

1215

**A Novel Method in Scam Detection and
Prevention using Data Mining Approaches**

Maryam Mokhtari, Mohammad Saraee¹⁰⁰, Alireza Haghshenas¹⁰¹

Abstract

‘Scam’ is a fraudulence message by criminal intent sent to internet user mailboxes. Many approaches have been proposed to filter out unsolicited messages known as ‘spam’ from legitimate messages known as ‘ham’.

However up to this date no suitable approach has been proposed to detect Scams. Almost all spam filters which use Machine Learning approaches, classify scams as hams when scam messages are more similar to the average ham than spam. But such fraudulence messages can be very harmful to users as many people in the world lose their funds by relying on scam messages.

In this paper we use Data Mining techniques for scam detection. Bayesian Classifier, Naïve Bayes and K-Nearest Neighbor which are mostly used in spam detection are experimented and the results are reported.

In addition, a new approach in scam detection is proposed. This approach uses K-Nearest Neighbor algorithm with modification to Document Similarity equation. Additionally, classification is not binary as ‘scam’ or ‘not scam’: a Fuzzy Decision is used instead of clear types of classes. Scam messages are successfully detected by applying this approach.

Keywords

Scam, Spam, Naïve Bayes, Nearest Neighbor, Fuzzy Decision, Clustering, Fraud

¹⁰⁰ Department of Electrical and Computer Engineering Isfahan University of Technology, Isfahan, Iran, Mokhtari@ec.iut.ac.ir, Saraee@cc.iut.ac.ir

¹⁰¹ Department of Computer Engineering Iran University of Science & Technology, Tehran, Iran, Haghshenas@comp.iust.ac.ir