2009, October 27-30, 2009, Lyon, France.

[1] http://www.blogsearchengine.com/
http://blogsearch.google.com/

have been adapted by participants in previous TREC conferences i.e. Lexicon based approaches [4, 5, 6, 7] and Machine Learning based approaches [8, 9].

- In Lexicon-based method, words are categorised according to their polarity i.e. positive or negative and a kind of approach is developed looking at the number of positive or negative words in the document text to conclude whether the document is opinionated or not [13, 14]. But the contextual polarities have not been taken into account for most of the time which are very important regarding the nature of the problem. For example, if a positive word like *Good* is preceded by a negation *not (i.e. not Good)*, it inverses its prior polarity. There are many other features that can be used to determine the contextual polarity of a word [10]. There are various lexical resources [11, 12] available for this task.
- In Machine learning methods, usually a machine learning classifier is trained on some already annotated corpus (annotated with sentiment words) and then tested on the original corpus to be processed. The details about the use of these and some other approaches can be consulted in overview papers of TREC 2006 [13], TREC 2007 [14] and TREC 2008 [15].

One of the main tasks of the TREC Blog track is to detect the opinionated documents having opinions about a given target (other task is *Blog Distillation* Task [14]). So the basic idea is that if documents contains some words or sentences about the topic (or let's say query words) and if these words or sentences are opinionated, the document is delivered to the user considering it a relevant opinionated document. Here, the question arises that, "Does a blog document only specifically talks about one thing?" The answer is not always "YES". The document might contain opinion about the topic, some characteristic of the topic not asked in query or any other topic or target. Like a review on a product review site may be positive as a whole but might contain some negative opinion about one of its features. For instance, a review about a laptop may be positive as a whole but can contain few negative words or sentences about its battery life. Another case that makes the task of *Opinion Detection* in blogs more challenging is the extension of one topic to another topic or sub-topic. The discussion (i.e. communication between blog/comment authors) may start from one topic but the subject changes with time and nature of the main topic of the blog post. For example, a blogpost on the topic of *"Afghan War"* may also be discussing the basic teachings of Islam, Iraq war, terrorism and lack of resources and justice in third world countries. It has been noted in TREC Blog qrels that if a large document contains even few opinionated phrases about the topic then it has been marked as a relevant opinionated document in the results. Now if an *Opinion Finding* approach is considering this document as a single monolithic document then the *Opinion Finding* results for this document (in fact the non-relevant or non-opinionated portions of the document) affect the overall ranking of the document. On the other hand, if we treat the document as a composition of small portions (like passages etc.), it might give us a more accurate and better ranking of the documents.

Concluding from above examples, we can say that documents should better be processed on smaller levels like sentence or passage (or paragraph) level. However we hypothesize that it is more practical to process blog documents on paragraph level because sentence splitting is a very challenging task especially when we are dealing with blogs. Lack of punctuations, capitalization and grammar mistakes make this task more difficult. Even if well-split, we may lose the context of the sentence if dealt on sentence-level. In addition, blog documents are more logically structured as paragraphs. A blogpost is split into many paragraphs and normally a comment is contained within the boundaries of a paragraph. Therefore, it becomes more feasible (and logical) for us to process blog documents on paragraph level.

This paper is an extended version of our previous work [4]. In this work, we revise the experiments of [4] under three different experimental setups with more details and little modifications in our approach. Also this paper gives more details of work that lacked in [4].

The remainder of this paper is organized as follows: Section 2 discusses our approach detailing the underlying phases. In section 3, we describe our experiments with analysis of their results. At end, we report some limitations of our work and outline some perspectives for our work.

2. OUR APPROACH

In this section, we describe each stage of our approach in detail:

2.1 Data Pre-processing

In this phase, we remove unnecessary tags like script, style etc. We also remove the hyperlinks present in a document because we think that most of the noisy data (like ads etc) lies in the form of links in a web document. Even there is a possibility that we can lose valuable data too but loss of valuable data is much less than getting rid of noisy data.

2.2 Topic-Relevance Retrieval

In this phase some kind of IR Model is used to retrieval top 1000 relevant documents to the query. In our case, we are using Terrier IR platform with Okapi BM25 [16] for retrieving top 1000 relevant documents. This baseline has an Opinion Finding MAP of *0.1689* over topics of TREC 2006 i.e. from topic 851 to 900.

