Exciting Adventures in Crime Linkage

Michael D. Porter

Engineering Systems and Environment
School of Data Science
Darden School of Business
University of Virginia

mdp2u@virginia.edu

Outline

1. Crime Linkage
2. Near Repeat Crime Patterns
3. Combining Linkage and Hawkes
4. (time permitting) Spatial event hotspot prediction using multivariate Hawkes features

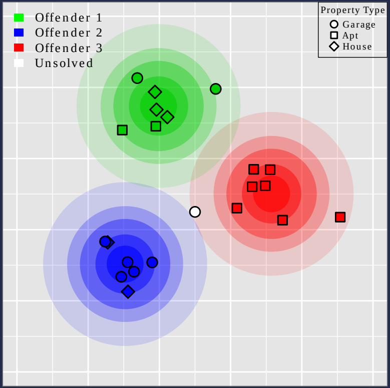
Crime Linkage

Crime Linkage

The objective of criminal linkage analysis is to group crime events that share a common offender (or group of offenders).

- Using the characteristics and features of the *crime*, *crime scene*, or *offender* to estimate linkage probability

- Combine evidence from multiple crime scenes
- Input to geographic profiling systems
- Input to next-event prediction systems
- Resource allocation (patrol routing)
- Interrogations
- Legal evidence



Types of Crime Linkage

Types of Linkage:

Pairwise Case Linkage: Determine if two crimes share a common offender

Crime Series Clustering: Discover groups of crimes that share a common offender.

Crime Series Identification: Discover other crimes that are part of an existing crime series.

Suspect Prioritization: Rank suspects for an existing crime series.

Burglary Problem

According to the FBI, in the US in 2010:

- An estimated 2,159,878 burglaries
- Victims of burglary offenses suffered an estimated $4.6 billion in lost property
- Arrests were made in only 12.4% of burglaries

FBI: http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/property-crime/burglarymain

Why Crime Linkage is Difficult

1. Too many crimes
   - In Seattle WA, $\binom{7102}{2} = 25,215,651$ burglary crime pairs for an analyst to compare in 2014
   - Burglars may also commit other crimes
   - Often crime linkage is a manual process (but see new NYPD system *Patternizr*)
2. Need to consider not only the similarity between crimes, but also the distinctiveness of the crimes
   - If all burglars had same M.O., then we couldn't distinguish their crimes

Why Crime Linkage is Difficult

Need to consider not only the similarity between crimes, but also the distinctiveness of the crimes

- If all burglars had same M.O., then we couldn't distinguish their crimes

Are these two crimes linked (share common offender)?

Incident Reports

Event	Crime Type	Date	Time Range	Address	Target	Items Stolen	POE	MOE
V_1	Burglary	3/15/2010	800-1100	310 Main St.	Apt-1st floor	Jewelry	Window	Forced-Broken Window
V_2	Burglary	3/17/2010	1100-1900	420 1st St.	Apt-1st floor	Cash	Window	Window Open

Pairwise Case Linkage Hypotheses

Casting the case linkage problem in terms of a hypothesis test

$$\mathcal{H}_L : O_i = O_j \quad \text{(Common Offender)}$$
$$\mathcal{H}_U : O_i \neq O_j \quad \text{(Different Offenders)}$$

we can formally quantify our uncertainty about the unknown model parameters using probability distributions.

Likelihood Ratios

The two competing hypotheses can be compared via the posterior odds

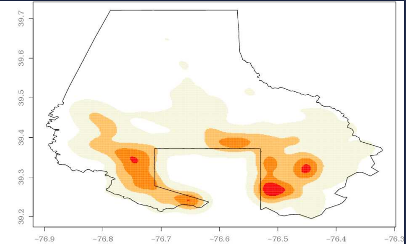
$$\underbrace{\frac{\Pr(\mathcal{H}_L \mid \text{Evidence})}{\Pr(\mathcal{H}_U \mid \text{Evidence})}}_{\text{Posterior Odds}} = \underbrace{\frac{\Pr(\text{Evidence} \mid \mathcal{H}_L)}{\Pr(\text{Evidence} \mid \mathcal{H}_U)}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{\Pr(\mathcal{H}_L)}{\Pr(\mathcal{H}_U)}}_{\text{Prior Odds}}$$

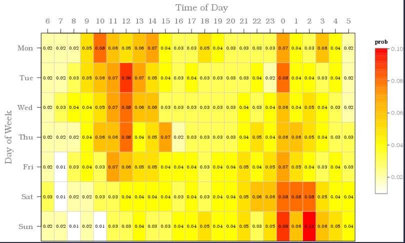
The Likelihood Ratio* offers a formal and explicit way to measure the similarity between events while accounting for the background crime process.

