SCHOEN Adrien

Rapport de réunion n°2

Date de la réunion : 5/02/2021

Date de la prochaine réunion : 12/02/2021

Ordre du jour : Lancement du stage

1. Ce que le stagiaire a dit

Nous avons pris connaissance de quatre documents ainsi que des deux datasets d'étude. Nous avons lu deux documents sur la méthodologie d'évaluation des IDS et sur la nécessité de pouvoir générer du trafic avec une labellisation fiable. Nous avons aussi lu deux documents présentant deux datasets intéressants.

Il y a un dataset avec des flux et un dataset pour lequel nous disposons des trames réseaux brut. Malheureusement, ce dernier dispose de moins de jours de récolte que celui juste avec des flux. Nous avons aussi des flux pour le second dataset mais notre objectif est bien de générer des trames réseaux.

Nous avons appris sur la méthode d'évaluation des IDS et sur les défauts et avantages des principaux dataset. Nous avons aussi terminé de télécharger le dataset de données bruts (CICIDS2018) et nous avons commencé son analyse

Cette pré-étude portait donc sur les évaluations d'IDS et les dataset pour détections d'intrusions réseau

2. Ce que les encadrants ont ajoutés

Le but de ces premières semaine est de se familiariser avec le sujet, et d'essayer de voir ce qui a été fait. À la fin de cette étude, nous dresserons des objectifs pour ce stage

Nous ne cherchons pas forcément à apprendre les caractéristiques des pcap directement, il peut être astucieux de passer par des sous-ensembles, si depuis ces sous-ensembles nous pouvons regenerer des pcap

Il faut utiliser de l'expertise en cybersécurité pour savoir quelles caractéristiques peuvent être extraite des pcap et lesquelles peuvent permettre de générer des pcap plausibles. Un exemple a été donné avec de la grammaire

Les flux du dataset sont créés depuis un ensemble de pcap passé par un outil appelé CICFlowMeter, les flux du premier dataset sur 4 ans sont eux directement extrait du routeur via NetFlow7/9. Il est possible de créer des fichiers de flux depuis les pcap, mais nous ne connaissons pas de moyen de créer des pcap plausible depuis des fichiers flux.

Il faut commencer à étudier la génération des trames réseaux, il y a beaucoup de méthode différente.

Ne pas hésiter à contacter les encadrants pour des références ou des indications.

3. Ce qu'il faut faire pour la prochaine séance

Lire les références sur la génération de données réseaux

Comprendre les règles de générations des différents dataset

Voir l'utilisation des GAN ou des Auto Encoder