A Point Process Model for Simulating Gang-on-Gang Violence

Kym Louie[*] Marina Masaki[†] Mark Allenby[‡]

Advisor: Tim Lucas[‡]

August 6, 2010

Abstract

Gang on gang violence is a prevalent problem in Los Angeles. In this paper a point process method is presented for simulating gang on gang crimes. The Hawkes process, a self exciting point process, is introduced as a temporal model for crimes between pairs of gangs. This model is expanded upon to simulate crimes not only temporally but also spatially. Following this, we introduce a method to assign directionality to these crimes. Finally, we present a model to simulate directed crimes in the 33 gang system of the policing district Hollenbeck, Los Angeles.

Introduction

Gang violence has been a significant problem in parts of Los Angeles since the 1940s. The policing district Hollenbeck, east of downtown Los Angeles is one of the most violent Los Angeles policing districts, despite having a total area of only 15.2 square miles[7].

Previous research has shown that inter-gang attack patterns follow a self-exciting model: a gang is likely to retaliate against a gang that has attacked it, and that retaliation is likely to be temporally near the inciting event [1].

We worked with a set of data from the Hollenbeck Policing Area in Los Angeles, CA containing a list of gang-related crimes from November 1999 to September 2002. Each crime was associated with a suspect, victim, location, and time. This data set contained information on 1206 crimes, but due to incomplete information about many of the crimes, the majority of our work was done on only 358 crimes.

[*]Harvey Mudd College
[†]University of California, Irvine
[‡]Pepperdine University

1 Temporal Hawkes Process

The Hawkes Process is a self-exciting point process developed to model seismic activity, with applications to other processes in which an event increases the likelihood of events following it[4]. In this paper, the events are inter-gang attacks, where one crime is likely to excite either retaliations or repeat victimizations. This activity is described by the rate function [3]

$$\lambda(t) = \mu + k_0 \sum_{t>t_i} g(t - t_i; \omega). \tag{1}$$

μ	Background Rate
k_0	Scaling Factor
ω	Rate of Decay

$g(t - t_i; \omega)$ is a function that decays as $t - t_i$ becomes greater. In modeling inter-gang crimes, Egesdal et al. used the function

$$g\left(t - t_i; \omega\right) = \omega e^{-\omega(t-t_i)}. \tag{2}$$

The rate function $\lambda(t)$ is the expected rate at which events occur. μ is the background rate, the rate at which events are expected to occur if there is no other activity. k_0 is a scaling factor for the effect of a crime, and ω is the rate of decay for a crime's effect. A higher μ value causes an overall higher activity rate, while a higher k_0 causes an event to generate more offspring. A lower ω value causes crimes to have a longer effect time and therefore causes less clustering and a more uniform distribution.

The crimes between pairs of gangs in strong rivalries (greater than 20 attacks over the 3 year period) often revealed temporal clustering as shown in Fig. 1. Plotted beneath the events is a rate function $\lambda(t)$. The parameters for μ, k_0, and

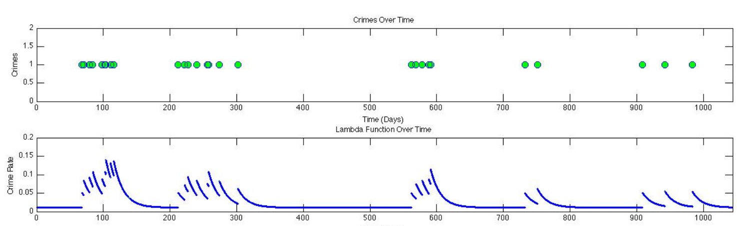


Figure 1: Crimes over time between El Sereno Locke Street and Lowell Street and value of $\lambda(t)$.

ω were approximated using maximum likelihood estimation. The log likelihood is

$$\log \mathrm{L} = \sum_c \log \lambda\left(t_c\right) - \int_T \lambda\left(t\right)\, dt, \tag{3}$$

where t_c is a time at which a crime was committed [5].

2 Extending the Temporal Hawkes Process

2.1 Spatio-Temporal Hawkes Process

To model locations as well as times of crimes, we added a spatial component to the rate function

$$\lambda(x, y, t) = \mu(x, y) + k_0 \sum_c g(t - t_c; \omega)\, h(x - x_c, y - y_c; \sigma). \qquad (4)$$

This adds two concepts to the temporal Hawkes Process. This new process has background rate $\mu(x, y)$ that varies by location but not by time. Additionally, a term $h(x, y; \sigma)$ is added causing the effect of a crime to decay over space is added. The function used for this was

$$h(x - x_c, y - y_c; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \cdot \frac{(x-x_c)^2 + (y-y_c)^2}{\sigma^2}}$$

This creates a Gaussian distribution centered at each crime. Because this distribution is scaled by $g(t - t_c; \omega)$, this makes the effect of a crime decay over both time and space. The value for σ represents the distance over which we expect a crime to influence future activity.