$$\text{LR} = \frac{\Pr(\text{Evidence} \mid \mathcal{H}_\text{L})}{\Pr(\text{Evidence} \mid \mathcal{H}_\text{U})} = \frac{\text{Similarity Measure}}{\text{Distinctiveness Measure}}$$

*Under certain conditions this is equivalent to the *Bayes Factor*

Crime Data Summary - Breaking & Entering





Property Type (34 levels)

Other	Single Home	Apt/ Condo	Yard	Row/ Town	Shed/ Garage
27%	24%	13%	13%	12%	11%

Point of Entry (8 levels)

Door	Window	None	Other	Missing
45%	21%	8%	6%	19%

Method of Entry (16 levels)

No Force	Other	Forced	Pried	Broke Glass	Missing
28%	20%	16%	10%	9%	17%

Censoring time interval (hours)

Exact	≤ 1	≤ 6	≤ 12	≤ 24	≤ 48	≤ 120
20%	32%	48%	67%	83%	87%	94%

Evidence Variables

Evidence variables are created that measure the similarities or dissimilarities between the attributes of two crimes.

Convert crime pairs to evidence variables

- `spatial`- Euclidean distance (km)
- `temporal` - temporal proximity (days)
- `tod` - time-of-day difference (hours)
- `dow` - day of week difference (days)
- `prop` - property type match indicator
- `poe` - point of entry match indicator
- `moe` - method of entry match indicator

Training Data: Evidence Variables

ID_i	ID_j	spatial	temporal	tod	dow	prop	poe	moe	label	weight
2459	2532	3.20	34.60	8.10	0.40	0	0	0	unlinked	1.00
33	35	7.10	1.10	3.50	1.10	1	1	0	linked	0.33
1689	1845	12.90	50.80	4.30	1.80	0	1	0	unlinked	1.00
159	947	14.10	256.40	6.00	1.90	0	1	1	linked	0.00
559	997	14.60	112.30	6.30	0.30	0	1	0	linked	0.00
306	1485	15.30	360.70	6.60	3.30	0	1	0	unlinked	1.00
...	...	...	...	...	...	...	...	...	...	...

*Used expected absolute difference for interval censored times.

Pairwise Crime Linkage

Use the *solved crimes* as training data to construct binary classification models.

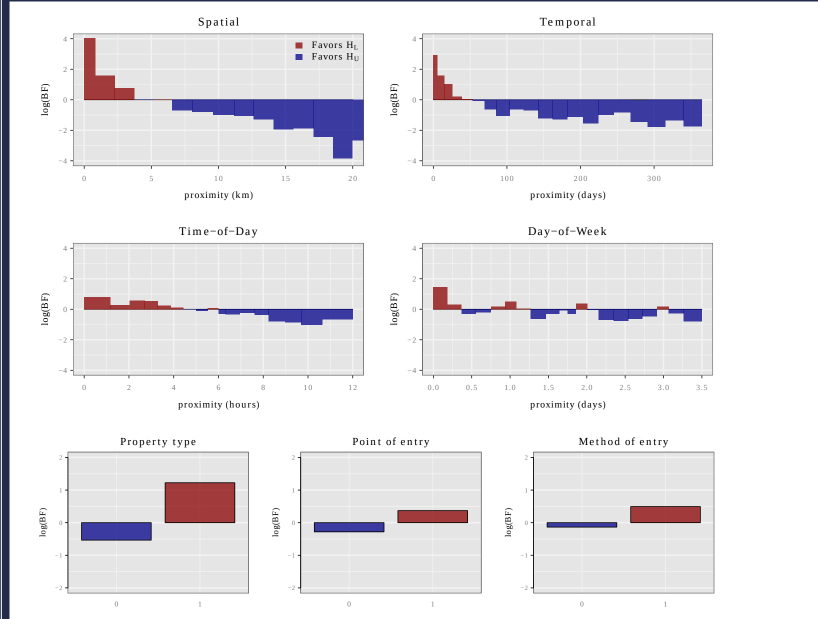
- Logistic Regression:

$$\text{logit}\left\{ \Pr(i, j \text{ are linked} \mid \text{Evidence})\right\} = \beta_0 + \beta_1 X_1(i,j) + \ldots + \beta_p X_p(i,j)$$

