

NetSpam a Network-based Spam Detection Framework for Reviews in Online Social Media

ABSTRACT

Online Social Media portals play an influential role in information propagation which is considered as an important source for producers in their advertising campaigns as well as for customers in selecting products and services. In this fake spammer's reviews designed to mislead customer's opinion for selecting products and services. For identify this fake spammer reviews, so that customer's can select the good products.

EXISTING SYSTEM

In the past years, people rely a lot on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services. In addition, written reviews also help service providers to enhance the quality of their products and services. These reviews thus have become an important factor in success of a business while positive reviews can bring benefits for a company, negative reviews can potentially impact credibility and cause economic losses. The fact that anyone with any identity can leave comments as review provides a tempting opportunity for spammers to write fake reviews designed to mislead users' opinion. These misleading reviews are then multiplied by the sharing function of social media and propagation over the web. The reviews written to change users' perception of how good a product or a service are considered as spam and are often written in exchange for money.

Disadvantages

- Negative reviews can potentially impact credibility and cause economic losses.
- Spammers to write fake reviews designed to mislead users' opinion.

PROPOSED SYSTEM

This study introduces a novel spam detection framework namely NetSpam based on a metapath concept as well as a new graph-based method to label reviews relying on a rank-based labeling approach. The performance of the proposed framework is evaluated by using two real-world labeled datasets of Yelp and Amazon websites. Our observations show that calculated weights by using this metapath concept can be very effective in identifying spam reviews and leads to a better performance. In addition, we found that even without a train set, NetSpam can calculate the importance of each feature and it yields better performance in the features' addition process, and performs better than previous works, with only a small number of features. Moreover, after defining four main categories for features our observations show that the reviews behavioral category performs better than other categories, in terms of AP, AUC as well as in the calculated weights. The results also confirm that using different supervisions, similar to the semi-supervised method, have no noticeable effect on determining most of the weighted features, just as in different datasets.

Advantages

- It can be very effective in identifying spam reviews.
- It leads to a better performance.

SOFTWARE REQUIREMENTS

Front-end	:	JSP
Back-End	:	MySQL
Server	:	Tomcat Server
OS	:	WINDOWS 7/above

HARDWARE REQUIREMENTS

PROCESSOR : CORE i3
RAM : 512MB-2GB
HARD DISK : 40GB

