Improving social harm indices with a modulated Hawkes process

George Mohler * Jeremy Carter† Rajeev Raje‡

Abstract

Communities are adversely affected by heterogeneous social harm events, e.g. crime, traffic crashes, medical emergencies, and drug use. Police, fire, health and social service departments are tasked with mitigating social harm through various types of interventions. While social harm indices have been proposed for allocating resources to spatially fixed hotspots, the risk of social harm events is dynamic and new algorithms and software systems capable of quickly identifying risks and triggering appropriate public safety responses are needed. We propose a novel modulated Hawkes process for this purpose that offers a flexible approach for both: i) incorporating spatial covariates and leading indicators for variance reduction in the case of more rare event categories and ii) capturing dynamic hotspot formation through self-excitation. We present an efficient l1-penalized EM algorithm for estimation of the model that simultaneously performs feature selection for spatial covariates of each incident type. We provide simulation results using data provided by the Indianapolis Metropolitan Police Department to illustrate the advantages of the modulated Hawkes process model of social harm over recently introduced social harm indices and property crime Hawkes processes.

1 Introduction

Crime is highly concentrated in urban communities and hotspot or "predictive" policing efforts aim to apply limited resources to high intensity geographic areas and time intervals to disrupt crime opportunities, leading to aggregate crime rate reductions [8, 28, 31, 42]. A number of algorithmic methods have been proposed for estimating crime hotspot risk including multivariate models [23, 25, 25, 40, 40], kernel density estimation [7, 9, 14, 20, 21] and spatio-temporal point processes [27, 29]. Point processes and density estimation have the advantage of capturing near-repeat effects and only require event data as input, whereas multivariate models gain variance reduction through the introduction of spatial covariates

*Computer and Information Science, Indiana University - Purdue University Indianapolis
†School of Public and Environmental Affairs, Indiana University - Purdue University Indianapolis
‡Computer and Information Science, Indiana University - Purdue University Indianapolis

(though variance can be increased if irrelevant covariates are included). Field trials of predictive policing using a property crime Hawkes process were conducted in [28] where patrols directed through the Hawkes process led to statistically significant crime rate reductions compared to analyst directed patrols [28].

However, police serve other roles in the community beyond crime response and prevention, including traffic enforcement, Emergency Medical Services (EMS) response, and more generally dealing with events related to social harm [30]. Despite these multiple and disparate daily challenges, existing hotspot and predictive policing algorithms and intervention strategies focus on single or groups of related sub-categories of social harm events. Scholars have recently called for the next evolution of hotspots policing to move beyond crime counts in space and time to the more expansive and hierarchical approach of policing "social harm" [30] [41] [35]. A recent approach to quantify the impact of crime on society has been the development of crime harm indices [12] [35] that attempt to weight crime offenses using sentencing guidelines as opposed to more simple count measures of crime occurrences. In this context, harm is operationalized as the impact upon society dependent on the qualitatively different level of severity across incidents of crime. This approach to quantifying crime has implications for developing more effective police interventions to reduce harm, as opposed to reducing crime counts. With this end-goal in mind, [30] extends the idea of a crime harm index to a social harm index that includes incidents to which police must respond that fall outside the traditional definition of crime but still inflict harm on society – such as vehicle crashes. The current study further progresses the notion of social harm to include additional incidents within the police purview that affect society. Here, harm is operationalized beyond crime impact to society to be more inclusive of the nature of police work. Put simply, a focus on social harms builds on hotspots policing by applying similar methodological approaches, but broadens the list of harm incidents to more accurately reflect day-to-day policing demands (e.g., crime, medical emergencies, vehicle crashes, etc.) while weighting these various incidents to reflect the degree of severity they may inflict upon society.

Preliminary findings in social harm research suggest that the inclusion, and weighting, of various harm incidents holds substantive promise for police practice and intervention. To date, the most common approach to weighting social harms is to map sentencing data to specific crime offenses. This method has taken the form of actual sentencing outcomes [6] [5] [12] as well as prescriptive sentencing guidelines often referred to as gravity or severity scales [30] [41] [35]. This method leverages suggested sentencing lengths to rank the "harm" of a given offense. For example, a criminal homicide may have a sentencing guideline of 24 years in prison, armed robbery may elicit a 12-year sentence, and residential burglary a 6-year sentence. In such a weighting scenario, criminal homicide would be twice as severe or harmful as armed robbery and four times more harmful than residential burglary. Weighting by sentencing guidelines can take many forms and the discussion presented here is limited to the importance of weighting crimes and other incidents by severity. Indeed, "neither criminology nor the adjacent social sciences have made a serious effort to systematically identify, evaluate or compare the harms associated with different crimes" [19] and that "focusing merely on counts, rather than on the severity or harm of crime is somewhat crude and imprecise" [41]. In [35] Sherman and his colleagues provide a robust discussion of varying weighting procedures using sentencing guidelines.

Studies employing this approach have concluded that social harm is variable across police patrol districts [30] and that a small proportion of crime victims are exposed to greater levels of social harm [13]. Most closely related to the current study, in [41] Weinborn et al employed the Cambridge Harm Index (CHI) [35] wherein crimes are weighted by the number of days in prison for a given offense as outlined in the Home Office Sentencing Guidelines to examine the spatio-temporal concentration of crime counts versus CHI social harm. Their results indicated social harm to be three times more concentrated when compared to crime counts alone across 15 councils in England and Wales during a 12-month period. Interestingly, and salient to call for scholars to consider a variety of social harms beyond traditional hotspot policing strategies, the authors observed that only 25% of crime count hotspots overlapped with social harm locations, or "harmspots". Thus, while conducting spatiotemporal analyses of crime counts alone can be insightful for focusing police strategies, it appears prudent to account for the severity of harm crime may cause as all crimes are not created equal and more harmful incidents may display spatiotemporal variation from less harmful events. Moreover, as harmspots exhibit different spatiotemporal patterns than hotspots, they too may have different corollary relationships with community structure than do hotspots; thus one focus of the present study.

