Abstract

Traditional and present methods to detect spam emails have been working quite well but they take no measures to detect and occlude the malicious actions of the spammers. In this paper a combination of certain parameters of an email is considered to cluster legit emails and spam emails. Initially, this approach tries to cluster spam emails. Based on their sources, the spam emails are clustered using their Message subjects, Attachments, Number of Hyperlinks, Message length, Stylistic and Semantic parameters. Since emails from same source have certain similarities, they are clustered together. These clusters are then mapped to their respective domains and their IP address is retrieved which is then reported to Anti-Spam Agencies.
Spammer Detection by Extracting Message Parameters from Spam Emails

References

- Marios Kokkodis and Ting-Kai Huang. 2006. An empirical study of spam and spammers behaviour, University of California, Riverside.
- Sudipto Guha, Rajeev Rastogi, Kyuseok Shim, 2001. Cure: An Efficient Clustering Algorithm For Large Databases

Index Terms

Computer Science  Information Sciences

Keywords

Detection  Email Parameters  Information Extraction  Spam