Forensic Vehicle Convoy Analysis Using ANPR Data

[1] Haiyue Yuan, Shujun Li, Anthony T.S. Ho
[2] Paul Palmer, Andy Lloyd and [3] Graham Head

[1]Department of Computer Science, University of Surrey, UK
[2]Surrey Police, UK [3]Independent advisor, UK

1. Introduction

- Growing need for using Automated Number Plate Recognition (ANPR) data to detect and prevent crimes
- 30 million ANPR reads per day across the UK (2015), and lack of software tools enabling efficient automatic processing data
- We are developing a digital forensic tool to identify unknown and suspicious vehicles traveling in convoy 'hidden' in ANPR data.
  - Enhance functionalities of current analysis tool
  - Combining data from multiple data sources
  - No known number plate is needed

2. Background information

2.1 Data sources

- ANPR data from Surrey Police: contains Vehicle Registration Mark (VRM), date, time, camera ID, and camera GPS location.

Table 1: *Example of ANPR reads (VRMs and cameras' geo-locations are anonymised)*

VRM	DATE	TIME	CAM ID	LATITUDE	LONGITUDE
A	01/01/2015	08:00:01	ANPR01.1	51.0000	1.1000
B	01/01/2015	08:00:02	ANPR02.1	52.0000	1.2000
C	01/01/2015	08:00:03	ANPR03.1	53.0000	1.3000

- Data from UK's Driver and Vehicle Licensing Agency (DVLA): vehicle's make, model, color, and tax code
- Crime map from police.uk: public crime statistics
- Goolge Map: travel time estimation

2.2 Concepts

- Single vehicle journey: describes all travel activities of a single vehicle during a period of time.

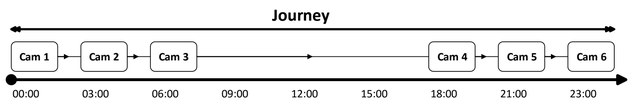

Figure 1: *Illustration of a journey*

- Single vehicle session: refers to a single vehicle continuously traveling for a period of time without a break.

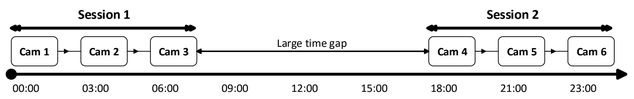

Figure 2: *Illustration of sessions*

- Convoy session: two or more vehicles traveling together for a period of time.

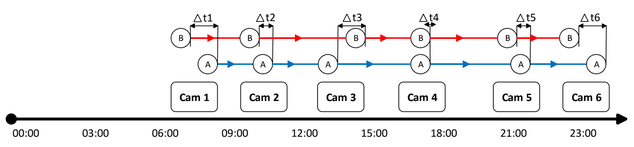

Figure 3: *Illustration of a convoy session*

3. Selected related work

- [1] uses data clustering techniques to extract and identify unusual patterns of multi-vehicle convoy activities using ANPR data.
- [2] investigates headway distance of two vehicles driving together.
- [3] investigates behavior of intentional following using ANPR data.
- [4] applies data mining techniques to discover vehicle activity patterns from ANPR data.

4. Simple rule based algorithm

The work presented in this section is an extended edition of a previous study [1].

4.1 Thresholds

- **Session break threshold**: the maximum time gap between two adjacent ANPR reads of a single vehicle.
- **Convoy session size threshold**: the minimum number of cameras that two vehicles passed together in a convoy session.
- **Search window threshold**: the maximum time difference between two vehicles in a genuine convoy session while passing the same camera.

4.2 Algorithm design

- Journey extraction (Raw ANPR ⇒ Journey data)
- Session extraction (Journey data ⇒ Session data)
- Convoy session extraction (Session data ⇒ Convoy sessions)

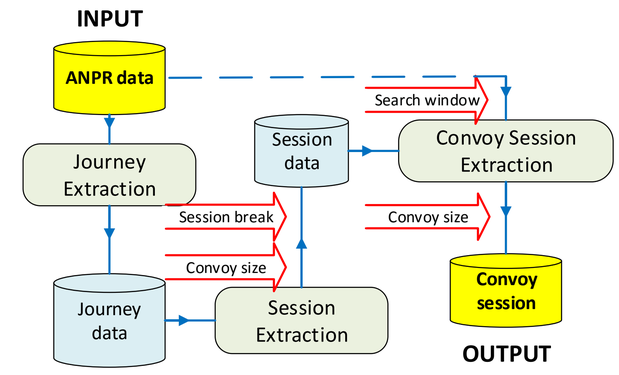

Figure 4: *Flow chart of the rule based algorithm*

4.3 Post processing

- Detect repeated 2-vehicle convoy sessions
- Detect multi-vehicle convoy sessions
- Retrieve convoy sessions using make/model/color and/or tax code
- Retrieve convoy sessions that share the whole/part of their traveling routes

4.4 Results

Results are generated using session break threshold of 5 hours, convoy session size threshold of 5, and search window threshold of 30 seconds.

