AN OPEN-DATA, AGENT BASED MODEL OF ALCOHOL-RELATED CRIME

JOSEPH REDFERN
DR KIRILL SIDOROV
PROF PAUL L. ROSIN
PROF SIMON C. MOORE
DR PADRAIG CORCORAN
PROF DAVID MARSHALL

MOTIVATION

>€10⁹ PER YEAR

~40,000 Nurses

~33,000 Police Officers

MOTIVATION

$>€10^9$ PER YEAR

~40,000 Nurses

~33,000 Police Officers

~800,000 Macbook Pros

>10,000 NVIDIA DGX-1 Systems

https://iea.org.uk/wp-content/uploads/2016/07/DP_Alcohol%20and%20the%20public%20purse_63_amended2_web.pdf

MOTIVATION

- Injuries incurred can be life-changing.
- Bad reputation could deter visitors/spending.
- Planning for novel scenarios or designing cities can be hard – no existing data.
- Assist police/city planning efforts to reduce risk.

RESEARCH QUESTION

"Without historical data, can we predict areas at-risk of alcohol-related violence in cities?"

EXISTING WORK: HOTSPOT ANALYSIS
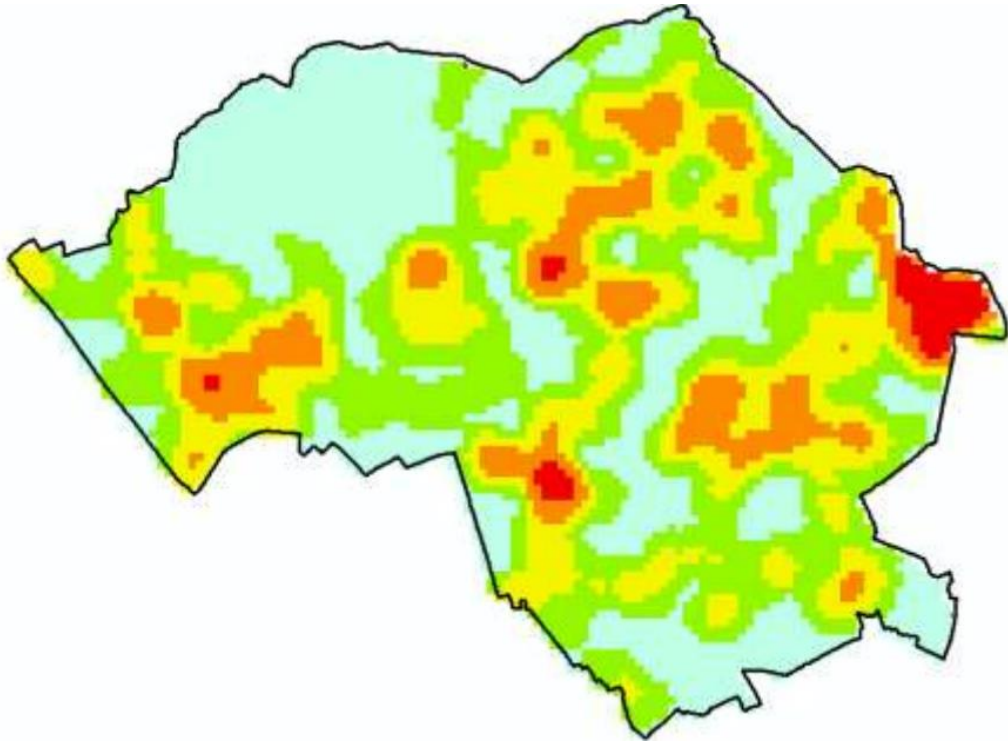
- Kernel Density estimation (or similar) of historical crimes.
- Commonly used in Crime Reduction strategies.
  - Increased police presence in high-risk areas to deter crime.

EXISTING WORK: HOTSPOT ANALYSIS

+ Can work well, given existing historical data.

+ Computationally inexpensive.

+ Evidence-based, easily interpreted.

