Burglary Pattern Prediction

Italo Sayan, Nathan Raw

April 2018

1 Problem

In the last two months, the Rochester Police Department has recorded two hundred burglaries. Sixty of them were illegal entries to single family homes [2]. The city of Rochester is 7% above the 2016 national average estimates of burglaries per capita [1]. One way to attack the issue is patrol allocation. Currently, precincts use a crime analyst to decide hot-spots and allocate units accordingly. However, learning algorithms can offer a more systematic approach to this problem.

On average, at both national and local levels, governments designate the third part of their budgets on policing[4]. We believe police departments should spend their budgets on evident-based solutions.

2 Introduction

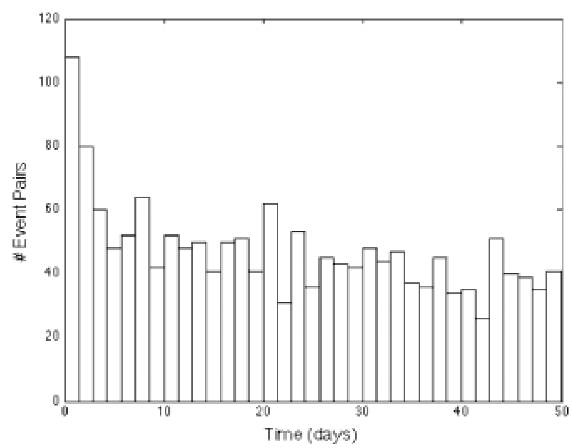
Burglaries, earthquakes, and tweets all have a particular characteristic in common. The occurrence of one event increases the probability of subsequent events. Earthquakes can produce aftershocks, tweets can produce subsequent re-tweets, and burglaries follow the same behavior. Academic efforts have shown that a burglary elevates the risk of another burglary occurring in nearby areas [5].

Epidemic type aftershock sequence(ETAS) models have already been adapted from seismology to produce habitation burglary prediction models [7]. In criminology, they are known as Self Exciting Point Processes(SEPP) models. Crimes, earthquakes, and tweets all follow self exciting behavior. Thus, it is possible to leverage previous literature to implement a tool for police departments. But we have not found an open source implementation of the new technology.

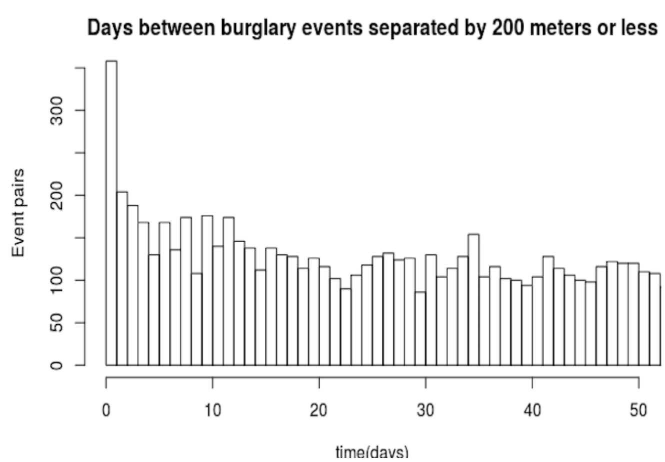
First, it is important to understand how SEPP models work. Then, we will breakdown how to apply them on burglary prediction using R code. Finally, we will produce a web app to serve a predictive crime map. Our main contribution is the programmatic application of George Mohler's SEPP model. For serving the map, we will use Google Maps API. Data from the San Antonio Police department will be used for modelling purposes. Randomized controlled field trials have been conducted on SEPP models [6]. The Los Angeles Police Department(USA) and the Kent Police Department(UK) collaborated with researchers to compare heatmaps and SEPP predictive powers. Results show that SEPP models predict 1.4-2.2 times as much crime compared to a dedicated crime analyst. Police patrols using SEPP forecasts led to a average 7.4% reduction in burglary volume as a function of patrol time [7].

3 Self-Exciting Point Process Models

Heatmaps are a common software tool used by crime analysts. Depending on the frequency of crime, heatmaps highlight high risk areas. Normally, count is expressed in a color scheme from green to red. It's not a bad approach, but we can do better than that. Heatmaps are not taking advantage of the self exciting nature of burglaries; meaning that an initial burglary can cause offspring. The repetition or 'self-exciting' property of crime has been previously visualized in other literature using data from the city of Los Angeles [7].



The histogram plots burglary pairs that are inside a 200 meter radius. As shown, the pairs happen mostly on short spans of time. We were able to replicate the results using data from San Antonio.



Both histograms reveal that burglaries indeed produce offspring. Locations close to recent burglaries have a higher risk of another burglary occurring [5]. In other words, crimes are indeed self exciting. Three concepts need to be clear to understand how to use SEPP models. The intensity or risk function, kernel density estimation, and stochastic declustering. We will give quick overview.

3.1 Intensity Function

The following function provides a new way to count the intensity of crime on a given location.

$$\lambda(t, x, y) = \mu(x, y) + \sum_{t^*=1}^{T} g(t - t^*, x - x^*, y - y^*)$$

λ= Intensity function [8].

x= Longitude

y= Latitude

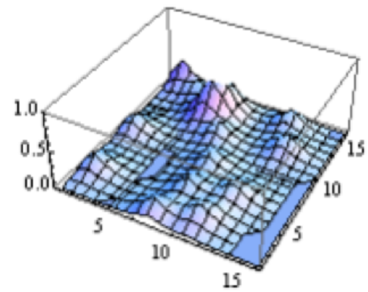
$*$= Datetime, longitude, and latitude of burglaries from the most recent(t^*) to the oldest(T).

