Abstract

Simultaneously with the development of networks, and with the increasing volume of unsolicited bulk e-mail especially advertising, indiscriminately has generated a need for reliable anti-spam filters. The problem for the traditional method of spam filtering cannot effectively identify the unknown and variation characteristics, therefore recently the researchers look at the artificial immune system exists diversity, immune memory, adaptive and self learning ability. The spam detection model describes an e-mail filtering is accomplished by extracting the characteristics of spam and ham (legitimate e-mail messages that is generally desired and isn't considered spam) that is been acquired from trained data set by feature extraction techniques. These techniques allowed to select subset of relevant, non redundant and most contributing features to have an added benefit in accuracy and reduced time complexity. The extracted features of spam and ham are then make a two types of antigen detectors, to enter then in series of cloning and mutation immune operations to built an immune memory of spam and ham. The experimental result confirms that the proposed model has a very high detection rate reach at 1 and a very low false alarm rate reach at 0 when using low numbers of feature extraction.
References


Index Terms

Computer Science  Artificial Intelligence

Keywords
Email, Spam, Ham (legitimate messages), Clonal selection, Information Gain, LDA, PCA