Predicting criminal behavior with truncated Lévy flights using real data from Bogota

Mateo Dulce Rubio
m.dulce10@uniandes.edu.co
Advisor: Álvaro J. Riascos Villegas

LXAI Workshop at NeurIPS 2018

December 8, 2018

Objective

I use residential burglaries data from Bogota, Colombia, to fit an agent-based model following truncated Lévy flights. I found a positive effect of repeat/near-repeat victimization and broken windows theory on the probability of future occurrence of crimes, without losing predictive accuracy against other state-of-the-art crime prediction models. This exercise sheds light on criminal behaviour, hotspots' dynamics, and efficient assignment of police resources.

Why bother at all?

- Crime is a persistent problem in modern cities and Bogota is not the exception. It is not uniformly distributed but presents spatio-temporal clustering patterns.
- Between 2012 and 2015, 2% of Bogota's streets accounted for all homicides and a quarter of all crimes reported, but they received less than 10% of police time and limited public services.

Why bother at all?

- *Predictive Policing* arose as the use of statistical inference and machine learning techniques to identify vulnerable crime areas.
- All of this models assume crime as a random event and attempt to estimate the underlying process of criminal distribution, but fails giving insights on the rational behavior of offenders.
- Agent-based models contribute to the literature of criminal rationality, quantifying the effect of repeat/near-repeat victimization and broken windows theory to the future nearby occurrence of crimes.

Literature

- Relies on Becker's economic approach to crime (1968), and two main mechanisms.

    1. Burglars often prefer repeat or near-repeat victimization partially because is where they already have good information of the types of properties that might be stolen and the routines of the police and their inhabitants.

Literature

- Relies on Becker's economic approach to crime (1968), and two main mechanisms.

  1. Burglars often prefer repeat or near-repeat victimization partially because is where they already have good information of the types of properties that might be stolen and the routines of the police and their inhabitants.
  2. Past occurrence of crime in a certain area creates an atmosphere of lawlessness and crime-tolerant region that encourages the occurrence of more crimes: *broken windows* effect.

Repeat victimization

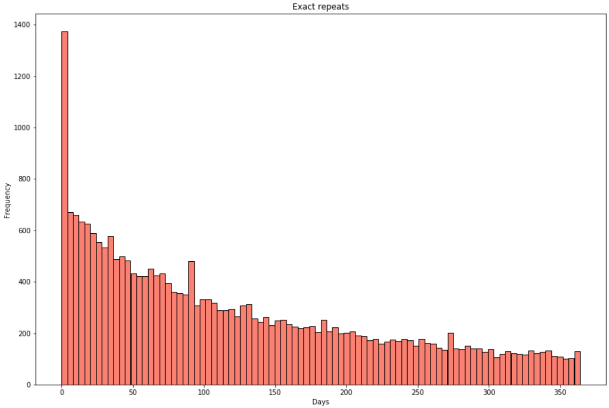

Figure: Histogram between burglary events (1 year)

Methodology

- Agent-based model where the city is represented as a two-dimensional lattice where each vertex $s = (i, j)$ represents a house with an *attractiveness level* $A_s(t)$.

- Attractiveness index displays the benefit of burglarize the house perceived by the criminal:

$$A_s(t) = A_s^0 + B_s(t),$$

where A_s^0 is a static, but possibly spatially heterogeneous, component and $B_s(t)$ varies with interactions of house s with burglars.

Methodology

- $B_s(t)$ increases θ each time the house is burglarized. This increment only affects the attractiveness level for a finite time:

  $$B_s(t + \Delta t) = B_s(t)(1 - \omega\Delta t) + \theta E_s(t).$$

  $E_s(t)$ is the total number of burglaries at site s in the time interval starting at t, and ω accounts for the time span repeat victimization is more likely to occur.

Methodology

- $B_s(t)$ spread spatially to adjacent sites modeling near-repeat victimization and broken windows theory:

$$B_s(t+\Delta t) = \left[(1-\eta)B_s(t) + \frac{\eta}{4}\sum_{s'\sim s} B_{s'}(t)\right](1-\omega\Delta t)+\theta E_s(t).$$

- $\eta \in [0,1]$ measures this spreading effect.

Methodology

- Burglars move in the lattice following a truncated Lévy flight biased by the attractiveness level of the houses.
- In each time interval criminals can either burglarize the house where they are located, or move to another place.
- Lévy flight: random walk with step-lengths obeying a heavy-tailed probability distribution.

Methodology

- Probability of a criminal burglarizing a house s in a time interval of size Δt follows a standard Poisson process:

$$p_s(t) = 1 - e^{-A_s(t)\Delta t}.$$

- Criminals that commit burglary are removed from the lattice. Moreover, burglars are generated with rate Γ at each grid vertex.

- If a criminal do not perpetrate a crime, he moves to another house following a truncated Lévy flight, with a limited jump range reflecting a limited traveling distance, biased toward regions with high attractiveness.

