Predictive Policing

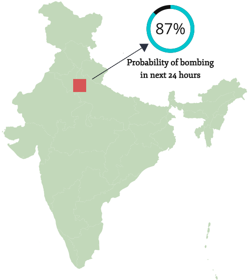
Marked Spatio-Temporal Point Processes for preventing crime

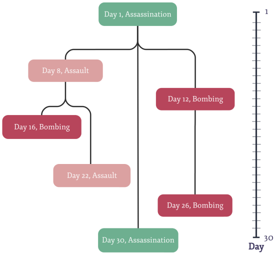
Santhosh Narayanan

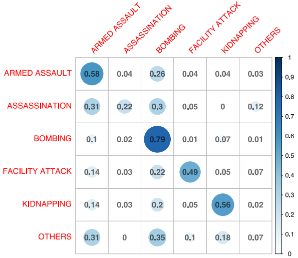
PredPoint Analytics

December 2022

Highlights







1

Prediction

Obtain probabilistic predictions for specific events over any region and time horizon

2

Event genealogy

Identify causal relationships between events by recovering the hidden branching structure

3

Dynamics

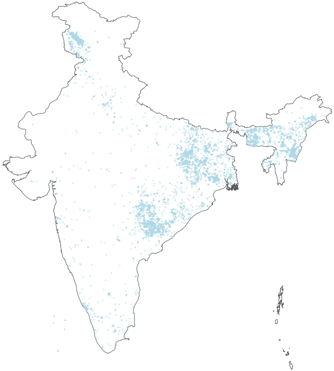
Learn the triggering probabilities between different types of events and how they evolve over time and space

Predictive policing

- Repeat victimisation theory suggests that crimes, like earthquakes exhibit clustering in both time and space
- Model provides probabilistic predictions for the occurrence of specific crime type over any region and time horizon
- Generate prediction maps that identify hot-spots where the probability of a particular type of crime is high
- The estimated event genealogy allows investigators to connect a new crime with previous events

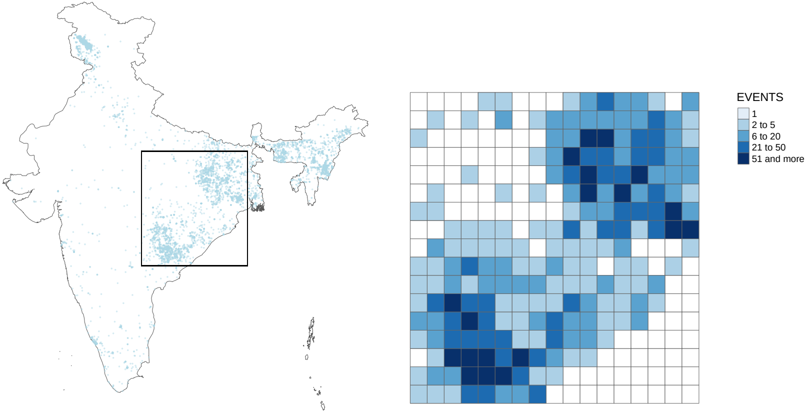
Dataset

	date	longitude	latitude	type
1	2008–04–01	84.371	24.750	ARMED ASSAULT
2	2008–04–01	85.279	23.075	ARMED ASSAULT
3	2008–04–02	84.371	24.750	ASSASSINATION
4	2008–04–03	76.576	33.778	ARMED ASSAULT
5	2008–04–03	74.351	34.552	KIDNAPPING
6	2008–04–04	86.221	24.924	ARMED ASSAULT
7	2008–04–04	91.351	26.524	ARMED ASSAULT
8	2008–04–04	73.698	24.583	KIDNAPPING
9	2008–04–04	91.833	24.197	ARMED ASSAULT
10	2008–04–04	81.892	18.364	ARMED ASSAULT



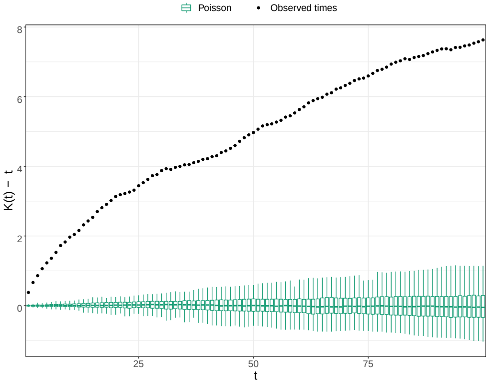
Source: National Consortium for the Study of Terrorism and Responses to Terrorism (START), University of Maryland. (2019). The Global Terrorism Database (GTD). Retrieved from `https://www.start.umd.edu/gtd`

