Computers, Environment and Urban Systems 39 (2013) 93–106

Contents lists available at SciVerse ScienceDirect

Computers, Environment and Urban Systems

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

Understanding the spatial distribution of crime based on its related variables using geospatial discriminative patterns

Dawei Wang [a], Wei Ding [a,*], Henry Lo [a], Melissa Morabito [b], Ping Chen [e], Josue Salazar [c], Tomasz Stepinski [d]

[a] Department of Computer Science, University of Massachusetts Boston, United States
[b] Department of Criminology and Criminal Justice, University of Massachusetts Lowell, United States
[c] Department of Computer Science, Rice University, United States
[d] Department of Geography, University of Cincinnati, United States
[e] Computer and Mathematical Sciences Department, University of Houston-Downtown, Texas, United States

ARTICLE INFO

Article history:
Received 23 April 2012
Received in revised form 25 January 2013
Accepted 31 January 2013
Available online 9 April 2013

Keywords:
Crime related variable
Geospatial Discriminative Pattern
Hotspot Optimization Tool
Footprint

ABSTRACT

Crime tends to cluster geographically. This has led to the wide usage of hotspot analysis to identify and visualize crime. Accurately identified crime hotspots can greatly benefit the public by creating accurate threat visualizations, more efficiently allocating police resources, and predicting crime. Yet existing mapping methods usually identify hotspots without considering the underlying correlates of crime. In this study, we introduce a spatial data mining framework to study crime hotspots through their related variables. We use *Geospatial Discriminative Patterns* (GDPatterns) to capture the significant difference between two classes (hotspots and normal areas) in a geo-spatial dataset. Utilizing GDPatterns, we develop a novel model—*Hotspot Optimization Tool* (HOT)—to improve the identification of crime hotspots. Finally, based on a similarity measure, we group GDPattern clusters and visualize the distribution and characteristics of crime related variables. We evaluate our approach using a real world dataset collected from a northeast city in the United States.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Crime is understood to be related to the interaction of victims and offenders, and to the strength of guardianship (Cornish & Clarke, 1986). In practice, these concepts can be measured using a variety of socio-economic and crime opportunity variables, such as population density, economic investment, and arrest rate.

Geographical studies reveal that crime is often concentrated in clusters, which in the literature are called hotspots. Hotspot mapping techniques for crimes draw continuous attention from researchers and public safety agencies. This is because accurately identified and clearly visualized crime hotspots, and understanding their relation to underlying crime related variables, can significantly benefit crime analysis and police practices by providing a solid basis for threat visualization, police resource allocation, and crime prediction.

Existing hotspot mapping methods can be essentially divided into three main categories: point mapping, choropleth mapping, and kernel density estimation (KDE) (Eck, Chainey, Cameron, Leitner, & Wilson, 2005; Williamson, McGuire, Ross, Mollenkopf, & Goldsmith, 2001; Boba, 2005). Usually, these methods aggregate the density of a target crime, which results in a net loss of information (Van Patten, McKeldin-Coner, & Cox, 2009). For example, in choropleth mapping, incident-level data is first aggregated into arbitrary administrative or political boundary areas. During this step, spatial details within and across the thematic areas are lost. Second, when hotspots are generated based on aggregated data, there is a further decline of precision in the resulting map. Because traditional methods mainly rely on target crime density, particular areas with relatively less crime may be left out of hotspots, even though crime related variables indicate they are under similar risks as those hotspots.

A reasonable way to reduce this accuracy and precision loss in choropleth mapping is to use more related information in the mapping process. Crime related variables can be aggregated and used along with target crime data in the hotspot identification process. Information carried by these variables can provide clues on whether the relatively high crime rate in a certain area happens by chance. Compared to traditional methods, the utilization of related information in hotspot mapping can reduce information loss during analysis.

Additionally, such an approach can benefit further analysis on the characteristics of crime related variables. Instead of just evaluating crime by itself, recent studies also integrate crime related data into a unified framework that assists the analysis and exploration of crime hotspots (Maciejewski et al., 2010). Using

* Corresponding author.
E-mail addresses: ding@cs.umb.edu (W. Ding), chenp@uhd.edu (P. Chen).

0198-9715/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compenvurbsys.2013.01.008

related variables in hotspot mapping can additionally benefit such visualization and analyzation processes by providing an intuitive linkage between target crime and its related data.