The present study further contributes to the social harm policing literature through the inclusion of multiple harm types that have yet to be examined in a single study. The present study includes a range of Part 1 (the most serious crimes that regularly occur across all jurisdictions and are likely to be reported to police) and Part 2 (other crimes) criminal offenses as well as vehicle crashes and drug overdoses – the latter of which is currently regarded as one of the most concerning social harms to society as drug overdose deaths across the United States have more than quadrupled since 1999 [32]. Part 1 and Part 2 criminal offenses are defined by the Federal Bureau of Investigation (2016) as a tiered classification system for the Uniform Crime Reporting and National Incident Based Reporting System. Furthermore, unlike static social harm indices that are estimated over a fixed historical window of observation, our methodology produces a dynamic harm index that incorporates new event data each day to account for spatio-temporal fluctuations of social harm risk.

In this work we introduce a modulated Hawkes process for modeling dynamic social harm hotspots. The model combines several advantageous aspects of both existing multivariate regression and point process models. In particular, the model is comprised of a background modulated Poisson process that links spatial covariates (census variables, average crime rate, etc.) to the risk of each social harm event category. Our estimation procedure also includes automatic variable selection to prevent over-fitting and determine important covariates for model explanation. Secondly, the point process approach allows for the incorporation of self-excitation present in some event categories. Because the output of the modulated Hawkes process is a conditional intensity for each event type, a dynamic social harm index can be easily defined through calculating the expected cost of a given spatial region and time interval.

The outline of the paper is as follows. In Section 2, we give an overview of the data set used in our study and the methods used to estimate the average societal cost of each event type. In Section 3, we provide the mathematical details of the modulated Hawkes process and a l1-penalized Expectation-Maximization (EM) algorithm for parameter estimation. In Section 4, we describe several experiments for validating the model, including validation tests on synthetic data and retrospective forecasts using data provided by the Indianapolis Metropolitan Police Department and Emergency Medical Services. In Section 5, we discuss the implications of dynamic social harm prediction and future research directions.

2 Indianapolis Social Harm Data

All crime, drug overdose, and vehicle crash data for years 2012-2013 for the city of Indianapolis were provided electronically by the appropriate government agencies and include time and date stamp as well as state-plane coordinates for each incident that were converted to WGS84 coordinates. Crime data was provided by the Indianapolis Metropolitan Police Department (IMPD), drug overdose from the Indianapolis Emergency Medical Services, Department of Public Safety, and vehicle crash data from the Indiana State Police using the Automated Reporting Information Exchange System (ARIES). Indiana motor vehicle collisions have two key characteristics that are used to determine whether or not an incident requires completion and submission of an Indiana crash report; if the incident resulted in personal injury or death, or property damage to an apparent extent greater than $1,000.

Rather than relying upon sentencing guidelines as a weighting mechanism to determine social harm, the present study employs monetary cost estimates. This decision was driven primarily by 1) a lack of variation in Indiana's sentencing guidelines that are restricted to four classifications within six larger levels of offenses (as compared to the 415 categories available used in [41]); and 2) monetary costs reflect tangible measures of harm impact on society as opposed to the offender alone and demonstrate the potential financial gains that could be achieved through improved interventions. Moreover, financial estimates have been argued to demonstrate an improved understanding of the relationship between criminal justice policy and beneficial interventions [24] [33]. Social harm weights were derived from established crime, drug, and vehicle crash cost estimation studies. Costs for homicide, rape, robbery, aggravated assault, arson, motor vehicle theft, residential burglary, larceny, embezzlement, forgery, fraud, and vandalism were gleaned from estimates of crime costs to society [26]. Vehicle crashes resulting from drugs or alcohol, simple assault, and driving while impaired costs were derived from monetary estimates of crime prevention [10]. To be clear, these prevention estimates are the estimated costs saved from preventing an incident, and not the cost of the intervention. Lastly, cost estimates based on per-incident occurrences in the United States were utilized for suicide attempts [34], vehicle crashes not related to drugs or alcohol [3], and drug overdoses [15]. Each of these latter three estimates were calculated by dividing the total annual costs for each incident type by the total number of each incident in a given year. Admittedly, crime cost estimates are not pristine and assume ubiquitous impact across individuals in a society. Cost estimates also do not capture the impact on

Table 1: Summary statistics for Indianapolis social harm 2012 & 2013

Type	Count	Cost Per Event	Total
Simple Assault	30802	$11,000.00	$338,822,000.00
Vehicle Crash No Influence	40718	$7,864.00	$320,206,352.00
Homicide	220	$1,278,424.00	$281,253,280.00
Aggravated Assault	11797	$19,537.00	$230,477,989.00
Larceny	53241	$3,523.00	$187,568,043.00
Robbery	6386	$21,398.00	$136,647,628.00
Residential Burglary	21468	$6,170.00	$132,457,560.00
Motor Vehicle Theft	9081	$10,534.00	$95,659,254.00
Vandalism	13641	$4,860.00	$66,295,260.00
Fraud	11371	$5,032.00	$57,218,872.00
Vehicle Crash Drugs or Alcohol	1610	$30,000.00	$48,300,000.00
Rape	1160	$41,247.00	$47,846,520.00
Drug Overdose	4112	$3,922.00	$16,127,264.00
Arson	723	$16,428.00	$11,877,444.00
Embezzlement	876	$5,480.00	$4,800,480.00
Forgery	481	$5,265.00	$2,532,465.00
DWI Arrest	3546	$500.00	$1,773,000.00
Suicide Attempt	134	$5,251.00	$703,634.00
Total	211367		$1,980,567,045.00

victims and communities. However, the cost estimates leveraged in the present study are validated to the extent they capture the financial severity of crime and harm costs to society.