Observation Region



Observation region for experiment and event counts between 1/4/08 and 31/12/19

Second-order analysis

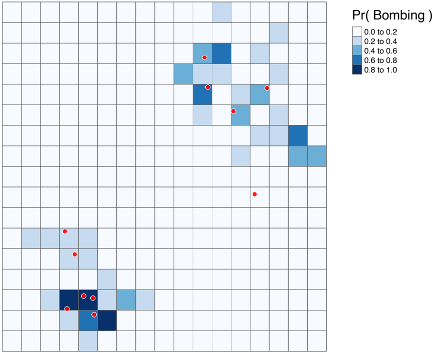


- The K function is an estimate of the surplus number of events compared to random occurrences
- The surplus is significant, confirming events cluster in time

Prediction Framework

- Generate S simulations of the process in the interval $(T, T+d)$
- Given the history upto T, iteratively simulate the next event, its occurrence time, (x, y) and event type
- Add the simulated event to the history as the most recent event
- Stop simulation when the time exceeds $T + d$
- Finally, aggregate event counts from the S simulations and validate against the observed counts

Validation



MODEL VALIDATION

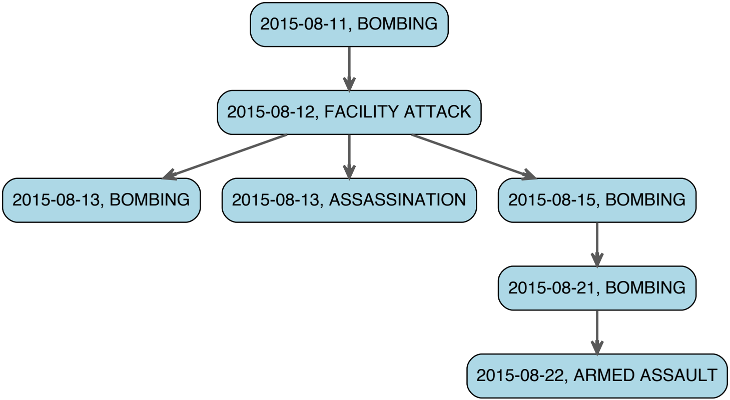
using scoring rules for count data predictions

Score	DTMSTPP	TSGLM	POISSON
logarithmic	10.51	10.72	11.27
brier	-596.53	-596.50	-596.47
spherical	-598.23	-598.22	-598.21
rankprob	1.91	1.92	1.94
dawseb	-3,345.09	-2,993.49	-2,705.75
normsq	562.48	1,035.96	1,286.80
sqerror	2.49	2.52	2.55

The discrete time marked spatio-temporal point process (DTMSTPP) model performs best across all 7 scoring rules

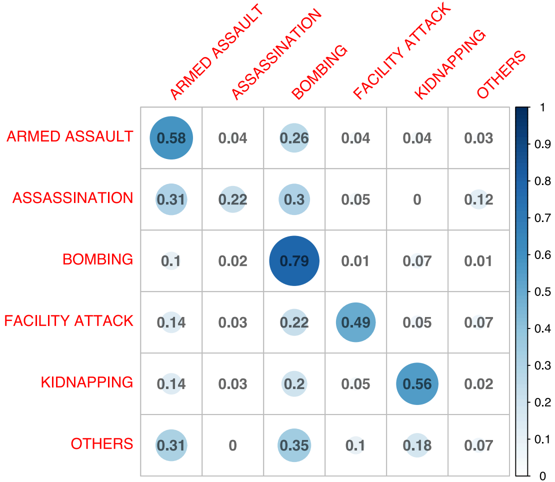
(Left) Prediction Map for the first week in 2018 between 1/1/18 and 7/1/18, based on history up to 31/12/17.

Event Genealogy



Investigating crimes using the triggering structure.

Learning dynamics



Cross-excitation matrix giving the triggering probabilities between event types.

Takeaways

WHAT IS IT?

DTMSTPP is a statistical model that gives probabilistic predictions for the number of occurrences of specific events over any region and time horizon

DOES IT WORK?

DTMSTPP performed better than the TSGLM and Poisson models in all 7 scoring rules validating one-day ahead model predictions over a period of 3.5 years. Also had a higher Prediction Accuracy Index than the Hotspot Mapping technique over the same time period.

WHY IS IT UNIQUE?

DTMSTPP offers more than just predictions. The model provides insight into the dynamics of the underlying phenomena that generated the data. It can also identify causal relationships and hence discover connected events.

References

01 **Time series generalised linear model (TSGLM)**

Kedem, B., & Fokianos, K. (2005). *Regression models for time series analysis.* John Wiley & Sons.

02 **Scoring rules for count data**

Czado, C., Gneiting, T., & Held, *L.* (2009). *Predictive model assessment for count data.* Biometrics, 65(4), 1254-1261.

03 **Prediction Accuracy Index**

Chainey, S., Tompson, L., & Uhlig, S. (2008). *The utility of hotspot mapping for predicting spatial patterns of crime.* Security journal, 21(1), 4-28.

04 **Hotspot Mapping**

Bowers, K. J., Johnson, S. D., & Pease, K. (2004). *Prospective hot-spotting: the future of crime mapping?.* British journal of criminology, 44(5), 641-658.

Thank You

Santhosh Narayanan

+91 85906 76703 stats.santhosh@gmail.com