BIG DATA

WAT IS HET EN WAT KAN DE LANDMACHT ER MEE IN HET MOBIELE DOMEIN?

Peter Güldenpfennig & Karin van Bodegraven,
Innovatiemanagers JIVC/KIXS

Iedereen heeft het over Big Data vandaag de dag. Data-analisten zijn niet aan te slepen en veel organisaties zijn er mee aan de slag gegaan om nieuwe bedrijfstoepassingen te realiseren. Maar wat is Big Data nu eigenlijk, wat kun je er mee en wat heeft de landmacht er aan? →

Wat is Big Data?

We spreken over Big Data wanneer de hoeveelheid digitale informatie uit een enkele dataset te veel en te divers is om op traditionele wijze, dus op een enkele server met klassieke (lees: relationele) databasetechnologie, te ontsluiten, verwerken en analyseren. Naast de typische wijze waarop data ontstaat zoals het opslaan van bestanden, het maken van foto's en video's leveren ook sensoren steeds meer data op. Naast Big Data is het *Internet of Things* (IoT) met een toenemende variëteit aan sensoren een belangrijke trend.

Data op zich betekent weinig; het krijgt pas waarde wanneer we informatie halen uit die data en ook over die data (bijvoorbeeld via metadata). Daarbij onderscheiden we gestructureerde- en ongestructureerde data. Gestructureerde data is data die weggeschreven is in een database en een bekende structuur en format volgen. Zie het als een Excel-sheet waarbij alle informatie keurig gelabeld is. Met andere woorden: het is data waarvan we weten wat voor informatie deze bevat. Ongestructureerde data is data waarvan we niet weten wat voor informatie deze allemaal precies bevat.

Onderzoeksinstituut Gartner classificeert Big Data op basis van drie v's: *volume, velocity, variety.*

- **Volume**: de hoeveelheid ongestructureerde data is groter dan gestructureerde data en vereist daarom ook meer opslag en een andere manier van verwerken;
- **Velocity**: nieuwe gegevens ontstaan zeer snel en gegevens veranderen ook waardoor ongestructureerde data significant sneller groeit en verandert dan gestructureerde data;
- **Variety**: de onderlinge variatie van ongestructureerde data is hoog door de veelheid aan bronnen van waaruit deze data afkomstig is.

Grofweg 80 procent van de data die gegenereerd wordt is ongestructureerde data. Dat is een probleem, want om waardevol te zijn voor een organisatie moet natuurlijk wel bekend zijn wat voor informatie deze data precies bevat. Daar zijn analisten voor nodig en specifieke tools die te integreren zijn in de nieuwe generatie informatiesystemen.

Informatiegestuurd optreden

Defensie en daarbinnen specifiek JIVC staan aan de vooravond van de komst van een nieuwe generatie informatiesystemen die zich kan richten op Big Data en *data analytics* met toepassingen die: data-intensief zijn (qua volume, complexiteit of variëteit), analytische routines bevatten die zwaar leunen op wiskundige technieken en geavanceerde vormen van visualisatie bevatten.
'Fact-based besluitvorming' of 'Informatiegestuurd optreden' is dan het business doel, waarbij het erom gaat om met actuele informatie, met grotere nauwkeurigheid en snelheid, het besluitvormingsproces te versnellen. Of het nu gaat om intelligence, beeldherkenning of preventief onderhoud aan materieel, vrijwel alle defensiedomeinen verwachten veel rendement te halen uit deze nieuwe generatie informatiesystemen met Big Data en data analyse. Een goed voorbeeld van waarde van data zien we al op grote schaal in de retail-sector. Zo weet Albert Heijn dankzij hun bonuskaart per filiaal bijvoorbeeld wie hun klanten zijn, welke producten goed verkopen en welke tijden het druk is in de winkel. Dit is slechts een greep uit de berg aan informatie die uit data gehaald kan worden. Vandaar ook dat een organisatie vooraf goed moet weten wat ze nu eigenlijk voor informatie uit de data wil halen. Op dit punt bevindt Defensie zich nu: zoeken naar de informatie die we uit data willen halen. Dat betekent niet dat we niet weten wat we willen weten, enkel dat we niet weten *wat we allemaal willen weten*. We maken echter flinke stappen op dit gebied, waarbij we in het vervolg van dit artikel een initiatief bij de landmacht uitlichten: Cyber Logging & Monitoring en Zebra Sword.

