International Journal of Forecasting 30 (2014) 491–497

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Marked point process hotspot maps for homicide and gun crime prediction in Chicago☆

George Mohler

Department of Mathematics and Computer Science, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053, USA

ARTICLE INFO

Keywords:
Marked point process
Expectation-maximization
Crime hotspot
Leading indicators

ABSTRACT

Crime hotspot maps are a widely used and successful method of displaying spatial crime patterns and allocating police resources. However, hotspot maps are often created over a single timescale using only one crime type. In the case of short-term hotspot maps that utilize several weeks of crime data, risk estimates suffer from a high variance, especially for low frequency crimes such as homicide. Long-term hotspot maps that utilize several years of data fail to take into account near-repeat effects and emerging hotspot trends. In this paper we show how point process models of crime can be extended to include leading indicator crime types, while capturing both short-term and long-term patterns of risk, through a marked point process approach. Several years of data and many different crime types are systematically combined to yield accurate hotspot maps that can be used for the purpose of predictive policing of gun-related crime. We apply the methodology to a large, open source data set which has been made available to the general public online by the Chicago Police Department.

1. Introduction

In this article we develop a methodology for the prediction of homicide, along with precursory gun crimes, with an application to predictive policing. The problem we consider is as follows: given all past homicide events (time and location), and all other gun-related crime reports in a police records management system (RMS), rank a list of geographic areas in a city according to the risk of homicide. We use homicide and gun crime in Chicago for illustration, but our methodology applies more generally to other crime types and their leading indicators.

Similar problems have been considered in a number of studies (Bowers, Johnson, & Pease, 2004; Chainey, Tompson, & Uhlig, 2008; Kennedy, Caplan, & Piza, 2010; Liu & Brown, 2003; Mohler, Short, Brantingham, Schoenberg, & Tita, 2011; Wang & Brown, 2012; Wang, Brown, & Gerber, 2012; Weisburd, Groff, & Yang, 2012), and their solutions can be used to allocate patrol resources each day according to risk. The algorithms typically fall into one of two broad categories, namely nonparametric methods utilizing only event data (kernel hotspot maps being the predominant choice), and multivariate models that explicitly incorporate additional variables such as demographics (Wang et al., 2012), income levels (Liu & Brown, 2003), distance from crime attractors (Kennedy et al., 2010; Liu & Brown, 2003; Wang et al., 2012), and leading-indicator crimes (Cohen, Gorr, & Olligschlaeger, 2007; Gorr, 2009). In multivariate models of hotspots, static variables such as demographics and distance to crime attractors are predictive of long-term crime hotspots, whereas recent event activity is predictive of short-term hotspots. In the case of kernel density estimation (KDE) and other event-based approaches, models usually do not reflect multiple timescales, and either long term hotspot maps are created using several

☆ This research was supported by ARO grant 58344-MA and NSF grant DMS-0968309.

E-mail address: gmohler@scu.edu.

http://dx.doi.org/10.1016/j.ijforecast.2014.01.004
0169-2070/

years of data (Weisburd et al., 2012), or short term hotspot maps are created using several weeks or months of data and a spatial bandwidth in the order of tens or hundreds of meters (Chainey et al., 2008; Kennedy et al., 2010; Wang & Brown, 2012). One of the goals of this article is to illustrate how event-based models of hotspots can be constructed to estimate both short and long term hotspots in a systematic way.

In recent articles which have focused on multivariate modeling (Kennedy et al., 2010; Wang & Brown, 2012), short-term hotspot maps have served as a method for accuracy comparison, due to their widespread operational use. Here, the short-term hotspot map serves as a straw man, as low event counts lead to a high variance in the risk assessment. On the other hand, multivariate models are well suited for handling low event counts by reducing the variance through the introduction of spatial variables which are correlated with crime rates. However, police departments' RMSs often contain crime reports going back years, if not decades, thus providing large data sets that can be used to reduce the variance and significantly improve the accuracy of crime hotspot maps. Here, we believe that the benefits of hotspot maps stand out, as for high event counts (1) the error due to variance is reduced significantly, (2) nonparametric estimates are potentially less biased than multivariate models, and (3) kernel-based hotspot maps facilitate the use of predictive analytics software that is robust and portable across various agencies without the need to gather data outside the RMS.

