Sentiment-based Ranking of Blog Posts using Rhetorical Structure Theory

Jose M. Chenlo[1], Alexander Hogenboom[2], and David E. Losada[1]

[1] Centro de Investigación en Tecnoloxías da Información (CITIUS)
Universidad de Santiago de Compostela, Spain
{josemanuel.gonzalez,david.losada}@usc.es
[2] Econometric Institute
Erasmus University Rotterdam, The Netherlands
hogenboom@ese.eur.nl

Abstract. Polarity estimation in large-scale and multi-topic domains is a difficult issue. Most state-of-the-art solutions essentially rely on frequencies of sentiment-carrying words (e.g., taken from a lexicon) when analyzing the sentiment conveyed by natural language text. These approaches ignore the structural aspects of a document, which contain valuable information. Rhetorical Structure Theory (RST) provides important information about the relative importance of the different text spans in a document. This knowledge could be useful for sentiment analysis and polarity classification. However, RST has only been studied for polarity classification problems in constrained and small scale scenarios. The main objective of this paper is to explore the usefulness of RST in large-scale polarity ranking of blog posts. We apply sentence-level methods to select the key sentences that convey the overall on-topic sentiment of a blog post. Then, we apply RST analysis to these core sentences in order to guide the classification of their polarity and thus to generate an overall estimation of the document's polarity with respect to a specific topic. Our results show that RST provides valuable information about the discourse structure of the texts that can be used to make a more accurate ranking of documents in terms of their estimated sentiment in multi-topic blogs.

Key words: Blog, Opinion Mining, Sentiment Analysis, Polarity Estimation, Discourse Structure, Rhetorical Structure Theory

1 Introduction

Social networks and blogs have rapidly emerged to become leading sources of opinions in the Web. These repositories of opinions have become one of the most effective ways to influence people's decisions. In fact, companies are aware of the power of social media and most enterprises try to monitor their reputation over Twitter, blogs, etc. to infer what people think about their products and to get early warnings about reputation issues. In this paper, we focus on one of the most important sources of opinions in social media, i.e., the blogosphere [1]. In this

scenario, classical information retrieval (IR) techniques are not enough to build an effective system that deals with the opinionated nature of these new sources of information. To mine opinions from blogs we need to design methodologies for detecting opinions and determining their polarity [2].

In recent years, several works have been conducted to detect opinions in blog posts [1]. Currently, the most popular approach is to consider this mining task as a two-stage process that involves a topic retrieval stage (i.e., retrieve relevant posts given a user query), and a re-ranking stage that takes into account opinion-based features [3]. This second stage can also be subdivided into two different substaks: an opinion-finding task, where the main aim is to find opinionated blog posts related to the query, and a subsequent polarity task to identify the orientation of a blog post with respect to the topic (e.g., positive or negative). For polarity estimation, researchers often apply naive methods (e.g., classifiers based on frequency of positive/negative terms) [4]. Polarity estimation is a really challenging task with many unresolved issues (e.g., irony, conflicting opinions, etc.). We argue that this difficult estimation problem cannot be solved with regular matching (or count-based) techniques alone. In fact, most lexicon-based polarity classification techniques fail to retrieve more positive/negative documents than baselines without polarity capabilities [3].

This phenomenon is caused by the polarity of a document being not so much conveyed by the sentiment-carrying words that people use, but rather by the way in which these words are used. Rhetorical roles of text segments and their relative importance should be accounted for when determining the overall sentiment of a text (e.g., an explanation may contribute differently to the overall sentiment than a contrasting text segment does) [5]. Rhetorical Structure Theory (RST) [6] is a linguistic method for describing natural text, characterizing its structure primarily in terms of relations that hold between parts of the text. Rhetorical relations (e.g., an explanation or a contrast) are very important for text understanding, because they give information about how the parts of a text are related to each other to form a coherent discourse.

Accounting for the rhetorical roles of text segments by means of a RST-based analysis has proven to be useful when classifying the overall document-level polarity of a limited set of movie reviews [5]. As this success comes at a cost of computational complexity, the application of a RST-based analysis in large-scale polarity ranking tasks in the field of IR is challenging. In this paper, we study how we can utilize RST in a large-scale polarity ranking task and how RST helps to understand the sentiment expressed by bloggers. More specifically, we aim to identify the rhetorical relations that give good guidance for understanding the sentiment conveyed by blog posts, as well as to quantify the advantage of exploiting these relations. We also compare our RST-based methods with conventional approaches for large-scale polarity ranking of blog posts.

