A Spatio-Temporal Model of Crime Dynamics

Workshop on Social Interactions and Crime

Alex Reinhart

Department of Statistics & Data Science
Carnegie Mellon University

October 20, 2017

Self-exciting point process models

- Statistical models of events in space and time
- The current rate of events depends on location *and* the previous history of events
- Past events can "trigger" new events, usually nearby in space and time

Self-exciting point process models

- Statistical models of events in space and time
- The current rate of events depends on location *and* the previous history of events
- Past events can "trigger" new events, usually nearby in space and time
- Widely used in seismology, now applied in epidemiology... and for crime

Self-exciting point process models

The rate of events λ at a location s and time t is:

$$\lambda(s, t) = \text{spatial factors at } s + \text{recent events near } s.$$

Fitting a model helps us learn the effects of the spatial factors and of self-excitation.

Explaining crime patterns

- Crime tends to concentrate

Explaining crime patterns

- Crime tends to concentrate
- We can find hotspots, but can we explain them?
  - Local spatial factors (poverty, demographics, policing patterns...)
  - Recent history of crime (retaliation, near-repeats)
  - Leading indicator events (minor offenses, calls for service...)

Explaining crime patterns

- Crime tends to concentrate
- We can find hotspots, but can we explain them?
  - Local spatial factors (poverty, demographics, policing patterns...)
  - Recent history of crime (retaliation, near-repeats)
  - Leading indicator events (minor offenses, calls for service...)
- These are usually modeled separately, if at all
- ...but they are confounded when modeled separately.

Explaining crime patterns

- Crime tends to concentrate
- We can find hotspots, but can we explain them?
  - Local spatial factors (poverty, demographics, policing patterns...)
  - Recent history of crime (retaliation, near-repeats)
  - Leading indicator events (minor offenses, calls for service...)
- These are usually modeled separately, if at all
- ...but they are confounded when modeled separately.
- Can we make a single unified model of crime dynamics?

Point process models of crime

- Self-exciting models have been used to model hotspots: events in the hotspot trigger more events nearby (Mohler et al, 2011)
- A conventional hotspot approach (kernel density estimation) finds chronic hotspots, and self-excitation models short-lived ones

Point process models of crime

- Self-exciting models have been used to model hotspots: events in the hotspot trigger more events nearby (Mohler et al, 2011)
- A conventional hotspot approach (kernel density estimation) finds chronic hotspots, and self-excitation models short-lived ones
- Extended to include leading indicators: other types of events which may increase crime rate (Mohler, 2014)
- Commercialized by PredPol, deployed by LAPD and tested in a randomized trial (Mohler et al, 2015)
- No spatial factors or inference tools

A new self-exciting point process model

The rate of crime λ at a location s and time t is:

$$\lambda(s,t) = \text{spatial factors at } s + \text{recent events near } s$$

A new self-exciting point process model

The rate of crime λ at a location s and time t is:

$$\lambda(s, t) = \exp\left(X_s \beta\right) + \sum_{\substack{\text{all events } i \\ \text{before time } t}} g(s - s_i, t - t_i, M_i),$$

where X_s is a vector of spatial covariates depending on the location s and M_i is the *type* of crime i.

$$g(s, t, M) = \frac{\theta_M}{2\pi\omega\sigma^2} \exp(-t/\omega) \exp\left(-\frac{\|s\|^2}{2\sigma^2}\right)$$

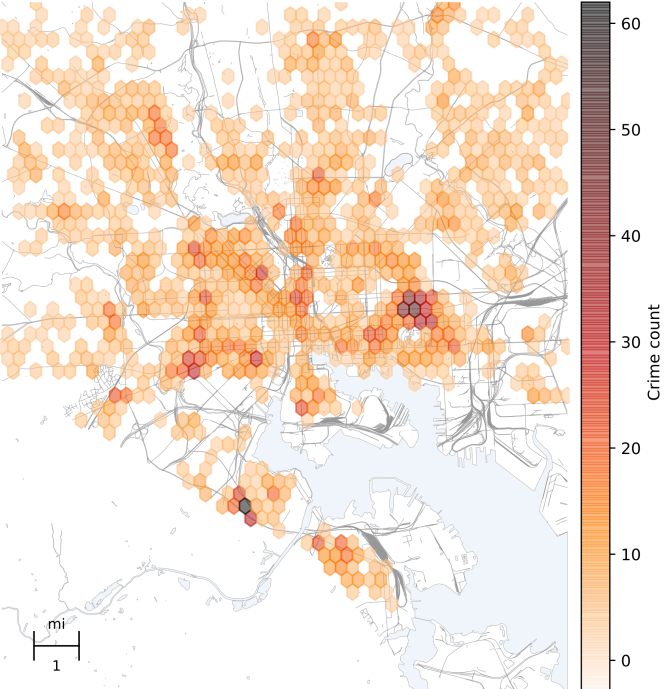
Fitting the model

- Parameters fit by maximum likelihood, via expectation-maximization
- Can obtain confidence intervals for each parameter

