

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

Abstract:

Nowadays, a big part of people rely on available content in social media in their decisions (e.g. reviews and feedback on a topic or product). The possibility that anybody can leave a review provide a golden opportunity for spammers to write spam reviews about products and services for different interests. Identifying these spammers and the spam content is a hot topic of research and although a considerable number of studies have been done recently toward this end, but so far the methodologies put forth still barely detect spam reviews, and none of them show the importance of each extracted feature type. In this study, we propose a novel framework, named NetSpam, which utilizes spam features for modeling review datasets as heterogeneous information networks to map spam detection procedure into a classification problem in such networks. Using the importance of spam features help us to obtain better results in terms of different metrics experimented on real-world review datasets from Yelp and Amazon websites.