In the blogosphere, the presence of spam, off-topic information, or *relevant* information that is non-opinionated introduces noise and this is a major issue that harms the effectiveness of opinion-finding techniques. Therefore, it would

not be wise to apply RST on the entire blog posts. We build on recent advances in extracting key opinionated sentences for polarity estimation in blog posts [4] and analyse the structure of the discourse only for selected passages. This is beneficial to avoid noisy chunks of text and it is also convenient from a computational complexity perspective because discourse processing is not lightweight.

2 Method

First, we present the methods to find *relevant* polar sentences in a blog post. Then, we show how to perform rhetorical analysis over these key evaluative sentences, in order to determine the relations between the different spans of text. Finally, we define the overall orientation of a blog post as positive (resp. negative) according to these key evaluative sentences. To this end, we take into account the information provided by rhetorical relations.

2.1 Finding Relevant Polar Sentences

Many efforts have been recently made to determine what are the important parts of a document for polarity purposes. Most authors [7–9] have studied this issue in a typical IR scenario, i.e., given a query, the system has to return a ranking of positive opinionated documents and a ranking of negative opinionated documents [3]. This task is approached as a re-ranking task in which systems first retrieve a list of relevant documents and then reorganize them according to their polarity. In this paper we follow the same approach.

We apply an effective and efficient approach [4] based on sentence retrieval and a well-know sentiment classifier, OpinionFinder (OF) [10]. OpinionFinder estimates what sentences are subjective and also marks various aspects of the subjectivity in the sentences, including the source (holder) of the opinions and the words that are included in phrases expressing positive or negative sentiments. The information provided by OF was very useful for both subjectivity and polarity estimation in numerous experimental validations [4, 8, 9, 11].

Basically, the terms tagged by OF as positive or negative are used to define the positive or negative polarity score of a sentence. Furthermore, to promote polar sentences that are on-topic (i.e., sentences that are relevant to the query topic), sentence retrieval is applied to determine the relatedness between the query terms and each polar sentence. To this end, we use the Lemur[3] implementation of tf-idf, with BM25-like weights[4] as our sentence retrieval method. BM25 [12] is a robust and effective IR model that has shown its merits in many search tasks.

Finally, the combination of relevance and polarity is done through linear interpolation:

$$pol(S, Q) = \beta \cdot rel_{norm}(S, Q) + (1 - \beta) \cdot pol(S), \tag{1}$$

---

[3] http://www.lemurproject.org/

[4] We build a sentence-level index and apply the well-known BM25 suggested configuration ($k1 = 1.2, b = 0.75$).

where Q is the query, $rel_{norm}(S, Q)$ is the Lemur's tf-idf score after a query-based normalization into [0,1] and $pol(S)$ represents the number of positive (resp. negative) terms tagged in the sentence S divided by the total number of terms in S[5]. $\beta \in [0, 1]$ is a free parameter.

Different aggregation methods were considered in [4] to compute the final polarity of a blog post based on its sentence-level scores, including the average score of all polar sentences, the first or the last k polar sentences and the sentences with the highest $pol(S, Q)$. This last method, *PolMeanBestN*, was shown to be very robust and, overall, it gives the best estimation of the polarity of a blog post. Therefore, in this paper, we use this approach to extract the key sentences that are injected to a RST module. The best configuration obtained in [4] for *PolMeanBestN* is $k = 1$, which means that we select just one sentence to estimate the overall polarity of a blog post.

Given an initial list of documents which is ranked by decreasing relevance score ($rel_{norm}(D, Q)$), we re-rank the list to promote on-topic blog posts that are positive (resp. negative) opinionated as follows:

$$pol(D, Q) = \gamma \cdot rel_{norm}(D, Q) + (1 - \gamma) \cdot pol_S(D, Q), \qquad (2)$$

where rel_{norm} is the document's relevance score after a query-based normalization in [0,1], $pol_S(D, Q) = max_{S \in D} pol(S, Q)$ (e.g., *PolMeanBestN*, with k=1), and $\gamma \in [0, 1]$ is a free parameter[6].

