

Ethics Committee Briefing Note

Project Reference: DAL_2019_0002_SOC_networks

Purpose of data analysis:

Context

West Midlands Police is the **second largest** police force in the country, covering an area of **348 square miles** and serving a population of almost **2.8 million** - it is the **3rd** most **densely populated** region in England. The region sits at the very heart of the country, with an average of **170,000** motorists travelling through the region on a daily basis making our motorways some of the busiest in Europe and making the West Midlands an attractive centre where Organised criminals operate.

Policing the West Midlands area is complex and the communities of the West Midlands are particularly vulnerable to crime and exploitation by organised criminals with over a quarter of communities live in the top 10% of the **most deprived** areas in the country and over a third of the population (35%) is under 25 years old, with **lower than average** levels of **qualifications** on leaving school and higher rates of **unemployment** compared with other young people in England.

Rises in **vulnerability and high harm offences** have been underpinned by an increase in total recorded crime (TRC) of 11% between 2017 and 2018 – the **highest volume recorded for a decade** (nationally, TRC has risen by 7%).

In the last year, in the WMP force area: **homicides** rose by 27% (to 66 offences), **robbery** increased by 11%. Nationally the increase in **knife crime** contributed to a rise in linked hospital admissions of 15%. While the number of **firearms discharges** reduced and recoveries of firearms remained constant, **gang tensions** remained unchanged. The carrying of knives also increased by 28% to a level double that in 2015 and we are seeing increasing numbers of injuries and **fatalities amongst our young people**.

There have also been increases in **Serious Acquisitive Crime** (SAC) with **Residential Burglary** and **Theft of Motor Vehicle** increasing significantly - crimes where a **car key was stolen** (2,752) have doubled in two years (2016-2018). **Fraud** has also increased and is the most prevalent crime type identified and Action Fraud has recorded an 11% increase in the last year.

The Home Office 2013 Serious and Organised Crime Strategy, revised in November 2018 recognises the **complexity of the problem**. The strategy advocates a **whole system approach** to ensure that issues are not tackled in isolation, as many problems cut across law enforcement boundaries and require the support of both statutory, voluntary and third sector organisations to **reduce harm** with the most appropriate legislation and powers.

To enable this we need to understand how **organised crime groups exploit people and harm communities to build assets and accumulate wealth**. The nature of **exploitation** varies across the whole organised crime business model. For example: Many individuals who are used to steal cars or car parts are likely to be **dependent on drugs**, with OCGs exploiting this **vulnerability** to keep people in a cycle of crime. Once the vehicle or parts are stolen, the nature of exploitation affects the

labourers needed to break the vehicles. In pop-up chop shops, there are likely to be **low-paid workers** or **modern slavery victims** exposed to **health and safety hazards** and poor working conditions.

The traditional model of organised crime is hierarchical, but it's not a simple top down structure. Instead it forms a **multifaceted network arrangement** that is resilient to disruption by being able to **adapt to changes in the market**. It's a form of **ecosystem** where different SOC criminals work together to generate profits when it suits them to do so. Their **strongest asset** is this **connectivity** and **cooperation** in building wealth and **asserting power and influence**. WMP recognises that by examining the problem from the perspective of the **commodities** being traded by organised criminals, there are **greater opportunities** to **prevent** crime, work better with partners, and **prevent** the **exploitation** of the vulnerable.

The **SOC ecosystem** relies upon **facilitators** and **enablers**, such as **accountants, solicitors, builders** and **estate agents** who help by providing **specialist services**, and it is important to understand how these groups and individuals link together

This project would seek to answer two (related) questions;

- Which serious organised crime (SOC) groups create the most harm and
- What are the SOC networks (including identifying centrality within those networks and potentially identifying links between different networks for further analysis via WMP's intelligence department)?

Source of analytical question / hypotheses to be examined:

The business question was posed by the FET.

Data to be used:

Level of analysis:

- Individual
Individuals aggregated?
- Yes
 No
- Specific Area:
 Output Areas
 Super Output Areas - Lower
 Super Output Areas - Mid
 Wards
 Districts
- West Midlands
 Other

Reliability of data:

The data are sourced from WMP systems. A major element of this project will involve making an assessment of the quality of the data, the robustness of the various systems, etc. Any data quality

issues will be noted and where applicable will be incorporated into the project (e.g. by excluding some data from a system if it is felt to be unreliable). These data, as part of the analytical project life cycle, will also be assessed for missing values, etc.

These systems are those currently used by WMP in their day-to-day business. Specifically in the case of intelligence data, these will be examined as to their veracity, source, etc. prior to inclusion (i.e. only intelligence considered to be credible from credible sources will be used).

Sample or entirety:

Entirety.

If sample:

Not applicable.

Method of sampling:

Not applicable.

Method of choosing sample size:

Not applicable.

Sample size:

Not applicable.

