WESTERN AUSTRALIA POLICE FORCE

Predicting Vehicle Location Using ANPR Camera Reads

1. Background

When a vehicle is reported as stolen by its owner, Police will instruct Automatic Number Plate Recognition (ANPR) cameras across Western Australia to send an alert should they capture that license plate. When this alert is received, officers are tasked with organising an intercept of the identified vehicle by attempting to manually predict the vehicle's location. The issue with this current method is that for each minute that officers spend trying to create a prediction of the vehicle's location, is another minute that the vehicle can have travelled further away from it was captured in its initial location, exponentially increasing the possible locations the vehicle may currently be. But since these ANPR cameras capture the license plate numbers of the vehicle, the direction the vehicle was travelling in, the previous location of the vehicle and the time since it captured the vehicle, there is potential to utilise this information alongside road networks data to automatically predict the locations a vehicle may potentially be, and thus assist in the interception of the stolen vehicle.

2. Objective

The primary aim of this project is to develop a system that can estimate the current geographic location of a vehicle based on ANPR camera read/s. This system will need to leverage the location of the vehicle, its direction of travel, and the time elapsed since the last read, with additional scope to consider environmental factors such as time of day, traffic data, or even other cameras in predictions.

3. Constraints

The output of this system must be understandable and interpretable by a non-technical audience and that this system must work efficiently such that it would be viable to run in real-time. Also, a simple machine learning recreation of the data provision is discouraged due to the synthetic nature of the data provided.

4. Potential Solution

Potential solutions include statistical analysis, time-series forecasting, graph-based algorithms, etc. The project may also explore the use of simulation techniques to model vehicle movements on the road network.

5. Data

The project will utilize synthetically generated ANPR camera location data, and synthetically generated ANPR situations involving the location of the read, the direction the vehicle was travelling in, the time since the read, details about the vehicle read, and a series of potential end locations for a given duration. There will be four datasets provided: three will be different electorates from around Perth, and one will be of the wider Metropolitan region of WA. Additionally, a set of ANPR incidents without potential pathways for the wider Metropolitan region of WA will also be provided. Road network data can be sourced from Main Roads, or any other participant preferred method.

6. Data Generation

Camera locations have been generated by randomly sampling street locations from the relevant geographic area for each dataset and assigning the bearing of the camera to match the direction of the road. Cameras can record a vehicle travelling in either the direction of its bearing or the opposite direction (one-way roads excepted). Each row of the dataset contains the ID of the camera, the geographic coordinate (long/lat) of the camera, the bearing the camera faces, the street name and suburb that the camera is located, and the carriageway of the road the camera is on.

Due to the sensitivity of real ANPR data, the ANPR data for this challenge has been synthetically generated by taking the Main Roads Road networks data and adapting that into a graph network. This allows road infrastructure to be represented as a collection of nodes and edges made from the segments of roads that model the connectivity and pathways within Perth, and by assigning metrics to edges we can find pathways based on those metrics. To explain this further, I will describe how the ANPR data has been generated:

Firstly, a camera node from the graph network is randomly selected as having read a car, recording a fake number plate and fake details about the recorded vehicle. A bearing representing the direction that the car is travelling in is generated alongside a duration in seconds indicating how long it has been since the car was read as being at that location. Next the camera location node and duration are use in order to find all possible pathways that the car may have taken. This prediction is made by using other locations (nodes) that are connected to the camera location node via road segments (edges), these edges can be used in order to find all potential locations within a certain traversal distance for example. The pathways in this data set have been weighted by a "time to travel" metric. This metric is calculated from the length and legal speed limit of a road segment, approximately representing how many seconds that road segment should take to traverse. Using this metric alongside the time since an ANPR read, all the potential pathways that a vehicle make have potentially taken can be generated. For the purpose of the data provided in this challenge the top 1% of pathways on the metric of geographic distance from the starting location, multiplied by a value ranging from 0 to 1 based on how in line the end location is with the bearing that the car was recorded as travelling in (as cars are more likely to keep going in the rough direction of travel they were recorded as going in, and not to just turn around immediately), have been extracted as the potential pathways for each simulated ANPR situation.

Each row of the dataset consists of a synthetic ANPR read with a fake license plate for the car read, a generated colour, make and model for the vehicle, the id of the camera that read the car, the bearing in which the car was travelling, the number of seconds since the car was read by the camera, and the series of potential pathways that the car may have taken. The additional synthetic metropolitan ANPR read dataset is the same except no potential pathways have been provided.

7. Data Dictionary

a. ANPR Camera Data

No	Column Name	Data Type	Description
	Camera_id	Integer	A numeric identifier of the camera.
	Camera_locations	Geographic	Geographic point data of where the segment of the road a camera is starts.
	Bearing	Float	The bearing of the road between its start and end
	Street_name	String	The name of the road that the camera is on
	Lg_name	String	The name of the road that the camera is on
	CWY	String	If the road is two-way (single) or one-way (left when the bearing is between is greater than 90 but less than 270, right when the bearing is between is greater than 270 but less than 90)

b. Synthetic ANPR Reads Data (with paths)

No	Column Name	Data Type	Description
	Car_id	String	Unique identifier for each car in the form of a fake license plate
	Colour	String	A generated colour for the read vehicle
	Make	String	A generated make for the read vehicle
	Model	String	A generated model for the read vehicle
	Camera_id	Integer	The name of the road that the camera is on
	Bearing	Float	The bearing of travel the car was read as heading
	Duration	Integer	The number of seconds since the read occurred

	Paths	Array of Linestrings	A series of algorithmically generated likely pathways the car may have potentially taken from the incident based on distance and direction

c. Synthetic ANPR Reads Data (without paths)

No	Column Name	Data Type	Description
	Car_id	String	Unique identifier for each car in the form of a fake license plate
	Colour	String	A generated colour for the read vehicle
	Make	String	A generated make for the read vehicle
	Model	String	A generated model for the read vehicle
	Camera_id	Integer	The name of the road that the camera is on
	Bearing	Float	The bearing of travel the car was read as heading
	Duration	Integer	The number of seconds since the read occurred