2.3 Opinion-Finding

This stage is composed of six individual components. The scores of all components are added to have a score of this stage i.e. DOC_{OPIN}. Later on scores of all components and DOC_{REL} are combined to re-rank topic relevant documents from phase 2.2. In the followings, we discuss these six components one by one:

2.3.1 Document Subjectivity Component

Generally it is assumed that more a document contains subjective terms more are the chances for it to be an opinionative document [17]. Using this heuristic, we have used the lexicon SentiWordNet (SWN) [18] to calculate subjectivity of a term. SWN contains the subjectivity score of a term divided between positive and negative scores. Average subjectivity score of a term is taken as a final score and finally we calculate the average document subjectivity score DOC_{SUBJ} [4].

$$Subj(w) = \frac{Neg(w, SWN) + Pos(w, SWN)}{\left|w_{sense}\right|} \quad \textbf{(1)}$$

$$DOC_{SUBJ} = \sum_{i=1}^{n} Subj(w_i, d).\left|d\right|^{-1} \quad \textbf{(2)}$$

Where $Subj(w_i,d)$ is the subjectivity score of a document term *i in SWN* , $|d|$ is the total number of words in document d and $\left|w_{sense}\right|$ is the total number of word senses found in SWN.

2.3.2 Emotiveness Component

[19] states that adverbs and adjectives represent emotiveness of a document. Assuming it an important clue of opinionativeness of a document, we calculate emotiveness of a document by counting the numbers of *adverbs* and *adjectives* in a POS (*Parts of Speech*) tagged document and is expressed as $DOC_{\ EMOT}$ [4].

$$DOC_{\ EMOT} = \frac{|Adjectives| + |Adverbs|}{|Verbs| + |Nouns|} \quad \textbf{(3)}$$

2.3.3 Reflexivity Component

Bloggers make a lot of use of reflexive pronouns like *I, me, myself* etc while writing blogs. For example, use of 'I' in *I think, as far as I am concerned, my point of view is that, its ours duty* etc. All these phrases give a sense of opinion to the words following them; therefore, we include *Reflexivity* measure as one of our components. The idea is that any document with larger number of such words will be more opinionative relative to the one with less number of such words. It is expressed as DOC_{REFL} [4] .

$$DOC_{REFL} = \sum_{i=1}^{n} C(w_i, d).|P|^{-1} \quad \textbf{(4)}$$

where $C(w_i,d)$ is the number of occurrences of reflexive pronouns w_i in document d with $|P|$=Total number Pronouns in the document.

2.3.4 Addressability Component

Most of the opinionated sentences are found in comments part of the blog about the topic being discussed in blogpost [1]. Users write their comments about topic, blogger or about others using by addressing them using words like 'you', 'yours', etc thus creating a discussion environment. Therefore, we consider the addressability (DOC_{ADD}) component [4] as part of our opinion finding approach and we calculate it as:

$$DOC_{ADD} = \sum_{i=1}^{n} C(w_i, d).|P|^{-1} \quad \textbf{(5)}$$

where $C(w_i,d)$ is the number of occurrences of addressive terms w_i in document d with $|P|$=Total number Pronouns in the document.

2.3.5 Common Opinion Phrases Component

Generally a subset of "opinion expressions" can be defined for a particular language. So we have defined such set for *English Language* phrases after analysing online blogs. We believe that any blogpost expressing opinions can contain a subset of this set of phrases and can be very simple and important clue for detecting documents containing opinions. This set contains phrases like *I think, I would like to, I should be, As for me/As to me* etc. A simple heuristic behind this metric would be to search and count the occurrences of this set in the blog document and assign a score $DOC_{\ PHRASE}$ to the document as a result of these findings [4].

$$DOC_{\ PHRASE} = \sum_{i=1}^{n} C(w_i, d).|d|^{-1} \quad \textbf{(6)}$$

where $C(w_i,d)$ is the number of occurrences of common phrase term w_i in document d and |d| is the total number of words in document d.

2.3.6 Opinion-Topic Association (OTA) Component

As described above, one of the major challenges in opinion detection is to retrieve documents that specifically contain opinions on the topic of the query. This problem can better be solved if we focus on sentences of a document rather than the document itself as a whole. A document having two or three opinionated sentences on the topic of the query can never be better than a document solely discussing the query topic or putting it other way, a document with one opinion sentence on the subject (it may be a quote or an example in that document) cannot be better than a document having two or three sentences specifically on the topic. In brief, we need to have such documents on top in the ranking that are not only relevant but they also contain people's opinion on the given topic. Therefore, we propose a component that combines topic relevance evidence and opinion evidence of a sentence to check the topic and opinion association within a sentence. We have named this component as *Opinion Topic Association* component. We make the assumption that the presence of subjectivity *verbs, adverbs or adjectives* in a sentence is the measure of its subjectivity. Similarly for topic relevance measures, we will check the presence of expanded query relevant terms within the sentence.