In Table 1 we provide summary statistics for Indianapolis social harm including the volume of incidents over 2012 and 2013, the estimated cost per event to society, and the total cost over the two year period attributed to each event category. In Indianapolis, the top three categories in terms of cost to society are simple assault, vehicle crash (no alcohol influence) and homicide respectively.

3 Modeling and estimation framework

3.1 The modulated Hawkes process

Following [28], we consider a Hawkes process defined on a grid G with conditional intensity determined by,

$$\lambda_{g,m}(t) = \mu_{g,m} + \sum_{\substack{t>t_i \\ \vec{x}_i \in g \\ m_i = m}} \theta_m \omega_m \exp(-\omega_m (t - t_i)). \tag{1}$$

Here the intensity is defined for each category m of event type and each grid cell $g \in G$, where m_i denotes the category (mark) of event i, t_i the time, and $\vec{x}_i$ the spatial location. When viewed as a branching process, the parameter θ_m determines the expected number of events triggered by each event and the expected waiting time between a parent-daughter event pair is given by ω_m^{-1}.

To introduce spatial covariates, we define μ_m as a modulated Poisson process [11,39],

$$\mu_{g,m} = \exp(\vec{a}_m \cdot \vec{z}_g), \tag{2}$$

where the background intensity $\mu_{g,m}$ in grid cell g for event category m is log-linear with coefficients $\vec{a}_m$ for event type m and spatial covariates z_g in each grid cell g. Here we use zipcode level variables provided by the American Community Survey along with the average historical number of events of each type to serve as a leading indicator. Other variables one might consider include locations of crime attractors (liquor stores, schools, etc.), locations of parolees, housing density, satellite imagery and other sensor data.

3.2 l1-penalized Expectation-Maximization

The Model given by Eq. 1 can be viewed as a branching process [29] [38] where events occur according to a stationary Poisson process $\mu_{g,m}$ and then each event of type m generates a Poisson process with intensity $\theta_m \omega_m \exp(-\omega_m(t - t_i))$. Let $u_{ij} = 1$ when event i is the direct offspring of event j and 0 otherwise and $u_{ii} = 1$ when event i is a "background" or "spontaneous" event generated by the background Poisson process (and 0 otherwise). Given knowledge of u_{ij}, the estimation problem decouples into several independent Poisson estimation problems. In particular, the complete data log-likelihood is given by,

$$L = \sum_{m=1}^{M} \sum_{g \in G} \Bigg\{ \sum_{\substack{i: \vec{x}_i \in g \\ m_i = m}} u_{ii} \vec{a}_m \cdot \vec{z}_g - \exp(\vec{a}_m \cdot \vec{z}_g) T |g| \Bigg\} - \chi_m \|\vec{a}_m\|_1 + \tag{3}$$

$$\sum_{m=1}^{M} \sum_{g \in G} \Bigg\{ \sum_{\substack{t_i > t_j: \\ \vec{x}_i, \vec{x}_j \in g \\ m_i = m_j = m}} u_{ij} \log(\theta_m \omega_m \exp(-\omega_m(t_i - t_j))) - \sum_{\substack{i: \vec{x}_i \in g \\ m_i = m}} \int_{t_i}^{T} \theta_m \omega_m \exp(-\omega_m(t - t_i) dt \Bigg\}, \tag{4}$$

where $|g|$ denotes the area of grid cell g and $\| \cdot \|_1$ denotes the l1 norm that we have added to enforce sparsity of the spatial covariate coefficients. Given an initial guess for model parameters, the EM algorithm then proceeds iteratively by alternating between the E-step of updating the current guess of the branching structure and then the M-step of maximizing the complete data log likelihood with respect to the model parameters. Given the estimated branching structure $\hat{u}_{ij}$, maximization in the M-step is decoupled for the modulated Poisson coefficients and the Hawkes parameters. The Hawkes MLE parameters are determined by maximizing Equation 4 which yields estimates that are weighted sample means for exponential and Poisson random variables [28].

At each iteration of the EM algorithm, the l1 maximization problem given by Equation 3 must be solved. For this purpose we use forward-backward splitting [16] to minimize the negative of the right hand side of 3. First we define the proximal operator for the l1 term,

$$\text{prox}(\vec{b}, \tau) = \arg\min_{\vec{a}} \tau\chi\|\vec{a}\|_1 + \frac{1}{2}\|\vec{a} - \vec{b}\|_2. \tag{5}$$

Letting $\frac{dH}{d\vec{a}}$ denote the negative gradient of the first term in Equation 3, then forward-backward splitting iteratively solves the minimization problem with the two-stage iteration:

$$\hat{a}^{k+1} = \vec{a}^k - \tau\frac{dH^k}{d\vec{a}} \tag{6}$$

$$\vec{a}^{k+1} = \text{prox}(\hat{a}^{k+1}, \tau). \tag{7}$$

Equation 5 is similar to LASSO [37], however the second term is simply the square error of $\vec{a}$ and $\hat{a}^{k+1}$. The proximal operator in this case is given by the shrinkage operator [16],

$$\text{prox}(\vec{a}, \tau)_i = \text{sign}(a_i)\max\{|a_i| - \chi\tau, 0\}. \tag{8}$$

The regularization parameter χ_m is selected at each EM iteration using 10-fold cross-validation to maximize the complete data log-likelihood conditioned on the current estimate of the branching structure.

