

Text mining police narratives of domestic violence events to identify coercive control behaviours

Min-Taec Kim and George Karystianis^a

^a School of Population Health, University of New South Wales

SUMMARY

CRIME AND JUSTICE BULLETIN
NUMBER 260 | SEPTEMBER 2023

BACKGROUND

The aim of this study was to construct a measure of coercive control behaviours from free-text narratives recorded by NSW Police Force, and assess the use of this measure for detecting and/or predicting which events are followed by domestic violence (DV) related violence within 12 months.

We developed a text-mining system to capture a (non-exhaustive) set of behaviours that could be used to exert coercive control. Our text mining system consists of a set of rules and dictionaries, with our definition of 'coercive control behaviours' drawn from Stark (2007).

We applied this text mining system across all police narratives of domestic violence events recorded between 1 January 2009 and 31 March 2020 (n = 526,787), and used this data to construct our measure of coercive control behaviour. We then compared the incidence of coercive control behaviours using our measure against existing incident categories.

Finally, we construct a gradient boosted decision tree model three times, predicting whether the POI commits a violent DV related event within 12 months - once with standard predictors, once with just our coercive control measure, and once with both sets of variables included. This allows us to assess the usefulness of the measure by comparing the performance of each model.

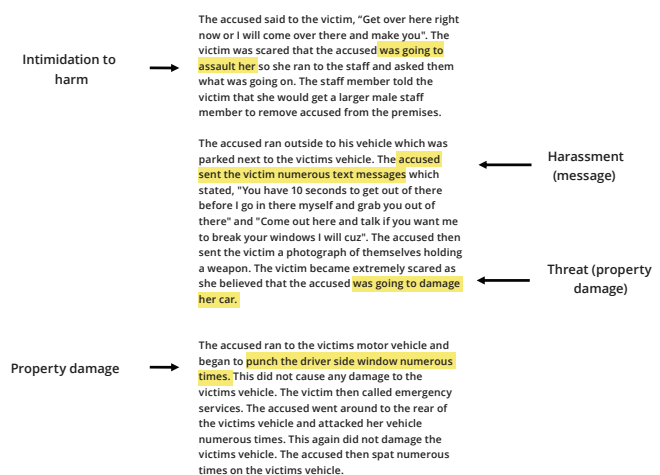
KEYWORDS

- Domestic violence
- Coercive control
- intimate partner violence
- abuse
- text mining

KEY FINDINGS

There were 852,162 behaviours extracted by the text-mining system across 526,787 domestic violence events, with 57% of these events having at least one coercive control behaviour detected and 8% with three or more distinct subcategories of coercive control behaviours. These events were associated with 223,645 unique persons of interest.

Figure 1. Example of marked up narrative



Comparing our text mining system and the pre-existing police incident categories (where they exist) agreed in 80 to 99% of cases, with the text mining approach identifying an additional 30 to 60% of events that were not identified using the police incident categories, depending on the behaviour.

The inclusion of the text mining variables provided essentially no improvement to our predictive models, with predictive metrics improving by 0.001– 0.015 depending on the measure.

CONCLUSION

The inherent difficulty in defining, operationalising, and capturing coercive control present challenges to evaluating the impact of the new coercive control legislation in NSW. Although our measure improves our ability to detect coercive control, we do not find that it improves our ability to predict future domestic violence events.