



# HET VOORSPELLEN VAN DE DREIGING VAN EEN DRONE MET BEHULP VAN EEN BAYESIAANS NETWERK

Stelt u zich voor, de COVID-pandemie is voorbij en u bezoekt een openluchtfestival. Plots hoort u een zoemend geluid boven u. U kijkt omhoog en ziet een drone. U besluit de politie in te schakelen. Die komt en moet ter plaatse beslissen of en hoe er actie ondernomen moet worden. Om de juiste actie te bepalen, is het van belang dat de intentie van de drone wordt achterhaald: gaat de drone een aanslag plegen of maakt de drone slechts foto's van het evenement? Voor mijn werkstage bij TNO heb ik een Bayesiaans netwerk opgesteld, dat gegeven de observaties die gemaakt worden van de drone en zijn omgeving, de meest waarschijnlijke dreiging van de drone kan voorspellen.

LAURA MIDDELDORP

Een drone is een onbemand luchtvaartuig. Hij kan zich voortbewegen met behulp van vleugels, propellers of straalmotoren. Drones maken steeds vaker deel uit van het straatbeeld, mede omdat de modellen die voor de consument zijn bedoeld almaar betaalbaarder worden.

Een drone geraakt mogelijk in handen van een individu of groep met minder goede intenties die kunnen variëren van privacyschendingen tot gerichte aanslagen. Dit vormt een uitdaging op het gebied van veiligheid. Het is daarom cruciaal dat veiligheids- en beveiligingsdiensten

zoals de politie zich wapenen tegen dronedreigingen. De meest waarschijnlijke dreiging kan achterhaald worden met behulp van een Bayesiaans netwerk.

## Wat is een Bayesiaans netwerk?

Een Bayesiaans netwerk is een gerichte acyclische graaf, waarin de nodes stochastische variabelen beschrijven en de pijlen de causale relaties tussen de stochasten weergeven. Een simpel voorbeeld van een Bayesiaans netwerk ziet u in figuur 1. Die laat zien dat node A en B een causale relatie hebben. Dit betekent dat de kansverdeling van node B afhangt van de uitkomst van node A.

Een Bayesiaans netwerk werkt aan de hand van de regel van Bayes: gegeven de a priori kansen van node A en de voorwaardelijke kansen van node B gegeven node A kunnen de a posteriori kansen van node A gegeven de uitkomst van de node B worden achterhaald. Stel dat node A en B ieder twee mogelijke uitkomsten kunnen aannemen: Ja en Nee. Als we erachter komen dat node B gelijk is aan Ja, dan kan de kans dat node A gelijk is aan Ja gegeven dat node B gelijk is aan Ja als volgt bepaald worden:

$$P(A = Ja|B = Ja) = \frac{P(B = Ja|A = Ja)P(A = Ja)}{P(B = Ja)}$$

Een Bayesiaans netwerk is in staat om de kansverdeling van nodes te updaten gegeven geobserveerde informatie.

## Hoe is een Bayesiaans netwerk toe te passen in deze context?

In het Bayesiaanse netwerk voor dronedreigingen worden de mogelijke dreigingen gemodelleerd door hypothesen. De waarnemingen die gemaakt kunnen worden van de drone en zijn omgeving worden beschreven door indica-

toren. Een hypothese en een indicator hebben een causale relatie als het optreden van de hypothese het waarschijnlijk maakt dat de indicator optreedt.

Een mogelijke dreiging is bijvoorbeeld:

*De drone gaat een bomaanslag plegen.*

Een mogelijke indicator is:

*Er hangt een pakketje onder de drone.*

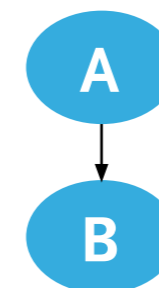
Deze hypothese en indicator hebben een causale relatie:

*gegeven dat de drone een bomaanslag gaat plegen, is het waarschijnlijk dat er een pakketje onder de drone hangt, waarbij het pakketje mogelijk de bom bevat.*

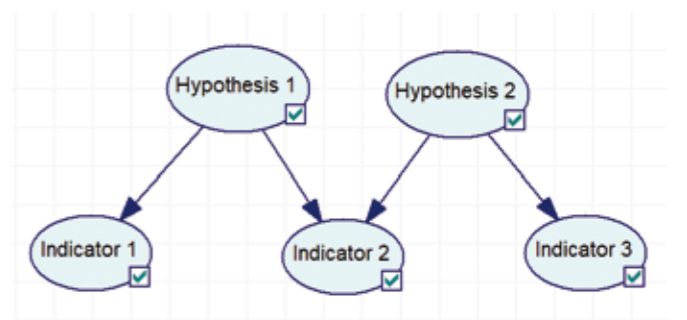
Figuur 2 laat een voorbeeld zien van het model bestaande uit 2 hypothesen en 3 indicatoren. Gegeven de a priori kansen op de hypothesen, de voorwaardelijke kansen op de indicatoren gegeven de hypothesen en de geobserveerde indicatoren kunnen de a posteriori kansen op de hypothesen gegeven de indicatoren worden berekend. De meest waarschijnlijke dreiging is de hypothese met de grootste a posteriori kans gegeven de geobserveerde indicatoren.