Type of analysis:

- Exploratory
- Explanatory
- Predictive
- Optimisation

Proposed methodology:

The first element of the project creates the network based upon identifying individuals who are members of gangs and / or organised crime groups (via intelligence logs). Intelligence and other systems are then used to identify if a (potential) relationship exists between individuals (e.g. have they ever been arrested together, etc.). Following this, a graph would be constructed of the linkages between individuals (focussing on those in SOC groups) but including edge values and temporal information. This would allow for the identification of links between the various sub-graphs, calculation of centrality within the various networks, calculation of the levels of harm created by these networks, etc. We would also aim to develop an algorithm to identify effective means of disabling networks (to avoid computationally expensive searches across the various nodes whilst identifying those nodes that would most quickly lead to the dissolution of the network if they could be removed from it).

Will the project eventually be automated:

- Yes
 No

Means of evaluation:

Not applicable.

ALGO-CARE considerations:**Advisory:****If applicable, are the outputs from the algorithm to be used in an advisory capacity?**

The project would enable a delineation of the SOC networks, the links within and between them and measures of centrality. Assessing the harm created by individuals through their criminal activities will also enable the extent of the harm created within the WMP area resulting from SOC to be assessed.

This information can then be used to feed into the risk and threat assessments undertaken by WMP to help in the allocation of tasks and resources.

Does a human officer retain decision-making discretion?

Not applicable.

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

SOC undoubtedly creates harm within society through involvement in a number of different crimes ranging from vehicle theft through to murder. This project would allow for the assessment of the harm created via the various SOC groups and as such will allow for extra information to be provided to the threat and risk assessments and so enable better decision making as to the focussing of resources.

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

The project only utilises data collected from WMP systems in their normal day to day activities. Assessing the priorities that should be tackled by the Police in their operations due to the risks and threats present is a legitimate policing purpose and this project would feed into that process.

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

There is currently no system or tool to undertake large scale network analysis or to delineate the

harm created via any such networks.

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 are gathered in the normal day to day operations of WMP in line with the aims of WMP.

Is the operation of the tool compliant with national guidance?

The analyses proposed would accord with DCMS Data Ethics Framework 2018.

Granularity:

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

The analyses would use information at the level of the individual to develop the networks and assess the overall levels of harm created by the networks. Suggestions would not be made by the associated algorithms per se with the exception of the means of best disabling a network.

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

The underlying data are not categorised.

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

The project could help focus resources of WMP and lead to enhanced assessments of threat and risk that are currently undertaken. Given the harm that SOC groups cause within society at large then uncertainties as to data quality (bearing in mind such issues have been investigated in the preliminary stages) are acceptable (no decisions would be made as to individuals per se as the aim is to provide further information for risk assessment).

Is the provenance and quality of the data sufficiently sound?

The data have been gathered during the day-to-day investigative work of WMP and so enable analyses of the type envisioned for this project.

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

The underlying data are refreshed on a sub-daily basis as they are part of WMP's core systems. The analyses may be run intermittently on an as and when needed basis but do not require specific data extracts.

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

<p>Not applicable.</p>
<p>Ownership:</p> <p>Who owns the algorithm and the data analysed?</p> <p>WMP owns the analyses and the underlying data.</p> <p>Does WMP need rights to access, use and amend the source code and data?</p> <p>Not applicable.</p>
<p>Are there any contractual or other restrictions which might limit accountability or evaluation?</p> <p>Not applicable.</p> <p>How is the operation of the algorithm kept secure?</p> <p>The data and the analyses are contained wholly within the WMP Hadoop system and the security measures employed therein.</p>
<p>Challenge:</p> <p>What are the post-implementation oversight and audit mechanisms, e.g. to identify any bias?</p> <p>Any findings relating to particular individuals that are deemed useful for further analysis would be subject to the normal intelligence processes of WMP and as such any incorrect findings could be notified to the Lab and incorporated into any future runs of the analyses.</p> <p>If the algorithm is to inform criminal justice disposals, how are individuals notified of its use?</p> <p>Not applicable.</p>
<p>Accuracy:</p> <p>Does the specification of the algorithm match the policing aim and decision policy?</p> <p>The nature of the analyses chosen have been determined to be the best means of addressing the research question.</p> <p>Can the accuracy of the algorithm be validated periodically?</p> <p>Not applicable.</p> <p>Can the percentage of false positives / negatives be justified?</p> <p>Not applicable.</p> <p>How was the method chosen as opposed to other available methods?</p>

Due to the nature of the research question (network analysis).

What are the (potential) consequences of inaccurate forecasts?

Not applicable.

Does this represent an acceptable risk?

Not applicable.

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

Not applicable (other than any intelligence analysis arising from findings as noted above).

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

If data were to be wholly inaccurate then the analyses would essentially provide inapplicable findings. The Lab has sought to minimise this potential through a thorough analysis of the data and their pitfalls, issues and overall nature; through discussions with SMEs.

Responsible:

Would the operation of the algorithm be considered fair?

The analyses will be fair in that each data point will be considered on its own merits.

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

The nature of the intended method(s) is such that the end-use is to provide information rather than predictions that feed into a process on an on-going basis.

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

In the face of current levels of harm generated through criminal activity and in the interest of the more efficient allocation of resources, this project would be in the public interest.

Explainable:

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

Information about the algorithms used would be available, however no decisions are made per se and the type of analyses are such that measures such as feature importance are not applicable.