ABSTRACT

In recent years, the proliferation of cyber threats has propelled the adoption of machine learning techniques in Intrusion Detection Systems (IDS). As the frequency and sophistication of cyber attacks continue to escalate, the demand for effective and intelligent intrusion detection solutions has never been greater. Unsupervised machine learning methods have gained prominence in the IDS domain due to their ability to detect both known and unknown attack types, including zero-day attacks. This thesis presents a novel approach that One-Class Support Vector Machine (OCSVM) algorithm to identify previously unknown cyber threats. The proposed methodology is rigorously evaluated using the CIC-IDS2017 dataset, a widely recognized benchmark in intrusion detection research. The critical aspect of this approach is feature selection, because clustering is done by feature extraction. On this research, KBest as feature selection method used to improve anomaly detection. We compare detection performance of OCSVM with KBest, OCSVM without feature selection, and OCSVM with another method. The research results show that OCSVM with KBest - 15 features has the best metric with silhouette score of 0.9978, 0.9989, 0.9978, and 0.9988 for the scenarios. In addition, it can be concluded that the feature selection process can reduce the training process time and testing or prediction processing time due to a reduction in the number of irrelevant features.

Keywords: anomaly detection, intrusion detection system (IDS), machine learning, unsupervised learning, OCSVM, feature selection, KBest