Abstract

As volume of internet traffic over last couple of years due to drastic rise in number of internet users, the area of IP traffic classification has gained significant importance for various internet service providers and other public and private sector organizations. In today’s scenario, traditional IP traffic classification techniques such as port number based and payload based techniques are rarely used because of their limitations of use of dynamic port number instead of well-known port number in packet headers and various cryptographic techniques which inhibit inspection of packet payload. In order to overcome these limitations, machine learning (ML) techniques are used for IP traffic classification. In this research paper, real time internet traffic
dataset has been developed using packet capturing tool and then using three different feature selection algorithms: Correlation based, Consistency based and Principal Components Analysis based feature selection algorithms, reduced feature datasets have been developed. After that, five popular ML algorithms MLP, RBF, C4.5, Bayes Net and Naïve Bayes are used for IP traffic classification with these datasets. This experimental evaluation shows that C4.5 Decision Tree Algorithm is an efficient ML technique for IP traffic classification with reduction in number of features characterizing each internet application using Correlation based Feature Selection Algorithm.

Reference

and System,” Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada.


Index Terms

Computer Science

Communications

Key words

MLP

RBF

C4.5

Bayes Net

Naïve Bayes