The background rate that was used was a scaled bimodal distribution comprised of two Gaussians. The background rate between gangs g and h is

$$\begin{aligned} \mu(x, y) = {} & C_a \frac{1}{2\pi\sigma_{a_x}\sigma_{a_y}} e^{-\frac{1}{2}\left(\frac{(x-\bar{a}_x)^2}{\sigma_{a_x}^2} + \frac{(y-\bar{a}_y)^2}{\sigma_{a_y}^2}\right)} \\ & + C_b \frac{1}{2\pi\sigma_{b_x}\sigma_{b_y}} e^{-\frac{1}{2}\left(\frac{(x-\bar{b}_x)^2}{\sigma_{b_x}^2} + \frac{(y-\bar{b}_y)^2}{\sigma_{b_y}^2}\right)}. \end{aligned} \qquad (5)$$

The values used for $(\bar{x}_g, \bar{y}_g)$ and $(\bar{x}_h, \bar{y}_h)$ were the mean locations of all crimes that gangs g and h were involved in (not necessarily with each other), respectively. $(\sigma_{gx}, \sigma_{gy})$ and $(\sigma_{hx}, \sigma_{hy})$ were the corresponding standard deviations. These were chosen such that the distributions would represent the regions in which the gangs were active. The values for C_1 and C_2 determine the rate of background activity. The ratio of these values describes how important one distribution is compared to the second; if most of the activity was located in the territory belonging to gang h, then C_h would be much larger than C_g. The parameters used in this are as follows:

2.1.1 Determining Parameters

We estimated parameters for this model using maximum likelihood estimation, as in the temporal Hawkes Process. The log likelihood is

$$\log \mathrm{L} = \sum_c \log \lambda(x_c, y_c, t_c) - \int_\Omega \lambda(x, y, t)\, dx dy dt.$$

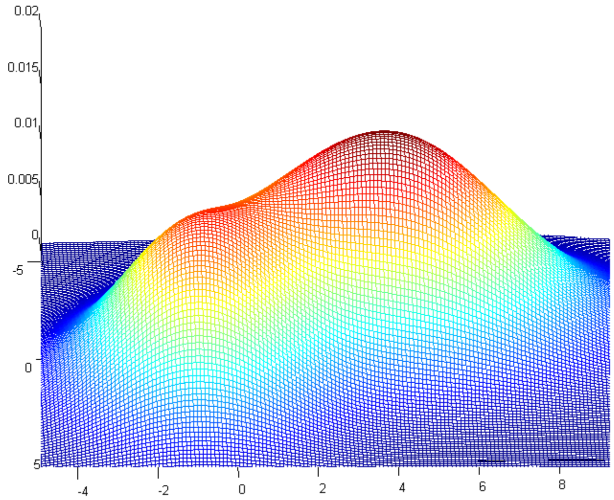


Figure 2: An example of a bimodal Gaussian distribution.

Parameter, Function, or Constant	Description
$\lambda(x, y, t)$	Rate Function
$\mu(x, y)$	Background Rate
k_0	Scaling Factor
ω	Decay Rate
σ	Range of Crime Effect
C_1 and C_2	Background Scaling Factors
$(\bar{x}_g, \bar{y}_g)$	Average Location of all g Crimes
$(\sigma_{gx}, \sigma_{gy})$	Standard Deviation of all g Crimes

Figure 3: Parameters and functions for spatio-temporal Hawkes Process.

However, if we simply estimate the values using this, the likelihood is maximized for vanishingly small values of σ. In this situation, the effect of a crime is confined to a negligible area and as such, almost all crimes are background events. This means that the events are not clustered and when simulating crimes using these numbers, the crimes are distributed spatially like the background rate μ and temporally as a uniform distribution.

However, our data does not support this sort of activity. To prevent this, we add a penalty term to the function we are maximizing [6]. This penalty term penalizes for high gradients in space (but not time):

$$\eta \int_T \iint_{X,Y} ||\nabla r\,(x, y)\,||^2 \,dxdy\,dt.$$

where

$$r(x,y) = k_0 \sum_c g(t - t_c; \omega)\, h(x - x_c, y - y_c; \sigma),$$

where η is simply a scaling factor determining to what extent this penalizing term is influential. Thus the equation we maximize to estimate parameters is

$$\sum_c \log \lambda(x_c, y_c, t_c) - \int_\Omega \lambda(x,y,t)\; dxdydt \\ - \eta \int_T \iint_{X,Y} ||\nabla r(x,y)||^2\, dxdy\, dt. \tag{6}$$

This penalizing term has the effect of making it unlikely that the estimated parameters describe crimes that sharply increase the rate at very small distances. Therefore, for sufficiently large values of η, the expression is no longer maximized with vanishingly small values of σ.

The following parameters were estimated with a η value of 100.

Estimated Parameter Summary Table					
Rivalry	k_0	w	σ	C_1	C_2
Loc-Low	0.122	6.117×10^{-2}	0.655	1.374×10^{-3}	1.971×10^{-2}
Clo-Eas	0.040	0.162	2.855	1.171×10^{-2}	1.865×10^{-2}
Lin-Eas	6.983	1.888×10^{-5}	0.310	1.315×10^{-2}	2.487×10^{-8}
KAM-Sta	1.288×10^{-3}	53.37	0.394	7.450×10^{-4}	1.780×10^{-2}
Tin-Sta	7.372×10^{-2}	4.966×10^{-2}	1.048	1.415×10^{-2}	1.794×10^{-10}
MCF-ELA	1.005×10^{-2}	0.308	0.294	1.890×10^{-2}	1.018×10^{-10}
VNE-Opa	5.450×10^{-2}	0.165	0.500	1.085×10^{-10}	1.778×10^{-2}
VNE-8th	8.159×10^{-3}	59.48	0.947	3.078×10^{-10}	1.640×10^{-2}
TMC-Cua	6.967×10^{-3}	0.677	0.327	8.082×10^{-4}	2.513×10^{-2}

Figure 4: Parameters for nine largest rivalries. Parameters were estimated with distance measured in hundredths of degrees latitude and longitude and time measured in days. 1/100 of a degree latitude is approximately 0.69 miles.