## Het opstellen van het Bayesiaanse netwerk

Om het Bayesiaanse netwerk op te kunnen stellen, moet er eerst een scenario worden gedefinieerd. Het scenario is de omgeving waarin de drone zich bevindt. De hypothesen en indicatoren in het Bayesiaanse netwerk hangen namelijk af van het scenario: als een drone in een oorlogsgebied vliegt in Afghanistan zijn de mogelijke dreigingen anders dan wanneer een drone in een drukke stad in Nederland wordt geobserveerd. Voor het scenario is gekozen voor een druk evenement in de stad waarbij er geen strenge toegangscontrole is.



Figuur 1. Een simpel voorbeeld van een Bayesiaans netwerk bestaande uit 2 nodes A en B



Figuur 2. Een voorbeeld van een Bayesiaans netwerk bestaande uit 2 hypothesen en 3 indicatoren

De hypothesen en indicatoren behorende bij dit scenario zijn in samenwerking met experts binnen de politie en defensie bepaald. In het uiteindelijke Bayesiaanse netwerk zijn 8 hypothesen en 23 indicatoren gedefinieerd, waarbij iedere indicator twee mogelijke uitkomsten heeft: Ja en Nee.

Om de meest waarschijnlijke dreiging gegeven de observaties van de drone te achterhalen, zijn de kansen op de hypothesen en de kansen op de indicatoren gegeven de hypothesen benodigd. Gezien er onvoldoende data beschikbaar waren om deze kansen in te schatten, is er gebruik gemaakt van meningen van experts om de kansen te bepalen. Hiervoor is een expert-sessie georganiseerd waaraan experts binnen zowel politie als defensie hebben deelgenomen.

### Robuustheid van het resultaat van het Bayesiaanse netwerk

Een nadeel aan het gebruik van meningen van experts voor het inschatten van de kansen is dat het onzekerheid met zich meebrengt. Als een andere groep experts zou zijn ingeschakeld om de kansen in te schatten, zouden de kansen in het Bayesiaanse netwerk mogelijk anders zijn waardoor het resultaat zou verschillen. Dit brengt ons op het vraagstuk van robuustheid: hoe zeker is het dat de meest waarschijnlijke dreiging volgens het Bayesiaanse netwerk ook daadwerkelijk de meest waarschijnlijke is?

Een van de mogelijke methoden die gebruikt kan worden om de robuustheid van het resultaat te bepalen is de methode van rangverschuivingen (Kipersztok & Wang, 2001). Deze methode is gebaseerd op een Monte Carlo simulatie. In iedere simulatie wordt er noise toegevoegd aan de kansen in het Bayesiaanse netwerk waarna de a posteriori kansen van de hypothesen berekend worden. Vervolgens wordt de rangorde van de hypothesen in de simulatie vergeleken met de originele rangorde. Op deze wijze wordt het aantal rangverschuivingen dat optreedt met de hypothesen inzichtelijk en wordt er meer duidelijkheid verkregen over de robuustheid van het resultaat.

Net als in Kipersztok & Wang (2001) wordt er aan de voorwaardelijke kansen noise toegevoegd uit de log-odds

normale verdeling. Hiervoor wordt de voorwaardelijke kans eerst getransformeerd naar de log-odds waarna er een noise component uit de normale verdeling wordt toegevoegd:

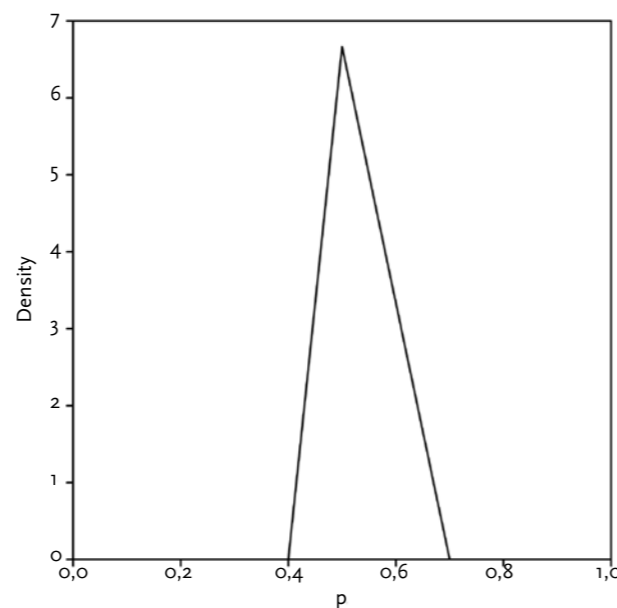
$$Y = \log_{10}\left(\frac{p}{1-p}\right) + \epsilon, \quad \epsilon \sim N(0, \sigma)$$

