

Ethics Committee Briefing Note

| |
|--|
| Project Reference: |
| Purpose of data analysis: To identify the organised crime gangs, their methodology and how and where they disposed of stolen property by correlating West Midlands Vehicle Crime Data with West Midlands Police ANPR data. |
| Source of analytical question / hypotheses to be examined: Solving car crime when committed by organised crime groups is difficult. Currently any use of ANPR data to track the movements of stolen cars and the offenders who steal them is a resource intensive task. Searching and analysing the many millions of lines of ANPR data generated annually by West Mid Police currently would be the work of many people and it would be laborious. This research project aims to automate the process of sifting through ANPR reads to uncover valuable intelligence leads which will enable West Midlands Police to recover stolen cars, identify organised criminals and prevent crime. When organised criminals steal cars, they use their own legitimate vehicle often referred to as "bandit " car(s) to plan for the crime and to help them dispose of their stolen property. West Midland Police currently recover about 12% of the vehicles stolen within the Force area. The other 88% are disposed of in one of three ways <ol style="list-style-type: none">1. They are shipped abroad in a container2. They are broken up and disposed of as valuable parts3. They are cloned, take on the identity of a genuine vehicle and remain in circulation. In each case the stolen vehicles are likely to be driven to a pre-arranged destination for cloning , dismantling or onward shipping. Having left the stolen vehicles at this disposal location the thieves will return home in the bandit car(s) that will have accompanied them to the disposal / transit site. This behaviour will leave traces , whenever the stolen cars and the bandit car(s) pass through an ANPR camera. The hypotheses to be examined are <u>Bandit Cars can be found through comparing lists of ANPR reads</u> <ul style="list-style-type: none">• The ANPR data read list, following a vehicle theft, will be searched. If the stolen vehicle is found a list of vehicles passing through the same camera(s) at say 15minutes either side of the stolen car will be created. This is (if you like) a unique "ANPR vehicle finger print" list for the stolen vehicle in question. In this list will probably be found the bandit car(s).• Similar ANPR vehicle finger print lists will be created for each stolen car. By comparing two |

or more lists the bandit cars will emerge as being the common factor in each vehicle crime. The more often one or more cars appear in different ANPR vehicle fingerprint lists the more likely it is that they are bandit cars.

- On identifying a bandit car, its journeys that can be associated stolen vehicles (passing through an ANPR camera in one direction and repassing through it in the opposite direction some -time later) will provide intelligence on the distance of the disposal location from the ANPR camera.
- The bandit cars and possible disposal locations can then be researched by the West Midlands Police

Cloned Cars Can Be found through mining the ANPR data set with high performance computers.

- When a vehicle is cloned the genuine car and the clone will continue to pass through ANPR cameras.
- When both vehicles are being used at the same (or more or less the same) time they will each pass through ANPR cameras which given their distance of one camera from the other would not be possible for a single car.
- By using high powered computer processing and comparing each ANPR read with every other ANPR read pairs of cloned vehicles will be identified.
- Further analysis of the ANPR reads will identify where the cloned vehicle is likely to be operating from. The genuine vehicle will have a pattern of regular behaviour and the clone a more erratic pattern.
- Once the clone has been identified WMP can research and take action against those in possession of it.

Data to be used:

1. 12 months of West Midlands Vehicle Crime Data.
 - The date , time and place of the theft
 - The registration mark of the stolen vehicle
 - The method of theft
 - Details and circumstances of where and how a stolen vehicle has been recovered
 - Details of any known criminal use of a stolen vehicle e.g .used in a burglary or robbery
2. West Midlands ANPR Camera Location Information

- Location of each ANPR camera
- Direction in which facing

3. 12 months of West Midlands ANPR data (contemporaneous with 1)

- Details of each ANPR read
 - Registration Mark of Read
 - Location of Camera
 - Time of Read
 - Direction of Travel of vehicle

Level of analysis:

Individual

Individuals aggregated?

