MOUNA: Mining Opinions to Unveil Neglected Arguments

Mouna Kacimi
Free University of Bozen-Bolzano
Italy
Mouna.Kacimi@unibz.it

Johann Gamper
Free University of Bozen-Bolzano
Italy
gamper@inf.unibz.it

ABSTRACT

A query topic can be subjective involving a variety of opinions, judgments, arguments, and many other debatable aspects. Typically, search engines process queries independently from the nature of their topics using a relevance-based retrieval strategy. Hence, search results about subjective topics are often biased towards a specific view point or version. In this demo, we shall present *MOUNA*, a novel approach for opinion diversification. Given a query on a subjective topic, *MOUNA* ranks search results based on three scores: (1) *relevance* of documents, (2) *semantic diversity* to avoid redundancy and capture the different arguments used to discuss the query topic, and (3) *sentiment diversity* to cover a balanced set of documents having positive, negative, and neutral sentiments about the query topic. Moreover, *MOUNA* enhances the representation of search results with a summary of the different arguments and sentiments related to the query topic. Thus, the user can navigate through the results and explore the links between them. We provide an example scenario in this demonstration to illustrate the inadequacy of relevance-based techniques for searching subjective topics and highlight the innovative aspects of *MOUNA*. A video showing the demo can be found in *http://www.youtube.com/user/mounakacimi/videos* .

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Information Search and Retrieval]

Keywords

Diversification, Ranking, Algorithms, Sentiment Analysis

1. INTRODUCTION

Querying subjective topics often returns results that are biased towards specific view points. Hence, diversification of search results is a crucial ingredient for improving the satisfaction of information seekers. While there is a considerable effort on diversifying search results of ambiguous queries, there is a little effort on covering different view points in the result set of queries on subjective topics, such as "assisted suicide", "human evolution", "UFOs existence", and queries looking for different versions of the same breaking news such us "Greece bailout". Users seeking for such type of knowledge are not necessarily interested in a specific way of looking at a given issue, but possibly learning about that topic and interested in finding all different view points. To this end, it is important to focus on how best to produce a set of diversified results that cover different arguments and sentiments.

Search result diversification has been investigated extensively [5, 1, 2, 8] aiming to produce a set of results that cover different interpretations of ambiguous queries, mainly based on relevance and novelty. Few attempts addressed the problem of queries on subjective topics [3] by introducing a diversification model based on positive, negative, and neutral opinions. However, identifying different opinions goes beyond just finding positive and negative sentiments. It involves additionally the arguments used to discuss a topic. For example, two documents can be both negative about "Greece bailout" but for completely different reasons. One is negative because *bailout forces Greece to accept unfair conditions*, while the other is negative because *Greece bailout is not fair to European taxpayers*. These two documents are not diverse with respect to their sentiments, but they are diverse with respect to the arguments they provide. To the best of our knowledge, we were the first to define an opinion as a combination of sentiments and arguments, and introduce a diversification model to find the different opinions about subjective topics [6].

Identifying diverse opinions with different sentiments and arguments faces two main challenges. First, finding documents discussing similar arguments without being mislead by the difference they might exhibit in presenting them. Second, diversifying a set of documents by maximizing the number of discussed arguments without being biased against specific sentiments. In this paper, we demonstrate *MOUNA* , a system that solves the above problems by using a novel ranking model based on (1) *relevance* to select relevant documents for the query topic , (2) *semantic diversity* to avoid redundancy and cover a diverse set of documents presenting different arguments, and (3) *sentiment diversity* to cover different types of sentiments that can be positive, negative, or neutral. Additionally, *MOUNA* provides a summary of opinions, including the different arguments and sentiments, covered by the result set of each query. Using this summary representation, the user can explore the links between the results sharing the same, opposing, or complementary view points. In this demonstration, we show a user scenario to illustrate the inadequacy of relevance-based techniques for searching subjective topics, and highlight the innovative aspects of our approach.

CIKM'12, October 29–November 2, 2012, Maui, HI, USA.
ACM 978-1-4503-1156-4/12/10.

