Identify the online product comments with suspicious Chinese content.

Lei, Ping Han; Hsu, Pingyu; Cheng, Ming Shien

Summary: Recently, people who need more information about the goods they planned to purchase will look for the online product reviews before the purchasing. This is how “Electronic Word-of-mouth” (eWOM) influences or even changes the purchasing decision. The purpose of this research is to identify the worst kind of the online product reviews: rumors. Rumors could cause serious damage to company’s goodwill and the sale of the product. In this study, we developed a new method that combined the research of rumors and the text mining techniques. Breaking the content of online product review into two components, and then use the “Keyword matching” technique to evaluate whether it is a rumor article. The result of this method shows that it could precisely identify those rumor articles from bunch of online product reviews. We could use it as a filter when we search for product information and make a better and more suitable buying decision. Based on the models developed in this study, the results show that the articles with more important attribute vocabulary and fuzzy vocabulary and fewer words are more likely to contain rumors. The results also show that rumor articles and articles containing normal responses to questions can be effectively separated. The collected training set results show: precision=71.43%, recall=73.5%, F-measure=72.45%; the testing set results show: precision=80%, recall=73.73%, F-measure=76.19%.

Keywords: Online word-of-mouth; Text Mining; Rumor; Internet Rumors; Vague Vocabulary; Online Product Review
doi:10.1007/978-3-319-06677-6_5