In our *Two Phase Query Expansion* method we use title of the query as a base to enrich with relevant and opinionated terms. In first phase, we use Wikipedia[2] for expanding the query with relevant terms. In Wikipedia based query expansion, we prepare a list of *proper nouns* and *named entities* (often found as hyperlinks text) related to the query

[2] if there is no Wikipage for the given query then Google search engine is used to extract the relavant terms

title from the Wikipage. It is to be noted that we filter this list of nouns and entities manually. . At the end of this phase of Query Expansion, we have a list of relevant terms including the title of the query. This list is later on used for selection of relevant passages in the phase "Relevant Passage Selection" from the top 1000 retrieved documents. TREC Blog results are used for second phase of query expansion by using documents labelled as 2, 3 or 4 in results. We assume it a particular case of Relevance Feedback in which the user identifies the documents that are relevant to his/her information need. We limit the number of chosen relevant (and opinionated) documents to 10[3]. First of all we remove *stop words* from these documents and then a list of *verbs*, *adjectives* and *adverbs* is prepared from these documents. Once we have this list of verbs, adjectives and adverbs prepared, we remove duplicates, manually filter for very common terms like (know, do, live etc) and calculate their *collection frequency (cf)* and *document frequency (df)* of all terms which have more than one occurrence in this small collection of top ten opinionated documents. Later on, we rank all of these terms according to the ranking function given in equation below and choose top *10* terms to be part of the final query with the terms already extracted from Wikipedia in 1st part of Query Expansion.

$$Subj(t) = df * cf \tag{7}$$

In the second step of this component phase, we used lexicon WordNet [20, 21] for the following purposes:

- To resolve the words' contextual sense ambiguities using glosses of the concepts present in WordNet [22, 23]. For example, differentiation between word *plant* (like flowers etc) and *plant* (like nuclear plant) is done using their contextual words.
- Extract compound words from sentence and query.
- To find any semantic relatedness between sentence words and the query words using different WordNet relations. We are using two semantic measures in our work. For example, one of the semantic similarity relation (*path* measure) is calculated by computing semantic relatedness of word senses by counting nodes in the noun and verb WordNet "IS-A" hierarchies like the path between shrub#n#1 and tree#n#1 is shrub#n#1 - woody_plant#n#1 - tree#n#1. Since a longer path length indicates less relatedness, the relatedness value returned is the multiplicative inverse of the path length (distance) between the two concepts: relatedness = 1 / distance. If the two concepts are identical, then the distance between them is one; therefore, their relatedness is also 1. The other one is the lesk measure [22, 23]. The lesk measure finds the overlap between glosses of the words being compared as well as words directly linked to them. We use path measure for comparing nouns to have more precise matches and lesk measure for rest of the parts of speech i.e. verbs and adjectives. We use different measures for different types of parts of speeches because for nouns, we want to have some precise matches between sentence words and query terms but for verbs and adjectives, we use lesk measure to accommodate all possible semantically related words. Also its to be noted that we use by default option of normalizing the lesk score [21, 23].

[3] It is to be noted that these *10* documents are removed when calculating results to avoid biasness in the results.

The OTA score of each sentence is calculated and normalized by the number of words in that sentence which leads to the OTA score of a document DOC_{OTA} [4] which is normalized by the number of sentences in that document.

$$DOC_{OTA} = \sum_{i=1}^{N} \frac{S_{SIM}(S_i, Q)}{|N|} \tag{8}$$

$$S_{SIM} = \sum_{i=1, j=1}^{Ws, Qs} \frac{Sem(S_i, Q_j)}{|Ws|} \tag{9}$$

Where S_{SIM} is the semantic similarity between sentence words, N=Total sentences and query words and *Sem(S,Q)* is calculated using WordNet as explained above.

3. EXPERIMENTAL SETUP

Experiments were performed in three different setups and using three different strategies on same data collection.

3.1. First Setup

First setup is as it has been described in section 2.