Yes

West Midlands

Reliability of data: [Reliable](#)

[Data set 1 is an extract of 12 months of the West Mids crime data set. The two most important items in this data set \(for the purposes of this trial\) are the registration mark of the stolen vehicle and the post code of the location of the theft. The registration mark might conceivably be wrongly recorded , the post code less likely so.](#)

[A wrongly recorded registration mark might cause the algorithm to hunt for a non-existent vehicle. A wrongly recorded post code would not significantly affect the trial. The algorithm will hunt for the ANPR camera located closest to the theft and work out from there. A wrongly recorded post code would simply slow down the process but would not affect its accuracy as the algorithm will try to find the stolen vehicle on each camera anyway.](#)

[Data Set 2 comprises the location and direction facing of each ANPR camera in the West Mids and is reliable](#)

[Data set 3 comprises the vehicle registration mark reads for 12 months . The data quality standards of these data set conform to the National ANPR Standards for Policing and Law Enforcement \(Part 3\)¹](#)

¹ Accessed at

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/936912/NASPLE_Version_2.1_November_2020.pdf

Formatted: Font: 8 pt

| |
|---|
| <p>Entire data set for 1,2,and 3 above:</p> |
| <p>Type of analysis:</p> <p>Exploratory</p> <p>Proposed methodology:</p> <p>This project involves the use of an algorithm, which uses data clustering techniques.</p> <p>The algorithm that will identify suspicion vehicle convoys (stolen car(s) plus bandit car(s) and has been developed by Surrey and Kent Universities in conjunction with Surrey Police. It has been laboratory tested and reported in a number of scientific publications². It has not yet been trialled with vehicle crime data.</p> <p>The methodology for this pilot will deploy the algorithm against the 12 months ANPR location and vehicle read data set and the contemporaneous vehicle crime data set (as described at 1 above) to identify</p> <ul style="list-style-type: none"> • suspicious vehicle convoys which include bandit cars driven by organised criminals and • vehicle clones. <p>The methodology deployed is described at pages 1 and 2 above. The process will be underpinned by high powered computer processing, a self -learning algorithm and data scientists who will in an iterative process refine the algorithm as each finding is made and confirmed.</p> |
| <p>Will the project eventually be automated:</p> <p>Yes</p> |
| <p>Means of evaluation:</p> <p>West Midlands Police will be supplied with the details of each suspicious vehicle convoy identified, and the behaviour of the constituent parts of the convoy (both stolen vehicles and bandit cars) as identified as being current during the research period. West Midlands Police will then be able to assess whether or not that information, had it been available at the time, would have enabled them to take operational action and conduct fruitful enquiries into the crime gangs involved in vehicle thefts. It may be that although the information is 12 months old or more it might be relevant to an ongoing serious case. It will be for the police to determine whether they should take operational</p> |

Formatted: Font: Not Bold

Formatted: Font: Not Bold

² Houmayounfar and others, University of Surrey 2011, "Multi-vehicle convoy analysis based on ANPR data"

~~action in relation to what may be new information, and will then be able to research operational intervention opportunities.~~

West Midlands Police will also be supplied with the details of identified clones of motor vehicles that were being driven during the period to which the research project applies. They will be able to assess whether or not, had they known of these clones at the time they could have taken positive action to recover stolen property. If an identified clone relates to a serious case that is ongoing it will be for the police to decide what action to take, and will be able to take appropriate operational action.

ALGO-CARE considerations:

Advisory:

If applicable, are the outputs from the algorithm to be used in an advisory capacity?

Yes.

Does a human officer retain decision-making discretion?

Yes.

Lawful:

What is the policing purpose justifying the use of the algorithm (means and ends)?

The policing purpose is the better understanding and improved operational effectiveness against organised vehicle theft. In the first instance, the project is one of research (does the algorithm do what is required?) and then following appropriate agreement it will be used for a purpose that is in the significant public interest – namely the recovery of stolen property and the prevention and detection of serious crime.

The way in which it is intended to process the data is **lawful**. It will be processed under

sec 19 (1) (b) of The Data Protection Act. (The scientific research purpose)

○ Article 89(1) of General Data Protection Regulation. And

section 8 (a) of the Act (The administration of justice condition) &

○ Article 6 of the General Data Protection Regulation .

