Paper ID: MISS21_58

Analysis of Various Machine Learning / Deep Learning Performance of Advanced Persistent Threat Attack

Indra Kumari[1,2*], Arjun Sarkar[3]

[1]Korea Institute of Science and Technology Information (KISTI),

[2]University of Science and Technology (UST), Daejeon, South Korea,

[3]Leibniz Institute for Natural Product Research and Infection Biology, Hans Knöll Institute, Jena, Germany

*Corresponding author

Abstract

Background: Advanced Persistent Threat (APT) attacks are noxious, intentional attacks for gaining the access to computer networks for an extended time-period [1]. The quantity of assaults and the size of their threats to associations, organizations, and state-run administrations is ever increasing [2]. This study intends to provide a unique method to APT early stage detection that is based on the life cycle of an APT assault and using high machine learning and deep learning algorithms [3].

Objective: The utilization of network traffic is one way to detect an APT attack unanimously. The High Detection Accuracy (HDA) machine learning and deep-learning model ensures an automatic removal of the critical characteristics of an attack and is a powerful approach for identifying an APT attack.

Methodology: This study utilized the data provided for knowledge discovery and data mining (NSL-KDD). We used two hybrid-algorithm techniques - machine learning algorithms such as Support Vector Machine (SVM), Naive Bayes, Decision Tree and Multi-Layer Perceptron with PCA, and Random Forest with Autoencoders [4]. To reduce the number of dimensions and data compression in the dataset, we used PCA and Autoencoders. The classification on the dataset was done using both binary (Normal and threats class) and multi-class (Normal, DoS, Probe, R2L, U2R classes) classification methods.

Result and Discussion: Most algorithms gave good results on the binary classification on the test set, but the results of SVM plus PCA yields more than 93% accuracy. The results of SVM even gives better results than the combination of Random Forest and Autoencoder (92% accuracy). The advantage of this model is that it achieves a highest classification result by identifying different cyber-attacks related to the APT assaults.

Future Work: This research was conducted to set up an APT model early warning mechanism to reduce the impact of APT attacks. To discover better models to learn machines in APT assaults, we intend to consider GAN based models. In addition, deep learning methods with both supervised and unsupervised techniques with auto-encoder could be utilized.

References

[1] Chu, Wen-Lin, Chih-Jer Lin, and Ke-Neng Chang. "Detection and classification of advanced persistent threats and attacks using the support vector machine." Applied Sciences 9.21 (2019): 4579.

[2] Alshamrani, Adel, et al. "A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities." IEEE Communications Surveys & Tutorials 21.2 (2019): 1851-1877.

[3] Joloudari, Javad Hassannataj, et al. "Early detection of the advanced persistent threat attack using performance analysis of deep learning." IEEE Access 8 (2020): 186125-186137.

[4] Eke, Hope Nkiruka, Andrei Petrovski, and Hatem Ahriz. "The use of machine learning algorithms for detecting advanced persistent threats." Proceedings of the 12th International Conference on Security of Information and Networks. 2019.

 DOI: 10.21467/abstracts.120

ISBN: 978-81-954993-2-8