In Fig. 4 we see that the estimated parameters are still sometimes unrealistic. For example, the estimated value for k_0 for the Lincoln Heights and Eastlake rivalry indicates that in the day following a crime, the rate of crimes per day per unit area is increased by nearly 7.

In Fig. 5 we see the estimated values for the parameters for this rivalry for multiple values of η. It is evident that in this case, while the penalty term affects σ in a positive way, it does not necessarily result in realistic values for the other parameters.

2.1.2 Simulating Data

We used a branching process to simulate data with the previously described parameters. While we were able to consistently produce a number of crimes

Estimated Parameter Between Lincoln Heights and Eastlake					
η	k_0	w	σ	C_1	C_2
.01	582.8	1.751×10^{-6}	0.184	8.343×10^{-3}	5.873×10^{-8}
1	0.103	9.081×10^{-2}	0.249	1.270×10^{-2}	2.911×10^{-10}
10	228.6	2.206×10^{-6}	0.265	1.031×10^{-2}	3.830×10^{-9}
1000	6.898	2.327×10^{-6}	0.319	1.422×10^{-2}	6.006×10^{-9}
10^{10}	1.905×10^{-9}	3.184×10^{-2}	0.523	1.437×10^{-2}	2.391×10^{-10}
10^{20}	8.052×10^{-5}	8.180×10^{-15}	1.219	1.437×10^{-2}	2.092×10^{-9}

Figure 5: Estimated parameters for Lincoln Heights and Eastlake rivalry for various η values.

similar to those in our data, the parameters often produced a very high number of background crimes and a very low number of offspring crimes.

Spatially, the distribution of simulated crimes reflected that of the actual crimes, as in Fig. 6. This is expected, because background crimes are distributed in the same form as the background rate $\mu(x, y)$, which was based on the distribution of the actual crimes.

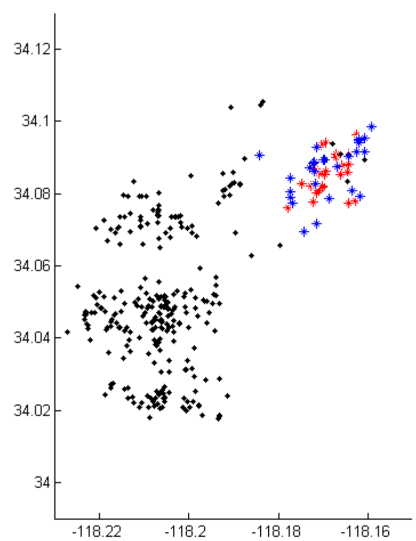


Figure 6: Simulated vs. actual crimes between El Sereno Locke Street and Lowell Street. Blue crimes are actual events, red are simulated. The black dots are all complete events in the data.

Temporally, the simulated crimes appear to be slightly less clustered than the actual crimes in Fig. 7. However, in fact the simulated data has 23 (out of a total of 28) crimes that are background events; only 5 of the events were offspring events. Thus the "clustering" we see is mostly the result of a Poisson process not distributing events equally in time.

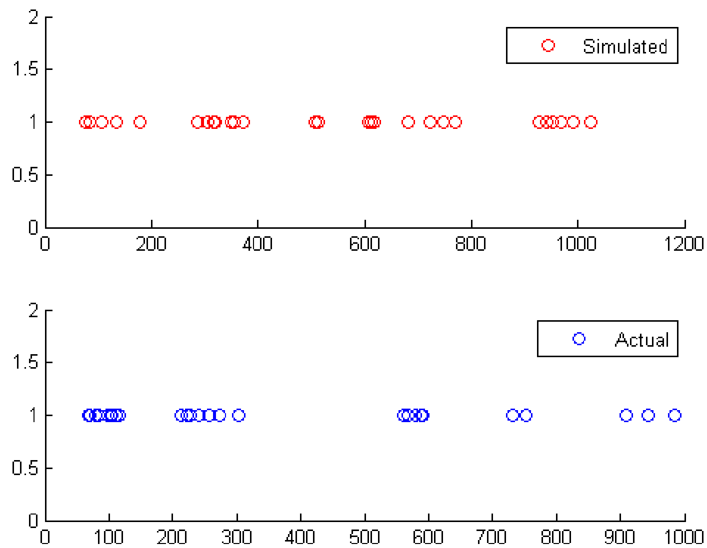


Figure 7: Simulated vs. actual crimes between El Sereno Locke Street and Lowell Street. Blue crimes are actual events, red are simulated. There are 28 simulated crimes (23 background) and 27 actual crimes. These crimes were simulated using the parameters from Fig. 4

This result seems to challenge the effectiveness of using a spatio-temporal Hawkes Process to model the gang-on-gang activity. However, this may be the result of $h(x, y; \sigma)$ not accurately representing the spatial effect of a crime. Additionally, while it appears that crimes are clustered in time, we currently do not have a way of determining whether points temporally in a cluster are necessarily related to the other events in the cluster, so some of the clustering in the data may not actually indicate the level of background and excited activity that is initially apparent.

