Host based machine learning approach to detect data exfiltration on computer networks

W.D. Samanwickrama[1*], I.P. Wickramasinghe[2], L.D.R.D Perera[1]

[1]*Faculty of Applied Sciences, Wayamba University of Sri Lanka, Kuliyapitiya, Sri Lanka,*
[2]*Department of Mathematics, Prairie View A&M University, USA.*

[*] *Corresponding author (email: dumindu@wyb.ac.lk)*

Data exfiltration, unauthorized transfer of sensitive information, is a most common threat on computer networks. It uses legitimate communication channels such as hypertext transfer protocol (HTTP) and hypertext transfer protocol secure (HTTPS). Differentiating the data exfiltration from normal network traffic is a challenge due to the regular exchange of data back-and-forth between networks, using traditional security solutions (e.g. firewalls). Due to its ability to adapt quickly to new and unknown complex situations, Machine learning (ML) offers a promise in this context that can be employed for resolving such challenges. In this study, we adopted an anomaly detection method to address the data exfiltration and examined the effectiveness of this approach under different network environments. An anomaly is an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism is the underlying hypothesis here.

Further, in the security domain, evaluation of novel algorithms against production network data is difficult. This is mainly due to the legal, security and privacy issues. Researchers often use simulation/emulation methods or some benchmark datasets to validate their novel algorithms. Thus, in this study, simulated number of different attacks on a network testbed and captured data for validation purpose. The advance persistent threat (APT) attacks are highly targeted in nature and often discovered years after the information has been stolen. As a result it's difficult to derive signatures to detect these attacks using signature based intrusion detection systems (SBIDSs) or to obtain labeled data to train supervised machine learning algorithms. Therefore our work employs unsupervised anomaly detection technique with Local Outlier Factor (LOF).

Experimental results are encouraging that the proposed method successfully isolates (Fig. 1) exfiltrated data flows (e.g. Remote Administration Tool traffic) from the rest of legitimate web traffic even when traffic is encrypted and uses the same channel (HTTPS). However, in order to generalize these findings, extensive validation mechanism is needed in future work.

Keywords: Anomaly detection, Data exfiltration, Machine learning, Network security, Remote administration

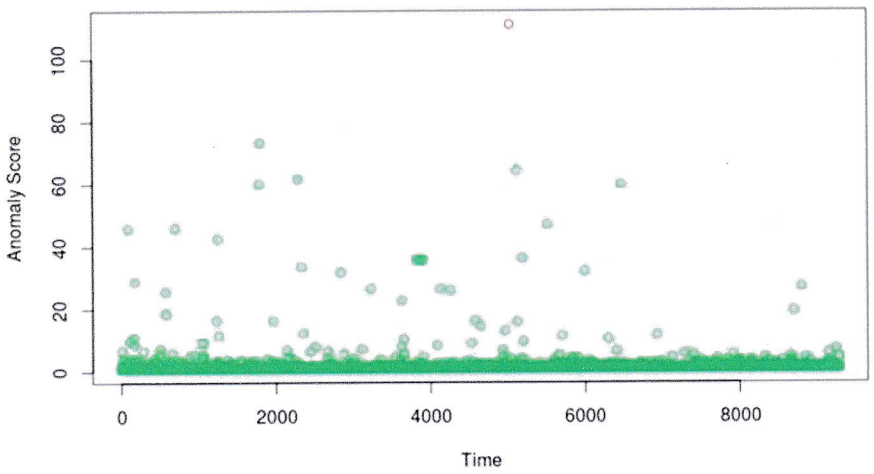

Figure 1: Anomaly scores vs time

Acknowledgement: Dr.Harsha Kalutarage at the Centre for Secure Information Technologies (CSIT) of Queen's University Belfast, UK.