Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

An approach to rank reviews by fusing and mining opinions based on review pertinence

Jun-ze Wang [a,*], Zheng Yan [b,c], Laurence T. Yang [d,e], Ben-xiong Huang [a]

[a] Non-Traditional Security Center, Huazhong University of Science and Technology, China
[b] The State Key Laboratory of ISN, Xidian University, China
[c] The Department of Comnet, Aalto University, Finland
[d] The School of Computer Science, Huazhong University of Science and Technology, China
[e] The Department of Computer Science, St. Francis Xavier University, Canada

ARTICLE INFO

ABSTRACT

Fusing and mining opinions from reviews posted in webs or social networks is becoming a popular research topic in recent years in order to analyze public opinions on a specific topic or product. Existing research has been focused on extraction, classification and summarization of opinions from reviews in news websites, forums and blogs. An important issue that has not been well studied is the degree of relevance between a review and its corresponding article. Prior work simply divides reviews into two classes: spam and non-spam, neglecting that the non-spam reviews could have different degrees of relevance to the article. In this paper, we propose a notion of "Review Pertinence" to study the degree of this relevance. Unlike usual methods, we measure the pertinence of review by considering not only the similarity between a review and its corresponding article, but also the correlation among reviews. Experiment results based on real data sets collected from a number of popular portal sites show the obvious effectiveness of our method in ranking reviews based on their pertinence, compared with three baseline methods. Thus, our method can be applied to efficiently retrieve reviews for opinion fusion and mining and filter review spam in practice.

1. Introduction

The fast growth of the Internet has dramatically changed the way that people express their opinions. Nowadays, people can freely post reviews on articles at numerous websites to express their personal opinions. They can also freely share their attitudes and comments in online and mobile social networking. As the reviews express the subjective attitudes, evaluations, and speculations of people in natural language, this kind of contents contributed by Internet users have been well recognized as valuable information. It can be exploited to analyze public opinions on a specific topic or product in order to figure out user like or dislike, etc. Opinion fusion and mining are the methods to analyze and summarize opinions from reviews in order to comprehend public perspectives on a specific topic or an entity.

Research has been conducted in opinion fusion and mining with regard to sentiment analysis and opinion extraction from reviews. For example, Kim and Hovy collected past election prediction messages from the Web and applied an SVM-based supervised learning method to predict election results [1]. Analogously, Lin et al. developed a statistical model to capture how perspectives are expressed at document and sentence levels, and evaluated the model using the articles about Israeli–Palestinian conflict [2].

An important issue that was neglected in the past research is the degree of relevance between a review and its corresponding article. Due to the openness of Internet forums, anyone can write anything on it. The online reviews are mostly not equally relevant to the article. The irrelevant and less relevant reviews are "noisy" to some extend in the collection of reviews. If we can estimate the degree of relevance between a review and its article, we can eliminate the irrelevant reviews and pay little attention to the less relevant ones. As a result, we won't suffer from the negative effect of noisy reviews and can focus on most relevant reviews in opinion fusion and mining, thus the performance of fusion and mining can be greatly improved. However, based on our knowledge, there is no published study on this research topic. The most related work is review spam detection, which simply divides the reviews into two classes: spam and non-spam [3–5]. It is particularly noted that the degrees of relevance between different reviews and the article are in fact different – even for the non-spam reviews.

* Corresponding author. Tel.: +86 15071230213.
  E-mail address: wangjunze@mail.hust.edu.cn (J.-z. Wang).

In this paper, we propose a notion of "Review Pertinence". It is the degree of relevance of a review to its corresponding article. Obviously, different reviews have different review pertinence. The higher the pertinence, the more the opinion expressed in the review relates the article. Thereby, the reviews with higher pertinence are more useful or valuable for fusing and mining opinions on the topic of the article; while the reviews with lower pertinence are obviously less helpful for this purpose. Particularly, the reviews that have no pertinence to the article are so-called review spam. Obviously, estimating the review pertinence can be used to rank the reviews in order to pick up valuable reviews and eliminate invaluable ones or review spam. It is useful for fusing and mining the reviews in order to analyze and comprehend public opinions on a specific topic and reviewers' personal interests and preferences.

However, review pertinence estimation is not as easy as we thought. First, the reviews are usually short and mostly contain several sentences. Thus, it is hard to capture their intrinsic meanings. Second, the reviews and the article may use different words to present same concepts. Although traditional similarity measures (such as Jaccard Coefficient and Overlap Coefficient [6]) have been widely used to estimate the relevance between documents or sentences, these measures work poorly in a situation that has little word overlap.