2.2 Rhetorical Structure Theory

Discourse analysis is concerned with how meaning is built up in the larger communicative process. Such an analysis can be applied on different levels of abstraction, i.e., within a sentence, within a paragraph, or – typically – within a document or conversation. The premise is that each part of a text has a specific role in conveying the message of a piece of natural language text. RST [6] is one of the leading discourse theories. The theory can be used to split texts into segments that are rhetorically related to one another. Each segment may in turn be split as well, thus yielding a hierarchical rhetorical structure. Within this structure, text segments can be either nuclei or satellites, with nuclei being assumed to be more significant than satellites with respect to understanding and interpreting a text. Many types of relations between text segments exist; the main paper on RST defines 23 types of relations [6]. A satellite may for instance be an elaboration on what is explained in a nucleus. It can also form a contrast with respect to matters presented in a nucleus.

---

[5] For positive document retrieval $pol(S)$ is the percentage of positive terms in the sentence, and for negative document retrieval $pol(S)$ is the ratio of negative terms in the sentence.

[6] We used the configuration provided in [4] for the parameters β and γ ($\beta = 0.6, \gamma = 0.6$ for negative polarity estimation, and $\beta = 0.2, \gamma = 0.5$ for positive polarity estimation). This configuration was shown to be very stable across different collections.

For an example of a RST-structured sentence, let us consider the sentence *"Although I like the characters, the book is horrible."*, which can be split into two segments. The core of the sentence, i.e., the nucleus, provides a negative sentiment with respect to a book (*"the book is horrible"*). The other segment is a satellite with contrasting information with respect to the nucleus, admitting to some positive aspects of the book (*"Although I like the characters"*). For a human reader, the polarity of this sentence is clearly negative, as the overall message has a negative polarity. However, in a classical (word-counting) sentiment analysis approach, all words would contribute equally to the total sentiment, thus yielding a verdict of a neutral or mixed polarity at best. Exploiting the information contained in the RST structure could result in the nucleus being given a higher weight than the satellite, thus shifting focus to the nucleus segment. We can thus get a more reliable sentiment score. As such, in order to exploit the rhetorical relations as imposed upon natural language text by a RST analysis, distinct rhetorical roles of individual text segments should be treated differently when aggregating the sentiment conveyed by these text segments. This could be accomplished by assigning different weights to distinct rhetorical roles, quantifying their contribution to the overall sentiment conveyed by a text [5].

2.3 Sentence-level Parsing of Discourse

In order to automatically structure our identified key evaluative sentences by means of a RST-based analysis, we used SPADE (Sentence-level PArsing of DiscoursE)[13], which creates RST trees for individual sentences. SPADE was trained and tested on the train and test set of the RST Discourse Treebank (RST-DT) [14], achieving a F1 score of 83.1% on identifying the right rhetorical relations and their right arguments [13]. The relations taken into account in our experiments are detailed in Table 1.

2.4 RST over On-Topic Polar Sentences

To include RST in our method, we compute $pol(S)$ as a weighted sum of the polar terms occurring in the nucleus and the satellite, respectively:

$$pol(S) = w_{nuc} \cdot pol_{nuc}(S) + w_{sat} \cdot pol_{sat}(S), \tag{3}$$

where nuc represents the nucleus of the sentence S, sat is the satellite of the sentence S, w_{nuc} is the weight for nucleus, w_{sat} is the weight for the concrete satellite and $pol_{nuc}(S)$ and $pol_{sat}(S)$ represent the ratio of positive (resp. negative) terms tagged in the nucleus and satellite respectively of sentence S. Observe that w_{sat} and w_{nuc} are free parameters that need to be trained for each different rhetorical relation. Finally, observe that despite the fact that RST is a computationally intensive task[7], this process can be done offline (at indexing time).

---

[7] SPADE software takes on average 3 seconds to compute each sentence in a regular desktop machine.

Table 1. RST relation types taken into account.

