Abstract

Characteristic in application while being communicate in the network has distintives characteristics. It means, before doing the communication between other application, there is a rule which must be obey. The rule is a protocol, used by application while communicate. Another characteristic is a packet size transfered by application while communicate. Based on this characteristic, we can use unsupervised machine learning method for grouping Internet traffic data.

This final task using two clustering algorithm unsupervised machine learning, K-Means algorithm and Expectation Maximization (EM) algorithm. Both of them will be compared by other method based on port-number in grouping Internet traffic data. Testing's results show K-Means algorithm and EM algorithm make a good result in grouping internet traffic data. EM algorithm produces a cluster and accuracy point better than K-Means algorithm, but K-Means algorithm works faster than EM algorithm.

Keywords: *Clustering, K-Means, Expectation Maximization, Internet Traffic*