Meeting

Axel Faes - 1334986

Mar 18, 2016

aanwezigen: Bram Bonne, Pieter Robyns, Axel Faes

Deze week is voornamelijk besteed aan implementatie. Er zijn verschillende algoritmes geïmplementeerd. Deze algoritmes zijn K-Nearest Neighbors, K-Means, Lineair Kernel Support Vector Machines met One vs All classification, Lineair Kernel Support Vector Machines met One vs One classification en Gaussian Kernel Support Vector Machines.

Er zijn verschillende experimenten uitgevoerd. De poort nummers zijn geïmplementeerd in 2 varianten. De eerste variant splitst de poorten in categoriën (veel gebruikte poort nummers en niet-veel gebruikte poort nummers) als een binaire feature. De andere variant maakt van elk poort nummer een nieuwe binaire feature (die stelt of de poort gebruikt is of niet). Deze variant geeft echter heel veel features, dit is enorm inefficient.

De gestelde feedback was om te kijken hoe goed het werkt als poort nummers als een continue feature voorgesteld wordt en om de categoriën op te spliten in >1024 en <1024. De IP-data can opgesplitst worden in aparte features voor IPv6, IPv4 en MAC. Deze data kan mogelijks voorgesteld worden als continue data. Er is voorgesteld om ook andere algoritmes te implementeren. Om gebruik te kunnen maken van datasets van Cegeka moet ik een presentatie maken en een afspraak maken met professor Quax.

Er is feedback gegeven op de thesistekst. Ik moet als eerste goed letten op spelfouten. De inleiding moet algemener uitgelegd worden. Ook niet gebruikte concepter zoals Signature-based IDS moet dieper uitgelegd worden. De attack classification maakt al teveel assumpties over wat er gedetecteerd kan worden. Dit zou eerst algemener uitgelegd moeten worden. Het machine learning hoofstuk bevat in het algemeen te weinig high-level beschrijvingen.

De actiepunten die gedaan zijn:

- Begin van hoe flowdata gebruikt kan worden
- Implementatie van K-Nearest Neighbors
- Implementatie van K-Means
- Implementatie van Lineair Kernel Support Vector Machines met One vs All classification
- Implementatie van Lineair Kernel Support Vector Machines met One vs One classification
- Implementatie van Gaussian Kernel Support Vector Machines.

Volgende actiepunten zijn besproken:

- Herschrijven en verwerken van feedback op de thesistekst
- Verdere implementatie: andere algoritmes
- Verdere implementatie: Ports indelen in >1024 en <1024
- Verdere implementatie: Ports indelen als continue data
- Verdere implementatie: IP indelen als continue data
- Verdere implementatie: Starttime instellen als unix time
- Maken presentatie voor Cegeka data set