Abstract

Most e-mail readers spend a significant amount of time regularly deleting junk e-mail (spam) messages, which are a part of marketing campaigning efforts of various companies wherein users normally signed in and it also results in increasing volume of storage space and consumes network bandwidth. A challenge, therefore, rests with the developers and improvement of automatic classifiers that can differentiate authentic e-mail from spam. Spam detectors normally use Naïve Bayesian approach and large feature sets of binary attributes that determine the existence of common keywords in spam emails. Spammers/Marketers recognize these approaches to impede their messages and have developed tactics to bypass these filters, but these ambiguous tactics are themselves patterns that human readers can often identify quickly. The preliminary study tests an alternative approach using a neural network (NN)
classifier to overcome drawbacks of Naïve Bayesian approach. This approach uses a feature set, which uses descriptive characteristics of words and messages similar in the way that users would use to identify spam.

References

- Declue, IP Lookup Against a List of All Known DNS-based Spam Databases.

Index Terms

Computer Science Information Sciences
Keywords
Spam  Html  Nsl  Nls  E-mail  Url  Neural Network