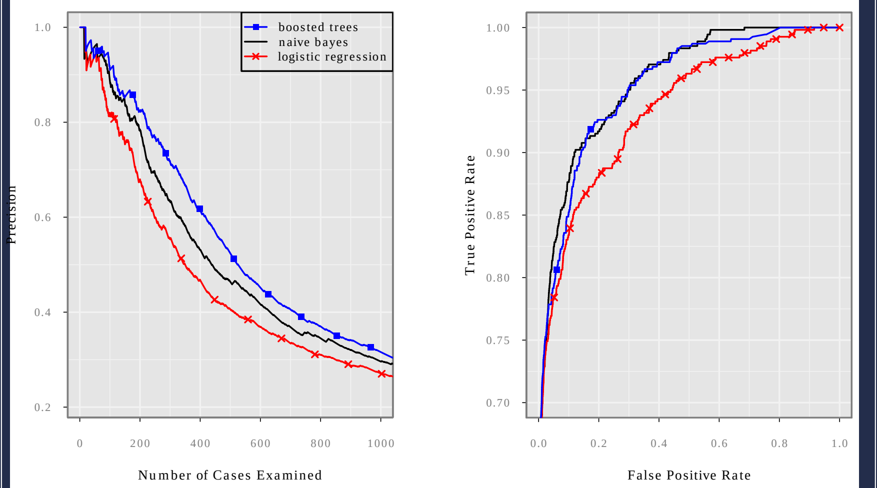
- Naive Bayes

$$\text{logit}\left\{ \Pr(i, j \text{ are linked} \mid \text{Evidence})\right\} = \alpha + \log(\text{LR}_1) + \log(\text{LR}_2) + \cdots + \log(\text{LR}_\text{p})$$

Naive Bayes Component Plots



Pairwise Case Linkage Results



Near-Repeat Crime Patterns

Near-Repeat Crime Patterns

The literature on near-repeats suggests two primary hypotheses for why so many crime pairs are close in space and time

- Flag Account Some locations are attractive (flagged) to a wide range of opportunistic offenders. So repeats and near-repeats are due to multiple offenders choosing their locations and times independently.
- Boost Account The risk of locations near recent crimes is boosted because the same offender (or associates) is likely to strike again in a nearby region (due to experience gained, foraging, etc.).

In reality, both of these concepts can help us model and predict future crime risk.

Modeling Near-Repeat Behavior

- The boost account suggests that after a crime, the risk of future crime in nearby regions will be elevated (boosted) for a short time.
- In other words, the occurrence of crime promotes more crime; the process *excites* itself
  - Self-exciting point process modeling of crime
- These models are developed to combine the Flag and Boost explanations into a single model

Self-Exciting Point Process

The self-exciting point process (sepp), or *Hawkes process*, is a two-component model for the conditional intensity of a Poisson process:

$$\lambda(s,t,m) = \mathrm{Flag}(s,t,m) + \mathrm{Boost}(s,t,m)$$

- The intensity $\lambda(s,t,m)$ of a marked space-time point process represents the event rate of event with characteristics m, at a specific time t and location s.
- This requires estimating two components (intensities), Flag and Boost, from historical crime data
  - The Flag process models the Flag component and is based on exogenous variables (characteristics of the location, seasonal effects, persistent hotspots, etc.) but it shouldn't be influenced by recent crimes
  - The Boost process produces near-repeat (or aftershock) events

Branching Process Perspective

$$\lambda(s,t,m) = \text{Flag}(s,t,m) + \text{Boost}(s,t,m)$$

The branching process perspective uses the superposition property to consider the Boost term as the sum of individual processes.

$$\text{Boost}(s,t,m) = \sum_{i:t_i<t} g_i(t-t_i, \|s-s_i\|, \gamma(m,m_i))$$

- Every event can create a child event
- The intensity of the i^{th} parent process is $g_i(t-t_i, ||s-s_i||, \gamma(m,m_i))$
- The probability that event i *caused* event j is

$$p_{ij} = \frac{g_i(t_j-t_i, \|s_j-s_i\|, \gamma(m,m_i))}{\lambda(t_j,s_j,m_j)}$$

Linkage informed Hawkes / Hawkes informed Linkage

Self-Excited Branching for Pairwise Crime Linkage

- Hawkes: The probability that event i *caused* event j

$$p_{ij} = \frac{g_i(t_j - t_i, \|s_j - s_i\|, \gamma(m, m_i))}{\lambda(t_j, s_j, m_j)}$$

- Linkage: The probability that event i is *linked* to event j (logistic regression)

$$\Pr(i, j \text{ are linked}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1(i,j) + \ldots + \beta_p X_p(i,j))}}$$

Two new questions

- Hawkes: The probability that event i *caused* event j

$$p_{ij} = \frac{g_i(t_j - t_i, \|s_j - s_i\|, \gamma(m, m_i))}{\lambda(t_j, s_j, m_j)}$$

- Linkage: The probability that event i is *linked* to event j (logistic regression)

$$\Pr(i, j \text{ are linked}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1(i,j) + \ldots + \beta_p X_p(i,j))}}$$

---

1. Can the self-exciting models help estimate linkage probability?
2. Can we use linkage to help inform the self-exciting models?