Relation	Description
attribution	Clauses containing reporting verbs or cognitive predicates related to reported messages presented in nuclei.
background	Information helping a reader to sufficiently comprehend matters presented in nuclei.
cause	An event leading to a result presented in the nucleus.
comparison	Clauses presenting matters which are examined along with matters presented in nuclei in order to establish similarities and dissimilarities.
condition	Hypothetical, future, or otherwise unrealized situations, the realization of which influences the realization of nucleus matters.
consequence	Information on the effects of events presented in nuclei.
contrast	Situations juxtaposed to situations in nuclei, where juxtaposed situations are considered as the same in many respects, yet differing in a few respects, and compared with respect to one or more differences.
elaboration	Rhetorical elements containing additional detail about matters presented in nuclei.
enablement	Rhetorical elements containing information increasing a readers potential ability of performing actions presented in nuclei.
evaluation	An evaluative comment about the situation presented in the associated nucleus.
explanation	Justifications or reasons for situations presented in nuclei.
joint	No specific relation is assumed to hold with the matters presented in the associated nucleus.
otherwise	A situation of which the realization is prevented by the realization of the situation presented in the nucleus.
temporal	Clauses describing events with a specific ordering in time with respect to events described in nuclei.

3 Experiments

In this section, we describe the experiments designed to determine the usefulness of RST in a large-scale multi-topic domain. Concretely, we work with the BLOGS06 text collection [15], which is one of the most renowned blog test collections with relevance, subjectivity, and polarity assessments.

3.1 Collection and Topics

We take into account the TREC 2006, TREC 2007, and TREC 2008 blog track's benchmarks. All these tracks have the BLOGS06 as the reference collection for experiments. Each year, a new set of 50 topics was provided and new judgments were made according to the documents retrieved by the participants. One of the core tasks of these tracks is the polarity task, i.e., given a query topic, systems have to return a ranking of positive (resp. negative) blog posts related to the query. Each query topic contains three different fields (i.e., title, description, and narrative). In this work we only utilise the title field, which is short and the best representation of real user web's queries, as reflected in the official TREC Blog track literature [3]. Documents and topics are pre-processed with Krovetz stemmer and we remove 733 English stopwords.

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 the on-topic documents contain explicit expression of opinion or sentiment about

the topic then the document is tagged as positive, negative, or mixed (if the opinion expressed is ambiguous, mixed, or unclear).

3.2 Retrieval and Polarity Baselines

In TREC 2008, to allow the study of the performance of a specific opinion-finding technique across a range of different topic-relevance baseline systems, a set of five topic-relevance baselines was provided. These standard baselines use a variety of different retrieval approaches, and have varying retrieval effectiveness[8].

Spam detection, topic retrieval in blogs, and subjectivity classification are out of the scope of this paper. We focus on the effect of RST on the set of subjective documents identified by the standard baseline runs. This means that the input to our methods is a set of opinionated documents with varied polarity orientations (positive, negative, or mixed polarity) and the objective is to distinguish the type of polarity that every document has (i.e., search for positive, and search for negative documents). This polarity task, per se, is quite challenging because there are many offtopic passages and conflicting opinions. The measures applied to evaluate performance are mean average precision (MAP), and precision at 10 documents (P@10). These measures are commonly applied to assess the performance of ranking algorithms.

3.3 Training and Testing

We have built a realistic and chronologically organised query dataset with the topics provided by TREC. We have optimised the parameters of our methods (e.g., satellite weights) on the TREC 2006 and TREC 2007 topics, while using the TREC 2008 topics as testing set. Two different training-testing processes focused on maximising MAP have been run, i.e., one for positive polarity retrieval and another for negative polarity retrieval. To train all the parameters of our models (including the satellite weights) we have used Particle Swarm Optimisation (PSO). PSO has shown its merits for the automatic tuning process of the parameters of IR methods [16].

3.4 Results

Table 2 shows the results of our polarity approaches. Each run is evaluated in terms of its ability to retrieve positive (resp. negative) documents higher up in the ranking. The best value in each column for each baseline is underlined. Statistical significance is assessed using the paired t-test at the 95% level. The symbols ▲ and ▼ indicate a significant improvement or decrease over the corresponding baseline. To specifically measure the benefits of RST techniques in the estimation of a ranking of positive (resp. negative) blog posts we compare its performance against the performance achieved by a very effective method for blog

---

[8] The baselines were selected by TREC from the runs submitted to the initial ad-hoc retrieval task in the TREC blog track.