Is the potential interference with the privacy of individuals necessary and proportionate for legitimate policing purposes?

The interference with the privacy of drivers of vehicles whose registrations have been read and stored in the ANPR data set will be non-existent, except for those whose cars are identified as bandit

or cloned cars.

The interference with their privacy is both necessary and proportionate.

In what way will the tool improve the current system and is this demonstrable?

Currently suspicious convoys involving stolen cars and the bandit cars associated with their theft can only be done manually and at great time and expense. The algorithm will increase the frequency and volume of analytical matches that will enable the West Midlands Police to act against those involved in organised car theft.

Currently the Police can only identify a vehicle clone one vehicle at a crime. They do not possess the computing power or the algorithm to identify clones simply from comparing millions of lines of data.

Are the data processed by the algorithm lawfully obtained, processed and retained, according to a genuine necessity with a rational connection to a policing aim?

[The data to be processed by the algorithm have been recorded by the police in accordance with their requirements to prevent and detect crime. This is a genuine policing aim and this extra processing with the use of an algorithm pursues this genuine aim.](#)

[The crime data is necessarily retained under the Criminal Procedures and Investigations Act 1996 \(Code of Practice\) .Yes](#)

[The processing of crime data is done in accordance with the Data Protection Act 2018 \(Schedule 9\) – public task](#)

[The ANPR data and the ANPR camera data is processed, and retained in accordance with National standards \(see footnote 1\)](#)

Is the operation of the tool compliant with national guidance?

[Yes, the DCMS Ethics Framework 2018 has been used to ensure that the use of this tool complies with relevant guidance.](#)

Granularity:

Does the algorithm make suggestions at a sufficient level of detail given it's purpose and the nature of the data processed?

Yes – the vehicles identified as bandit cars will be presented with the surrounding data that causes them to be identified as bandit cars.

(Which stolen cars they are associated with and when ,and their movements before and after the theft including details of how far the disposal site is likely to be away from any ANPR camera it travels through in the outward and inward journeys).

Are data categorised to avoid broad-brush grouping and results and therefore issues of potential

bias?

There is no broad-brush grouping.

Do the potential benefits outweigh any data quality uncertainties or gaps?

Yes; [- Organised Criminals do not just steal cars. They are involved in a wide range of offending and the stolen vehicle aspect of their criminality is sometimes a facilitator of other more serious crime. Detecting and calling to account organised criminals is a painstaking business. This policing tool will enable law enforcement to make serious offending more difficult.](#)

[Detecting organised criminals acts as a significant public benefit.](#)

Is the provenance and quality of the data sufficiently sound?

~~Yes~~ [The data has been recorded, processed and retained by West Midlands Police in accordance with the data standards that apply to police data recording.](#)

If applicable, how often are the data to be refreshed?

In the first stage of this project the ANPR and Vehicle Crime data sets are historic data sets and will not be refreshed.

If the tool takes a precautionary approach in setting trade-offs, what are the justifications for the approach taken?

N/A

Ownership:

Who owns the algorithm and the data analysed?

The data controller for this project is the West Midlands Police and they remain the **owners** of the data and the project. However, the National Police Chiefs Council Lead for ANPR (CC Hall ; Hertfordshire Constabulary) has approved the use of the data in this manner. Transtekniq and Surrey and Kent Universities are the **owners** of the algorithm.

Does WMP need rights to access, use and amend the source code and data?

WMP have these rights over these data. [WMP can see the code](#)

Are there any contractual or other restrictions which might limit accountability or evaluation?

No

How is the operation of the algorithm kept secure?

By Transtekniq in conjunction with the University of Kent and West Midlands Police

Challenge:

What are the post-implementation oversight and audit mechanisms, e.g. to identify any bias?

There are no demographic data being analysed- so no biases should be present in the working of the algorithm.

West Midlands Police will assess the effectiveness of this project in terms of the extent to which it is likely to contribute to the policing of car crime.

If the algorithm is to inform criminal justice disposals, how are individuals notified of its use?

