

Predictive Policing – wens of werkelijkheid?

drs. Dick Willems is data miner bij de RIO in Amsterdam en ontwerper van CAS. ir. Reinder Doelman is chef van de RIO in Amsterdam en landelijk portefeuillehouder Business Intelligence.

In de film Minority Report uit 2002 van Steven Spielberg komt aan de orde hoe de politie in de toekomst criminaliteit voorspelt. Vlak voordat de woedende echtgenoot met een schaar zijn vreemdgaande vrouw doodsteekt, valt een arrestatieteam binnen en houdt de aanstaande moordenaar aan. Het misdrijf wordt voorkomen en de bijna-dader wordt achter slot en grendel gezet.



http://www.dvdbeaver.com/film3/blu-ray_reviews51/minority_report_blu-ray.htm

» *Ca. 40% van de woninginbraken en 60% van de straatroven in Amsterdam kunnen worden voorspeld*

De gebruikte techniek in het voorbeeld is vooralsnog niet erg realistisch: Tom Cruise gebruikt voor de voorspelling beelden uit de hersenen van drie op elkaar aangesloten helderzienden die samen in een bad liggen (zie afbeelding). Maar de film bevat ook meer weldoordachte technieken die de komende jaren zo maar eens werkelijkheid zouden kunnen worden. Het voorspellen van criminaliteit, in het Engels aangeduid met de term *predictive policing*, ligt dichterbij handbereik dan velen zich zullen realiseren.

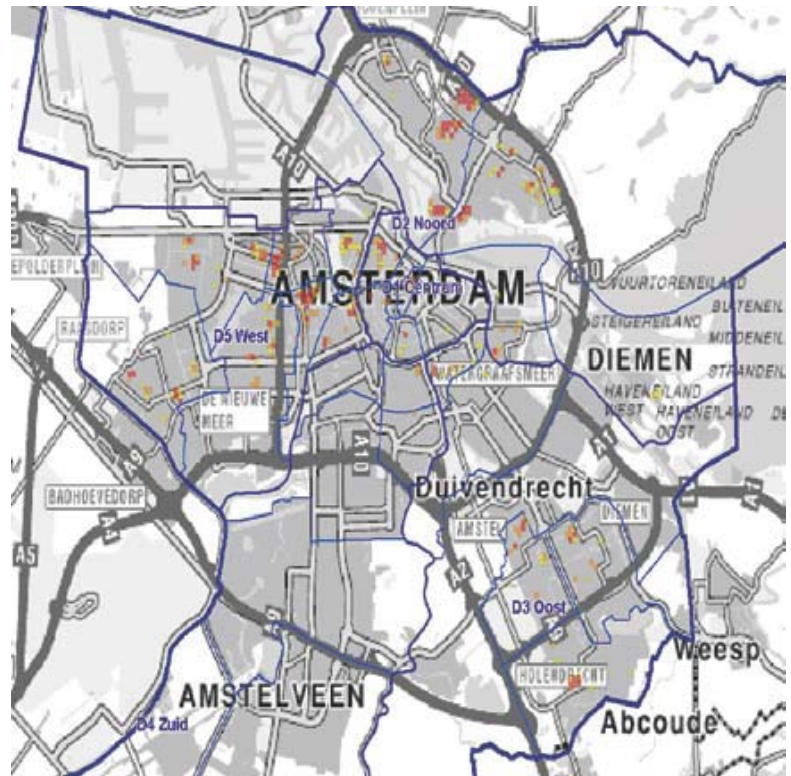
Capaciteitsmanagement

In tegenstelling tot het hierboven genoemde voorbeeld, gaat het voorspellen van criminaliteit niet zozeer over het indivi-