After performing experiments using first setup and analyzing its results, few modifications were made in our approach and first experimental setup to check their impact on the results.

3.1 Second Setup

We performed the following modifications for second setup. 1) We removed *Emotiveness* component from this setup because of its nominal impact on results. 2) A Sentence Selection component was introduced in our Opinion Finding approach (in OTA component) i.e. only subjective sentences were given as input to our OTA component. A sentence was declared as subjective if it contains one or more *adjectives* [24].

3.2 Third Setup

For third setup, 1) we remove Emotiveness component as we did for second setup and 2) we propose a relevant passage selection strategy for this setup of experiments which is explained below.

3.3.1. Relevant Passage Selection. This stage performs three tasks basically:

3.3.1.1. Passage Identification. There are three ways passages can be identified in documents [25, 26]: discourse passage (passages based on document mark-up), semantic passages (based on shift of topics within a document) and window passages (based on fixed or variable number of words). Regarding the structure of the blog documents in our collection, we decided to identify passages based on their mark-up.

3.3.1.2. Remove unnecessary and noisy paragraphs from the documents.

- *Presence of noisy words.* In this module, we check each passage for some words which are not important for us. A list of words like *copyright, reserved, email, home, contacts etc* is prepared and percentage of their presence in the paragraph is calculated. If this percentage is more than a threshold we fixed using a rough estimate, such passage is removed.
- *Presence of Days/Months.* Most of the blogs contain the calendar or archive of its blogposts which is not of our interest. Therefore, we check for months and days names within a certain passage, calculate its percentage and if this percentage is more than a threshold we fixed using a rough estimate, such passage is removed.

3.3.2. Selecting Relevant Passages. Deciding criteria for selection of a relevant passage is not an easy task because we have few options like "to select all passages having title of the query in it" or "to select all passages having any query term of the expanded query through Wikipedia". Thinking not to miss any relevant passage, we decide to go with second option. Therefore, we choose all such passages in relevant opinionated documents which have at least one occurrence of any of the query term which is part of the expanded query through Wikipedia.

4. RESULTS AND CONCLUSIONS

After calculating scores and bringing them on same scale for each individual component, we add Document Relevance score DOC_{REL} and Document Opinion score DOC_{OPIN} score to have final score of a document. Finally documents are re-ranked using their final scores.

Table1 shows the MAP *(Mean Average Precision)* and P@10 *(Precision at 10 documents)* results for our own baseline obtained using Terrier standard configuration with okapi BM25. An improvement of almost 29% is observed in opinion finding MAP over baseline results.

TABLE I. RESULTS

Run	MAP	P@10	Improvement
BaseLine Run	0.1689	0.3420	N/A
1st Setup Opinion Run	0.2177	0.5120	28.89%
2nd Setup Opinion Run	0.2198	0.5127	30.13%
3rd Setup Opinion Run	0.2243	0.52	32.26%

Table I lists the results for our three setups. Surprisingly the results for second setup did not make a big difference as was expected. The major cause of such results may be less than expected performance of *POS Tagger* which might have not performed well because of absence of proper punctuations, good use of grammar rules and capitalization etc. within the sentences. Results for third setup are the best result that proves our point that processing blog documents on passage level can be more helpful for this task. As we note that we got some improvements in 2nd and 3rd case but these are very marginal improvements and most probably reason behind this is that most of the model being used is same in all setups (i.e. the opinion finding components). However these results encourage us to explore the passage-level front in detail.

Our approach has performed well but still need to be improved on many fronts. For example, we need to automate the process of Query Expansion used in this work. We can also use some external data sources (like Amazon or Wize[4] product review data, etc) for having a collection of opinionated terms. As we see that passage-based model is working well so we have plan to extend it further using more robust IR models. One of the major issues that concerns opinionated document retrieval is its consistency in performance on baselines of different strengths. It has already been said that even the best opinion detection approach might fail on a stronger baseline [15]. We consider this issue a big problem and we are working to propose such approach that adapts itself according to the strength of the baseline and performs accordingly. An approach with a good balance of weights between opinion and topic relevance is our need.