We extend the point process model of burglary introduced by Mohler et al. (2011) to a marked point process that allows for several years of crime data, and multiple crime types, being utilized by hotspot maps. The model incorporates both fixed risk heterogeneity across the city and temporally dynamic risk. While the model of Mohler et al. (2011) is fully non-parametric, we consider a parametric version of the triggering kernel in order to balance the added computational cost associated with the incorporation of leading indicator crimes.

Our model is related to the decomposition of hotspots into chronic and temporary hotspots by Gorr and Lee (2012). Chronic hotspots are long term in duration and necessitate problem-oriented policing strategies in order to address the root causes of crime (Clarke & Eck, 2005; Weisburd et al., 2012). Chronic hotspots, defined by a high crime volume over several years rather than several weeks, can capture a large percentage of all crime within a small percentage of the area of a city (Weisburd et al., 2012). Temporary hotspots, on the other hand, last on the time scale of weeks or months. Thus, models and policing strategies must be able to detect and react to emerging trends in order to identify temporary hotspots, otherwise the hotspots may have moved before maps can be generated and patrols deployed. For example, the Los Angeles Police Department updates predictive policing models for every 8 h shift as new crimes are reported, and directs patrols accordingly.

One technical issue that arises in the creation of kernel hotspot maps is the selection of the bandwidth or "search radius". In this work, an Expectation–Maximization (EM) algorithm is developed that allows for the automated selection of model parameters, thus avoiding the need for hotspot bandwidth parameters to be tuned manually by the crime analyst. We apply the methodology to gun crime and homicide data in Chicago, and illustrate that the predictive accuracy of hotspot maps is improved significantly when utilizing large data sets over several years. In particular, we find that dynamic hotspots account for only 12%–15% of gun crime in Chicago, and that chronic hotspots are the most dominant component of the estimated point process model.

The outline of the article is as follows. In Section 2, we briefly review the mathematics of hotspot maps and draw the connections between kernel hotspot maps, self-exciting point processes, and mixture models. In Section 3, we provide details of the marked point process model, and in Section 4 we provide details of the EM algorithm for estimating model parameters. In Section 5, we provide results on the application of the model to homicide prediction in Chicago.

2. Hotspot maps, self-exciting point processes, and mixture models

2.1. Hotspot maps

Crime hotspot maps are a widely used method for visualizing spatial crime patterns, where spatial maps are color coded based upon levels of criminal activity. Our focus here is on kernel-based hotspot maps, which exhibit the important features of hotspot mapping, and we refer the reader to Chainey and Ratcliffe (2005) for a more comprehensive treatment of the subject.

Given a spatial–temporal crime data set of event locations (x_i, y_i) and times t_i, a common method of constructing a kernel hotspot map is to use a subset of the data consisting of all crimes occurring within a specified time interval $[T_1, T_2]$:

$$\lambda(x, y) = \sum_{t_i \in [T_1, T_2]} g(x - x_i, y - y_i). \quad (1)$$

In practice, this time window is often chosen to be the past several weeks or months leading up to the present, though it could also be years (Gorr & Lee, 2012) when estimating chronic hotspots. The kernel g is often a 2D function that decays from the origin.

While the kernels are defined irrespective of a discretization of space, evaluating $\lambda(x, y)$ on a grid is still necessary for visualization. Furthermore, hotspot maps can be used to flag high crime areas for policing intervention (Bowers et al., 2004; Chainey et al., 2008), in which case the values of $\lambda(x, y)$ in different discrete regions of a city are used to rank those areas in terms of priorities for the receipt of policing attention.

From a mathematical perspective, a hotspot map can be viewed as a nonparametric estimate of a stationary Poisson process over the time interval $[T_1, T_2]$. Furthermore, by taking the time interval to be moving, $[t - T, t]$, with t being the present day, hotspot maps can also function as a nonparametric estimator of a non-stationary point