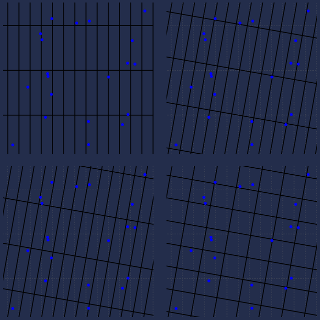
Exciting Adventures in Crime Linkage

Michael D. Porter
Engineering Systems and Environment
School of Data Science
Darden School of Business
University of Virginia
mdp2u@virginia.edu

Thanks!

Spatial Event Hotspot Prediction

- Methods developed for *2017 NIJ Crime Forecasting Challenge*
- G. Mohler & M. Porter (2018) “Rotational Grid, PAI-Maximizing Crime Forecasts”, *Statistical Analysis and Data Mining*.
- Goal of contest was to forecast (grid-based) *hotspots* for several crime types and forecasting windows (1 week to 3 months)



Threshold Modeling

- Instead of modeling the event rate (or counts) in each cell, we model the probability that the event rate (or equivalently the number of events) exceeds a threshold.
- The threshold is set so that if the event count in the cell reaches the threshold then the cell would be part of the *optimal* hotspot region.
- Using the historical event data, we found the threshold, $\phi(\tau, m)$, that a grid cell would need to be part of the optimal hotspot region (subject to the minimum size constraints) for a forecast period of length τ and crime type m.
- This creates a *binary classification problem*: grid cell {member, not member} of hotspot.

Mutually Exciting Hawkes Features

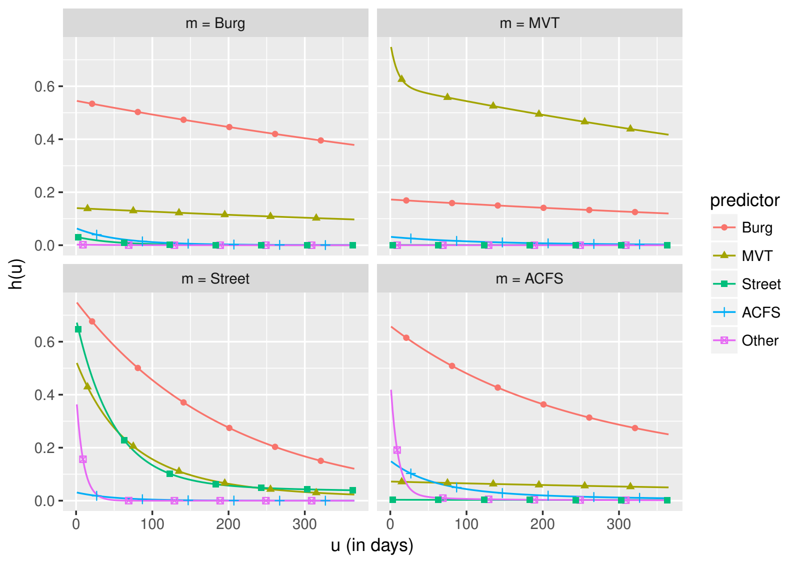
Let p_{jm} be the probability that grid cell j is part of the hotspot for crime type m.

$$\begin{aligned}\text{logit}(p_{jm}) &= \beta_m + \sum_{i:c_i=j} h_{m,m_i}(t - t_i; \vec{\alpha}, \vec{\omega}) \\ &= \beta_m + \sum_{i:c_i=j} \sum_{k=1}^{K} \alpha_k(m_i, m) g(t - t_i; \omega_k)\end{aligned}$$

where:

- m_i is the crime type for event i
- $h_{m,m_i}(u)$ is the contagion from event of type m_i to type m (*mutual excitation*).
- $h_{m,m_i}(u) = \sum_k h_{m,m_i}(u; \alpha_k, \omega_k)$ is a sum of K different decay rates.
  - $\alpha_k(m_i, m)$ is the k^{th} branching ratio for a crime of type m_i producing a crime of type m.
  - $g(u; \omega_k)$ is the k^{th} decay function

Estimated Contagion Functions



The contagion functions $h_{m,l}(\cdot)$ for predicting crime type m using the past events of type l.