World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:16, No:08, 2022

A Survey in Techniques for Imbalanced Intrusion Detection System Datasets

Authors : Najmeh Abedzadeh, Matthew Jacobs

Abstract : An intrusion detection system (IDS) is a software application that monitors malicious activities and generates alerts if any are detected. However, most network activities in IDS datasets are normal, and the relatively few numbers of attacks make the available data imbalanced. Consequently, cyber-attacks can hide inside a large number of normal activities, and machine learning algorithms have difficulty learning and classifying the data correctly. In this paper, a comprehensive literature review is conducted on different types of algorithms for both implementing the IDS and methods in correcting the imbalanced IDS dataset. The most famous algorithms are machine learning (ML), deep learning (DL), synthetic minority over-sampling technique (SMOTE), and reinforcement learning (RL). Most of the research use the CSE-CIC-IDS2017, CSE-CIC-IDS2018, and NSL-KDD datasets for evaluating their algorithms.

Keywords : IDS, imbalanced datasets, sampling algorithms, big data
Conference Title : ICICCS 2022 : International Conference on Intelligent Computing and Control Systems
Conference Location : Montreal, Canada
Conference Dates : August 08-09, 2022