- Requires historical crime data for given area.

- Less effective for novel scenarios/events.

EXISTING WORK: SIMDRINK

SimDrink, Scott et al (2016)

- Simulates drinking in Melbourne.

- Agents move between bars and consume alcohol.

- Simulates alcohol consumption, tracks agent finances, models availability of transport.

- Measures various "harms" incurred by agents.
  - Includes violence, drinking too much, having difficulty getting home.

EXISTING WORK: SIMDRINK

+ No historic crime data required.

+ Simulation produces realistic output.

+ Possible to run "what if" scenarios.

- Produces aggregate figures.

- True geography not considered, cannot localise hotpots.

- New cities/simulation areas require new set of parameters.

OUR SOLUTION: OVERVIEW

Graph-constrained, agent-based model.

- Agents (representing people) move around road network.

- Consume alcohol at pubs/bars/clubs.

- Track blood-alcohol concentration (BAC).

- Use pedestrian density and BAC as a proxy to predict risk of violence – common assumption, likely valid.

- Implemented using Python (graph-tool) and Postgres (PostGIS).

OUR SOLUTION: DATA SOURCES (I)

OpenStreetMap

- Open, collaboratively generated map of the world.

- Data includes nodes (like vertices), and ways (like edges) forming roads.

- Easily parsed into a graph/network.

- Agents traverse this network when moving around the city.

- Easy to modify OSM data to add new roads, venues etc

OUR SOLUTION: DATA SOURCES (II)

FourSquare and Google Places

- OSM road data typically excellent – venue data can be poor.
  - Out-of-date venue names, incorrect opening times, missing data.
- Foursquare and Google provide APIs for their own data.
- Includes venue location, opening times.
- Query these APIs to supplement OpenStreetMap data.

OUR SOLUTION: AGENT DESCRIPTION (I)

Each agent has a number of properties, some fixed (calculated once for each agent at run-time), some dynamic (changing throughout the simulation)

Name	Type	Description
Gender	Fixed	P(male) == P(female) == 0.5
Mass	Fixed	Sampled from Normal Distribution
Home Location	Fixed	Randomly sampled from OSM Residential Streets
Drinking Rate	Fixed	Randomly sampled from cut-off normal distribution.
Time spent per venue	Fixed	Time spent per venue, sampled from cut-off normal distribution.
Session Duration	Fixed	Overall duration of drinking session, sampled from cut-off normal distribution.

OUR SOLUTION: AGENT DESCRIPTION (II)

Each agent has a number of properties, some fixed (calculated once for each agent at run-time), some dynamic (changing throughout the simulation)

| Name	Type	Description
Previous Drinks	Dynamic	Timestamped list of drinks consumed
Current Route	Dynamic	The current route being followed (if any)
BAC	Dynamic (Derived)	Derived from Gender, Mass and Previous Drinks (using Widmark Formula)

OUR SOLUTION: AGENT BEHAVIOUR

- Agents move from their home locations (determined from OSM data).

- Visit a venue, consume drinks at a particular rate, and move to a new venue (or go home) according to individual agent preferences.

- Shortest path considered (avoiding major roads).

- When a venue closes, the agents leave the venue.
 - Opening/closing time data from Google Places & Foursquare.