The branching structure u_{ij} can be estimated in the E-step using the ratio of the background rate and triggering kernel components to the overall intensity at each event as is done in [28]. We have included Matlab code for simulation and estimation of the modulated Hawkes process on Github [1].

3.3 Social harm index and model evaluation

For each event type m we have a secondary mark $c(m)$ representing the average societal cost of an event of type m. Given this cost mark, we can then define a dynamic social harm index $SI_g(t)$ in each grid cell g as the expected cost per unit time,

$$SI_g(t) = \sum_{m=1}^{M} c(m)\lambda_{g,m}(t). \tag{9}$$

The dynamic social harm index can then be used to rank hotspots over a given time interval, where the top k hotspots are flagged for intervention. Because this type of ranking is common in hotspot analysis and policing, a popular accuracy metric is the Predictive Accuracy Index (PAI). The PAI is the percentage of events captured in the top k hotspots divided by the percentage of city area that the k hotspots comprise. We therefore propose a social harm variant of PAI for assessing social harm indices that we will refer to as S-PAI,

$$\text{S-PAI@k} = \frac{\text{\% societal cost captured in top k hotspots}}{\text{\% city area covered by k hotspots}}. \tag{10}$$

Alternative metrics have been proposed for evaluating crime forecasts and hotspot selection methods that could also be extended to social and crime harm indices. Mean Square Error and Mean Absolute Percent Error are used in evaluating crime forecasts over larger space-time windows [17] and qualitative metrics measuring crime hotspot compactness and variability are proposed in [2]. Because our goal here is to measure the potential social harm reduction under limited resources in the highest risk hotspots, we restrict our attention to the S-PAI.

4 Results

4.1 Synthetic Data

We first validate the EM algorithm for Equation 1 on simulated data from a modulated Hawkes process. We define a 50x50 grid where each cell has 100 covariates and coefficients drawn from independent uniform random variables: $z \sim U[0,1]$ and $a \sim U[-1,1]$. For the coefficients a, we then set half equal to zero to simulate sparsity. We note that the regularization parameter χ_m is selected automatically using cross-validation within the EM algorithm.

We simulate two examples where the parameters are chosen to be similar to the values estimated from Indianapolis crime data. In the first example we let $\theta = .1$ and $\omega = .1$ and simulate the process for $T = 1000$ time units where the expected number of events is O(4000). In the second example we let $\theta = .1$ and $\omega = 1$ and simulate the process for $T = 100$ time units where the expected number of events is O(500).

In Figure 1 we plot parameter estimates for 100 simulations of each example along with the true parameter values in red. We find good agreement between the estimated parameters and the true values. On the left we plot the 90% pointwise confidence range in gray for the estimated spatial covariate coefficients. We note that there is some bias of the coefficients towards zero, which is expected due to the l1 penalization. While in each simulation a percentage of the coefficients are zero, over 100 simulations each individual coefficient is estimated to be non-zero in a percentage of simulations (which is why the 90% range extends on either side of zero). The variance of the estimators increase as the number of events in the sample decrease.

4.2 IMPD Social Harm Data

Next we apply our methodology to Indianapolis social harm data from 2012 and 2013. We use a 100x100 grid to cover Indianapolis and for spatial covariates we use 46 demographic and economic population variables from the American Community Survey at the zipcode level, along with 18 variables defined as the grid cell crime rate over the first half of 2012 for each of the 18 event categories. We note that the crime rate covariates are highly correlated with the self-exciting component of the model, however this is consistent with self-exciting point process models where the background intensity μ and triggering kernel are both estimated