2. OPINION DIVERSIFICATION MODEL

From documents about subjective topics, we extract a set of *topic elements*, E=$\{e_1, e_2, ..., e_n\}$, which reflect the different arguments of the topic discussion. Concretely, topic elements are assumed to be the noun phrases that occur in the text. Each topic element occurs within a set of sentences. These sentences give the context in which the topic element was discussed. We define the *context* C_i of a topic element e_i by the set of words composing the sentences to which it belongs. We call these words, *context words* and we write $C_i = \{cw_1, cw_2, ..., cw_m\}$ where cw_j is a context word of topic element e_i . We restrict context words only to nouns, adjectives, and adverbs. For each context word, we define its sentiment orientation, which will be used to predict the sentiment assigned to the corresponding topic element. Combining topic elements, context words, and sentiment orientations, we define a document as a set of triples:

$$d = \{< e_i, cw_j, o_j >\}$$

where e_i is a topic element, cw_j is a context word associated with e_i, and o_j is the sentiment orientation of cw_j.

As a first component of opinion diversification, we need to compute semantic dissimilarity to detect documents that have dissimilar arguments. To this end, we do not use the whole vocabulary of a document to compute the semantic distance.We rather use only the topic elements to compute the distance between two documents. The main reason is that two pages can discuss the same argument of a subjective topic but using different context words. Therefore, our approach can capture similar opinions even if they are presented in two different ways as shown in [6] . Another key aspect is that we can effectively combine semantic dissimilarity that uses only topic elements (i.e., nouns) with sentiment dissimilarity that uses context words with adverbs and adjectives as described further. Formally, we define an opinion-based Jaccard similarity function $O\_Jaccard$ as follows:

$$O\_Jaccard(a, b) = \frac{|E(a) \cap E(b)|}{|E(a) \cup E(b)|}$$

where $E(a)$ and $E(b)$ are the set of topic elements of document a and document b respectively. As a second component of opinion diversification, we need to compute sentiment dissimilarity between documents. To this end, documents are classified into one of the three classes *positive*, *negative*, or *neutral* based on the sentiment orientations of their context words. Thus, two documents belonging to two different classes are dissimilar.

Based on the previously defined semantic and sentiment components, we propose an opinion diversification model as an extension to the Max-sum model presented in [5]. The goal of this model is to maximize the sum of the relevance, the semantic dissimilarity, and the sentiment dissimilarity of the selected set. The function we aim at maximizing can be formalized as follows:

$$f(L) = \alpha(k-1)\sum_{a\in L} r(a) + 2\beta \sum_{a,b\in L} d(a,b) + 2\gamma \sum_{a,b\in L} s(a,b)$$

where $|L| = k$, and $\alpha, \beta, \gamma > 0$ are parameters specifying the trade-off between relevance, semantic diversity, and sentiment diversity. $r(a)$ is a BM25 scoring function, d(a,b) is the $O\_Jaccard$ distance, and $s(a, b)$ is binary distance function that returns 1 if a and b belong to the same class and 0 otherwise. Note that we need to scale up the three terms of the function. The reason is that there are $\frac{k(k-1)}{2}$ numbers in the semantic similarity sum, and $\frac{k(k-1)}{2}$ in the sentiment sum as opposed to k numbers in the relevance sum.

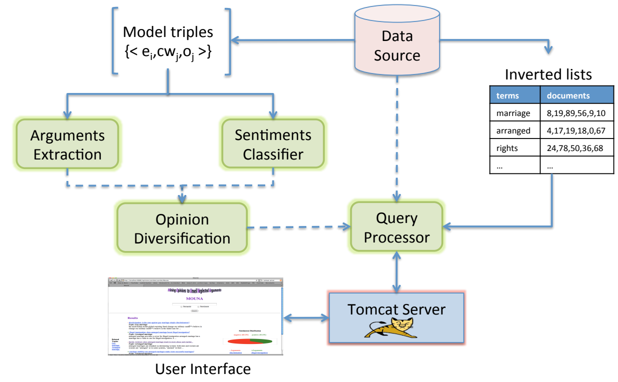

Figure 1: Architecture of MOUNA

3. SYSTEM ARCHITECTURE

Figure 1 shows an overview of the architecture of *MOUNA*. Data from a search engine or a database are imported and precomputed into data structures used for query processing. Given a query entered by a user through a Tomcat servlet, the query processor computes the topk relevant documents. Then, it applies the opinion diversification model to obtain a set of diverse documents which is returned through Tomcat. The system consists of four major components:

- *Arguments Extraction*: obtains the set of topic elements from the pre-computed (alternatively on the fly) triple representation of documents. The arguments are then ranked using *tf-idf* scheme. Note that we have used OpenNLP[7] to extract the different types of words from the text.

- *Sentiments Classifier*: classifies documents into one of the following classes: *positive*, *negative*, or *neutral* using *Logistic Regression* provided by WEKA[9].The features used for classification are the set of *context words* that have either *positive* or *negative* sentiment orientation. Sentiment orientations of context words are determined using the opinion lexicon provided in [4].