In this paper, we present a framework that uses spatial data mining concepts to map hotspots and investigate the relationship between socio-economic and criminal variables. Recently, spatial data mining has emerged as an active research area in studies of spatial relationships that try to answer the questions like "why" and "where" (Ester, Kriegel, & Sander, 1997; Mu, Ding, Morabito, & Tao, 2011). It has been proven to be very powerful in identifying the linkage between target objects and its related factors. The components of our method are shown in Fig. 1. In particular, we:

- Introduce a spatial data mining concept, *Geospatial Discriminative Patterns* (GDPatterns), to study the relationship between target crime hotspots and their underlying related variables.
- Introduce a model, *Hotspot Optimization Tool* (HOT), to identify crime hotspots through their related variables.
- Use a similarity based method to cluster the crime related variables that contribute to hotspots into groups.
- Visualize the locations of those clusters in a rational way to assist domain scientists in further analysis, using the footprints of GDPatterns.

Utilizing the proposed framework, a case study is conducted using a 6-year crime dataset from a city in northeast United States. We compare our mapping tool with a widely used hotspot evaluating technique,the G_i^* statistics (Getis & Ord, 2010), and demonstrate the potential in assisting crime analysis using related variable clusters.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 introduces the data representation and formal definition of the research problems. HOT model and the implementation of the similarity measure are also presented in this section. Section 4 evaluates the proposed framework in a real-world case study. We conclude the paper in Section 5.

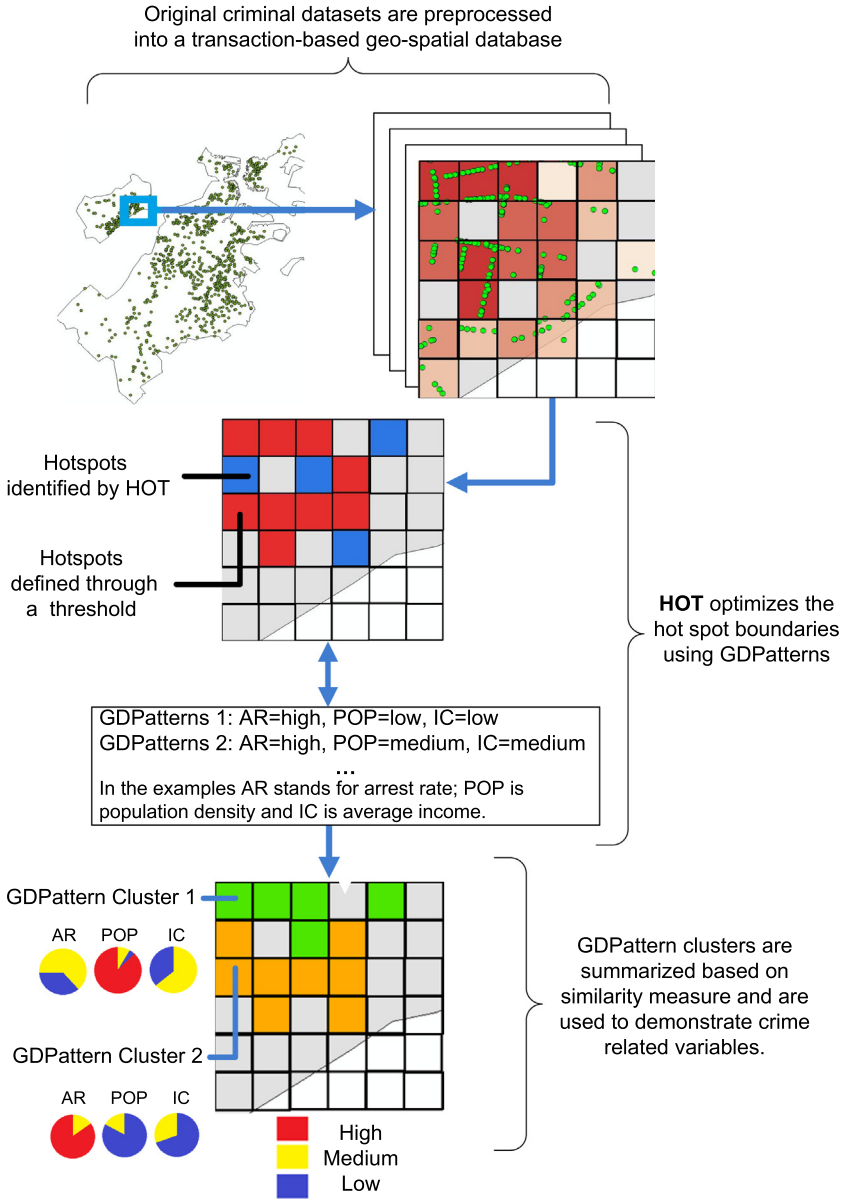

Fig. 1. The framework of our methods. With the help of GDPatterns, criminal hotspot maps are generated using HOT. By applying a similarity measure method, GDPatterns are clustered and visualized for domain scientists.

Download English Version:

https://daneshyari.com/en/article/6922031

Download Persian Version:

https://daneshyari.com/article/6922031

Daneshyari.com