dule misdrijf. Eén van de uitdagingen waar de politie als organisatie immers voor staat, is de capaciteit die zij tot haar beschikking heeft zo goed mogelijk te gebruiken. Om dat te doen, is ten eerste informatie nodig. Waar moeten de agenten lopen en wanneer, en wat kunnen ze tijdens de dienst verwachten? Met dit soort informatie kan een schatting worden gemaakt hoeveel collega's er nodig zijn om een dienst te lopen en waar de aandacht dan op gericht zou moeten zijn. Ten tweede moet deze informatie op tijd beschikbaar zijn. Daarnaast is het ook belangrijk dat er geprioriteerd wordt. Er is nu eenmaal veel werk voor de politie, maar voor bepaalde gebeurtenissen wordt redelijkerwijs een grotere inspanning verricht dan voor andere. Het komt de veiligheid in een eenheid meer ten goede als een inbrekersbende wordt opgepakt dan als iemand een bon krijgt voor rijden zonder licht.

Predictive policing

Het bepalen wáár en wanneer politieinzet er toe doet is het terrein van het genoemde predictive policing. Predictive policing behelst in het kort het gebruik van statistische voorspellingen om te kunnen anticiperen op criminele incidenten. Deze informatie kan, afhankelijk van de acties die erop worden gebaseerd, worden gebruikt ter preventie van misdaad of het vergroten van de heterdaadkracht. Een voorwaarde voor het succesvol kunnen inzetten van predictive policing is dat gegevens over criminele activiteiten ontsloten zijn in een datawarehouse, en dat andere data hieraan gekoppeld kunnen worden. Veel gegevens die potentieel relevant zijn voor criminaliteit zijn immers niet opgenomen in een systeem als BVH. Zo is het bijvoorbeeld redelijk te veronderstellen dat gegevens over vluchtroutes



relevant zijn voor ram- en plofkraak, maar om dit aan te tonen moet van elke locatie binnen BVH bekend worden gemaakt hoe snel men bijvoorbeeld met een auto of scooter een bepaalde afstand kan overbruggen. Daarnaast kunnen allerlei gegevens over de bevolkingssamenstelling of type bedrijven in een gebied relevante indicatoren zijn voor patronen van criminaliteit.

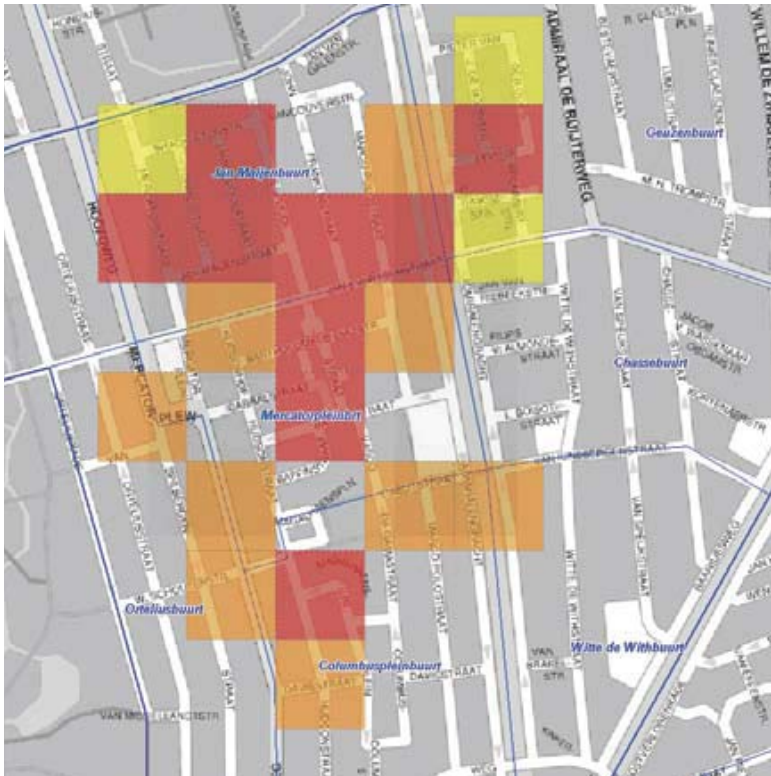
Er is ook een manier nodig om de relevante patronen in de gegevens te ontdekken. De hoeveelheid gegevens is potentieel zó immens dat de kans klein is dat een menselijke analist er alle nuttige informatie uit zal kunnen halen. Voor dit soort taken zijn computers sneller en nauwkeuriger. Het geautomatiseerd ontdekken van relevante patronen in grote hoeveelheden data is bij uitstek het terrein van het vakgebied dat datamining heet. Binnen het repertoire van datamining valt een grote verscheidenheid aan methoden om deze patroonherkenning uit te voeren.

