Stochastic Domain Transfer Multiple Kernel Boosting: Application to anomaly Detection in Encrypted Network Traffic

Noah Reef
Cal Poly Pomona, College of Science
Cal Poly Pomona Cybersecurity and Awareness Fair 2022
Cybersecurity Problem-Solving Category

Problem

- In Cybersecurity it would be useful to flag network traffic if it contains malware
- According to WatchGaurd Technologies, 91% of malware is sent over encrypted Network Traffic
- Encryption schemes make it difficult to view what data is being sent over the network
- Hence it is difficult to label network traffic as containing malware or not.

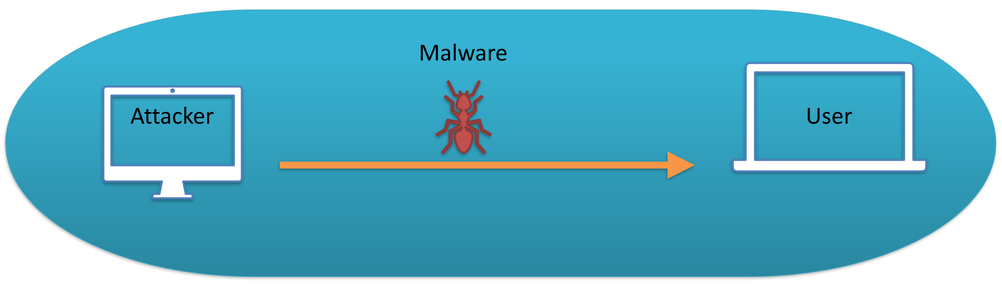

Figure1: Graphic of Malware Deployment

Approach

- Utilize metadata from unencrypted network traffic to train ML model to make predictions over encrypted network traffic.
- We use a Domain-Transfer Multiple Kernel Learning Approach to develop our model

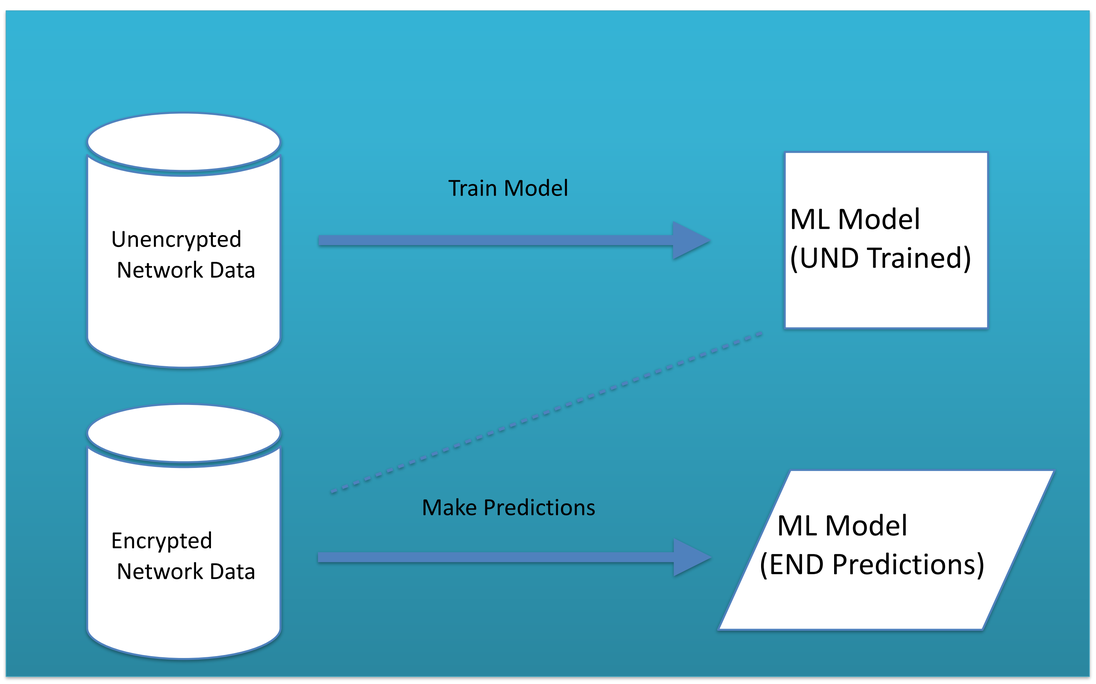

Figure 2: Flow Chart of ML Model

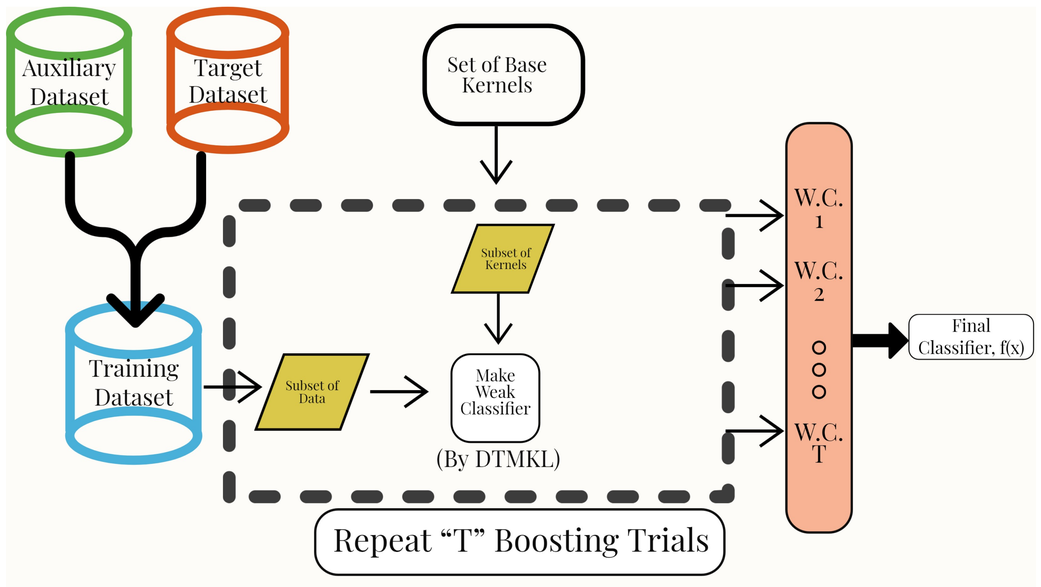

Figure 3: Machine Learning Model

Preliminary Results

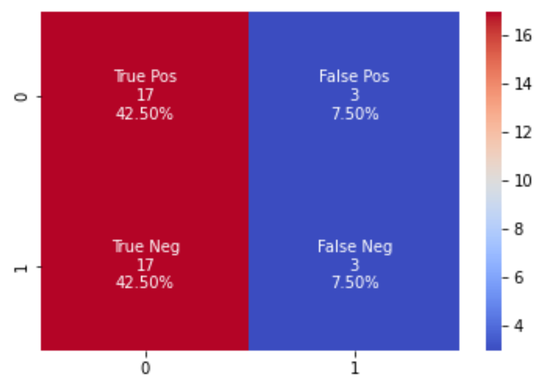

Figure 4: Confusion Matrix of models prediction

- We trained the model with 350 training samples (50 on Target Domain data and 300 on Source Data)
- The model was tested over testing set of 40 samples evenly split on labels from the Target Domain only
- Had a net accuracy of 0.85

Sources

Internet security report - Q2 2021. WatchGuard Technologies. (n.d.). Retrieved September 13, 2022, from https://www.watchguard.com/wgrd-resource-center/security-report-q2-2021