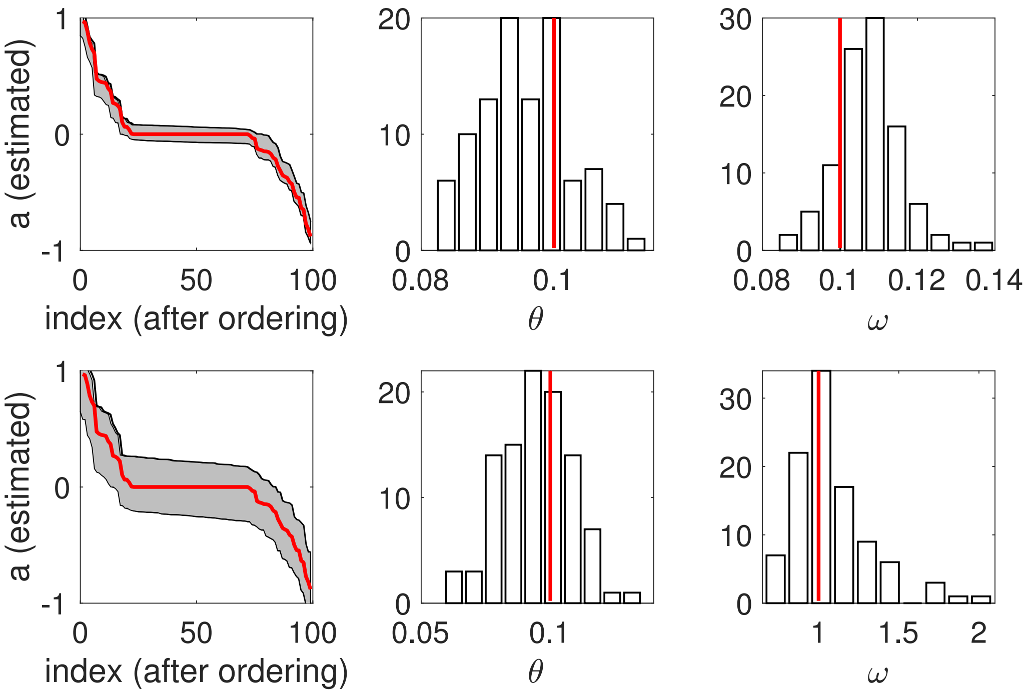

Figure 1: Parameter estimates for 100 realizations of a modulated Hawkes process with true values in red. On the left, 90% pointwise range in gray for estimated spatial covariate coefficients. Top corresponds to simulation with $O(4000)$ events and $\theta = .1$, $\omega = .1$. Bottom corresponds to simulation with $O(500)$ events and $\theta = .1$, $\omega = 1$.

from the same event data. While KDE is normally used to estimate the background intensity, we have incorporated this modeling step into the regression framework along with the census covariates. The census variables are assigned to each grid cell using the zipcode of the cell centroid and all predictor variables are whitened to have mean zero and variance one.

For cross-validation, we first train the model over the second half of 2012. In Table 2 we list the Hawkes parameter values for each event type along with the number of non-zero spatial covariate coefficients selected by the l1-penalized EM algorithm. For testing, we apply the trained model to each 4-hour interval for each day in 2013. We use the social harm Hawkes intensity given in Equation 9 to rank the top 50, 100 and 200 hotspots (0.5%, 1% and 2% of the city) and compute the S-PAI over 2013. We note that S-PAI is not computed for each 4-hour interval and averaged, but instead it is calculated as the percentage of harm captured in the 4-hour interval hotspots over the course of 2013 divided by the percent area flagged. We bootstrap the events in 2013 to estimate standard errors of S-PAI.

We compare to a social harm index of the form,

$$SI_g = \sum_{i \in g} c(m_i), \tag{11}$$

Table 2: Hawkes parameters (ω in units of days^{-1})

type	θ	ω	# non-zero a
DWI Arrest	0.0032	10.0647	19
Drug Overdose	0.0477	0.0452	20
Vandalism	0.0522	0.1652	14
Fraud	0.0369	0.1387	34
Suicide Att	0.0237	12.000	8
Forgery	0.0105	0.5244	16
Embezzlement	0.2253	0.0221	7
Larceny	0.1651	0.074	16
Res Burg	0.0648	0.0988	13
Veh Crash No Inf	0.2018	0.0521	15
MVT	0.0294	0.1237	17
Simple Assault	0.0585	0.0867	12
Arson	0.0267	0.3322	13
Agg Assault	0.0165	0.1574	13
Robbery	0.0377	0.0537	17
Veh Crash Drug/Alc	0.0024	23.999	16
Rape	0.0106	0.0641	24
Homicide	0.0188	1.7143	39

that is the total cost in each cell over 2012, similar to recently introduced social harm indices that sum prison sentence length. We also compare our methodology to the Hawkes process model used in [28]. The model is a Hawkes process of the form,

$$\lambda_g(t) = \mu_g + \sum_{\substack{t>t_i \\ \vec{x}_i \in g}} \theta\omega \exp(-\omega(t - t_i)), \tag{12}$$

only trained on property crime (burglary, theft from vehicle and auto-theft) aggregated together.

In Table 3 we compare the S-PAI values of the three methods for 50, 100, and 200 hotspots selected. The static social harm index is 4x better than random chance at ranking the top 100 hotspots each day, where 4% of the cost of social harm is captured in 1% of the city each day. However, the property Hawkes process is significantly better capturing 8% of social harm cost in the same number of hotspots. The S-PAI for the social harm Hawkes process is over 12, meaning that over \$120 million of the total annual \$1billion in social harm cost to Indianapolis can be captured in 1% of the city.

In Figure 2, we plot an example hotspot map from a day in 2013. We color code each hotspot by the most frequent event type occurring in the grid cell. Here we see four main types of hotspots, namely vehicle crash, burglary, larceny, and assault hotspots. Police interventions would need to be tailored to each event type across these disparate types of social harm.

Table 3: S-PAI and standard errors for varying numbers of hotspots.

# Hotspots	50	100	200
Dynamic Soc. Harm Hawkes (Eq 9)	15.0 (0.8)	12.5 (0.4)	9.9 (0.3)
Property Hawkes (Eq 12)	9.4 (0.8)	8.4 (0.6)	7.1 (0.3)
Soc. Harm Index (Eq 11)	5.0 (0.4)	3.7 (0.2)	2.7 (0.1)

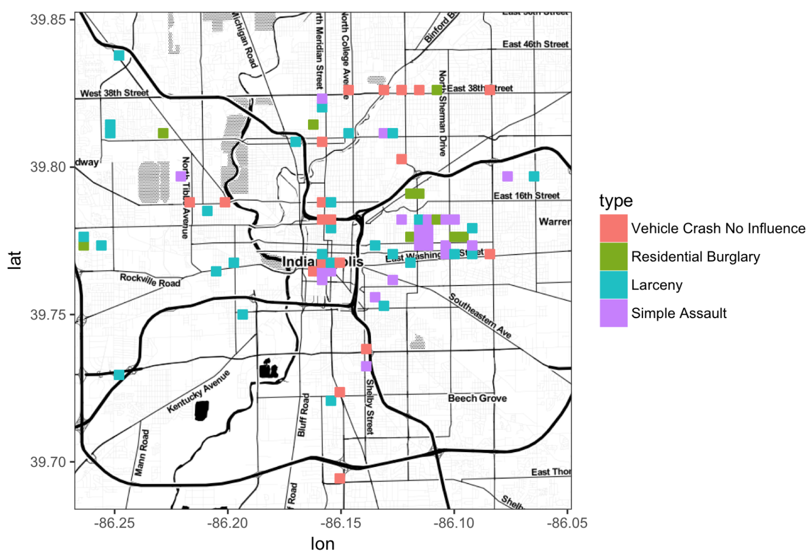

Figure 2: Example social harm hotspot map for a day in 2013 in Indianapolis. Hotspots are color coded by the most frequent event type in the cell.

