Abstract

Generally intrusion detection systems (IDS) use all the data features to classify normal and anomaly packet. It has been observed in the studies that some of the data features may be redundant or are less important in this classification process. Authors have studied NSL KDD dataset with different feature selected from Gain Ratio and Chi-Square feature selection methods and carried out the experiments with single Decision Tree and then applied ensemble with Random Forests and Decision Tree with Bagging. Results show that significant feature selection is very important in the design of a lightweight and efficient intrusion detection system. Random Forests are better than Single Decision Tree and Decision Tree with Bagging for the current dataset. Performance of Gain Ratio is better than Chi square feature selection method for this dataset.

References


Index Terms

Computer Science  Security
Keywords

Network security, NSL KDD, classifier, ensembles, Decision trees, Random Forests, Chi Square, Gain Ratio.