Rapport d’activité journée 4 Hackathon

Aujourd’hui nos efforts se sont principalement concentrés sur la finalisation des requêtes nécessaires à l’interface graphique d’une part, et au layout des pages web d’autre part. nous avons aussi pensé à tester plusieurs méthodes de clusterisation et de détection d’outlier pour valider notre fonction heuristique.

**Finalisation de l’interface graphique**

Diverses fonctions ont été ajouté à notre driver neo4j afin notamment de calculer le score d’anomalie pour chaque processus et de le stocker dans la base. Cette solution nous est apparue plus pratique que de simplement stocker ces valeurs sur une machine hôte.

Nous avons bien pris en considération vos remarques sur le côté ‘white box’ de notre heuristique et le détail des valeurs devrait apparaître lors de l’affichage d’un processus particulier dans l’interface.



Nous avons aussi un « score board » qui permet de visualiser rapidement les processus avec les plus gros scores d’anomalies, ainsi qu’une barre de recherche qui va permettre de renseigner une chaine de caractère que l’on veut voir apparaître soit dans l’exécutable du processus, soit dans son procees\_path.

D’un point de vue interface, nous avons avancé en ajoutant de nouvelle visualisation de données et permettant de naviguer entre les processus détectés comme menaces.

**Clusterisation et détection d’outliers**

Nous avons tester quelques méthodes de clusterisation, l’idée étant d’utiliser les valeurs des features que nous avons calculé.

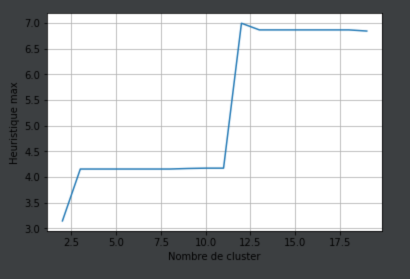
KMeans : Cette méthode donne des résultats intéressants.

Forêt d’isolations : Mauvais résultats

SVM : Mauvais résultats

KNN : Mauvais résultats

Pour la clusterisation avec les Kmeans: on a utilisé la fonction heuristique de menace que l’on a calculé statistiquement pour évaluer la clusterisation. Nous obtenons 2 plateaux :



La fonction étant le plus haut score de menace pour le cluster comprenant le plus de menaces.

Je vous joins dans le mail les clusters identifiés par le KMeans et les processus détecté par notre fonction heuristique.

**Quelques interrogations :**

Souhaitez-vous participer à la présentation de nos résultats demain ?

Quel type de livrable soutez-vous (code source, rapport,etc...) ?