We note that while these categories are the most prevalent, all 18 categories are captured to some degree within the top 100 hotspots. In Figure 3, we plot the number of crimes captured by each of the three methods disaggregated by event type. As expected, the property crime Hawkes process captures the most property crime, whereas the social harm Hawkes process captures significantly more larceny, vehicle crashes, and assaults. Thus there is a tradeoff when using a social harm based model for hotspot policing.

In Table 4, we display the top three spatial covariates for each event category selected as the covariates with the largest (magnitude) estimated coefficients in the background intensity. Here we find several patterns that emerge. Vacant housing is a strong indicator for crimes such as vandalism, arson, and burglary. Vehicle crashes where alcohol is not involved are strongly correlated with spatial areas where a large number of the population leaves to work between 7 and 7:30am. Some unexpected leading indicators also emerge, for example the average rate of motor vehicle theft in a hotspot is a good predictor for fraud, larceny and simple assault. These covariates may provide some insight into long-term problem-oriented solutions to social harm because they are based on census variables reflecting long-term characteristics of a spatial region of the city.

Table 4: Top three spatial covariates for each event type

Event type	z_1	z_2	z_3
DWI Arrest	DWI Arrest	White	MVT
Drug Overdose	Income.10k-15k	Drug Overdose	White
Vandalism	Vacant.Housing	MVT	Vandalism
Fraud	MVT	Unemployment.Rate	Embezzlement
Suicide Att	Simple Assault	Mean.Travel.Time.Min	Asian
Forgery	Forgery	MVT	Vacant.Housing
Embezzlement	Robbery	Veh Crash No Inf	Leave.to.Work.7.730am
Larceny	MVT	Res Burg	Vandalism
Res Burg	Vacant.Housing	Res Burg	Black
Veh Crash No Inf	Leave.to.Work.7.730am	MVT	DWI Arrest
MVT	MVT	Res Burg	Income.75k-100k
Simple Assault	MVT	Res Burg	Income.150k.200k
Arson	Vacant.Housing	MVT	Res Burg
Agg Assault	Income.75k-100k	MVT	Res Burg
Robbery	Robbery	Vacant.Housing	Income.75k-100k
Veh Crash Drug/Alc	White	DWI Arrest	Veh Crash No Inf
Rape	Rape	Income.10k-15k	MVT
Homicide	Income.10k-15k	Poverty.Rate	Hispanic

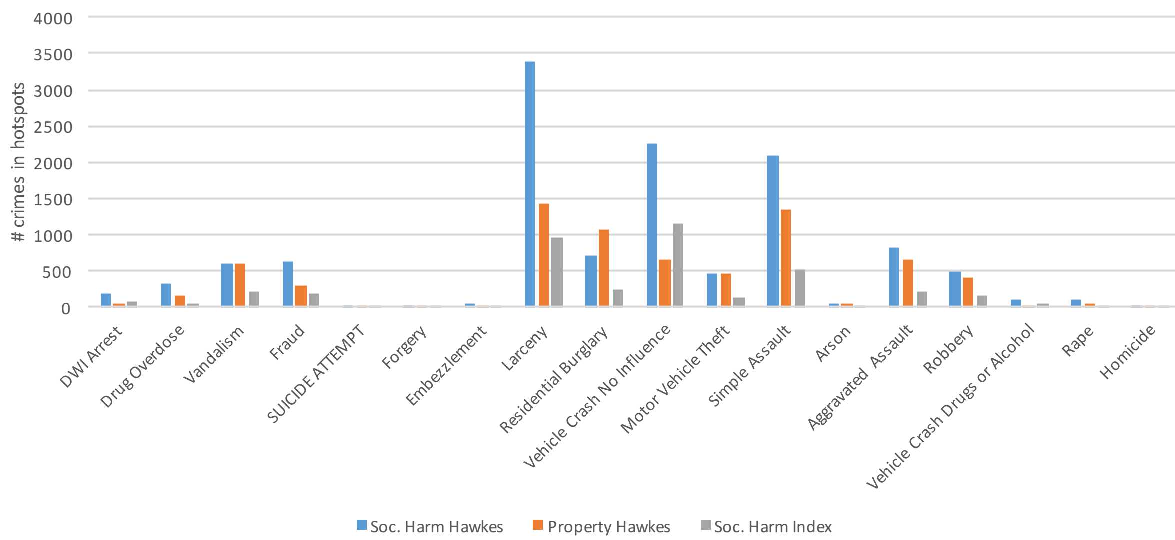

Figure 3: Number of crimes captured by each of the three methods disaggregated by event type.

