

Trendsetters

A data mining and crime mapping tool developed with three Dutch police forces is helping officers to predict trends and identify the most likely suspects in the Netherlands. **Gary Mason** reports

A crime mapping and data mining tool that combines several automated techniques is helping Dutch police officers to find possible suspects for a crime, or a series of crimes. The key to the system's success is its simplicity, according to the three police forces involved in developing the system. With little training, officers can quickly discover patterns and trends, make forecasts, map criminal networks and identify suspects.

Rob van der Veer from Sentient Information Systems BV, which developed the system with Amsterdam Police, West Brabant Police and North Brabant Police, says Dutch forces had already used data mining systems in the past, but these were limited by the high level of expertise needed to operate the systems.

'Expert knowledge on data mining or statistics is not required [with the new system],' he says. 'A much larger audience for data mining results is reached through weekly reports containing statistics, prediction maps, crime clusters, trends and lists of suspects. These reports are automatically produced by the data mining system.'

Automatic matches

The software suite combines automated techniques to help find suspects for a crime or a series of crimes. In this 'multi-aspect suspect matching', incidents are clustered together based on spatial and contextual elements, such as modus operandi and offender description.

The method applies several techniques to find possible suspects, looking at aspects such as address, career, spatial activity and description. It helps officers create a geographical profile and match it to the addresses of possible suspects. It also finds approximate or 'fuzzy' matches to discover similar cases using spatial and contextual data.

The system also uses spatial analysis to match activity areas of possible suspects with the crimes. These can, for example, be based on automatic number plate recognition, which will identify whether a suspect's car has been detected in an area of interest. It also applies automated fuzzy matching to witness

descriptions of possible suspects, resulting in a match-score between the series and personal characteristics of each suspect. According to Van der Veer, research has shown a 50 per cent gain in hit rate for this fuzzy technique as opposed to conventional techniques.

The combined outcome of the analysis is a list of suspects, each with a calculated match score based on all of the data mining techniques. The highest scores are the most likely suspects and statistically represent the typical offender for the crime series. Therefore, the typical characteristics of this group can be used as a 'profile' of the unknown offender.

Dutch police systems have traditionally focused on small parts of the available data, such as year, month and type of crime, according to Van der Veer. Without data mining, the amount of data used in traditional crime pattern analysis is limited by the time analysts have to go through it.

The system has been used by Dutch police to tackle a number of linked crime problems. For example, link analysis on robberies in the Tilburg district produced a list of suspects, whose details could be checked for outstanding warrants and photographs distributed at police briefings. Within two weeks of the analysis, officers arrested and charged a number of suspects. Also by increasing police presence in the hotspot areas identified by the mapping, it acted as a deterrent to would-be thieves.

Caught in the act

The system was also used to tackle a spike in vehicle thefts in a district of Amsterdam. A profile analysis of the offences showed most of the increase in thefts were from private garages. Spatio-temporal cluster analysis and a list of the most prolific car thieves in the area were also given to patrolling officers.

As a result, a repeat offender was caught in the act within one hour of the data mining

'A repeat offender was caught in the act within one hour of the data mining analysis being distributed'

'With this in mind, our design philosophy for the data mining system has been to just require users to know their domain and to have analytical skills,' he says. 'The developed system brings together various techniques from business intelligence, statistics, machine learning and geographic information systems in a comprehensive data mining structure.'

By integrating most of the commonly used crime mapping tools into the new system, officers are not required to install and learn the tools separately or exchange results between different tools by importing and exporting them onto different systems. Popular tools have been linked into the system so results are automatically exchanged between applications such as Excel, MapInfo, Microsoft Word, Cognos Reportnet, Analyst's Notebook, Weka and Google Maps.

analysis being distributed. As a result of that one arrest vehicle crime in the area fell by 90 per cent within a week.

While the system is now a well-established analysis tool within the three police forces who helped develop it, Van der Veer says it is constantly being developed to improve its accuracy. For example, the company is developing more models to predict where crimes are likely to take place and to explain trends or behaviour.

Research is under way to expand these models to predict who has the highest risk of becoming a repeat offender, the likelihood of suspects carrying weapons and the ability to predict crime rates based on local infrastructure, buildings and the socio-demographic make-up of specific neighbourhoods. ■