Vervolgens wordt de verstoorde voorwaardelijke kans verkregen door de inverse te nemen van bovenstaande:

$$\tilde{p} = \frac{10^Y}{1 + 10^Y}$$

Voor iedere voorwaardelijke kans moet er een standaarddeviatie worden gekozen. Hiervoor is gebruik gemaakt van de resultaten van de expert-sessie: hoe meer de experts het eens waren over de waarde van de voorwaardelijke kans, hoe kleiner de standaarddeviatie van de normale verdeling.

Naast de voorwaardelijke kansen moet er ook noise worden toegevoegd aan de a priori kansen van het Bayesiaanse netwerk. Aanvankelijk wilde ik ook de log-odds normale verdeling gebruiken voor de a priori kansen, maar dit gaf enkele problemen. Uiteindelijk bleek de driehoeksverdeling de meest geschikte verdeling om noise toe te voegen aan de a priori kansen. Figuur 3 toont een voorbeeld van een driehoeksverdeling.



Figuur 3. Plot van de dichtheid van een driehoeksverdeling.

Voor iedere a priori kans is het interval waarop de driehoeksverdeling positieve massa heeft bepaald aan de hand van de expert-sessie: voor iedere hypothese zijn de kleinste en grootste a priori kans ingeschat door de experts gebruikt als grenzen voor het interval. De mode van

de driehoeksverdeling is gelijk aan de a priori kans van de hypothese.

De robuustheid van het resultaat van het Bayesiaanse netwerk wordt bepaald door 2 dingen: de kansen in het Bayesiaanse netwerk en de mate waarin de experts het met

<b>CATEGORIE 1</b>		
I1	<input checked="" type="checkbox"/> Ja	<input type="checkbox"/> Nee
I2	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
I3	<input type="checkbox"/> Ja	<input checked="" type="checkbox"/> Nee
I4	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
<b>CATEGORIE 2</b>		
I5	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
I6	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
<b>CATEGORIE 3</b>		
I7	<input checked="" type="checkbox"/> Ja	<input type="checkbox"/> Nee
I8	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
I9	<input type="checkbox"/> Ja	<input checked="" type="checkbox"/> Nee
I10	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
I11	<input checked="" type="checkbox"/> Ja	<input type="checkbox"/> Nee
I12	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
I13	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
<b>CATEGORIE 4</b>		
I14	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
I15	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee
I16	<input type="checkbox"/> Ja	<input type="checkbox"/> Nee

**Instructie voor het gebruik van de tool:**

Voor iedere geobserveerde indicator kunt u de uitkomst aanvinken in de checkboxes. Om een checkbox aan te vinken, is het voldoende om met uw muis te klikken op de cel waarin de checkbox staat. Indien de uitkomst van een indicator onbekend is, kunt u de checkboxes behorende bij deze indicator leeg laten.

Als u alle geobserveerde indicatoren heeft aangevinkt, kunt u op een van de onderstaande knoppen drukken. De knop 'Bereken het resultaat' geeft u alleen de kansen op de hypothesen. Wanneer u op de knop 'Bereken de robuustheid van het resultaat' klikt, krijgt u, naast de kansen op de hypothesen, ook de robuustheid van de uitkomst te zien.

Wanneer u op een van de twee knoppen drukt, ziet u een scherm verschijnen wat betekent dat het programma aan het runnen is. Als het scherm verdwijnt kunt u de resultaten aflezen in de sheet genaamd 'Resultaat'.

In het geval dat u extra indicatoren observeert, kunt u deze toevoegen door de desbetreffende states aan te vinken en opnieuw op een van de twee knoppen drukken.

Het runnen van de robuustheidsanalyse duurt ongeveer 25 à 30 seconden. In het geval dat een snel resultaat gewenst is, is het aangeraden om op de 'Bereken het resultaat' knop te drukken

BEREKEN HET RESULTAAT
BEREKEN DE ROBUUSTHEID VAN HET RESULTAAT

Figuur 4. Prototype van de tool voor het voorspellen van drone dreigingen

	Kans(%)	1	2	3	4	5	6	7	8
H5	62,6	80,6	16,6	2,7	0,1	0	0	0	0
H1	23,6	17,7	67,5	13,3	1,4	0,1	0	0	0
H2	8,8	1,7	13,8	59,6	21,8	2,7	0,4	0	0
H6	3,4	0	2	22,6	56,6	14,1	4,3	0,4	0
H7	1	0	0,1	1,7	12,6	50,9	29,9	4,2	0,6
H8	0,6	0	0	0,1	7,5	29,7	53	8,4	1,3
H3	0	0	0	0	0	0,5	3,5	42,4	53,6
H4	0	0	0	0	0	2	8,9	44,6	44,5