- *Opinion Diversification*: gets, from the query processor, the topk results as well as the parameters α, β and γ. Depending on the values of these parameters, it computes a set of diverse documents based on arguments, sentiments, or both. For an efficient processing, we use the 2-approximation algorithm presented in [6] to maximize our objective function f.

- *Query Processor*: applies the Threshold algorithm for top-k processing based on BM25 scoring. In case of search with semantic and/or sentiment diversification, the query processor sends the topk results to the *Opinion Diversification* component, gets the diverse results back, and then return them to the user.

(a) without opinion diversification

(b) with opinion diversification

Figure 2: Screenshot of the search interface

4. DEMO DESCRIPTION

The demo presents a full implementation of *MOUNA*, a hybrid search system for subjective topics, where the user can perform diversification of search results to cover various opinions based on arguments, sentiments or hybrid combination of them. We showcase a dataset crawled from *http://debatepedia.idebate.org/*, an encyclopedia of pro and con arguments and quotes on critical issues. We have gathered the different sections of each debate and all the web pages of the external links forming a dataset of 17637 documents on 621 topics. Alternatively, we have the possibility to pose queries to a search engine and then apply the diversification on the result set.

A visitor of the demo can provide a query consisting of one or multiple terms using the interface shown in Figure 2. The visitor can choose the diversification mode to be either semantic, sentiment, or both (using checkboxes in the interface that set the parameters to predefined values). The visitor can also search without any diversification, an important feature that helps to demonstrate the impact of our approach. The system evaluates the query, performs the requested diversification, and then produces a set of results (i.e., documents) together with a description of their characteristics. First, for each document, it shows its sentiment with a positive or a negative symbol, the topic discussed in the document, and few sentences extracted from its content. Second, the interface provides a summary of the opinions covered by the returned results. The summary contains a pie chart showing the distribution of sentiments of the retrieved documents. This chart demonstrates a possible bias in the results and the impact of diversification. Additionally, the topk arguments are displayed togther with the context in which they were discussed (positive or negative). Using the arguments summary, the visitor can browse the documents related to each of the arguments and sentiments. This feature helps moving away from a row list of results to a more structured way of presenting the resulting documents and how they relate to each other. Moreover, it demonstrates the difference between the various diversification methods in covering which and how many arguments and sentiments. Third, the interface provides the list of related topics to the submitted query based on the topk results. In this way, the visitor can navigate through other topics to have a broader view. This feature demonstrates the effectiveness of the search strategy. The more related are these topics the more the ranking model is effective.

Figure 2 shows an example query on *human evolution* without and with diversification. We can observe the change in opinion distribution from Figure 2(a) that shows biased results with 20% of negative documents vs 80% positive documents to Figure 2(b) where the percentage of negative documents is increased to 40%. Additionally, more relevant arguments were covered in both negative and positive contexts. A key observation is that without diversification, some topics and arguments are not related to *human evolution* like *punishment of adults* and *capitalism*, whereas with diversification all the topics and arguments are related to *human evolution*. A video of the demo can be found in *http://www.youtube.com/user/mounakacimi/videos* .

5. REFERENCES

[1] R. Agrawal, S. Gollapudi, A. Halverson, and S. Ieong. Diversifying search results. In *WSDM*, pages 5–14, 2009.

[2] C. L. A. Clarke, M. Kolla, G. V. Cormack, O. Vechtomova, A. Ashkan, S. Büttcher, and I. MacKinnon. Novelty and diversity in information retrieval evaluation. In *SIGIR*, pages 659–666, 2008.

[3] G. Demartini. Ares: A retrieval engine based on sentiments - sentiment-based search result annotation and diversification. In *ECIR*, pages 772–775, 2011.

[4] X. Ding, B. Liu, and P. S. Yu. A holistic lexicon-based approach to opinion mining. In *WSDM*, pages 231–240, 2008.

[5] S. Gollapudi and A. Sharma. An axiomatic approach for result diversification. In *WWW*, pages 381–390, 2009.

[6] M. Kacimi and J. Gamper. Diversifying search results of controversial queries. In *CIKM*, pages 93–98, 2011.

[7] openNLP. http://opennlp.sourceforge.net/.

[8] R. L. T. Santos, C. Macdonald, and I. Ounis. Selectively diversifying web search results. In *CIKM*, pages 1179–1188, 2010.

[9] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations (The Morgan Kaufmann Series in Data Management Systems)*. Morgan Kaufmann, 1st edition, October 1999.