- Processed ANPR data for 185 days (around 200 million reads in total)
- 35 repeated 2-vehicle convoy sessions
- 260 multi-vehicle convoy sessions
- Genuine convoy sessions confirmed by Surrey Police involving taxis, family vehicles, business vehicles, etc.
- Successfully detected a multi-vehicle convoy session created by our project partner Thales UK

5. Adaptive threshold based algorithm

- Aims to replace hard thresholds with adaptive ones using historical ANPR data and public traffic data.
- Aims to adjust all parameters to control the number of outputs to make them manageable.

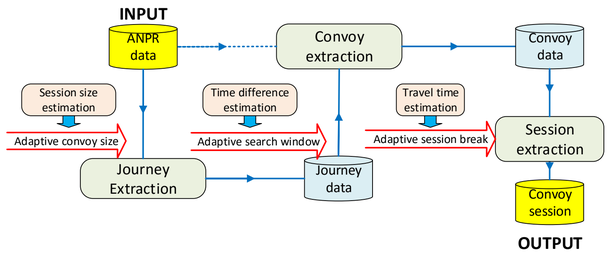

Figure 5: *Flow chart of the algorithm with adaptive thresholds*

5.1 Adaptive session break threshold

This algorithm is used to estimate a distribution of travel times of all vehicles passing two ANPR cameras, given the time of a day and the day of a week. The idea is to use the estimated travel time distribution to determine the session break threshold adaptively to define sessions.

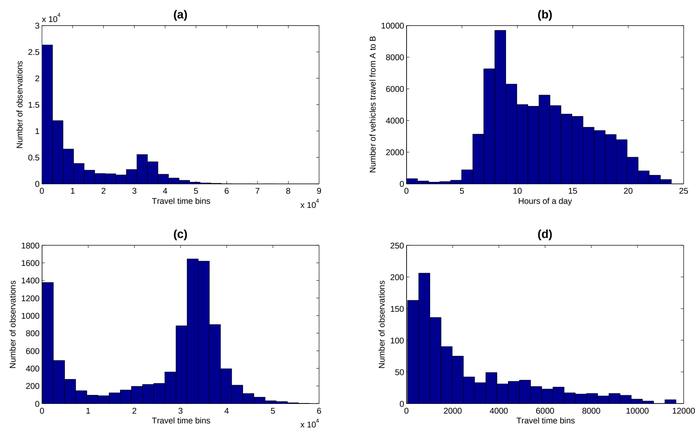

Figure 6: *(a) Distribution of travel times of all vehicles passing A and B on Monday (b) Distribution of the number of vehicles passing between A and B at different hours on Monday (c) Distribution of travel time of all vehicles passing between A and B from 08:00 to 09:00 on Monday (d) Distribution of travel time of all vehicles passing between A and B from 21:00 to 22:00 on Monday*

- Figure 6(a): the shorter travel times at the first peak represent that vehicles pass A, turn around and pass B; The longer travel times at the second peak represent that vehicles leave early to work passing A, and return home after work passing B.
- Figure 6(b): most of the vehicles are likely to pass A, turn around and pass B within a short time window during night.

6. Micro-behaviors of drivers in convoy sessions

6.1 Related driving behaviors

- Time difference (headway, daytime/night time, criminal behavior)
- Change order of driving (unusual overtaking behavior)
- Change of lanes
- Number of intervening vehicles

6.2 Machine learning based analysis

- Features based on micro-behaviors
- Refinement and clustering of convoy sessions (see Figure 7).
- Prioritising/Ranking convoy sessions based on behavioral features

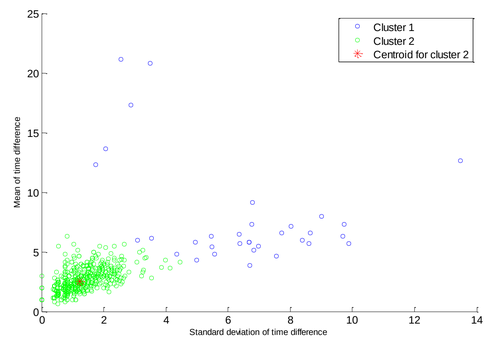

Figure 7: *Clustering convoy sessions using two features derived from time differences*

7. Acknowledgments

This work was co-funded by the Innovate UK, under the Application Number 41698-293424. We also thank our project partners Thales UK Ltd for providing input to our work.

References

[1] A Homayounfar, ATS Ho, N Zhu, G Head & P Palmer, "Multi-vehicle convoy analysis based on ANPR data," in *Proeedings of 4th International Conference on Imaging for Crime Detection and Prevention* (ICDP 2011), 2011

[2] A Heino, HH van der Molen & GJS Wilde, "Differences in risk experience between sensation avoiders and sensation seekers," *Personality and Individual Differences*, 20(1):71-79, 1996

[3] B Gunay, "Using automatic number plate recognition technology to observe drivers," *Journal of Advanced Transportation*, 46(4):305-317, 2012

[4] Y Sun, X Zhou, L Sun & S Chen, "Vehicle Activity Analysis Based on ANPR System," in *Proceedings of 12th IEEE International Conference on Embedded and Ubiquitous Computing* (ITQM 2014), *Procedia Computer Science*, 31:48-57, 2014

Corresponding Authors: [1]{haiyue.yuan; shujun.li; a.ho}@surrey.ac.uk