It will not inform CJ disposals. However, it will provide investigators with actionable intelligence and some evidence that could help with a crime investigation.

Accuracy:

Does the specification of the algorithm match the policing aim and decision policy?

Yes; [the algorithm is designed](#)

[1- to identify bandit cars , involved in vehicle thefts , which would enable West Mids Police to conduct enhanced investigations into organised crime groups. Investigating and disabling organised crime groups is a policing aim.](#)

[2- to identify previously stolen vehicles , which are \(now\) clones. Recovering stolen property is a policing aim.](#)

Can the accuracy of the algorithm be validated periodically?

Yes – by comparing relevant stolen vehicle recoveries and detected vehicle theft crimes with what the algorithm has identified in relation to those thefts and recoveries.

Can the percentage of false positives / negatives be justified?

The algorithm will provide the police with intelligence and options. When a vehicle is identified as having been associated with a series of vehicle thefts and no other vehicle has the same pattern of behaviour it is a strong lead for the police to follow. It is justified for the police to follow up such a lead, irrespective of whether it results in a crime detection. How the police follow it up will be a matter for them.

This is justifiable – the police following all leads in investigating a crime as required of them by the Criminal Procedures and Investigations Act.

[Until the project is complete it will not be possible to say how many false positives/ negatives are generated by the algorithm.](#)

How was the method chosen as opposed to other available methods?

This is a unique algorithm and given the low recovery of stolen vehicles in the West Midlands Police Area it has been chosen to establish whether it can help with vehicle crime detection and stolen vehicle recovery. It offers an opportunity to work with data sources that are rich in potential but so

far not as well used as they could be. There are no known other comparable algorithms available.

What are the (potential) consequences of inaccurate forecasts?

Some vehicle owners may be questioned or stopped by the police because their vehicle has been associated with a number of stolen vehicles. It is unlikely that this is coincidental (but not impossible) given the irregular behaviour of most serious crime gangs. Not applicable as this is not a predictive tool.

Does this represent an acceptable risk?

N/A~~Yes~~

How are the results checked for accuracy and how is historic accuracy fed back into the algorithm for the future?

West Midlands Police will feed back to Transtekniq details of their research into identified bandit cars and vehicle clones. This feed-back will be used to improve and refine the algorithm.

How would inaccurate or out-of-date data affect the result?

In this first phase this will not be an issue as the data are all historic and are being offered up as a research opportunity. It is unlikely that vehicle theft or ANPR data being used will be inaccurate. It is more likely to be incomplete.

If (say) the time and date of a theft were wrongly reported then any unique "ANPR vehicle fingerprint" list for that crime would be unlikely to contain any bandit cars and so an intelligence opportunity would be lost.

Responsible:

Would the operation of the algorithm be considered fair?

Yes~~This algorithm simply does at high speed with superior computing power what investigators and analysts currently do manually.~~

The successful management of organised crime groups (though the identification of bandit cars) and the recovery of stolen vehicles disguised as clones would be considered fair.

Is the use of the algorithm transparent (taking account of the context of its use), accountable and placed under review?

Transparency

It needs to be considered whether public exposure of the use of the algorithm to identify bandit cars is in the public interest as publicity could cause unintended consequences if it proves successful.

vandalism of ANPR cameras, criminals changing behaviour to avoid ANPR cameras.

Accountability

This project is overseen by a senior WMP officer and is known to the NPCC lead for ANPR.

Review

This is a research project and before the algorithm is used for a live policing purpose the lessons learned for the project will be captured in a review. Yes

Formatted: Font: Bold

Would it be considered to be used in the public interest and to be ethical?

Public Interest

Yes – the capture of serious criminals and the recovery of stolen property through the use of low resource tactics is in the public interest. It is much less intrusive than human surveillance, which is one of the tactics it could replace or make more accurate.

Ethical

The output of the algorithm will be intelligence. The police will need to investigate that intelligence to corroborate it before taking action. Yes

Formatted: Font: Bold

Explainable:

Is information available about the algorithm / decision-making rules and the impact of each feature?

Yes ; once the experiment is concluded a technical document will be published which will describe how the algorithm performs.