Applied Geography 99 (2018) 89–97

Contents lists available at ScienceDirect

Applied Geography

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

A spatio-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation

Yujie Hu[a], Fahui Wang[b,*], Cecile Guin[c], Haojie Zhu[b]

[a] School of Geosciences, University of South Florida, Tampa, FL 33620-5550, USA
[b] Department of Geography & Anthropology, Louisiana State University, Baton Rouge, LA 70803, USA
[c] Office of Social Service Research and Development (OSSRD), School of Social Work, Louisiana State University, Baton Rouge, LA 70803, USA

ARTICLE INFO

Keywords:
Spatio-temporal kernel density estimation (STKDE)
Optimal bandwidth
Significance test
Predictive accuracy index (PAI) curve
Crime hotspot prediction
Residential burglary
Baton rouge

ABSTRACT

Predictive hotspot mapping plays a critical role in hotspot policing. Existing methods such as the popular kernel density estimation (KDE) do not consider the temporal dimension of crime. Building upon recent works in related fields, this article proposes a spatio-temporal framework for predictive hotspot mapping and evaluation. Comparing to existing work in this scope, the proposed framework has four major features: (1) a spatio-temporal kernel density estimation (STKDE) method is applied to include the temporal component in predictive hotspot mapping, (2) a data-driven optimization technique, the likelihood cross-validation, is used to select the most appropriate bandwidths, (3) a statistical significance test is designed to filter out false positives in the density estimates, and (4) a new metric, the predictive accuracy index (PAI) curve, is proposed to evaluate predictive hotspots at multiple areal scales. The framework is illustrated in a case study of residential burglaries in Baton Rouge, Louisiana in 2011, and the results validate its utility.

1. Introduction

It is well-known to researchers and law enforcement agencies that crime tends to be concentrated in certain areas (e.g., Bernasco, Johnson, & Ruiter, 2015; Bowers, Johnson, & Pease, 2004; Chainey, Tompson, & Uhlig, 2008). Various spatial analysis methods have been applied to detect spatial concentrations of past crime incidents and predict patterns of future crime. The first type of methods aggregates crime incidents to counts and then calculates rates by geographic boundaries such as census units. It then employs the multivariate regression analysis to study the relationship between crime rates and a wide range of crime attraction and inhibition variables such as socio-economic conditions, neighborhood demographics, land use types, cultural values, and substance abuse histories (e.g., Bushman, Wang, & Anderson, 2005; Kikuchi & Desmond, 2010; Peterson & Krivo, 2010). From these factors, predictions of future crime rates in certain areas can be made. Another line of approaches focuses on identifying spatial clustering patterns of incidents and relying on the clustering locations (i.e., hotspots) for crime predictions. These approaches are usually known as predictive hotspot mapping and include spatial ellipses, grid thematic mapping, and kernel density estimation (KDE) among others (Chainey et al., 2008; Ratcliffe, 2010). This article specifically focuses on using predictive hotspot mapping techniques to identify and predict where crime is most likely to take place. By this means, the police can focus its resources on predicted crime hotspots, a practice termed "hotspot policing" (Braga, 2007; Chainey, 2013; Eck, Chainey, Cameron, Leitner, & Wilson, 2005; Ratcliffe, Taniguchi, Groff, & Wood, 2011). Predictive hotspot mapping is widely used by large law enforcement agencies, especially those large ones serving more than 500,000 population in the U.S. (Hart & Zandbergen, 2014; Reaves, 2010).