This paper proposes a novel method for opinion fusion and mining by considering both the similarity between a review and its corresponding article and the correlation among reviews. Different from the prior art, we consider the correlations between reviews in the review pertinence estimation in order to overcome the shortcoming of similarity measures. We hold such a hypothesis that if review r_1 has high pertinence, review r_2 that is similar to review r_1 should also have high pertinence, even though the degree of similarity between review r_2 and the article is low. The effectiveness of our method is verified through a number of experiments based on real data sets collected from a number of popular portal sites by comparing it with three baseline methods.

The rest of the paper is organized as follows. Section 2 gives a brief review of related work. In Section 3, we analyze the issue of review pertinence estimation in details. Section 4 describes the proposed method. We show our experiment results in Section 5, followed by additional analysis and discussion in Section 6. Finally, conclusions and future work are presented in the last section.

2. Related Work

2.1. Opinion fusion and mining

It has been well recognized that user-generated contents contain valuable information about users. Fusing and mining opinions (positive or negative) from reviews has become a popular research topic in recent years [7–10].

Opinion fusion and mining are applied to extract public opinions on a product or specific topic. Researchers attempted to identify sentiments (i.e., the affective parts of opinions) in reviews [11,12], or classified online product reviews into positive and negative classes [13]. In [14], the authors presented a system that, given a topic, automatically finds the people who hold opinions on that topic and the sentiment of each opinion.

Some researches focused on decomposing or summarizing opinions from reviews. Lu et al. studied the problem of decomposing the overall ratings of a large number of short comments into ratings on some major aspects, so that a user can gain different perspectives of a target product [15]. Wang et al. analyzed opinions on an entity in an online review at the level of topical aspects to discover the latent opinion of each individual reviewer on an aspect, as well as relative emphasis on different aspects when forming the overall judgment of an entity [16]. Hu and Liu adopted semantic analysis techniques to mine and summarize all the customer reviews of a product [8].

The reviews were also used to analyze online public opinions to predict political events. Kim and Hovy presented an election prediction system named Crystal based on user opinions posted on an election prediction website [1]. Given a prediction request, Crystal first identifies which party is requested to predict, and then aggregates a large amount of opinions to provide election prediction.

Many methods have been used in the opinion fusion and mining field. Choi et al. used a global sequence model to classify and assign sources to opinions [17,18]. Mao and Lebanon used a sequential CRF (Conditional Random Fields) regression model to measure the polarity on a sentence level and determine the *sentiment flow* of authors in reviews [19]. Wei and Gulla proposed an approach to label the attributes of a product and their associated sentiments in product reviews through a Hierarchical Learning process with a Sentiment Ontology Tree [20]. Kazutaka et al. employed an unsupervised approach to extract the opinions on the aspects of a product (e.g., comfort and portability) in a summarization process [21]. A graph-based summarization framework (Opinosis) was proposed to generate concise abstractive summaries of highly redundant opinions [22]. A probabilistic rating inference framework, known as *Pref*, was proposed to mine user preferences from reviews and map such preferences into a numerical rating scale [23]. Potthast and Becker introduced OPINIONCLOUD, a technology to summarize and visualize opinions that are expressed in the form of Web comments [24]. Lin et al. investigated the problem of identifying the perspective expressed in a document [2]. They proposed a number of models to learn perspectives from the words used in a document with high accuracy.

Although the above techniques were applied and examined in different domains, an important issue that has been neglected so far is the degree of relevance between a review and its corresponding article. With review pertinence estimation we can improve the performance of opinion fusion and mining by focusing on relevant reviews and at the same time filter potential review spam.

2.2. Review pertinence

For opinion fusion and mining base on reviews, an implicit demand is that the reviews and the corresponding articles should be related.

Review spam detection can be used to filter out unrelated reviews. Review spam is an activity to introduce irrelevant information into reviews. It was firstly introduced by Jindal and Liu in [3]. They presented a supervised learning approach to detect the review spam. Jindal and Liu also studied the problem of opinion spam and the trustworthiness of online opinions in the context of product reviews [5].

However, current researches on review spam detection simply treat this issue as a binary classification task: either spam or non-spam decision is made. Almost all existing methods and models were proposed based on this classification. Few researchers pay attention to the task that rank the reviews that are not spam.

We argue that, even if the reviews are non-spam ones, the degrees of their relevance to the corresponding article are different. Obviously, prior arts lack incisive study on the issue of review pertinence. In this paper, we rank the reviews depending on their relevance with corresponding articles.