Fitting the model

- Parameters fit by maximum likelihood, via expectation-maximization
- Can obtain confidence intervals for each parameter
- Residual maps compare predicted crime rates to actual crime rates
- Animations illustrate the model fit

Baltimore burglary rate



Burglary in Baltimore

- Fit to 7,565 burglaries over one year in Baltimore, using larceny/theft and motor vehicle theft as leading indicators
- Included household density, population age 18–24, poverty, unemployment, and several education variables, across each neighborhood of Baltimore

Burglary in Baltimore

Results:

- The average burglary is part of a spurt of about 2.2 burglaries
- Related burglaries occur within a spatial bandwidth of 375 feet and a temporal decay of 40 days

Burglary in Baltimore

Results:

- The average burglary is part of a spurt of about 2.2 burglaries
- Related burglaries occur within a spatial bandwidth of 375 feet and a temporal decay of 40 days
- Motor vehicle theft is a stronger predictor of burglary ($\theta = 0.16$) than is larceny/theft ($\theta = 0.075$).

Burglary in Baltimore

Results:

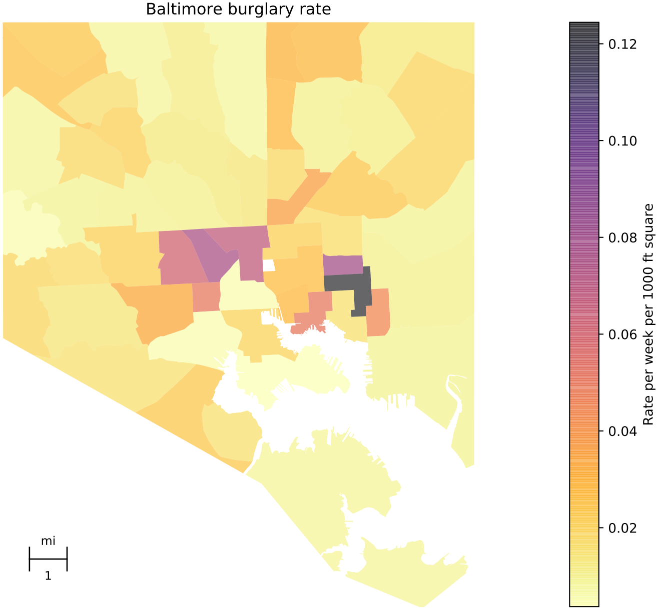
- The average burglary is part of a spurt of about 2.2 burglaries
- Related burglaries occur within a spatial bandwidth of 375 feet and a temporal decay of 40 days
- Motor vehicle theft is a stronger predictor of burglary ($\theta = 0.16$) than is larceny/theft ($\theta = 0.075$).
- Curiously, these results are very similar to those in Pittsburgh—and Atlanta!

Burglary in Baltimore

Covariate	Coefficient	S.E.	exp(Coef)
Intercept	−33.41	0.231	3.1×10^{-15}
Household density	0.037	0.002	1.04
Age 18–24	0.967	0.130	2.63
Household poverty	−0.228	0.065	0.80
Unemployment	0.043	0.120	1.04
$<$ H.S. education	0.476	0.075	1.61
H.S. dropout rate	−0.105	0.246	0.90
H.S. chronic absence	−0.278	0.085	0.76

(Household density unit is 100 households/sq. mi. Other slopes are per 10 percentage points.)

Burglary rate predicted by covariates



Animated hotspots

- Animated hotspot maps can show chronic and acute hotspots, plus model predictions of crime intensity
- Animations reveal very interesting behavior

Animated hotspots

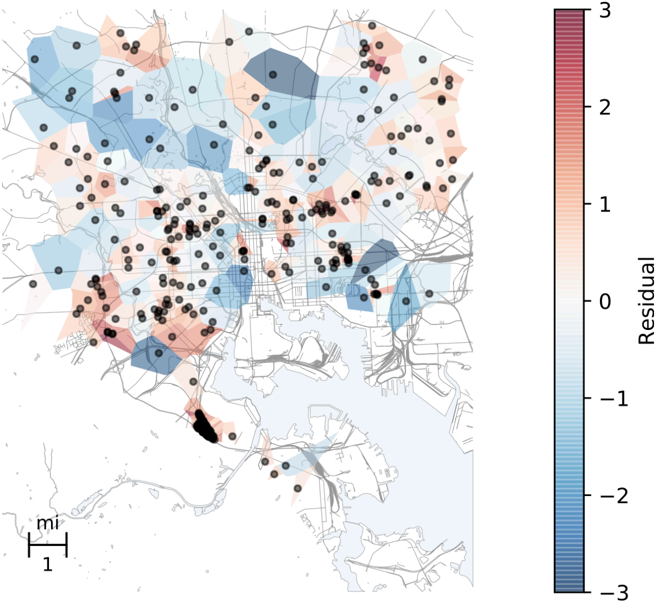
Diagnostics

- Just like in regression, we want to know: does the model fit the data?
- Difficult to understand fit of a spatio-temporal model

