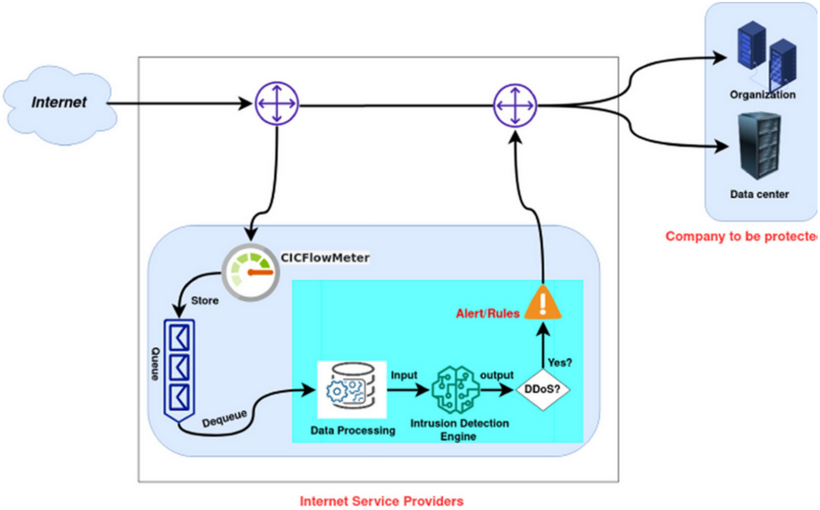
Detecting DDoS Attacks Using Adversarial Neural Network

Ali Mustapha

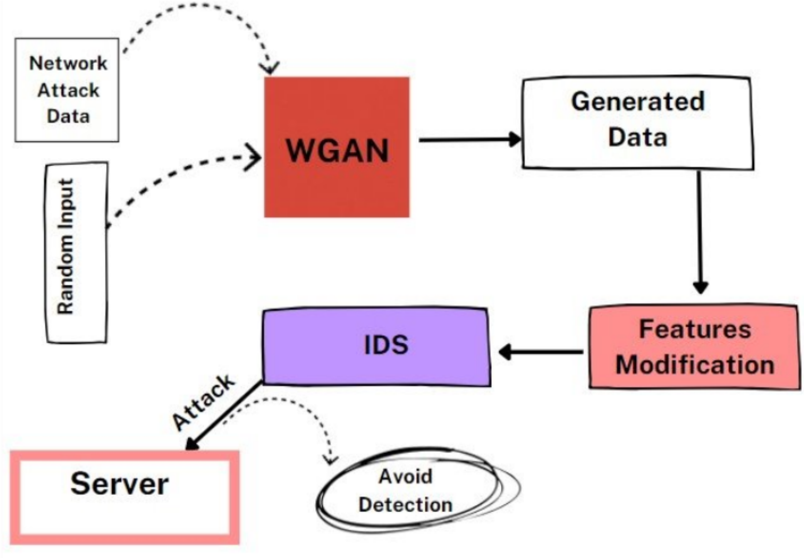
Télécom Paris, Polytechnic Institute of Paris, France

September, 2022

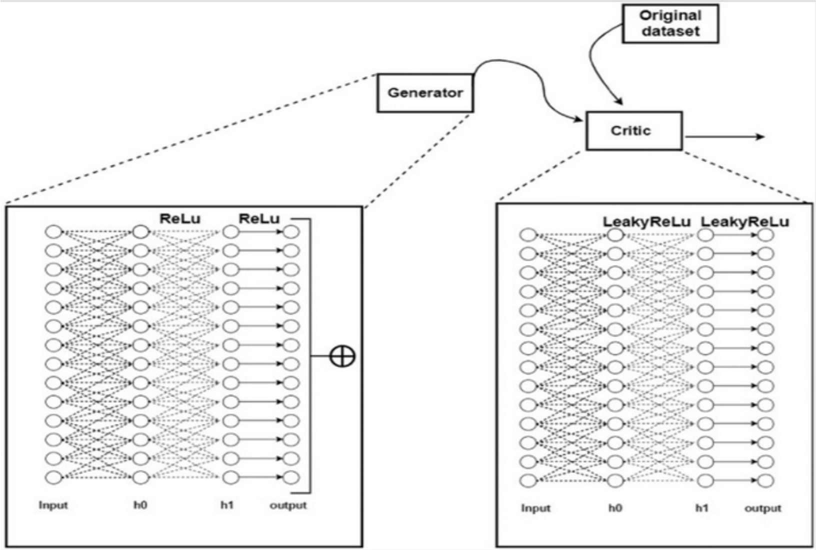
Intrusion Detection System IDS



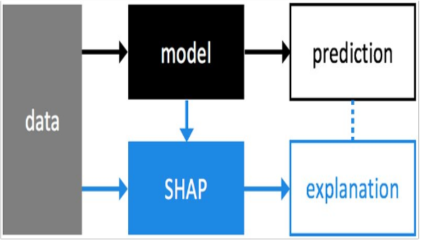
Adversarial DDoS Attack

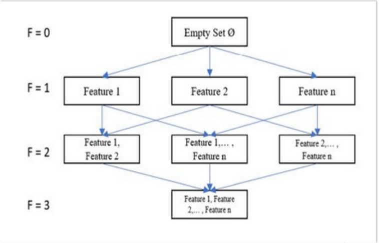


WGAN Architecture



SHapley Additive exPlanations(SHAP)