OUR SOLUTION: MODEL OUTPUT
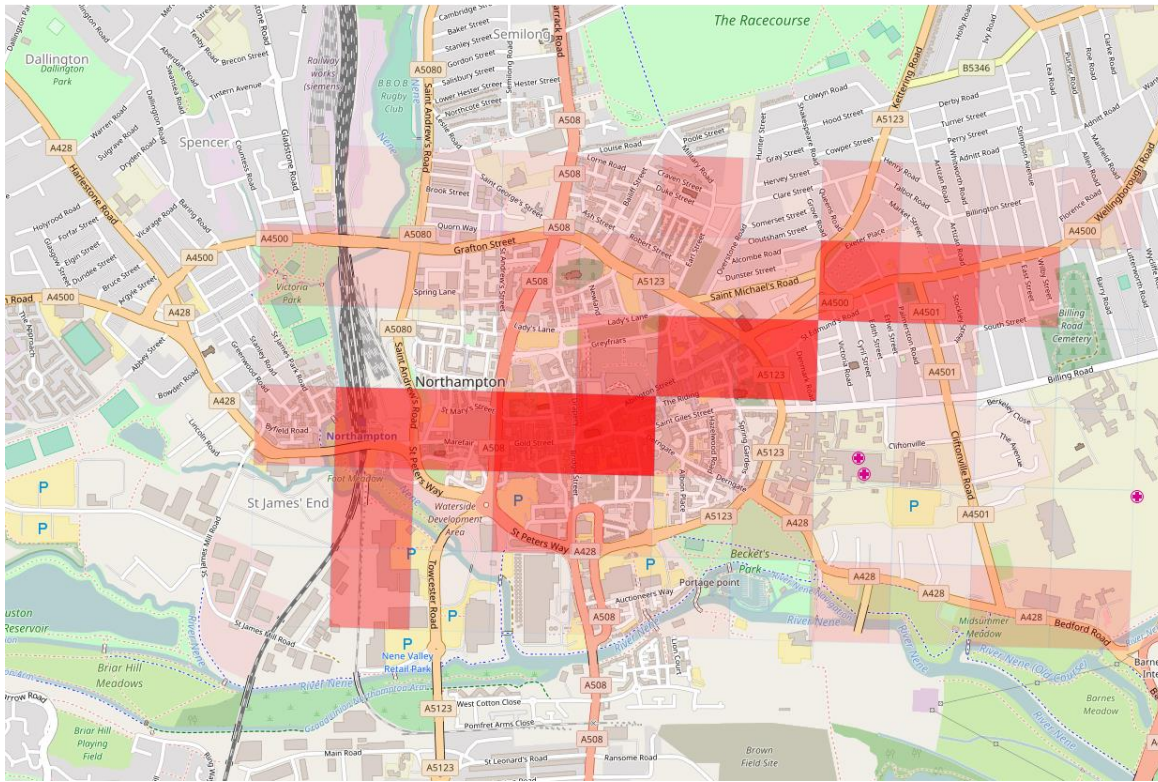
Risk

- We divide the simulation area into grid cells (typically 250x250m – a reasonable area to patrol).
- For each cell, we calculate the total number of seconds that each agent has been present.
- We then multiply agent-second count by average BAC of the agents within the cell.
- This value is used as our simulated risk level.