[4] www.amazon.com
www.wize.com

5. REFERENCES

[1] B. Liu, *Web Data Mining: Exploring Hyperlinks, contents and Using Data, Chapter 11*, Data-Center Systems and Applications, Springer , 2007

[2] C.J. Gray, “Adolscent Blogging: A Comparison of Developmental Psychology and Self-Depiction in Adolesent Blogs”, Master’s Thesis, University of North Carolina Nov 2005

[3] I. Ounis, C. Macdonald, and I. Soboroff, “On the TREC Blog”, ICWSM’08: In Proceedings of International Conference on Weblogs and Social Media, AAAI Press, Seattle USA, March 30-April 2, 2008

[4] M.M.S. Missen, and M. Boughanem, “Sentence-level topic association for opinion detection in blogs”, MAW09: In Proceedings of IEEE International Symposium for Web and Data Mining, IEEE Computer Society, Bradford, UK, May 2009

[5] M.M.S. Missen, and M. Boughanem, “Using WordNet’s Semantic relations for Opinion Detection in Blogs”, ECIR 2009: European Conference

[6] B. Ernsting, W. Weerkamp, and M. de Rijke, “The University of Amsterdam at the TREC 2007 Blog Track”, TREC 2007 Blog Track

[7] K. Yang, N. Yu, A. Valerio, and H. Zhang, “WIDIT in TREC-2006 Blog track”, TREC 2006 Blog Track

[8] G. Zhou, H. Joshi and C.Bayrak, “Topic Categorization for Relevancy and Opinion Detection”, TREC 2007 Blog Track

[9] S. Gerani, M. Carman, and F. Crestani, “Investigating Learning Approaches for Blog Post Opinion Retrieval”, ECIR 2009: In Proceedings of European Conference on Information Retrieval, Springer, Toulouse France, 1-3 April 2009

[10] T. Wilson, J. Wiebe, and Hoffmann, “Recognizing contextual polarity in phrase-level sentiment analysis”, HLT/EMNLP 2005: In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Vancouver, British Columbia, Canada, October 06 - 08, 2005, pp. 347-354

[11] C. Strapparava, and A. Valitutti, “WordNet-affect: An affective extension of WordNet”, LREC 2004: In Proceedings of Language Resources and Evaluation Conference, European Language Resources Association , Lisabon, Portugal, 2004

[12] A. Esuli, and F. Sebastiani, “SentiWordNet: A publicly available lexical resourcefor opinion mining”, LREC-06: in Proceedings of Language Resources and Evaluation Conference, European Language Resources Association, Genova, 2006

[13] I. Ounis, M. Rijke, C. Macdonald, G. Mishne, and I. Soboroff, “Overview of the TREC-2006 Blog Track”, TREC 2006 Blog Track

[14] C. Macdonald, I. Ounis, and I. Soboroff, “Overview of the TREC-2007 Blog Track”

[15] I. Ounis , C. Macdonald, and I. Soboroff, “Overview of the TREC-2008 Blog Track”

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

[17] H. Yang, J. Callan, and L. Si. Knowledge Transfer and Opinion Detection in the TREC 2006 Blog Track. In Proceedings of TREC 2006

[18] A. Esuli, and F. Sebastiani, “SentiWordNet: A publicly available lexical resourcefor opinion mining”, in Proceedings of LREC-06, the 5th Conference on Language Resources and Evaluation, Genova

[19] L. Zhou, D.P. Twitchell, T. Qin,J.K. Burgoon, and J.F. Nunamaker, “An exploratory study in deception detection in text-based computer mediated communication”, Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS’03)

[20] C. Fellbaum, WordNet: An Electronic Lexical Database, MIT Press

[21] T Pedersen, S Patwardhan and J Michelizzi, “WordNet:: Similarity – Measuring the Relatedness of Concepts”, In Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)

[22] S. Banerjee and T. Pedersen, “Extended gloss overlaps as a measure of semantic relatedness”, In Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, 805–810.

[23] S. Banerjee and T. Pedersen. An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet. In *Proceedings of the Third international Conference on Computational Linguistics and intelligent Text Processing* (February 17 - 23, 2002). A. F. Gelbukh, Ed. Lecture Notes In Computer Science, vol. 2276. Springer-Verlag, London, 136-145.

[24] M. Hu, B. Liu, “Mining and Summarizing Cutomer Reviews”, in Proceedings of the 10th ACM SIGKDD International Conference on Knowledge and Data Mining. August 22-25, 2004, Seattle, WA, USA

[25] M. Kaszkiel and J. Zobel. Effective ranking with arbitrary passages. Journal of theAmerican Society for Information Science, 52(4):344–364, November 2001

[26] James P. Callan. Passage-level evidence in document retrieval. In Proceedings of SIGIR,pages 302–310, 1994