Methodology

- The probability to move from a house s to a house r is equal to

$$q_{s\to r} = \frac{w_{s\to r}}{\sum_{s'\in\mathbb{Z}^2, s'\neq s} w_{s\to s'}},$$

with

$$w_{s\to r} = \begin{cases} \frac{A_r}{||s-r||^{\mu}} & ||s-r|| < L \\ 0 & ||s-r|| \geq L \end{cases}$$

for a given L (larger jump allowed) and μ the exponent of the underlying Lévy flight.

Methodology

- Finally, the model is described by two equations:

$$B_s(t+\Delta t) = \left[(1-\eta)B_s(t) + \frac{\eta}{4} \sum_{s' \sim s} B_{s'}(t) \right] (1-\omega\Delta t)+\theta E_s(t),$$

$$n_s(t+\Delta t) = \sum_{\substack{r \in \mathbb{Z}^2 \\ ||s-r||<L}} [n_r(t) - E_r(t)]q_{r \to s} + \Gamma \Delta t.$$

- The former gives the dynamics of the attractiveness index while the later the criminal activity.

Effect of L

- $L = 1$: Random (Gaussian) walk. Could led to a high number of small hotspots due to the impossibility of criminals to visit farther sites of the city, and limited knowldedge only of neighboring sites.

Effect of L

- $L = 1$: Random (Gaussian) walk. Could led to a high number of small hotspots due to the impossibility of criminals to visit farther sites of the city, and limited knowldedge only of neighboring sites.

- $L = \infty$: Pure Lévy flight. Could led to an opposite regime of aggregation: due to the possibility of traveling long distances, all criminals might displace to the most attractive neighborhoods generating few static hotspots.

Effect of L

- $L = 1$: Random (Gaussian) walk. Could led to a high number of small hotspots due to the impossibility of criminals to visit farther sites of the city, and limited knowldedge only of neighboring sites.
- $L = \infty$: Pure Lévy flight. Could led to an opposite regime of aggregation: due to the possibility of traveling long distances, all criminals might displace to the most attractive neighborhoods generating few static hotspots.
- $1 < L < \infty$: Truncated Lévy flights model.

Computer simulations

Parameter	Definition
l	Grid space.
Δt	Time step.
μ	Exponent underlying Lévy flight.
L	Larger jump allowed.
A^0	Static intrinsic attractiveness.
Γ	Rate of burglar generation.
ω	Dynamic attractiveness decay rate.
θ	Increase in attractiveness due to crimes.
η	Triggering effects ($\in [0, 1]$).

Table: Parameters of the discrete model.

Data assimilation

- The goal is to find the parameters (ω, θ, η) that maximize the probability of obtaining the known crime sample $\{(s_k, t_k)\}_{k=1}^{N}$.
- The dynamic quantities of interest (burglars' location and house attractiveness) are unobserved. Instead, we have data on the times and locations of actual crimes, which are a function of these two quantities.

Data assimilation

- Log-likelihood function:

$$\mathbb{L}(\text{data}|\Theta) = \sum_{k=1}^{N} \log(n_{s_k}(t_k)A_{s_k}(t_k)) - \int_0^T \int_X \int_Y nA \; dxdydt$$

- Interpretation: the product nA must be large where crimes actually occurred at the time they happened, but not with an arbitrarily large number of expected crime (spatio-temporal integral) over the lattice.

Data

- The dataset correspond to georeferenced residential burglaries in Bogota in 2012-13 collected by SIEDCO system from Metropolitan Police.

- The model was trained with six months data, and its predictive accuracy tested using the corresponding following month.

Training procedure

Parameter	Definition	Value
Δx	Grid x space	180 meters
Δy	Grid y space	320 meters
Δt	Time step	1 day
μ	Exponent Lévy flight	2.5
L	Larger jump allowed	10 Km.

Table: Model parameters

Training procedure

- Six months training window was divided in two subgroups:

  1. First three months were used to estimate the static attractiveness component and the burglars generation rate, $(A_s^0,\ \Gamma)$.
  2. The following three months were used to estimate the parameters (ω, θ, η) using MLE.

Training procedure

- In a steady state with constant number of criminals, the number of burglars removed from the system each time step due to burglary events must equals the number of criminals generated. Γ was set to the average daily number of crime events in Bogota in the first three months of the training set.
- A_s^0 was estimated using KDE fitted to the locations of the crime events occurring during the first three months of the training dataset, such that each house has an static attractiveness proportional to its historical criminal rate.

Training procedure

- Given $(A_s^0,\ \Gamma)$ and a criminal located at each vertex, the system evolves with $E_s(t)$ coming from the training data via:

$$B_s(t+\Delta t) = \left[(1-\eta)B_s(t) + \frac{\eta}{4}\sum_{s'\sim s} B_{s'}(t)\right](1-\omega\Delta t) + \theta E_s(t),$$

$$n_s(t+\Delta t) = \sum_{\substack{r\in\mathbb{Z}^2 \\ ||s-r||<L}} [n_r(t) - E_r(t)]q_{r\to s} + \Gamma\Delta t,$$

- Then, use MLE to find the parameters (ω, θ, η).