EVAULATION DATA
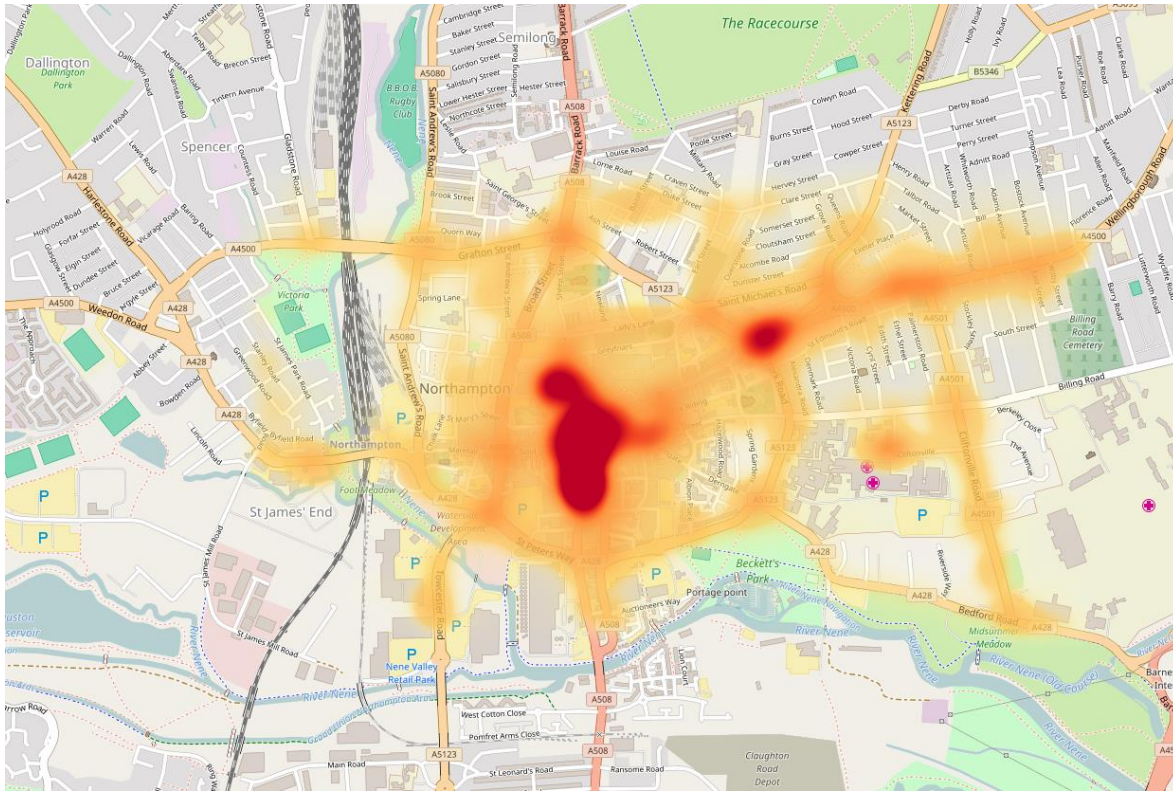
Heatmap of Violence against a Person and Rowdy Behaviour for Saturdays in Northampton.

Crime Data provided by Northampton Police Force (UK).

- Officer location history over 3 years.
  - 2 minute-intervals, includes lat/lon, officer status, crime type
- We consider violence against a person & rowdy behavior (not **necessarily** caused by alcohol).

EVAULUATION METHOD

- Apply same 250x250m grid to ground-truth data.
- Count instances of violence against the person and rowdy behaviour occurring during given weekday and given time window.
- Run Pearson's R and Spearman's ρ correlation measures on cell-wise basis.

OUR SOLUTION: RESULTS

Weekday	Pearson's r	Spearman's ρ	P-value
Mon	0.558	0.565	<0.001
Tue	0.584	0.610	<0.001
Weds	**0.683**	**0.647**	**<0.001**
Thurs	0.651	0.609	<0.001
Fri	**0.664**	**0.632**	**<0.001**
Sat	**0.733**	**0.696**	**<0.001**
Sun	0.565	0.517	<0.001

Results for Northampton City Centre, 10,000 agents, 18:00-04:00

- Reasonable correlation measures
  - $\bar{r}$: 0.634
  - $\bar{\rho}$: 0.611
- Varies by day of week
  - Particular peaks on Wednesday, Friday and Saturday
  - Likely due to Wednesday, Friday and Saturday being busier nights – higher SNR.
    - Friday and Saturday weekend nights, more drinking.
    - Wednesday often social night at Universities.

CONCLUSION

- Our results show a reasonable level of correlation between ground-truth data and simulation output.
- Suggests that agent-based modelling has promise for predicting levels of alcohol-related crime without requiring historical data, with minimal manual data input required.
- Potential for use by police, city planners and Alcohol licensing officials as preventative measure.

FUTURE WORK

- Group Behaviour & Demographics
  - Different drinking styles with different groups/venues, "pre-drinking"
- Routing Strategy
- Venue Capacities
- Taxis, Public Transport (and associated queues)
- Fast-food outlets/hotels
- Incorporate some ideas presented in SimDrink:
  - Influence of Peers, Spending limits
- Learning of agent parameters
  - Need to ensure model still generalises well
- Improve performance evaluation method