2.2 Directional Hawkes Process

Another extension of the temporal Hawkes Process is identifying a suspect and victim with the crime. Here we address this for crimes between a single pair of gangs.

When a gang has been attacked, they have a tendency to retaliate against the offender. Additionally, some gangs repeatedly victimize other gangs. To model these trends, we add directionality to the crimes between two gangs. Instead of viewing events between two gangs as identical, we separate them into events where gang a attacked gang b and gang b attacked gang a. This enables us to gain more insight into the dynamics of the two gangs in order to simulate the crimes and shows us which gang commits crimes more frequently. The rate function for gang a attacking gang b is as follows:

$$\lambda_{a\to b}(t) = \mu_{a\to b} + r_0 \sum_{t_c<t}^{N} g\left(t - \underset{b\to a}{t_c}; \omega\right) + s_0 \sum_{t_d<t}^{M} g\left(t - \underset{a\to b}{t_d}; \nu\right). \qquad (7)$$

The self-excited component of the original Hawkes process is replaced with terms accounting for retaliatory attacks and repeat victimizations. The second term is the retaliatory component. This boosts the crime rate by a r_0 when a gang has been attacked, and decays exponentially over time at a rate of ω. The third term is the repeat victimization component. This increases the crime rate by s_0 when a gang has attacked the other and decays exponentially over time at a rate of ν. A high r_0 corresponds to a situation in which gang a often commits crimes in retaliation to crimes by gang b, and a high s_0 corresponds to a situation in which gang a repeatedly victimizes gang b. A small ω and ν cause the crime rate to decay slowly over time and therefore crimes will continue to occur after the attacks. To obtain the parameters for the Directional Hawkes process, we fit the parameters to the actual data using the maximum log likelihood estimation.

Estimated Parameter Table					
Rivalry	r_0	s_0	ω	ν	μ
Loc-Low	8.568×10^{-2}	0.343	0.665	2.980×10^{-2}	2.764×10^{-3}
Low-Loc	0.462	0.463	8.738×10^{-2}	6.834×10^{-2}	7.199×10^{-3}
Clo-Eas	0.273	5.804×10^{-2}	0.391	4.265	1.228×10^{-2}
Eas-Clo	5.114×10^{-2}	0.396	4.491	5.419×10^{-2}	7.720×10^{-3}
VNE-8th	0.326	0.135	3.452×10^{-2}	4.502	2.588×10^{-3}
8th-VNE	0.179	0.116	1.742×10^{-2}	2.539×10^{-2}	7.275×10^{-3}

Figure 8: Estimated parameters for three large rivalries.

Using the parameters that were obtained by the maximum log likelihood estimation, we simulated the directional Hawkes process using the marked point process as in Fig. 9. The simulation accurately reflects the relationship between the two gangs. Locke Street has a higher crime rate than Lowell Street in the actual data, and is reflected in the simulation. The self-exciting nature of the Hawkes process is also apparent in the short inter-attack times in both diagrams. This simulation is a useful tool in modeling gang activity between pairs of gangs because we are able to obtain accurate parameter estimations to produce a model that resembles the actual data.

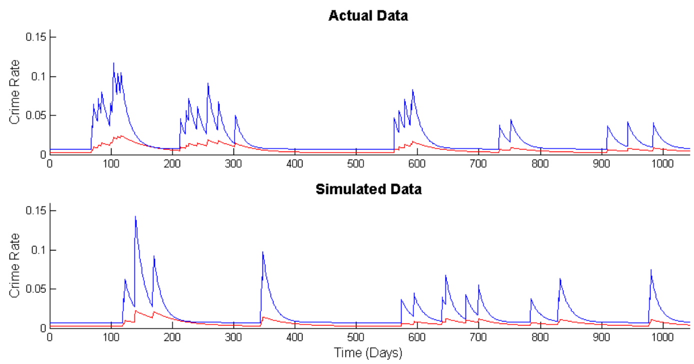


Figure 9: Actual vs simulated crime rates for El Sereno Locke St. and Lowell St. The blue represents crimes committed by Lowell St., red represents crimes committed by El Sereno Locke St.

3 Combining Spatial and Directional Models

3.1 Sequential Combination

Now that we are able to simulate crimes with location and crimes with directionality, we combine the ideas to attempt to simulate crimes with time, location, and direction for pairs of gangs. One way to combine these ideas is to simulate crimes in time and space, and then give these crimes directionality.

To do this, we first estimate parameters and simulate crimes as described in Section 2.1. This process provides us with a set of crimes, each having a time and location. We then identify for each crime which gang was the suspect and which was the victim.

In order to assign a perpetrator to each crime, we consider the rate at which each gang is likely to attack the other. Here we use the rate function described in Eq. 7 and do not use our spatial data. Using this equation, we have a rate for crimes in either direction at any time for which we know the direction of all previous crimes. We start at the initial simulated crime and probabilistically assign a directionality to it using the rate functions:

$$\mathrm{P}\left(a \text{ attacked } b \middle| \, t; H\right) = \frac{\lambda_{a\to b}\left(t\right)}{\lambda_{a\to b}\left(t\right) + \lambda_{b\to a}\left(t\right)}, \tag{8}$$

where H is the history of crimes between gangs a and b. For the initial crime, there are no previous crimes, so the values of the rate functions are simply the background rates. We then inspect the second crime and find the rates for each direction and assign a perpetrator for that crime. We continue to go through the crimes chronologically until all crimes have been assigned a direction.