Table 2. Polarity Results. Mean average precision (MAP) and precision at 10 (P10) for positive and negative rankings of blog posts. The symbols △ (▽) and ▲(▼) indicate a significant improvement(decrease) over the original baselines provided by TREC and the *polMeanBestN* method, respectively.

	Negative		positive	
	MAP	P10	MAP	P10
baseline1	.2402	.2960	.2662	.3680
+polMeanBestN	.2408	.3000	.2698	.3720
+polMeanBestN(RST)	.2516	.3180△ ▲	.2733	.3740△ ▲
baseline2	.2165	.2780	.2390	.3340
+polMeanBestN	.2222	.2820	.2368	.3160
+polMeanBestN(RST)	.2261▲	.3100△ ▲	.2423△	.3560△ ▲
baseline3	.2488	.2840	.2758	.3500
+polMeanBest	.2524	.2760	.2755	.3420
+polMeanBestN(RST)	.2584△ ▲	.2820	.2770△	.3380▼
baseline4	.2636	.2740	.2731	.3580
+polMeanBestN	.2730	.2840	.2705	.3500
+polMeanBestN(RST)	.2825△	.3240△ ▲	.2716	.3620△ ▲
baseline5	.2238	.3000	.2390	.3600
+polMeanBestN	.2279	.3120	.2404	.3580
+polMeanBestN(RST)	.2393	.3420△ ▲	.2786△ ▲	.4380△ ▲

polarity estimation (*PolMeanBestN* [4], presented in Section 2). *PolMeanBestN* estimates the overall recommendation of a blog post by taking into account the on-topic sentence in the blog post that has the highest polarity score (e.g., the most controversial contents of the post). This configuration leads to a performance comparable to the best performing approach at the TREC 2008 Blog track (KLE system) [1, 4]. Observe that the RST technique proposed in our paper is an evolution over *PolMeanBestN*, in which the estimation of polarity is also done with the highest polarity sentence but we take into account its RST structure (eq. 2). The symbols △ and ▽ indicate a significant improvement or decrease over this polarity method.

Polarity retrieval performance. The technique that performs the best across all different baselines is the RST-based method, showing usually significant improvements with respect to both the baseline and *PolMeanBestN*. Another important finding is that *PolMeanBestN* never significantly outperforms the baselines.

Positive vs Negative results. Another observation is that the performance of negative document rankings is lower than the performance of positive document rankings. This may be caused by negative documents being harder to find. As a matter of fact, there are more positive documents than negative ones in the polarity judgements (3,338 against 2,789). Additionally, the lexicon-based identification of negative documents may be thwarted by people having a tendency of using rather positive words to express negative opinions [5].

Optimised weights for relations. Table 3 shows the weights learnt for the different RST elements. The weight of the nucleus was fixed to one. Weights of satellites are real numbers in the interval $[-2, 2]$. Having been assigned a weight of 1, nuclei are assumed to play a more or less important role in conveying the overall sentiment of a piece of natural language text. Yet, some types of satel-

Table 3. Optimised weights for RST relation types trained with PSO over positive and negative rankings and the percentage of presence of different relations in the training

	Positive		Negative	
Relation	% of Presence	Weight	% of Presence	Weight
attribution	.183	0.531	.177	2.000
background	.034	-0.219	.038	-2.000
cause	.009	1.218	.009	-0.011
comparison	.003	-1.219	.003	-2.000
condition	.029	-0.886	.025	-2.000
consequence	.001	0.846	.001	1.530
contrast	.016	-1.232	.017	-2.000
elaboration	.207	2.000	.219	2.000
enablement	.038	2.000	.038	1.221
evaluation	.001	0.939	.001	-2.000
explanation	.007	2.000	.008	2.000
joint	.009	-1.583	.010	1.880
otherwise	.001	-1.494	.001	-0.428
temporal	.003	-2.000	.003	-0.448

lites appear to play an important role as well in conveying the overall sentiment of a document. For instance, the most salient relations (highest percentage of appearance in the collection) in our training set appear to be the *elaboration* and the *attribution* relation. For both positive and negative documents, satellite segments elaborating on matters presented in nuclei are typically assigned relatively high weights, exceeding those assigned to nuclei. Bloggers may, therefore, tend to express their sentiment in a more apparent fashion in elaborating segments rather than in the core of the text itself. A similar pattern emerges for attributing satellites as well as for persuasive text segments, i.e., those involved in *enablement* relations, albeit to a more limited extent (lower frequency of occurrence). Interestingly, however, the information in attributing satellites appears to be more important in negative documents than in positive documents. Another important observation is that the sentiment conveyed by elements in contrast satellites gets a negative weight. This permits to appropriately estimate the polarity of sentences such as the one we introduced in Section 2 (*"Although I like the characters, the book is horrible."*).

