Data mining for intelligence led policing

ir. R.C.P. van der Veer

Sentient Information Systems, Singel 160 Amsterdam

Abstract

This is an abstract of the paper presented at KDD2009, written in co-operation with the Amsterdam Police. It can be found at www.sentient.nl/docs/data_mining_for_intelligence_led_policing.pdf.

The benefits of data mining for police applications seem tremendous, yet police data mining appears to be hardly used in practice, judging by the few and limited documented examples. This paper discusses the implementation problems of police data mining and introduces a new approach that tries to overcome these problems in the form of a data mining system with associative memory as the main technique. The system has been developed in a group effort of data mining company Sentient and several Dutch police forces. It consists of an integrated data mining tool called DataDetective, which is being developed and applied by Sentient since 1992 and an extensive datawarehouse containing data from various police systems and external sources, such as weather data, geographical data and socio-demographics. A number of Dutch police forces have already been using this system for several years with over 30 users. Experiments have shown a factor 20 efficiency gain, a factor 2 prediction accuracy increase, a 15% drop in crime rate, and 50% more suspect recognition. The paper discusses the benefit of police data mining, the design of the system, a number of practical applications, best practices and success stories.

The described system addresses several shortcomings of traditional crime analysis systems. Because of the diverse information need, crime analysts typically work with a complex combination of tools that require special training and technical skills. These tools are based on just a few variables, typically generate static reports, and are unable to find complex patterns. Furthermore, it is typically hard to extract data from police source systems because of old and diverse database systems with quality issues and data models based on transactions instead of analysis. As a result, traditional crime analysis requires technical skills in the area of databases, statistics, let alone data mining. Such skills are rare within the police because of the already high demands on the level of domain expertise. This is why the design of the described system tries to reduce the need for technical skills as much as possible by working with one standard datawarehouse, techniques that can be configured automatically and active user guidance. The ease of use is also increased by integrating many tools and techniques from statistics, business intelligence and data mining into a single interactive environment. The user interface allows for analysis to be performed through step by step interaction, instead of requiring the user to design a workflow beforehand.

Associative memory is used as the main technique for prediction, clustering and matching in the described system. This means that input data is matched to a representation of training data using a similarity principle, as in Self Organizing Maps but with a much wider acceptance of different data types. Because of this, the mining process becomes easier and data usage richer:

1. Hardly any data preparation is required because the associative memory can handle a wide range of data types. As long as similarity between values can be calculated, a data type can be used. Furthermore, missing values do not need to be removed or guessed since similarity metrics are able to leave out the missing values in the calculation.
2. Because the technology can handle a wide range of data types, much more information is included in the mining process than would be feasible with other techniques with more strict data requirements.

3. The associative memory is able to explain how it reached a result by displaying relevant cases or persons from the memory - which is an intuitive way of explaining a decision to any user.

4. Building associative memory models converges well and fast, compared to for example back propagation neural networks.

5. Associative memories are robust against suboptimal parameters and therefore suitable to use in a situation where parameters are set automatically and the user is not an expert on optimizing the technique.

In other words: the user does not need to worry about fine tuning parameters, retraining, solving missing values, variable selection, and decoding variables into a usable form. Using associative memory makes the system easier to use, allows uncomplicated data handling and supports many different data types. Consequently, data preparation becomes easier and results contain more information.

The paper discusses various ways in which the described system is applied in practice. The main applications are:

- Spatio-temporal clustering: cluster recent incidents on co-ordinates, time of day and day of the week. Every cluster is reported with a list of crime characteristics, serving as tactical suggestions for patrol.
- Cluster sub problems: by involving contextual information in clustering incidents, the data mining system is able to detect clusters of similar incidents that each define a sub problem.
- Associative spatial prediction: creating a forecast map of a specific type of criminal behaviour in a given context: taking into account the time frame, weekday, season, period (school holidays), the weather and recent trends.
- Finding series of incidents with similar MO and suspect descriptions, to link behaviour to a person or a group.
- Determination of possible suspects by associating cases with incidents from the past and linking to the suspects of those incidents.
- Finding relations to explain a trend or certain behaviour using rule induction. Example: finding causes for an increase in burglary, or finding combinations of factors that account for a person becoming a repeat offender.
- Optimal visualization of graphs containing networks of criminals and finding the ‘spiders in the web’.
- Case association: find matches by not requiring a perfect match on all aspects, but rather find an overall ‘best fit’. Example: search persons similar to a witness description.
- Geographic profiling: use locations of crimes to find the most likely areas of the offender's home address by applying theories about the activity radius of offenders combined with theories about offenders usually not operating in the vicinity of their own home address.

The paper shows that data mining is important for police and that it has been made a continuous activity, performed by police employees without extensive expertise in databases or data mining. This has been realized by developing a data mining system in co-operation with police, bringing together interaction, visualization, tool integration, automated algorithms, a large and diverse datawarehouse and ease of use.