Using this method we may attempt to simulate data that mimics the patterns we see in the real data. An example simulation is shown in Fig. 11.

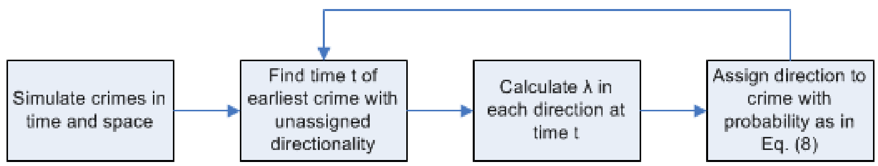


Figure 10: Method for simulating crimes over time and space with directionality.

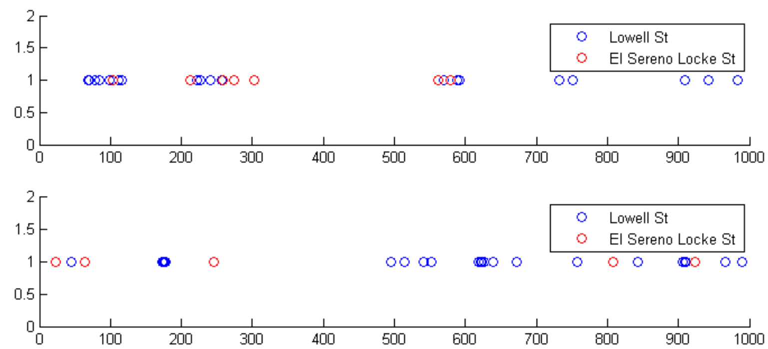


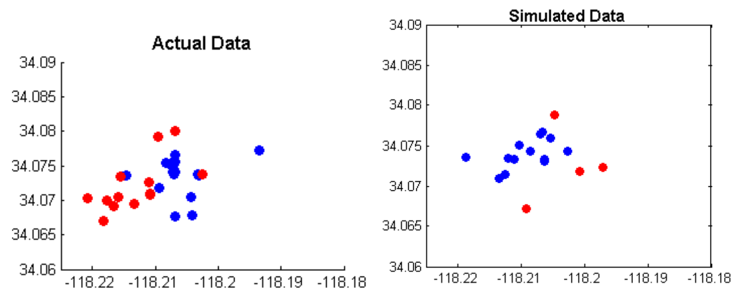
Figure 11: Directions added to simulated crimes between El Sereno Locke St. and Lowell St. The actual crimes appear on the top and simulated crimes on the bottom. Crimes in which Lowell St. was the suspect are shown in blue, while those in which El Sereno Locke St. was the suspect are shown in red.

3.2 Simultaneous Combination

To further investigate the effects of retaliatory and repeated crimes, we incorporate the spatial component in 4 to the temporal directional Hawkes Process.

$$\lambda_{a\to b}(x,y,t) = \mu_{a\to b}(x,y) + r_0 \sum_{t_c<t}^{N} g\left(t - \underset{b\to a}{t_c}; \omega\right) \frac{1}{2\pi\sigma^2} e^{-\frac{(x-x_c)^2+(y-y_c)^2}{2\sigma^2}}$$
$$+ s_0 \sum_{t_d<t}^{M} g\left(t - \underset{a\to b}{t_d}; \nu\right) \frac{1}{2\pi\gamma^2} e^{-\frac{(x-x_d)^2+(y-y_d)^2}{2\gamma^2}}. \quad (9)$$

The variable background rate used in this model is similar to as in Eq. 5. The values for $(\bar{x}_a, \bar{y}_a)$ and $(\bar{x}_b, \bar{y}_b)$ are the mean locations of the crimes gangs a and b were involved in. However, $(\sigma_{ax}, \sigma_{ay})$ is the standard deviation of the crimes that gang a committed against gang b, and $(\sigma_{bx}, \sigma_{by})$ is the standard deviation of the crimes that gang b committed against gang a. This reduces the influence on the background rate to only crimes in specific direction. The retaliatory



(a) Actual locations of crimes (b) Simulated locations of crimes

Figure 12: Actual vs simulated crime locations of Clover and Eastlake. Blue dots are events in which Clover was the suspect, and red dots are events in which Eastlake was the suspect.

and repeated victimization terms now decay over space around the previous crime locations. σ is the standard deviation of the Gaussian distribution that defines the increased crime rate around locations where gang b attacked gang a. Similarly, γ is the standard deviation of the Gaussian distribution that defines the increased crime rate around locations where gang a attacked gang b. Higher values for σ and γ cause events to have influence at greater distances. Using the maximum penalized likelihood estimation we determined values for the parameters. The locations of crimes between Clover and Eastlake were simulated using the marked point process.

Fig. 13 shows histograms of the distances between Clover's mean location $(\bar{x}, \bar{y})$ and each event where Clover attacked Eastlake. The left is the histogram for the actual data and the right is the simulated data. It is apparent that they share the same pattern. The most crimes are committed near Clover's gang center and as they travel further away the crimes decrease, and increase again as they travel even further away. This simulation works great for gangs which we have many data points for like these two gangs. However due to the small amount of data for each pair of gangs, the parameter estimations were not completely accurate which skewed the simulations to some extent. With more data the parameters will better fit the actual data and better simulate the data.