Test procedure

- From the final state of the training process, the system evolves according to:

$$B_s(t+\Delta t) = \left[(1-\eta)B_s(t) + \frac{\eta}{4}\sum_{s'\sim s} B_{s'}(t)\right](1-\omega\Delta t)+\theta \boldsymbol{p_s(t)n_s(t)},$$

$$n_s(t+\Delta t) = \sum_{\substack{r\in\mathbb{Z}^2 \\ ||s-r||<L}} [\boldsymbol{1-p_r(t)}]\boldsymbol{n_r(t)}q_{r\to s} + \Gamma\Delta t$$

- Replace $E_s(t)$ by $p_s(t)n_s(t)$.

Test procedure

- At each time interval of the test setting, the $\mathrm{x}\%$ of the cells with higher expected numbers of crimes (nA) are marked as hotspots.
- Hit Rate as a measure of predictive accuracy:

$$\text{Hit Rate} = \frac{\#\text{ of crimes in predicted HS}}{\text{Total }\#\text{ of crimes}}.$$

- Cumulative Accuracy Profile: number of positive outcomes vs. the classifying parameter.

Results: Cumulative Accuracy Profile

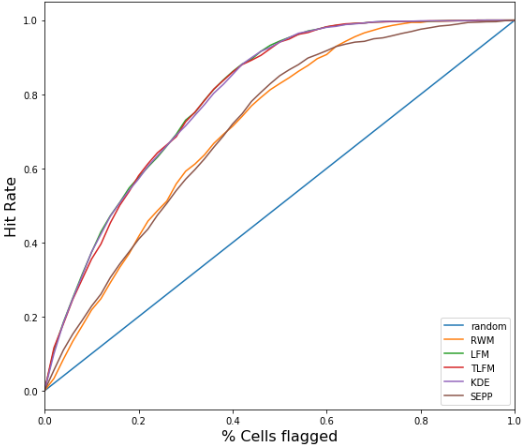

Figure: Average Cumulative Accuracy Profile curve for different crime prediction models.

Results: Cumulative Accuracy Profile

	RWM	**LFM**	**TLFM**	**KDE**	SEPP
AUC	0.7152 (3.3e-4)	**0.7993** (5.3e-4)	**0.7969** (4.8e-4)	**0.7973** (5.8e-4)	0.7132 (0.0015)

Table: Area under average CAP curve for different crime prediction models. Standard deviation in parentheses.

Results

ω	θ	η
0.004 (1.4e-5)	5.5133 (0.8974)	0.9488 (7.4e-5)

Table: Estimated parameters Truncated Lévy flights model. Standard deviation in parentheses.

Results

- Decay rate $\omega = 0.004$ evidences a long time span when repeat/near-repeat victimization has an increased likelihood of occurrence: each day the dynamic attractiveness component decays 0.4%.

- Increase in attractiveness due to crimes parameter $\theta = 5.5133$ exhibits a bias of criminals to return to previously burglarized houses.

- Triggering effect $\eta = 0.9488$ shows a strong and quickly spreading of attractiveness to neighboring sites: validates broken windows theory and self-exciting nature of crime.

Conclusions

- Repeat/near-repeat victimization and broken windows theory increase the probability of new crimes in a spatio-temporal neighborhood of an initial crime.
- Agent-based models elucidate criminal rationality without losing predictive accuracy against other state-of-the-art crime prediction models.
- Public policies seeking to reduce criminal activity and its negative consequences must take into account these mechanisms and the self-exciting nature of crime, to effectively make criminal hotspots safer.

(Principal) References:

- Martin B Short, Maria R D'Orsogna, Virginia B Pasour, George E Tita, Paul J Brantingham, Andrea L Bertozzi, and Lincoln B Chayes. A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences*, 18(supp01):1249–1267, 2008.
- Sorathan Chaturapruek, Jonah Breslau, Daniel Yazdi, Theodore Kolokolnikov, and Scott G McCalla. Crime modeling with Lévy flights. *SIAM Journal on Applied Mathematics*, 73(4):1703–1720, 2013.
- Pan Chaohao, Li Bo, Wang Chuntian, Zhang Yuqi, Geldner Nathan, Wang Li, & Bertozzi Andrea L. Crime Modeling with Truncated Levy Flights for Residential Burglary Models. *Mathematical Models and Methods in Applied Sciences*, 2018.
- Lloyd David J.B., Santitissadeekorn Naratip, & Short Martin B. Exploring data assimilation and forecasting issues for an urban crime model. *European Journal of Applied Mathematics*, 27(3), 451-478, 2016.