

Anomaly Detection in Network Traffic Using Unsupervised Machine Learning Approach

Harish Lakhangire¹, Abhishek Prasad², Shriyash Kamble³, Rupesh Singh⁴, Prof. Vishal Walunj⁵
Students, Department of Computer Science and Engineering^{1,2,3,4}
Professor, Department of Computer Science and Engineering⁵
D.Y. Patil School of Engineering Technology, Pune, India

Abstract: *The advent of IoT technology and the increase in wireless networking devices has led to an enormous increase in network attacks from different sources. To maintain networks as safe and secure, the Intrusion Detection System (IDS) has become very critical. Intrusion Detection Systems (IDS) are designed to protect the network by identifying anomaly behaviors or improper uses. Intrusion Detection systems provide more meticulous security functionality than access control barriers by detecting attempted and successful attacks at the endpoint of within the network. Intrusion prevention systems are the next logical step to this approach as they can take real-time action against breaches. To have an accurate IDS, detailed visibility is required into the network traffic. The intrusion detection system should be able to detect inside the network threats as well as access control breaches. IDS has been around for a very long time now. These traditional IDS were rules and signature based. Though they were able to reduce false positives they were not able to detect new attacks. In today's world due to the growth of connectivity, attacks have increased at an exponential rate, and it has become essential to use a data-driven approach to tackle these issues. In this paper, the KDD data set was used to train the unsupervised machine learning algorithm called Isolation Forest. The data set is highly imbalanced and contains various attacks such as DOS, Probe, U2R, R2L. Since this data set suffers from redundancy of values and class imbalance, the data preprocessing will be performed first and also used unsupervised learning. For this network traffic-based anomaly detection model isolation forest was used to detect outliers and probable attack the results were evaluated using the anomaly score*

Keywords: anomaly detection, isolation forest, machine learning, intrusion detection system, KDD Cup, NSL-KDD

REFERENCES

- [1]. G. Karatas et al., "Deep Learning in Intrusion Detection Systems" 2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT), Turkey, 2018.
- [2]. H. Azwar et al., "Intrusion Detection in secure network for Cybersecurity systems using Machine Learning" 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences, Bangkok, Thailand, 2018.
- [3]. Y. Chang et al., "Network Intrusion Detection Based on Random Forest and Support Vector Machine," IEEE International Conference on Computational Science and Engineering (CSE), Guangzhou, 2017
- [4]. Brao, Bobba et al., "Fast kNN Classifiers for Network Intrusion Detection System", Indian Journal of Science and Technology. 2017.
- [5]. S.M. Z. Alom et al., "Network intrusion detection for cyber security using unsupervised deep learning approaches", 2017 IEEE National Aerospace and Electronics Conference (NAECON), Dayton, OH, 2017.
- [6]. Mukkamala et al., "Intrusion detection using neural networks and support vector machines", International Joint Conference 2012.
- [7]. Azwar, Hassan et al., "Intrusion Detection in secure network for Cybersecurity systems using Machine Learning and Data Mining", 2018.
- [8]. Jeya, P et al., "Efficient Classifier for R2L and U2R Attacks", International Journal Comput. Appl. (2012)

- [9]. Mohana, NK Srinath “Trust Based Routing Algorithms for Mobile Adhoc Network”, International Journal of Emerging Technologies and Advanced Engineering (IJETA), volume 2, issue 8, pp. 218-224, IJETA.