4 Pairwise Model Comparison

We discussed two models for simulating directed events between pairs of gangs.

The model discussed in Section 3.2 combines both spatial and directional ideas into a single equation, which is preferable for estimating parameters when we expect all parameters to have an effect on each other. However, this model requires estimating 8 parameters for each direction, for a total of 16 parameters. Unfortunately, our data only contains 7 pairs of gangs with more than 16

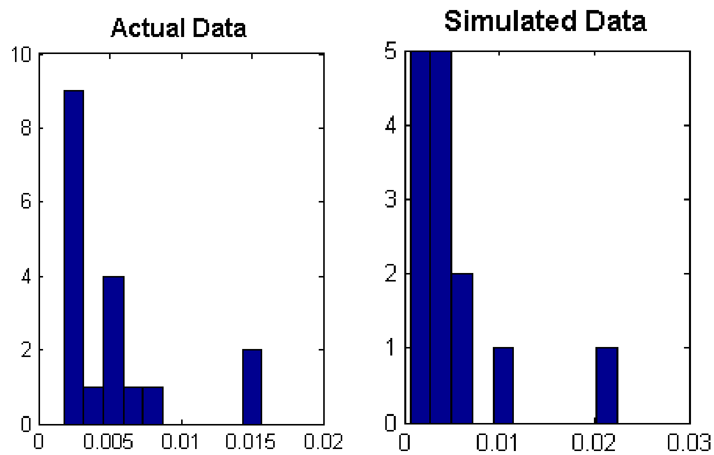


Figure 13: Histograms of the distances between Clover's mean location and each crime committed by Clover. Left is the actual inter-attack distances, and the right is the simulated.

attacks between them, only 3 of which have greater than 8 attacks in both directions. Given this, we cannot assume that the parameters we estimate accurately describe the system.

The model discussed in Section 3.1, while not allowing for interaction between directional and spatial components, does not require simultaneous estimation of as many parameters: the spatial process requires 5 parameters and each of the directional processes require 5 additional parameters. While this is still a large number of parameters given our very limited data set, we are more likely to get meaningful spatial parameters for the largest rivalries.

For very limited data sets, the first model is preferable due to its requiring fewer parameters. However, if we were able to work with a much larger data set, the second model may produce better estimations due to its ability to allow location to influence direction.

Multi-Gang System

In order to simulate crimes over all of Hollenbeck, a system of rivalries must be constructed. As shown in equation (4), our Hawkes Process is going to be one of space and time, with a spatially-variant background rate and a spatio-temporal self exciting function.

Crimes are produced at the crime rate $\lambda(x, y, t)$. Each crime is distributed throughout space and time and is associated with two gangs; an attacking gang and a victim gang.

In order to model this system of crimes, we implement a cluster or branching process as proposed by Hawkes and Oakes [2]. In simulating events, a branching process first generates background crimes at the background rate over the entire time domain. The number of background crimes simulated is a Poisson random variable with parameter $\int_\Omega \mu(x,y)$. These crimes are distributed uniformly through time. We then inspect each background crime to determine whether additional retaliatory or repeat crimes occur due to it. Each crime generates a Poisson random variable with parameter k_0 of offspring crimes. After inspecting each background crime, these 'offspring crimes' are themselves inspected to determine if more offspring crimes are created in the same way.

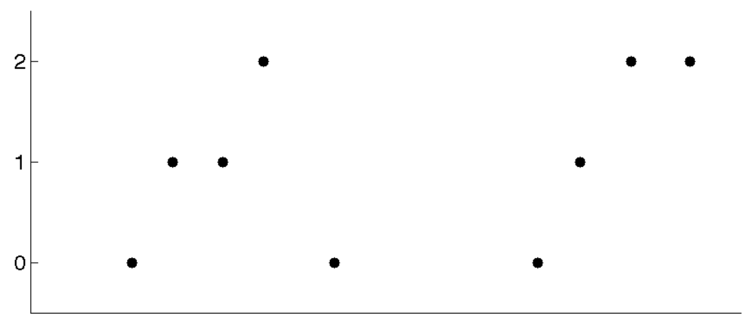


Figure 14: An example of a temporal branching point process.

Figure 14, shows an example of such a branching process over time. For different points through time, random background or 'level 0' events are generated. Each 'level 0' event then has the opportunity to create 'level 1' offspring events later in time, and then each 'level 1' event has the ability to create additional offspring events. Likewise, our model's crimes may branch out to create more crimes, which can create more crimes, and so forth. For realistic parameters, this process terminates: a Poisson random variable with parameter k_0 will usually be 0 and very rarely greater than one and any offspring crimes generated outside the time domain are disregarded.