5 Discussion

We showed how static models for crime prediction using crime or social harm indices may be improved using dynamic point process models of social harm. Social harm indices suffer from high variance, as high severity/low volume events may dominate the risk estimate of hotspots where they occur. On the other hand, single crime type models (or those focusing solely on property or violent crime) fail to properly weight social harm events by their severity. To address these problems, we introduced a novel sparse modulated Hawkes process for modeling disparate social harm event categories, incorporating spatial covariates, near-repeat effects (self-excitation), and periodic trends. This methodology significantly improves upon existing uses of social harm indices and single crime type point processes in terms of the S-PAI.

While this methodology shows promise, field trials are needed to assess the efficacy of such an approach similar to predictive policing trials focusing on property crime [28]. Predictive policing trials focusing on social harm will present challenges in that a wider range of interventions will be necessary given the wider range of event types and collaborations with other city agencies and community stake holders may be necessary. In practice, not all event categories may be easily prevented by police, for example fraud and embezzlement, as such events occur largely outside the reach of day-to-day police operations. Crime deterrence can be achieved through an offender's increased perception of apprehension, thus police presence and activity. As such, effective deterrence interventions should follow empirical evidence that suggests a focus on social harm events that occur primarily "on the streets" where police are most likely to generate crime prevention benefits. It may be plausible to assume that increased – or focused – police activity in a high social harm risk area could translate to crime prevention or displacement of "off the street" events such as fraud and embezzlement as such offenders may seek offending locations away from potential police contact. Though this specific type of displacement has not been empirically tested, existing evidence demonstrates that displacement and diffusion of benefits (additional deterrence gains outside of the intervention focus) are likely to occur; however displacement is likely to occur at a lower rate compared to the intervention focus [22]. Thus, focused police deterrence is likely to yield a net crime prevention gain. Future research employing an operational social harm experiment should seek to capture this potential crime prevention benefit. More specifically, future inquiries should attempt to compare the displacement and diffusion of benefits from social harm-focused police activities to those of more traditional hotspots policing approaches.

Future research may also focus on improving dynamic models of social harm. The models considered in this work do not account for spatial auto-correlation that may be present in the spatial covariates as well as through endogenous, self-exciting effects in the point patterns. Methods for controlling for spatial auto-correlation in spatial regression models are discussed in [18] and [4] and can be adapted for the Poisson regression in Equation 2. Methods for modeling spatial correlations in Hawkes processes are discussed in [27]. Daily, weekly and seasonal effects can also be incorporated into point process models of crime [36] and may lead to improvements in accuracy. Cost estimates are employed in the present study and demonstrate the potential to be incorporated with harm indices that leverage sentencing data and guidelines. Scholars should seek methods to merge the two weighting techniques to determine potential gains in the identification of geographic-specific harms and development of effective interventions. Beyond point processes, machine learning methods may be able to further improve the accuracy of dynamic social harm indices and should be tested on historical data. Software applications will also need to be re-envisioned that effectively communicate the information contained in dynamic social harm indices in near real time to officers in the field. This is especially true if collaborations with other city agencies are reflected in social harm based predictive analytics applications. These questions will be addressed in future research.

Acknowledgements

This work was supported in part by NSF grants S&CC-1737585, SES-1343123, and ATD-1737996. G.M. is a co-founder and director of PredPol, a company offering predictive policing services to law enforcement agencies.

References

[1] GitHub Repository: https://github.com/gomohler/, 2017.

[2] Monsuru Adepeju, Gabriel Rosser, and Tao Cheng. Novel evaluation metrics for sparse spatio-temporal point process hotspot predictions-a crime case study. *International Journal of Geographical Information Science*, 30(11):2133–2154, 2016.

[3] National Highway Traffic Safety Administration et al. The economic and societal impact of motor vehicle crashes, 2010. *Report DOT HS*, 812:013, 2014.

[4] Luc Anselin. Under the hood issues in the specification and interpretation of spatial regression models. *Agricultural economics*, 27(3):247–267, 2002.

[5] Colin Babyak, Asma Alavi, Krista Collins, Amanda Halladay, and Dawn Tapper. The methodology of the police-reported crime severity index. *Ottawa: Statistics Canada*, 2009.

[6] Mark Bangs. Developing a crime severity score for england and wales using data on crimes recorded by the police. *Office for National Statistics*, 2016.

[7] Kate J Bowers, Shane D Johnson, and Ken Pease. Prospective hot-spotting the future of crime mapping? *British Journal of Criminology*, 44(5):641–658, 2004.

[8] Anthony A Braga and Brenda J Bond. Policing crime and disorder hot spots: A randomized controlled trial. *Criminology*, 46(3):577–607, 2008.

[9] Spencer Chainey, Lisa Tompson, and Sebastian Uhlig. The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, 21(1):4–28, 2008.

[10] Mark A Cohen and Alex R Piquero. New evidence on the monetary value of saving a high risk youth. *Journal of Quantitative Criminology*, 25(1):25–49, 2009.

[11] David R Cox. The statistical analysis of dependencies in point processes. *Stochastic point processes*, pages 55–66, 1972.

[12] Sophie Curtis-Ham and Darren Walton. The new zealand crime harm index: Quantifying harm using sentencing data. *Policing: A Journal of Policy and Practice*, 2017.

[13] Gavin Dudfield, Caroline Angel, Lawrence W Sherman, and Sarah Torrence. The "power curve" of victim harm: Targeting the distribution of crime harm index values across all victims and repeat victims over 1 year. *Cambridge Journal of Evidence-Based Policing*, pages 1–21, 2017.

[14] Matthew Fielding and Vincent Jones. 'disrupting the optimal forager': predictive risk mapping and domestic burglary reduction in trafford, greater manchester. *International Journal of Police Science & Management*, 14(1):30–41, 2012.