KDE is a popular hotspot mapping method. It converts point incidents to a density surface that summarizes the point distribution. Specifically, this technique estimates the concentration of events at each sample location by (1) placing a kernel over a predefined area around that location, (2) assigning more weights to nearby events than distant ones, and (3) summing up the weighted events within the kernel. Areas on the surface with high density values above a predefined threshold are defined as hotspots (Hu, Miller, & Li, 2014). For example, Chainey et al. (2008) compared KDE to other techniques in predictive hotspot mapping and found that KDE significantly outperformed others. Based on KDE, Maciejewski et al. (2010) proposed a visual analytical approach to detecting and visualizing crime hotspots. Later, Maciejewski et al. (2011) and Malik, Maciejewski, Towers, McCullough, and Ebert (2014) applied that approach to forecast crime hotspots. Some researchers took a step further by testing whether

* Corresponding author.
E-mail address: fwang@lsu.edu (F. Wang).

https://doi.org/10.1016/j.apgeog.2018.08.001
Received 16 January 2018; Received in revised form 2 August 2018; Accepted 4 August 2018
0143-6228/

incorporating auxiliary data into KDE could improve its performance. For instance, Gerber (2014) blended crime related Twitter records with KDE and found that the addition of Twitter data improved prediction accuracy over the plain KDE in some crime types. Justification for hotspot policing is the belief that areas with high crime incidents in the past will remain so for some time. For example, environmental criminologists attribute the spatial clustering of crimes to the presence of motivated offenders, availability of potential targets, and lack of sufficient guardianship or deterrence in those areas (Brantingham & Brantingham, 1981, pp. 27–54; Clarke & Eck, 2003; Cohen & Felson, 1979; Felson & Clarke, 1998). As these factors remain largely stable for a period of time, the spatial pattern persists. That is also supported by the phenomena of repeat victimization and near repeats. In repeat victimization, targets victimized in a recent crime are more likely to become targets of new crimes again in near future (Farrell & Pease, 2014). In near repeats, suitable targets in close proximity to the location of a recent crime will experience higher risk of victimization in near future (Bower and Johnson 2005; Johnson, 2008; Ratcliffe & Rengert, 2008). As the traditional KDE considers only the spatial dimension, hereafter it is referred to as "*spatial kernel density estimation (SKDE)*."

However, crime tends to cluster temporally in addition to spatially. In other words, crime is concentrated in certain parts of the city during certain times of the day (Eck & Weisburd, 1995; Farrell & Pease, 1994; Nelson, Bromley, & Thomas, 2001; Ratcliffe, 2010). Ignoring the temporal component of crime deprives researchers and practitioners of the opportunity to target specific time periods with elevated crime risks. Temporal dimension is also integral to the study of crime displacement and diffusion of benefits, as crimes may displace to different areas, time, or even types when law enforcements adopt hotspot patrolling (Eck, 1993; Eck & Weisburd, 2015). Most recently, increasing efforts have been made to integrate temporal data into Geographic Information Systems (GIS) and build models that can predict when and where future crimes will occur (Jefferson, 2018; Rummens, Hardyns, & Pauwels, 2017). In the work by Maciejewski et al. (2010), they combined a time series analysis termed the cumulative summation model into KDE to allow for an additional temporal view of crime patterns. Maciejewski et al. (2011) and Malik et al. (2014) further advanced their work by introducing a seasonal trend decomposition method to account for the seasonality impact. Further, Lukasczyk, Maciejewski, Garth, and Hagen (2015) invented a topological visual analytical approach, a combination of Reeb graph and KDE, to detecting and visualizing crime hotspots in Chicago. Another relevant method used several contour intervals mapped to a rainbow color scheme to highlight the spatio-temporal changes of event density patterns, although it was applied to a lightning dataset (Peters & Meng, 2014). Another body of research focused on revising the structure of KDE to account for the temporal component. Similar to the notion of distance decay in SKDE to capture higher probability of new crimes at locations closer to past crimes, temporal decay can be introduced to model that new crimes are more likely to occur around a crime event in more recent past. Following this line of reasoning, Bowers et al. (2004) developed the prospective hotspot mapping (ProMap), which is a product of spatial and temporal weighting functions. Instead of kernel-based functions, they used a simple inverse distance weighting function in the model. Later, Brunsdon, Corcoran, and Higgs (2007) proposed a *spatio-temporal kernel density estimation (STKDE)* by multiplying SKDE by a temporal kernel function. It is a space-time cube method that extends the 2-D grid used in SKDE to a 3-D cube and computes density values at cube centroids with overlapping space-time cylinders. It has been applied in visualizing crime (Nakaya & Yano, 2010) and disease patterns (Delmelle, Dony, Casas, Jia, & Tang, 2014). Based on the "generalized product kernels" proposed by Li and Racine (2007), Zhang et al. (2011) designed a slightly different STKDE. Instead of using a bivariate kernel, they utilized two univariate kernel functions to capture possibly different distribution patterns along x and y dimensions. Their model was applied in disease risk estimation. In addition to the above works, some other spatio-temporal models have been developed, especially in the fields of epidemiology and transportation. These popular models include the Bayesian spatio-temporal model (Flaxman, 2014), spatio-temporal scan statistics such as SATSCAN by Kulldorff (1997) and among others. They are commonly used to detect spatio-temporal clustering patterns of a disease or a traffic accident, or to forecast the rates of these events in the near future. Given the prevalence of KDE in predictive crime hotspot studies, we refrain from discussing these models in more detail.