To simulate events, we first generate the background crimes for the region, at a rate equal to the integral of $\mu(x,y)$ over all space and time. These background crimes are then uniformly distributed throughout time, and distributed with respect to a multimodal normal distribution in space as in Eq. 10.

$$\mu(x,y) = \sum_g \frac{C_g}{2\pi\sigma_{gx}\sigma_{gy}} e^{-\frac{1}{2}\left(\frac{(x-\bar{x}_g)^2}{\sigma_{gx}^2}+\frac{(y-\bar{y}_g)^2}{\sigma_{gy}^2}\right)} \tag{10}$$

The multimodal normal distribution is made up of a sum of scaled normal distributions with a mode at each gang center. These distributions are placed such that they center around each gang's center. These gang centers are either gang bases supplied to us by the L.A.P.D. or approximated as the average of all the locations of crimes that that gang was involved in. Each distribution also had a standard deviation equal to the standard deviation of all the crimes it was involved in. A scaling factor C_g was estimated for each distribution as well, which affected the number of background crimes associated with that gang's 'territory'. The quantity of background crimes generated, the integral of $\mu(x, y)$ over all space and time, is reduced to the sum of all scales C_g as each normal distribution integrates to 1.

Parameter	Explanation
C_g	Normal distribution scaling factors
$(\bar{x}_g, \bar{y}_g)$	Gang centers
$(\sigma_{gx}, \sigma_{gy})$	Gang standard deviations

Figure 15: $\mu(x, y)$ Equation Parameters

In some instances, data given identified a single gang with two gang bases. In such a case, two normal distributions were used instead of one, with separate centers and scaling factors but the same previously described standard deviation.

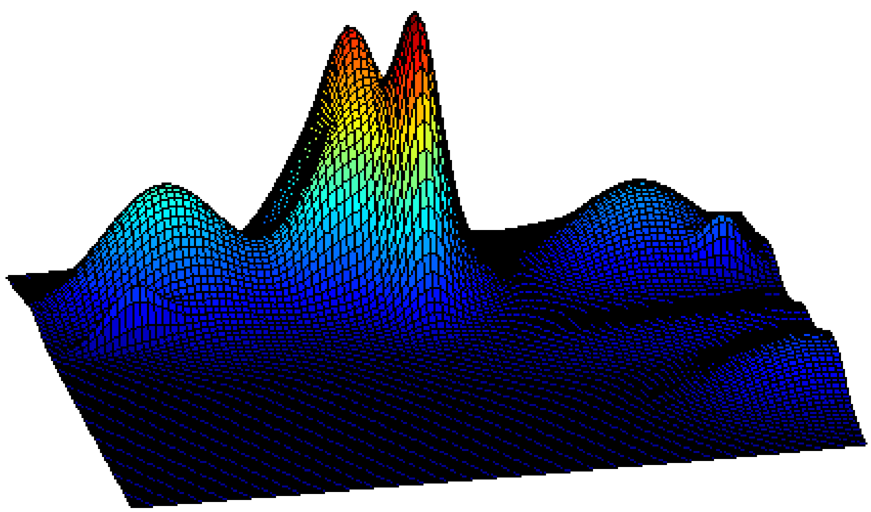


Figure 16: A probability density map of the background crime rate $\mu(x, y)$ of Hollenbeck.

Background crimes have now been simulated and distributed in space and time. Now these background crimes must be assigned to a pair of gangs. Each possible pairing of gangs is examined and a we determine a likelihood that a crime is identified with that pair.

$$\underset{a,b}{\ell}(x,y) = \frac{C_a}{2\pi\sigma_{xa}\sigma_{ya}} e^{-\frac{1}{2}\left(\frac{(x-\bar{x}_a)^2}{\sigma_{xa}^2} + \frac{(y-\bar{y}_a)^2}{\sigma_{ya}^2}\right)} + \frac{C_b}{2\pi\sigma_{xb}\sigma_{yb}} e^{-\frac{1}{2}\left(\frac{(x-\bar{x}_b)^2}{\sigma_{xb}^2} + \frac{(y-\bar{y}_b)^2}{\sigma_{yb}^2}\right)} \quad (11)$$

The likelihood function, $\underset{a,b}{\ell}(x,y)$ is comprised of two scaled bimodal distribution. This distribution uses the same centers, standard deviations and scaling factors that identified with the two gangs in equation (10). A pair of gangs is assigned probabilistically, such that

$$\mathrm{P}\left(c \text{ is identified with } (a,b) \middle| x_c, y_c\right) = \frac{\underset{a,b}{\ell}(x_c, y_c)}{\sum_g \sum_{h<g} \underset{g,h}{\ell}(x_c, y_c)}. \quad (12)$$

Here, the nested sums in the denominator sum over all possible pairings of gangs.

The background step of our process is now complete. All of of the background crimes have been generated, distributed in space and time and assigned gangs. Each background crime is now generates offspring crimes at a rate k_0 in equation (4). The offspring crimes are then distributed spatially and temporally with respect to equation (13).

$$g(x,y,t) = \frac{\omega}{2\pi\sigma^2} e^{-\frac{(x-\bar{x}_c)^2 + (y-\bar{y}_c)^2}{2\sigma^2} - \omega(t-t_i)} \quad (13)$$

As described in Eq. (13), offspring crimes are distributed with an exponential decay in time, and a normal distribution in space centered around their ancestor crime. These offspring crimes are assigned the same pair of gangs as their ancestor, as repeat and retaliatory crimes would be between the same pair of gangs. This process continues after all background crimes are inspected for offspring crimes, and inspects all offspring crimes for offspring crimes, until all crimes have been inspected.

All crimes have been generated, distributed in time and space and are associated with a pair of gangs. The final step in this model is to incorporate a directionality to all crimes. That is, if a background crime has been established to be between gang i and gang j, it is not known whether gang i attacked gang j or gang j attacked gang i. We apply equation (8) for each crime to get this directionality.