The intensity or risk score(λ) is divided in two parts. The μ function represents counts of crimes, just like a heatmap. The g function increases the risk on points where nearby burglaries have low time-distance, latitude-distance, and longitude-distance. If we try to assess the risk of a crime occurring in a location, we sum over the values of g across all of the crimes in our dataset. We still have to show how μ and g are estimated but lets start by analyzing the form of the λ function.

The difference with heatmaps is that SEPP models take into consideration if nearby burglaries happened **recently**. Another key improvement is that SEPP models allow us to separate between **background** events and **offspring** events. If we put this in terms of seismology, it compares to differentiating between new earthquakes and aftershocks.

3.2 Kernel Density Estimation

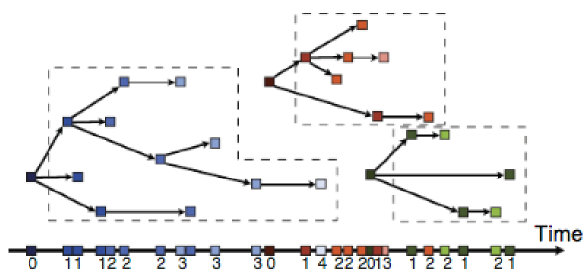
Part of the challenge of using SEPP models is the estimation of μ and g. Kernel density estimation(KDE) methods fit our purposes. While histograms stack rectangles to indicate the frequency of an observation, KDE stacks standard normal distributions to have a smoothed density estimate of a random variable. In our case, we are estimating densities given location data(μ) and densities given inter-point distances(g).



The graph above shows an example of KDE on a 2D visualization. In the case of μ, the axes are latitude and longitude. In the case of g, time-distance, latitude-distance, and longitude-distance are the axes. Both density lumps are added together over time to calculate the λ intensity function.

3.3 Stochastic Declustering

In order to separate background events from offspring events we use stochastic declustering [9]. The following figure displays the branching structure of example events [3].



Over time some events cause offspring. The declustering procedure is achieved as follows. Given a matrix P, we define Pii as the probability of burglary i being a background event. We also define Pij as the probability that burglary j triggered event burglary i.

$$Pii = \frac{\mu(t_i, x_i, y_i)}{\lambda(t_i, x_i, y_i)}$$

$$Pji = \frac{g(t_i - t_j, x_i - x_j, y_i - y_j)}{\lambda(t_i, x_i, y_i)}$$

Next, using an iterative procedure we execute the following n times:

Step 1) Sample background events and offspring/parent inter-point distances.

Step 2) Estimate μ and g from the sampled data using KDE.

Step 3) Update P_n from μ and g using the $P_{n=0}$.

We repeat this procedure until the branching structure converges. Mohler performed 75 iterations on the LAPD dataset [7].

4 Modelling Using R

Our implementation of the intensity function, stochastic declustering, and adaptive kernel density estimation is available online. We couldn't find a comprehensive package that applied self exciting models to crime. Therefore, we decided to implement the three components from scratch. The implementation is directly translated from SEPP literature to R.

5 Web App Visualization

Our current implementation is also available online. The visualization shows how Google maps is used. The R code models the burglary. Then, the MapApp uses Python's Flask and the Maps API to display the burglary predictions from the model.

6 Impact on Public Safety

As RIT students, we pride ourselves in being a part of an institution that stays on the bleeding edge of research. Moreover, we find it imperative to apply this research in the real world to help law enforcement officers make our community safer. This project leverages new analytic methods for burglary analysis and allows police departments to better predict crime. In future work, we want to replicate this modelling methodology using Rochester's data. We also aim to stream live data from the Rochester Police Department's API to provide current estimates of future crimes. By doing so, the expected impact is that RPD will be able to predict 1.4 times more burglaries than with previous efforts. The world is a dynamic and ever-changing place. Unfortunately, so is the world of crime. This technology will help law enforcement officers adapt to changes in crime dynamically. We hope our work will help continue to make Rochester a better, and safer, place to live.

References

[1] "Crime in United States". In: *Crime in United States* (Feb. 2017). URL: `https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016`.

[2] Rochester Police Department. "RPD Open data portal". In: *RPD Open data portal* (Apr. 2018). URL: `http://data-rpdny.opendata.arcgis.com/`.

[3] Vladimir Filimonov and Didier Sornette. "Quantifying reflexivity in financial markets: Toward a prediction of flash crashes". In: *Physical Review E* 85.5 (2012), p. 056108.

[4] "Freedom to Thrive: Reimagining Safety & Security in Our Communities". In: *The Center for Popular Democracy* (2018). URL: http://populardemocracy.org/news/publications/freedom-thrive-reimagining-safety-security-our-communities.

[5] Shane D Johnson. "Repeat burglary victimization: a tale of two theories". In: *Journal of Experimental Criminology* 4.3 (2008), pp. 215–240.

[6] George O Mohler et al. "Randomized controlled field trials of predictive policing". In: *Journal of the American Statistical Association* 110.512 (2015), pp. 1399–1411.

[7] George O Mohler et al. "Self-exciting point process modeling of crime". In: *Journal of the American Statistical Association* 106.493 (2011), pp. 100–108.

[8] Christopher Stover. "Intensity Function". In: *From Math World–A Wolfram Web Resource* (). URL: http://mathworld.wolfram.com/IntensityFunction.html.

[9] Jiancang Zhuang, Yosihiko Ogata, and David Vere-Jones. "Stochastic declustering of space-time earthquake occurrences". In: *Journal of the American Statistical Association* 97.458 (2002), pp. 369–380.