Figuur 5. Output van het prototype inclusief robuustheid van het resultaat

elkaar eens zijn over de waarden van de kansen. De kansen in het Bayesiaanse netwerk bepalen de a posteriori kansen. Hoe dichter de a posteriori kansen bij elkaar liggen, hoe groter de kans dat er een rangverschuiving optreedt als er noise wordt toegevoegd aan de kansen. Verder geldt dat de noise die wordt toegevoegd aan de kansen groter is als de experts het minder met elkaar eens zijn. Hierdoor treedt er meer variatie op in de a priori en voorwaardelijke kansen en als gevolg de a posteriori kansen waardoor er meer rangverschuivingen zullen optreden.

### Het communiceren van het resultaat van het Bayesiaanse netwerk naar de gebruiker

Het uiteindelijke doel van het Bayesiaanse netwerk is om de politie te ondersteunen bij de keus welke actie er ondernomen moet worden tegen de drone. Om de resultaten van het Bayesiaanse netwerk te communiceren naar de politie toe, is er een prototype van de tool ontwikkeld in Excel. Dit prototype is een eerste opzet voor de interface die de politie mogelijkwijs in de toekomst kan gaan gebruiken. Figuur 4 toont het prototype.

In figuur 5 is te zien dat de meest waarschijnlijke hypothese H5 in 80,6 procent van de gevallen ook de meest waarschijnlijke hypothese was in de Monte Carlo simulatie. Dit impliceert dat het resultaat redelijk robuust is.

### Conclusie

Een Bayesiaans netwerk biedt een goede manier om de dreiging van een drone te voorspellen. Een nadeel echter is dat het gebruik maken van mening van experts onzekerheid veroorzaakt in het resultaat van het netwerk. De robuustheid kan worden vergroot als er meer zekerheid is over de waarden van de kansen. Hierbij kan een drone incident database uitkomst bieden: gegeven voldoende data kunnen de a priori en voorwaardelijke kansen accurater worden ingeschat. Een ander nadeel is dat het huidige Bayesiaanse netwerk opgesteld is rondom een scenario. Als er een grote verandering optreedt in het scenario, moeten er nieuwe hypothesen en indicatoren gedefinieerd worden als ook de kansen opnieuw ingeschat worden wat erg veel werk is. Meer onderzoek naar het efficiënt uitbreiden van het Bayesiaanse netwerk voor meerdere scenario's is gewenst.

### LITERATUUR

Kipersztok, O., & Wang, H. (2001). Another look at sensitivity of Bayesian networks to imprecise probabilities. In *Proceedings of Machine Learning Research*, R3 (pp.149–155)

LAURA MIDDELDORP is masterstudent Applied Mathematics aan de Technische Universiteit Delft. Tijdens haar stage bij de afdeling Military Operations van TNO heeft ze een Bayesiaans netwerk opgesteld. E-mail: lmiddeldorp@xs4all.nl



Sven Kramer tijdens de Olympische Spelen in 2010. Foto: David Rosen CC

# DE MATRIX VAN OLYMPISCH SCHAATSGOUD

GERARD SIERKSMA

De hoge prestatiedichtheid aan de top van veel traditionele sporten heeft geleid tot een aantal interessante uitdagingen betreffende onder meer het vergelijken van prestaties van atleten en het beslissen over winnaars. Zo liggen de onderlinge verschillen steeds vaker binnen de foutmarges van de meetsystemen en als het dan gaat om 'goud' is het in die gevallen onmogelijk de winnaar te bepalen. Die uitdagingen betreffen niet alleen het vergelijken van wedstrijdresultaten, ook het bepalen van eenduidige selectieprocessen, waar wordt vastgesteld op welke wijze atleten worden geselecteerd voor belangrijke toernooien, is, gegeven die grote prestatiedichtheid aan de top, een tamelijk precaire aangelegenheid. In het geval van het Nederlandse schaatsen is de pool van competitieve topatleten groot en zullen potentiële medaillewinnaars

niet altijd kunnen worden geselecteerd. In het geval van de Olympische Winterspelen gelden strikte quota, waardoor slechts een beperkt aantal atleten het eigen land mag vertegenwoordigen. Geen wonder dat in zulke gevallen de selectieprocessen, die in het verleden werden uitgevoerd door goedwillende experts, met argusogen werden gevolgd en de beslissingen juridisch werden aangevochten.

### De top-10 ambitie

In aanloop naar de Olympische Winterspelen 2010 in Vancouver werden in de media de selectieperikelen van de Koninklijke Nederlandse Schaatsenrijdersbond