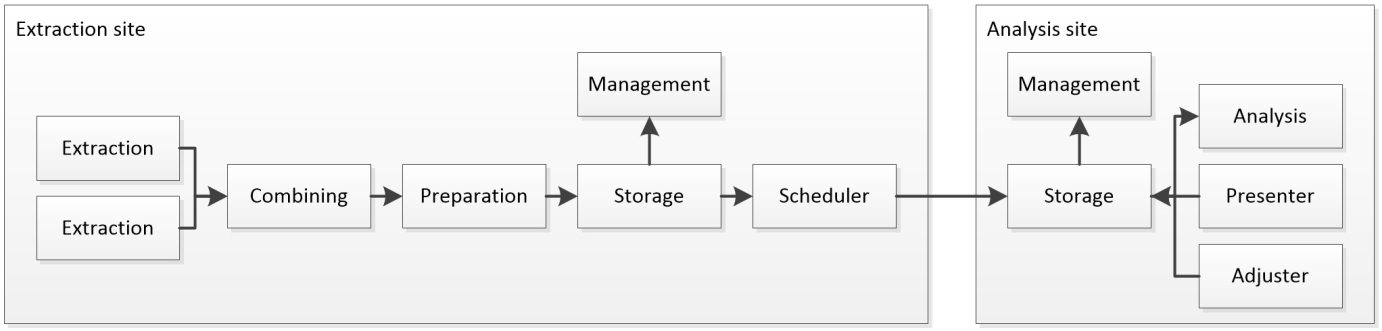
Cyber Logging & Monitoring (CyLoM) en Zebra Sword

De landmacht beschikt over wapensystemen die vol IT zitten en is daarmee kwetsbaar voor cyberaanvallen van een tegenstander. Maar hoe krijgt een commandant inzicht in de status van de wapensystemen? Is er malware aanwezig op een van de IT systemen waarmee de effectiviteit van het wapensysteem kan worden aangegrepen? Er is tot op heden geen oplossing beschikbaar voor logging & monitoring in het mobiele domein van de landmacht.
CyLoM is een generieke en flexibele infrastructuur voor het kunnen loggen van data van wapensystemen. Deze data bestaat uit platformdata, doorgaans van het *Controller Area Network* (CAN-bus) van het wapensysteem en C2-data zoals vanuit BMS en TITAAN en ELIAS. Met de verzamelde data kan Defensie vervolgens cyberanalyses uitvoeren en threat hunting-activiteiten ontplooien.

CyLoM is een nationaal technologieproject (gestart maart 2018) van TNO en Technolution, begeleid door KIXS, waarbij als use case is gekozen voor het landgebonden optreden omdat dit voor de technologie veel uitdagingen biedt op het gebied van flexibiliteit en schaalbaarheid:

- Er is over het algemeen weinig tot geen bandbreedte over om data te kunnen versturen via de organieke transmissiesystemen zoals VHF radio;
- De wapensystemen bij CLAS hebben beperkt ruimte en stroom voor extra apparatuur;
- Er worden per inzet verschillende samenstellingen van eenheden en wapensystemen gebruikt (diverse configuraties).