All crimes have now been generated, distributed and fully labeled. The simulation is complete.

We were able to simulate crimes for all pairs of gangs in Hollenbeck by creating a global system. This global system modeled the Hawkes Point Process through a branching process. We used a multimodal probability distribution, $\mu(x,y)$ over space as our background rate. Every offspring crime was exponentially distributed in time and normally distributed in space. These crimes were associated with a pair of gangs: a victim and an attacker.

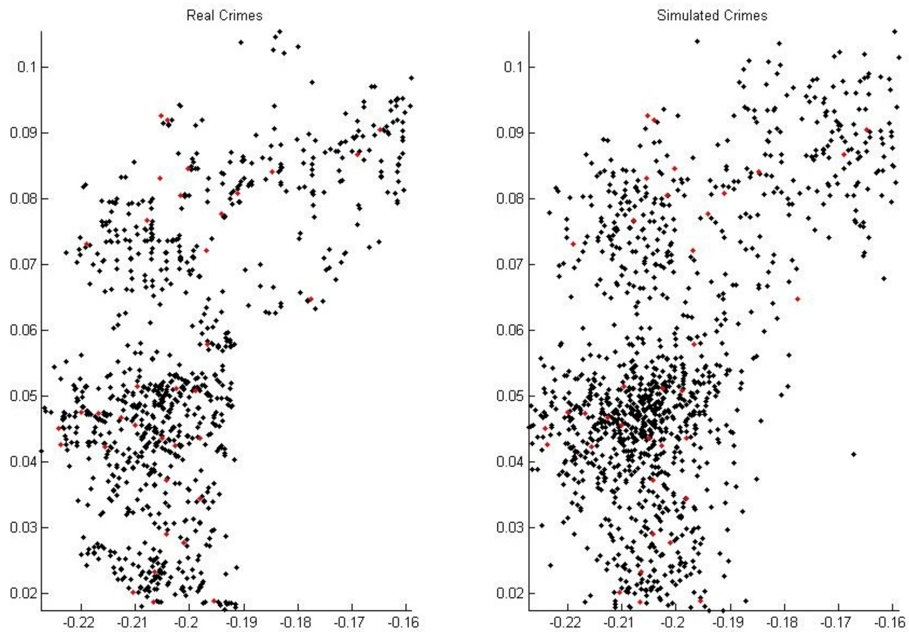


Figure 17: A simulation over all of Hollenbeck.

Table 1: Crimes Generated

Type	Amount
Total Real	1206
Total Simulated	1233
Background Simulated	1208
Self-Excited Simulated	25

Conclusion

We extended the temporal Hawkes Process to model crimes with locations and times and to model crimes with directionality and times. We used maximum likelihood estimation to find parameters for both of these models, but included in the spatio-temporal parameter estimation a penalizing term to ensure non-negligible σ values. We then combined these two models to produce two different models which both simulated crimes with location, direction, and time. Unfortunately, due to the very limited size of our data set, we were unable to reliably estimate accurate parameters for this model.

Additionally, we created a global system to simulate all pairs of gangs in Hollenbeck. This model similarly had difficulties with the amount of available data.

There was a persistent problem with estimating a low σ value, which suggests that crimes may not be related spatially or may only be related very rarely spatially. If this model with spatial dependence is pursued, it may be advisable to consider other ways that a crime would affect the rate at various locations.

Future work on this subject could include a different spatial effect of crimes as mentioned above. In addition, a new method for determining the background rate may allow for better determination of the remaining parameters. Any and all of these methods could be more accurately implemented and tested with access to additional data.

References

[1] M. Egesdal, C. Fathauer, K. Louie, and J. Neuman. Statistical and stochastic modeling of gang rivalries in los angeles. *SIAM Undergraduate Research Online*, 3:1, 2010. Research from UCLA REU last summer in modeling gang activity. Covers agent model and Hawkes process.

[2] A. G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11:493–503, 1974. Discusses self-exciting point processes as Poisson cluster processes.

[3] Alan G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83–90, 1971. Stationary Point Processes. (1) Poisson Process. (1) Poisson Covariance Densities. (1) Hawkes Self-Exciting Point Process. (2) Hawkes Covariance Density. (2) Hawkes w/ Exponential Decay. (3) Mutually Exciting Point Processes for non-interaction and one-way interaction. (4) Process with known exciter. (5).

[4] Yosihiko Ogata. Space-time point-process models for earthquake occurances. *The Institute of Statistical Mathematics*, 50:379–402, 1998. Occurance rate of aftershocks Stationary Poisson process using Omori formula Hawkes process Diffusion fuction Conditional intensity function Linear fault model log liklihood comparing models using AIC.

[5] T. Ozaki. Maximum likelihood estimation of hawkes' self-exciting point-process. *Annual Institute of Statistical Math*, 31:145–155, 1979. Background of Hawkes, uses an exponential decay response function Very straightforward calculation of the Log-Likelihood of the Hawkes Process Obtaining the maximization and then examples.

[6] B. W. Silverman. On the estimation of a probability density function by the maximum penalized likelihood method. *The Annals of Statistics*, 10(3):795–810, 1982.

[7] George E. Tita. *Reducing gun violence: results from an intervention in East Los Angeles*. RAND Corporation, 2003.