Diagnostics

- Just like in regression, we want to know: does the model fit the data?
- Difficult to understand fit of a spatio-temporal model
- But if $\lambda(s, t)$ is the predicted rate of crime, we can compare the predicted rate against the observed incidents
- Maps reveal where crime is systematically mispredicted, suggesting missing covariates
- Animations reveal the appearance and disappearance of hotspots

Burglary residuals
2016-06-01 to 2016-06-15



Modeling multiple cities

- I'm building a Bayesian hierarchical model of several cities
- Allows comparisons of crime dynamics between cities

Modeling multiple cities

- I'm building a Bayesian hierarchical model of several cities
- Allows comparisons of crime dynamics between cities
- I have data from Atlanta, Pittsburgh, and Baltimore
- Interesting results so far, much work still remaining

Where next?

- Self-exciting processes can model the birth and death of crime hotspots
- They unify previously separate parts of crime analysis
- Can test hypotheses about factors associated with crime or events which trigger it
- Can compare factors between cities

Where next?

- Self-exciting processes can model the birth and death of crime hotspots
- They unify previously separate parts of crime analysis
- Can test hypotheses about factors associated with crime or events which trigger it
- Can compare factors between cities
- ...But I am a statistician, not a criminologist or sociologist
- What questions are most interesting?

Thank you

- Thanks to Joel Greenhouse and Daniel Nagin

- Alex Reinhart and Joel Greenhouse, "Self-exciting point processes with spatial covariates: modeling the dynamics of crime." `https://arxiv.org/abs/1708.03579`
- For more on self-exciting point processes: Alex Reinhart, "A Review of Self-Exciting Spatio-Temporal Point Processes and Their Applications," *Statistical Science*. `https://arxiv.org/abs/1708.02647`
- areinhar@stat.cmu.edu

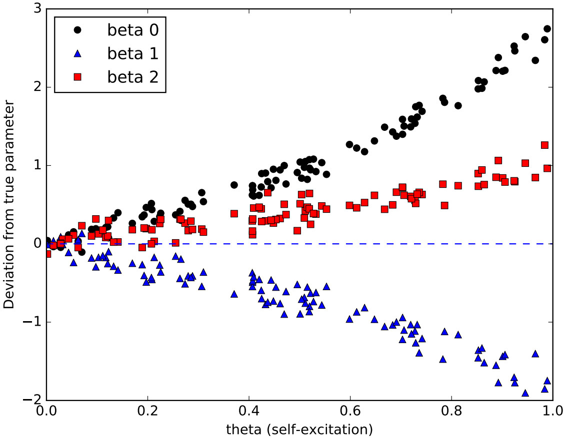
Confounding and regression

Spatial covariates are generically confounded with self-excitation; you cannot ignore one and model the other:

$$\begin{tikzcd}
\text{Covariate 1} \arrow[r] \arrow[d] \arrow[dr] & \text{Covariate 2} \arrow[dl] \arrow[d] \\
\text{Events in } i \text{ at } t-1 \arrow[r] & \text{Events in } i \text{ at } t
\end{tikzcd}$$

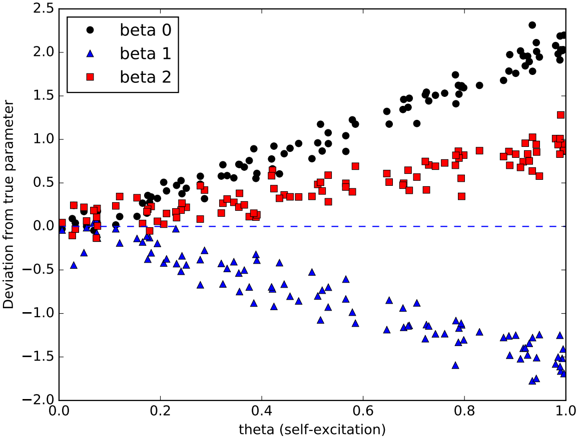
Regression models including lagged crime rates can work, but discretizing the event history throws away information, and bias remains

Self-excitation must be accounted for



Both regression coefficients shrink towards zero as self-excitation increases. The intercept increases.

Lags aren't enough



Even including several lagged crime counts, the coefficients are still biased towards zero as self-excitation increases.