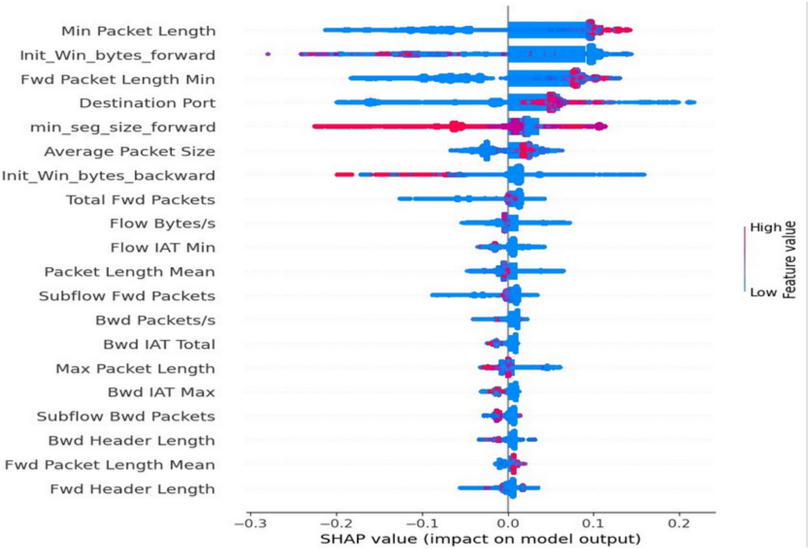


$$g(z') = \phi_0 + \sum_{j=1}^{M} \phi_j z'_j$$

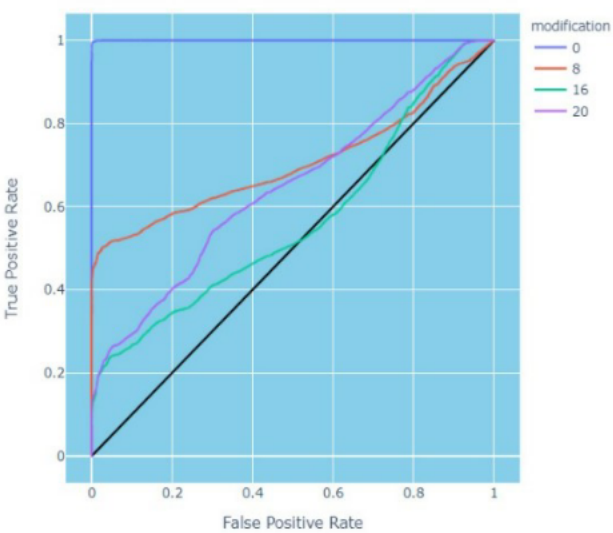
Global Interpretability: This technique provides essential features from a dataset and a contribution of each feature for a target result and effect of the feature.

Local Interpretability: With this method, we can get an impact of an individual feature across the whole dataset.

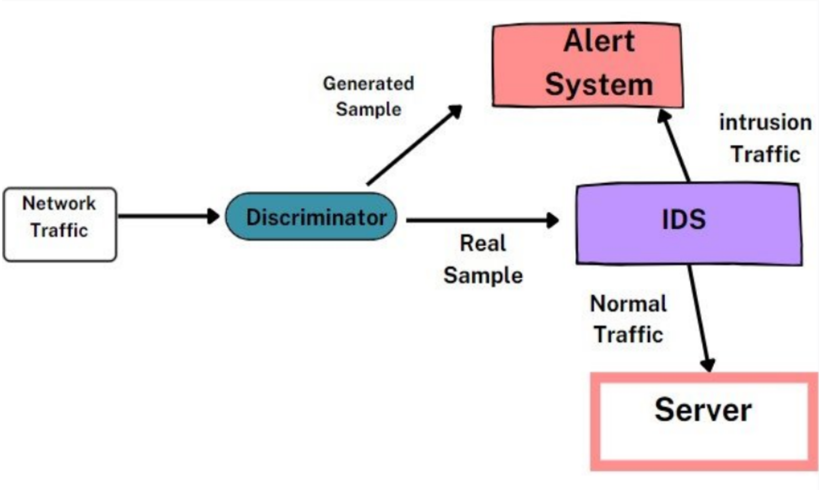
SHAP Results



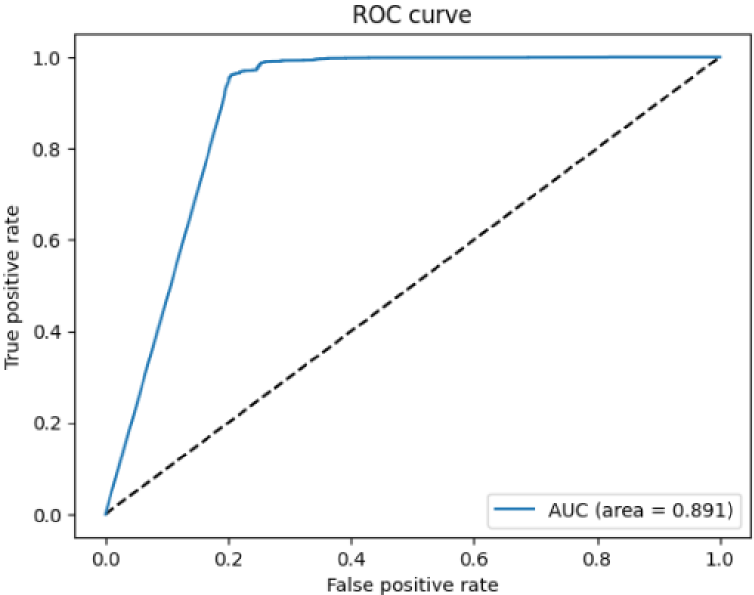
Model Performances



Proposed Solution



ROC-AUC Curve for the enhanced IDS


ROC curve
True positive rate
False positive rate
AUC (area = 0.891)