4   Related Work

Numerous studies have been conducted to determine opinions in blog posts. In large-scale scenarios the search for subjective documents (regardless of their polarity) has been studied in detail [7–9]. Most successful studies in this area try to find document that are both opinionated and on-topic [7, 8]. To perform this task, some authors consider positional information as the best guidance to find opinions related to the query. For example, Santos et al. [8] used the proximity of query terms to subjective sentences in a document to detect on-topic opinions. In a similar way, Gerani et al. [7] proposed a proximity-based opinion propagation method to calculate the aggregated opinion at the position of each query term in a document.

Pang and Lee [17] considered the use of the location of the opinionated sentences on the accuracy of two state-of-the art polarity classifiers of film reviews. They built polarity classifiers based on sentences from different parts of a document (e.g. first sentences, last sentences), however these classifiers were not able to overcome local-unigram state-of-the-art systems. Nevertheless, the results obtained showed that the last sentences of a document might be a good indicator of the overall polarity of the review.

In [18], Zirn et. al. presented an automatic framework for fine-grained sentiment analysis at sub-sentence level in a product review scenario. They combined several sentiment lexicons with neighborhood information and discourse relations to enhance polarity performance. Concretely, they used Markov logic to integrate polarity scores from different sentiment lexicons with information about relations between neighboring segments of texts. They demonstrated that the use of structural features improves the accuracy of polarity predictions achieving accuracy scores of up to 69%. In our paper we have studied the impact of structural information in a more demanding multi-topic scenario, the blogosphere.

Somasundaran et al. [19] demonstrated the importance of general discourse analysis in polarity classification of multi-party meetings. The importance of RST for the classification of ambiguous sentences (i.e., sentences with conflicting opinions) was studied in [20]. Closer to our work, Heerschop et. al. [5] worked with film reviews and used RST to determine the importance of every piece of text in the review for polarity classification. By dividing the text into important and less important parts, depending on their rhetorical role according to a sentence-level RST-analysis, they were able to outperform a whole-document approach based on polarity lexicons. One of the main issues that the authors found in their experiments was the processing time required for identifying/classifying discourse structure in natural language text. This problem prevents the application of these methods in large-scale scenarios. In our work we have revisited this issue and we have studied and successfully applied rhetorical relations in a large-scale scenario.

In [21], Lioma et. al. designed a Language Model (LM) that takes into account RST information to estimate the relevance of a document to a query in a web search scenario. Their experiments showed that some rhetorical relations lead to important gains in performance over state-of-the-art retrieval methods.

Chenlo and Losada [4] proposed some effective and efficient methods to find the opinionated passages of a blog post that are on-topic. By combining simple sentence retrieval methods and polarity evidence, the authors were able to represent the overall opinion of a blog post by selecting just a few sentences from the beginning, from the end or from the set of most subjective and on-topic sentences of the document. In our current endeavours, we have used this method to focus our RST analysis on the core parts of documents and also to avoid the problems related to the use of a computationally expensive method such as RST-based polarity analysis.

5   Conclusions and Future Work

In this paper we have taken the first steps towards studying the usefulness of RST-based polarity analysis in the blogosphere. We found that the use of discourse structure significantly improves polarity detection in blogs. We have applied an effective and efficient strategy to select and analyse key opinion sentences in a blog post and we have found some trends related to the way in which people express their opinions in blogs. Concretely, there is a clear predominance of attribution and elaboration rhetorical relations. Bloggers tend to express their sentiment in a more apparent fashion in elaborating and attributing text segments rather than in the core of the text itself.