CyLoM bestaat kortweg uit twee typen bouwblokken: data-extractie en data-analyse (Figuur 1). Met deze blokken zal Defensie in staat zijn om zelf te bepalen waar een cyberanalyse uitgevoerd wordt: lokaal bij de gebruiker van een wapensysteem, op de commandopost bij een analist, in Nederland bij DCSC of zelfs op alle locaties indien gewenst. Zo kan De-



Figuur 1: de 'bouwblokken' van CyLoM. Bron: TNO en Technolution

fensie bijvoorbeeld kiezen om een *intrusion detection system* aan boord van het wapensysteem te plaatsen, waarbij voor de gebruiker alleen die meldingen worden gegenereerd die te maken hebben met de directe inzetbaarheid van het wapensysteem (real-time analyse) en de grondige analyses (non real-time) over meerdere wapensystemen uit te laten voeren door experts bij DCSC.

De extractie van data uit een netwerk zal altijd plaatsvinden via datadiodes zodat gegarandeerd kan worden dat CyLoM de netwerken niet verstoort maar vooral ook om te garanderen dat CyLoM niet uiteindelijk zelf een cyberdreiging kan worden. Daar waar de extractie en analyse gescheiden zijn van elkaar zal er ook altijd een datadiode tussen zitten.
Op dit moment is het project zover dat de bouwblokken in een experimenteeromgeving met fictieve datastromen zijn getoetst. Die testen waren succesvol. De volgende stap is dat de technologie getoetst wordt op een echt wapensysteem. Voor deze toetsing is gekozen voor de Boxer (Figuur 2).

Vanuit het experiment *Forward Command Post* kreeg het CyLom de mogelijkheid om tien CP-Boxers eind dit jaar uit te rusten met de CyLoM bouwblokken. Deze Boxers zullen ingezet worden voor de oefening Zebra Sword in november 2020. Op dit moment vinden alle voorbereidingen hiervoor plaats met Rheinmetall, DMO Matlog, JIVC Landgebonden IT, JIVC GIT & INFRA en wordt er nagedacht over de nodige ontwerpaanpassingen om CyLoM te laten

Figuur 2: Boxer

passen, ervoor te zorgen dat alles ruggedized genoeg is om een heel jaar te kunnen blijven zitten en worden maatregelen genomen om met gerubriceerde data om te kunnen gaan. Omdat er nog veel onzekerheden overblijven, zoals bijvoorbeeld hoeveel netwerkverkeer er per tijdseenheid over deze netwerken gaat, zal de plaatsing van CyLoM in verschillende iteraties plaatsvinden, waarbij er twee hardware-iteraties plaatsvinden en vier software-iteraties op eerst een enkele Boxer en dan uitgebreid naar tien Boxers. CyLoM zal het netwerkverkeer gaan loggen van zowel het CAN-netwerk als het Ethernet (BMS & TITAAN). Gaat alles goed, dan weten we eind 2020 in hoeverre CyLoM goed genoeg is voor gebruik door Defensie.

Naast een stuk doorontwikkeling levert de plaatsing van CyLoM aan boord van tien Boxers nog een groot voordeel op en dat is een hele grote database van ruwe netwerkdata. De opbouw van CyLoM is in principe zo generiek dat deze niet alleen gebruikt kan worden voor cyberanalyses, maar ook voor analyses voor het onderhoud van netwerken, materieelonderhoud, normstelling, inzet, etc. De gegenereerde data zal in een secure omgeving van JIVC KIXS worden opgeslagen en ter beschikking worden gesteld voor eenieder bij Defensie die analyses wil gaan uitvoeren. Daarbij biedt JIVC KIXS tevens de mogelijkheid om te ondersteunen bij dit proces door een pool van in totaal twintig data-analisten en –beheerders beschikbaar te stellen voor Defensie.

Afsluiting

De inzet van Big Data is voor Defensie een relatief nieuwe ontwikkeling, zeker in het mobiele domein van de landmacht. Defensie gaat in samenwerking met twee marktpartijen een experiment uitvoeren om meer te leren over de Big Data toepassing logging & monitoring in mobiele wapensystemen.
Voor meer informatie over (big) data en data analyse en wat hier zoal mee mogelijk is kun je terecht bij KIXS: kixs@mindef.nl