Concrete voorbeelden

Een van de bekendste voorbeelden van predictive policing is het inzetten van het softwarepakket PredPol door het Los Angeles Police Department, om locaties te ontdekken waar het risico op misdaad (woninginbraken, autokraken en autodiefstal) hoger is. Vervolgens krijgen deze locaties op de diensten meer aandacht, en er worden indrukwekkende resultaten gerapporteerd.

Ook binnen Nederland wordt predictive policing ingezet. De Eenheid Amsterdam heeft zelfstandig een systeem ontwikkeld om misdaad mee te voorspellen. Dit Criminaliteits Anticipatie Systeem (CAS) wordt vooralsnog toegepast

» *Los van de effectiviteit van de inzet creëert CAS ook duidelijkheid*



om High Impact Crimes mee te voorspellen (woninginbraak, straatroof en overval). In Amsterdam wordt CAS ook gebruikt om inzet op te baseren. Het systeem is leidend in de planning voor de flexteams, die over de hele eenheid ingezet worden, en is in de verschillende districten van de eenheid een belangrijke informatiebron voor het inzetadvies. Dit systeem zal hieronder nader worden omschreven.

CAS

Het Criminaliteits Anticipatie Systeem deelt de eenheid Amsterdam op in vakjes van 125 bij 125 meter. Gebiedjes waarvan de kans op een incident vooraf al laag kan worden ingeschat, zoals weilanden en open water, worden verwijderd. Van de overblijvende vakjes wordt een grote hoeveelheid gegevens verzameld: criminaliteitshistorie, afstand tot bekende verdachten, afstand tot de dichtstbijzijnde snelwegoprit, soort en aantal bedrijven bekend bij de politie, en daarnaast ook demografische en socio-economische gegevens via het CBS.

Van ieder vakje wordt op verschillende peilmomenten geregistreerd welke gegevens er op dat moment bekend zijn. Vervolgens wordt vastgelegd wat er in de twee weken na de peiling aan incidenten heeft plaatsgevonden. Per vakje wordt drie jaar historie gemeten, onderscheiden in tweewekelijkse peilmomenten. Dit resulteert in 76 peilingen per gebiedje. De totale set aan gegevens is erg groot te noemen, te groot voor een mens om te overzien.

Om een uitspraak te doen over welke combinaties van kenmerken indicatief zijn voor criminaliteit in de nabije toekomst, wordt gebruikt gemaakt van kunstmatige neurale

netwerken. Dat zijn algoritmes uit de kunstmatige intelligentie, die in staat zijn om patronen te leren herkennen. Dit gebeurt analoog aan hoe in de wetenschap wordt verondersteld dat mensen dit doen. Het is dus illustratief om eerst iets uit te leggen over de werking van de menselijke hersenen.

Kennis in het netwerk

Zoals bekend zijn de hersenen verantwoordelijk voor het verwerken van informatie. De hersenen bestaan voor een belangrijk deel uit *zenuwcellen*, neuronen genaamd. Deze neuronen bestaan grofweg uit een cel-lichaam (het *soma*), een *axon* en een aantal *dendrieten*. Een axon is een uitloper van een neuron waarlangs informatie wordt verzonden. Dendrieten zijn uitlopers van neuronen waarlangs informatie wordt ontvangen. Via axonen en dendrieten kunnen neuronen door middel van elektrische signalen met elkaar informatie uitwisselen. Globaal is het zo dat een neuron een signaal afvuurt via het axon als het voldoende geactiveerd is via de dendrieten.