Finally, most of the methods proposed on this work are based on a simple combination of scores. As future work, we would like to study more formal combination methods. Related to this, we are also interested in more refined representations of rhetorical relations (e.g., LMs [21]). Another problem to take into account is that we are using only one sentence to evaluate the polarity of the blog post. Under these conditions the benefits of applying rhetorical relations have some limitations (e.g., the sentence selected may not be a good representative for the blog post). In the near future, we plan to explore the benefits of discourse structure while taking more sentences into account in our analysis. Related to this, one of the core problems derived to the use of RST is the processing time required for identifying discourse structure in natural language text. Therefore, we would like to explore more efficient methods of identifying the discourse structure of texts.

Acknowledgments

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. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval **2**(1-2) (2007) 1–135
3. Ounis, I., Macdonald, C., Soboroff, I.: Overview of the TREC 2008 blog track. In: Proc. TREC 2008, the 17th Text Retrieval Conference, NIST (2008)
4. Chenlo, J.M., Losada, D.: Effective and efficient polarity estimation in blogs based on sentence-level evidence. In: Proc. 20th ACM Int. Conf. on Information and Knowledge Management. CIKM'11, Glasgow, UK (2011) 365–374
5. Heerschop, B., Goossen, F., Hogenboom, A., Frasincar, F., Kaymak, U., de Jong, F.: Polarity analysis of texts using discourse structure. In: Proc. 20th ACM Int. Conf. on Inf. and Knowledge Manag. CIKM '11, Glasgow, UK (2011) 1061–1070

6. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: Toward a functional theory of text organization. Text **8**(3) (1988) 243–281

7. 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

8. Santos, R.L.T., He, B., Macdonald, C., Ounis, I.: Integrating proximity to subjective sentences for blog opinion retrieval. In: Proc. 31th European Conf. on IR Research on Advances in Information Retrieval. ECIR '09, Berlin, Heidelberg, Springer-Verlag (2009) 325–336

9. He, B., Macdonald, C., He, J., Ounis, I.: An effective statistical approach to blog post opinion retrieval. In: Proc. 17th ACM Int. Conf. on Information and knowledge management. CIKM '08, New York, NY, USA, ACM (2008) 1063–1072

10. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: Proc. Conf. on Human Language Technology and Empirical Methods in Natural Language Processing. HLT '05, ACL (2005) 347–354

11. He, B., Macdonald, C., Ounis, I.: Ranking opinionated blog posts using opinion-finder. In: SIGIR. (2008) 727–728

12. Robertson, S.: How okapi came to TREC. E.M. Voorhees and D.K. Harman (eds.), TREC: Experiments and Evaluation in Information Retrieval (2005) 287–299

13. Soricut, R., Marcu, D.: Sentence level discourse parsing using syntactic and lexical information. In: Proc. 2003 conf. of the North American Chapter of the ACL on Human Language Technology - Volume 1. NAACL '03, Stroudsburg, PA, USA, ACL (2003) 149–156

14. Carlson, L., Marcu, D., Okurowski, M.E.: Building a discourse-tagged corpus in the framework of rhetorical structure theory. In: Proc. 2nd SIGdial Workshop on Discourse and Dialogue - Volume 16. SIGDIAL '01, ACL (2001) 1–10

15. 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)

16. Parapar, J., Vidal, M., Santos, J.: Finding the best parameter setting: Particle swarm optimisation. In: 2nd Spanish conf. on IR, CERI'12. (2012) 49–60

17. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Pr. of the ACL. (2004) 271–278

18. Zirn, C., Niepert, M., Stuckenschmidt, H., Strube, M.: Fine-grained sentiment analysis with structural features. Number 12, Asian Federation of Natural Language Processing (2011)

19. Somasundaran, S., Namata, G., Wiebe, J., Getoor, L.: Supervised and unsupervised methods in employing discourse relations for improving opinion polarity classification. In: Proc. 2009 conf. on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1. EMNLP '09, ACL (2009) 170–179

20. Zhou, L., Li, B., Gao, W., Wei, Z., Wong, K.F.: Unsupervised discovery of discourse relations for eliminating intra-sentence polarity ambiguities. In: Proc. Conf. on Empirical Methods in Natural Language Processing. EMNLP '11, Stroudsburg, PA, USA, ACL (2011) 162–171

21. Lioma, C., Larsen, B., Lu, W.: Rhetorical relations for information retrieval. In: Proc. 35th Int. Conf. ACM SIGIR on Research and development in information retrieval. SIGIR '12, New York, NY, USA, ACM (2012) 931–940