Abstract

Now a days internet traffic classification is an emerging research field since 1990’s because of its use in a large number of network activities. Traditional techniques of internet traffic classification that relied on well known TCP/UDP port numbers or payload based are rarely used because of use of dynamic port numbers instead of fixed port numbers and due to various cryptographic techniques which inhibit inspection of packet payload. Recent trends are use of ML (machine learning) algorithms for internet traffic classification. In our research work we propose a technique to classify the internet traffic into two classes, one for educational websites
and another for non-educational websites. In educational institutes for the optimum use of network resources and for the welfare of the students, the use of non-educational websites should be banned while only the educational websites should be allowed to open. To classify the internet traffic we propose a technique to capture data packets first, related with various educational and non-educational websites, using a packet capturing tool Wireshark. Then using feature selection algorithm, a reduced feature dataset will be developed. After that training and testing of various ML algorithms will have to be performed. Finally comparative analysis of the different classifiers from the obtained results is to be performed.

Reference


Index Terms

Computer Science Communications

Key words

Machine Learning
Features