Als een mens een nieuwe taak leert, is het niet zo dat de hersenen nieuwe neuronen aanmaken om deze taak op zich te nemen, maar dat de reeds bestaande neuronen zich op een andere manier gaan organiseren; er worden nieuwe verbindingen aangemaakt, oudere verbindingen worden verbroken of veranderd. De nieuwe kennis wordt dus niet opgeslagen in de afzonderlijke neuronen, maar wordt gerepresenteerd door de manier waarop deze zijn georganiseerd: de kennis zit in het netwerk.

Dit proces kan worden gesimuleerd met een kunstmatig neuraal netwerk, dat op een computer draait. Hierbij worden virtuele neuronen aan elkaar gekoppeld en deze koppelingen worden stapsgewijs geoptimaliseerd. In iedere stap wordt informatie van een vakje die bekend is op een bepaald peilmoment aan het netwerk aangeboden. Het netwerk maakt met deze informatie een voorspelling, waarna het wordt geconfronteerd met de feiten: hebben er daadwerkelijk incidenten plaatsgevonden in het vakje in de twee weken na het peilmoment?

Vervolgens bepaalt het netwerk waar de grootste afwijkingen zaten, welke koppelingen hiervoor verantwoordelijk waren en hoe deze bijgesteld moeten worden om het een volgende keer beter te doen. Dit proces neemt enige tijd in beslag en gaat door totdat het wijzigen van de koppelingen in het netwerk niet meer resulteren in verbetering van de voorspelling.

Dit netwerk kan nu gebruikt worden om iets te zeggen over de waarschijnlijkheid van incidenten in de toekomst met behulp van informatie over de huidige stand van zaken. Aan ieder vakje in Amsterdam kan zo'n risico-score worden toegekend. Deze scores kunnen worden gebruikt om *heat maps* mee te maken: hogere scores krijgen een 'warmere' kleur, lagere scores een 'koudere'. Aangezien het systeem tot doel heeft om kleine besurveillanceerbare gebieden met een hoog risico op incidenten aan te wijzen, is ervoor gekozen slechts de top-3% van de vakjes te kleuren. Het hoogste

percentiel wordt hierbij rood gekleurd, het daaropvolgende oranje en het daaropvolgende wordt geel.

Als deze hoog-risicolocaties zijn geselecteerd wordt op een vergelijkbare manier bepaald wanneer het risico op een incident het grootst is. Zo kan op basis van de verwachting van CAS een rooster worden opgesteld, zodat inzet gepleegd kan worden waar en wanneer het ertoe doet. Op deze manier kan ongeveer 40% van de woninginbraken en 60% van de straatroven in Amsterdam worden voorspeld. Niet gek voor een oppervlak van slechts 3%.

De praktijk

In de eenheid Amsterdam wordt de roostering en briefing van de flexteams gedaan op basis van de adviezen van CAS. Iedere twee weken worden de CAS-kaarten automatisch ververs en wordt door een team analisten van de DRIO een inzetadvies vastgesteld. In dit advies wordt verwerkt waar men wanneer moet zijn, maar ook wat men op die tijdstippen kan verwachten en op wie gelet moet worden. Het is immers zo dat er veel informatie bekend is tussen de oren van de collega's op straat, en de auteurs van de inzetadviezen proberen die zo goed mogelijk vast te leggen en te communiceren zodat ook deze kennis daar kan worden toegepast waar het ertoe doet.

