Abstract

Intrusions pose a serious risk in a network environment. New types of attack, of which detection system are unaware, are the most difficult to detect. Current signature based methods and learning algorithms which rely on labeled data to train, generally cannot detect these new intrusions. In addition, labeled training data in order to train the misuse anomaly detection system is typically very expensive. We present a new types of clustering based intrusion detection algorithm for mixed type data, unsupervised anomaly detection, which trains on unlabelled data in order to detect new intrusion. In our approach, the algorithm for clustering is based on modified K-mean algorithm. The labeling is refined by constructing minimum spanning tree (MST) based upon distance between every two clusters coming from modified k-means algorithm. Our algorithm is able to detect many different types of intrusions and maintain a low false alarm as verified over KDD cup 1999 daraset.

Keywords—Outlier, Unsupervised, K-means, MST, ANIDS.