[15] Curtis S Florence, Chao Zhou, Feijun Luo, and Likang Xu. The economic burden of prescription opioid overdose, abuse, and dependence in the united states, 2013. *Medical care*, 54(10):901–906, 2016.

[16] Tom Goldstein, Christoph Studer, and Richard Baraniuk. A field guide to forward-backward splitting with a fasta implementation. *arXiv preprint arXiv:1411.3406*, 2014.

[17] Wilpen Gorr, Andreas Olligschlaeger, and Yvonne Thompson. Short-term forecasting of crime. *International Journal of Forecasting*, 19(4):579–594, 2003.

[18] Carol A Gotway and Walter W Stroup. A generalized linear model approach to spatial data analysis and prediction. *Journal of Agricultural, Biological, and Environmental Statistics*, pages 157–178, 1997.

[19] Victoria A. Greenfield and Letizia Paoli. A framework to assess the harms of crimes. *The British Journal of Criminology*, 53(5):864–885, 2013.

[20] Shane D Johnson. *Prospective crime mapping in operational context: Final report.* 2007.

[21] Shane D Johnson, Kate J Bowers, Dan J Birks, and Ken Pease. Predictive mapping of crime by promap: accuracy, units of analysis, and the environmental backcloth. In *Putting crime in its place*, pages 171–198. Springer, 2009.

[22] Shane D Johnson, Rob T Guerette, and Kate Bowers. Crime displacement: what we know, what we don't know, and what it means for crime reduction. *Journal of Experimental Criminology*, 10(4):549–571, 2014.

[23] Leslie W Kennedy, Joel M Caplan, and Eric Piza. Risk clusters, hotspots, and spatial intelligence: risk terrain modeling as an algorithm for police resource allocation strategies. *Journal of Quantitative Criminology*, 27(3):339–362, 2011.

[24] Steven D Levitt and Thomas J Miles. Economic contributions to the understanding of crime. *Annu. Rev. Law Soc. Sci.*, 2:147–164, 2006.

[25] Hua Liu and Donald E Brown. Criminal incident prediction using a point-pattern-based density model. *International Journal of Forecasting*, 19(4):603–622, 2003.

[26] Kathryn E McCollister, Michael T French, and Hai Fang. The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and alcohol dependence*, 108(1):98–109, 2010.

[27] George Mohler. Marked point process hotspot maps for homicide and gun crime prediction in chicago. *International Journal of Forecasting*, 30(3):491–497, 2014.

[28] George O Mohler, Martin B Short, Sean Malinowski, Mark Johnson, George E Tita, Andrea L Bertozzi, and P Jeffrey Brantingham. Randomized controlled field trials of predictive policing. *Journal of the American Statistical Association*, 110(512):1399–1411, 2015.

[29] GO Mohler, MB Short, P Jeffrey Brantingham, FP Schoenberg, and GE Tita. Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108, 2011.

[30] Jerry H Ratcliffe. Towards an index for harm-focused policing. *Policing: A Journal of Policy and Practice*, 9(2):164–182, 2015.

[31] Jerry H Ratcliffe, Travis Taniguchi, Elizabeth R Groff, and Jennifer D Wood. The philadelphia foot patrol experiment: a randomized controlled trial of police patrol effectiveness in violent crime hotspots*. *Criminology*, 49(3):795–831, 2011.

[32] Rose A Rudd, Noah Aleshire, Jon E Zibbell, and R Matthew Gladden. Increases in drug and opioid overdose deaths—united states, 2000–2014. *American Journal of Transplantation*, 16(4):1323–1327, 2016.

[33] Robert J Shapiro and Kevin A. Hassett. The economic benefits of reducing violent crime. a case study of 8 america cities. *Center for American Congress*, 2012.

[34] Donald S Shepard, Deborah Gurewich, Aung K Lwin, Gerald A Reed, and Morton M Silverman. Suicide and suicidal attempts in the united states: costs and policy implications. *Suicide and life-threatening behavior*, 46(3):352–362, 2016.

[35] Lawrence Sherman, Peter William Neyroud, and Eleanor Neyroud. The cambridge crime harm index: measuring total harm from crime based on sentencing guidelines. *Policing: A Journal of Policy and Practice*, 10(3):171–183, 2016.

[36] Shinichiro Shirota, Alan E Gelfand, et al. Space and circular time log gaussian cox processes with application to crime event data. *The Annals of Applied Statistics*, 11(2):481–503, 2017.

[37] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[38] Alejandro Veen and Frederic P Schoenberg. Estimation of space–time branching process models in seismology using an em–type algorithm. *Journal of the American Statistical Association*, 103(482):614–624, 2008.

[39] Rasmus Waagepetersen. Estimating functions for inhomogeneous spatial point processes with incomplete covariate data. *Biometrika*, pages 351–363, 2008.

[40] Xiaofeng Wang, Donald E Brown, and Matthew S Gerber. Spatio-temporal modeling of criminal incidents using geographic, demographic, and twitter-derived information. In *2012 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pages 36–41. IEEE, 2012.

[41] Cristobal Weinborn, Barak Ariel, Lawrence W Sherman, and Emma O'Dwyer. Hotspots vs. harmspots: Shifting the focus from counts to harm in the criminology of place. *Applied Geography*, 2017.

[42] David Weisburd, Laura A Wyckoff, Justin Ready, John E Eck, Joshua C Hinkle, and Frank Gajewski. Does crime just move around the corner? a controlled study of spatial displacement and diffusion of crime control benefits*. *Criminology*, 44(3):549–592, 2006.