De resultaten zijn goed. In gebieden waar is ingezet is te zien dat er minder incidenten hebben plaatsgevonden, en ook tijdens het donkere-dagenoffensief van 2013/2014 is een duidelijke daling van het aantal woninginbraken geconstateerd. Uiteraard kan dit succes niet alleen maar aan CAS worden toegeschreven, de Eenheid Amsterdam heeft veel meer inspanningen geleverd om het aantal woningbraken terug te dringen, maar dit is niet los te zien van de inzet van CAS.

Behalve effectiviteit is een gevolg van het gebruik van CAS ook dat het duidelijkheid schept: inroostering wordt gedaan op basis van informatie. Collega's wordt geadviseerd om meer aandacht te schenken aan bepaalde gebieden dan aan andere, omdat de berekende kans dat daar incidenten plaats gaan vinden nu eenmaal hoger is.

Uiteraard is het wel zo dat CAS een statistische verwachting weergeeft. Het is geen kristallen bol en moet niet als zoda-

nig worden geïnterpreteerd. Als men de beschikking heeft over specifieke en relevante informatie moet deze zeker meespelen in het inzetadvies.

Toekomstbeeld

Momenteel is het nog zo dat CAS om de twee weken wordt ververs. De kaarten zijn dus gemaakt om twee weken lang een zo goed mogelijke verwachting af te geven. Verder wordt CAS aan de gebruiker beschikbaar gesteld met behulp van specifieke kaartensoftware die binnen een werkomgeving beschikbaar is. De gebruiker hoeft niet bekend te zijn met het modelleerproces om de kaarten te kunnen gebruiken.

Ideaal zou natuurlijk zijn als CAS *real time* beschikbaar was. Niet alleen kan dan worden gewerkt met actuele gegevens, maar het is dan ook zo dat andersoortige gegevens kunnen worden gebruikt. Zo zouden weersinvloeden bij de voorspelling kunnen worden betrokken, en incidentele gebeurtenissen zoals uitslagen van voetbalwedstrijden. Het zou ook zeer fraai zijn als CAS beschikbaar was op mobiele apparatuur, zodat collega's te allen tijde de kaarten zouden kunnen raadplegen, en zodat zij ter plekke de kaarten zouden kunnen bewerken en verrijken. Om dit te realiseren zullen een flink aantal hordes genomen moeten worden, maar het is duidelijk dat dit de moeite zal lonen.

Predictive policing landelijk?

Op dit moment worden predictive policing-systemen zoals CAS alleen nog gebruikt in Amsterdam. De potentie voor de nationale politie is natuurlijk enorm en deze ontwikkeling past ook prima binnen de strategie van Business Intelligence en Big Data. Het centraal bijeenbrengen van grote hoeveelheden data en hierop technieken als CAS loslaten kan de nationale politie patronen en inzichten geven die tot dan toe nog niet mogelijk waren. Hoewel voor het bouwen en uitrollen van de systemen wel specifieke competenties vereist zijn, is na implementatie geen kennis van exotische algoritmes meer vereist. Wel is het nodig dat specifieke software en hardware voorhanden is om de benodigde berekeningen uit te voeren.

In theorie zou dus iedere eenheid gebruik kunnen maken van predictive policing, mits de dataverzameling op orde is en er wordt geïnvesteerd in een werkproces binnen de DRIO om de verwachting op de CAS-kaarten ook daadwerkelijk te operationaliseren. Deze laatste voorwaarde is essentieel: effectieve inzet van predictive policing kan niet vrijblijvend zijn. Het doel van predictive policing is om relevante en bruikbare informatie te verschaffen om de inzet op straat te optimaliseren. Maar het interpreteren van deze informatie en het formuleren van geschikte acties is noodzakelijkerwijs mensenwerk, en hiervoor moet capaciteit worden vrijgemaakt en er zal gehandeld moeten worden naar de inzetadviezen. Hoe graag we ook willen, het is nu eenmaal niet mogelijk de boeven de gevangenis in te analyseren. <<

»» *Effectieve inzet van predictive policing kan niet vrijblijvend zijn*