The aforementioned STKDE studies exclusively focused on visualizing, not forecasting, incident clustering patterns. Also, the search bandwidths (both spatial and temporal in our case), considered critical parameters in kernel-based methods, were usually chosen arbitrarily. For example, Nakaya and Yano (2010) defined the bandwidths as the average distance between the twentieth nearest neighbors. Another point often neglected in existing studies is the prevalence of false-positive hotspots, which can significantly compromise the efficacy of hotspot policing. Comparing to existing work in this scope, the proposed framework has four major features. Firstly, a spatio-temporal kernel density estimation (STKDE) method is proposed to include the temporal component in predictive hotspot mapping, possibly the first attempt of extending STKDE to crime prediction. Secondly, as a data-driven optimization process, the likelihood cross-validation can help detect the most appropriate bandwidths. Thirdly, a statistical significance test is developed to filter out false-positive hotspots. Last, a new metric, the predictive accuracy index (PAI) curve, is developed to evaluate the predictive accuracy of crime hotspots at multiple areal scales and provide more consistent and meaningful comparisons between methods. The proposed framework is illustrated in a case study of residential burglary crimes in Baton Rouge, Louisiana in 2011.

2. Method

2.1. Refined spatio-temporal kernel density estimation (STKDE)

The STKDE designed by Brunsdon et al. (2007) multiplies a bivariate kernel placed over the x-y (spatial) domain with a univariate kernel along the temporal dimension t to estimate the density of an event. They applied it to detect and visualize crime patterns. It was later used by Nakaya and Yano (2010) to visualize crime clustering patterns in Japan. As formulated below,

$$\hat{f}(x, y, t) = \frac{1}{nh_s^2 h_t} \sum_{i=1}^{n} K_s\left(\frac{x - X_i}{h_s}, \frac{y - Y_i}{h_s}\right) K_t\left(\frac{t - T_i}{h_t}\right) \tag{1}$$

where (X_i, Y_i, T_i) represents each crime incident $i \in (1, \ldots, n)$, (x, y, t) is the location in the space-time domain where the density $\hat{f}$ is being estimated, $K_s(\cdot)$ is a bivariate kernel function for the spatial domain, $K_t(\cdot)$ is a univariate kernel for the temporal domain, and h_s and h_t are the spatial and temporal bandwidths, respectively.

In the field of econometrics, Li and Racine (2007) developed the so-called "generalized product kernels" such as,

$$\hat{f}(x) = \frac{1}{nh_1 \ldots h_q} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

$$K\left(\frac{x - X_i}{h}\right) = k\left(\frac{x_1 - X_{i1}}{h_1}\right) \times \cdots \times k\left(\frac{x_q - X_{iq}}{h_q}\right) \tag{2}$$

where $k(\cdot)$ is a univariate kernel function that may vary with a specific dimension $\in (1, \ldots, q)$. The difference between Equations (1) and (2) is that the former has three dimensions (and the kernels along the spatial, two dimensions have the same distribution), and the latter has q dimensions and thus more general.

Equation (2) has several advantages. Firstly, it can be readily applied to data of multiple dimensions (more than three). Secondly, each dimension is considered separately in the model. This second feature is