A Crime Prevention System in Spatiotemporal Principles With Repeat, Near-Repeat Analysis and Crime Density Mapping: Case Study Turkey, Trabzon

Gamze Bediroglu¹, Sevket Bediroglu¹, H. Ebru Colak¹, and Tahsin Yomralioglu²

Abstract
In this study, we investigated crime events with repeat and near-repeat analysis for Turkey’s Trabzon city’s crime data after standardization process on raw crime data. First, a new crime geodatabase model was created. All types of recorded crime data for events between the years 2010 and 2014 were standardized, generalized, and Geo-referenced. We gave certain locations to crime events with geocoding techniques. Then, we created density maps of crime events with Kernel method in Geographic Information Systems (GIS). Repeat and near-repeat methods were tested on Burglary crime type in this geodatabase. Studies focused to applying prediction analysis besides showing current situation. These predictive analyses may be applied for all the security, intelligence, or defense departments at local, national, or international levels.

¹Karadeniz Teknik Üniversitesi, Trabzon, Turkey
²İstanbul Teknik Üniversitesi, Turkey

Corresponding Author:
Gamze Bediroglu, Karadeniz Teknik Üniversitesi, Trabzon, Ortahisar 61080, Turkey.
Email: gamze.yilmaz@ktu.edu.tr
Keywords
crime prevention, repeat and near-repeat, crime statistics, interpolation

Introduction
Crime analysis describes the qualitative and quantitative study of crime and law enforcement information in combination with sociodemographic and spatial factors to apprehend criminals, prevent crime, reduce disorder, and evaluate organizational procedures (Boba, 2001). Empirical studies show that crime events are spatially concentrated (Glasner & Leitner, 2017; Guerry, 1833; Quetelet, 1835). The goal of crime analysis is the unlocking of valuable insights from the collected crime information to assist law enforcement with criminal apprehension and crime prevention, to the end of improving the overall quality of life for community residents (O’Shea & Nicholls, 2003; Roth, Ross, Finch, Luo, & MacEachren, 2013).

The geography of crime focuses on the relationship between crime, space, and the social environment by analyzing crime behaviors, criminals, and crime influences (Feng, Dong, & Song, 2016; Johnston, Gregory, Pratt, & Watts, 2000). Place-based strategy is not enough for law enforcement agencies and crime mitigation activities because it only focuses on the relationship between crime and place; besides this, time phenomenon is important. It is important to know where crime patterns cluster in both space and time as it has significant effects on strategic action toward crime prevention (Glasner & Leitner, 2017). Where and when crime is likely to occur and risky areas can be determined through temporal and spatial analysis so that crime reduction actions can be planned in a more focused way. Spatiotemporal analyses lend support to the claim that crime events indeed do occur discretely not only in space but also in time (e.g., Bowers & Johnson, 2003; Johnson, Bowers, & Hirschfield, 1997; Morgan, 2000; Ratcliffe, 2002, 2004, 2005; Townsley et al., 2000, 2003; Wells, Ling, & Xinyue, 2012; Wyant, Taylor, Ratcliffe, & Wood, 2012; Zhang, Zhao, Ren, & Hoover, 2015).

Geographic Information Systems (GIS) is one of the most important information systems for recording, analyzing, and visualizing crime data. In addition, a temporal component can be added to spatial data and spatiotemporal analysis can be done in GIS. By transferring the spatial data of crimes into GIS, its relations with other spatial objects can be observed more easily and allows the right precautions to be taken.

There are several utilities of a wide variety of techniques for analyzing crime events through space and time. In recent studies, the utility of repeat and near-repeat patterns, one of the spatiotemporal analysis methods, has
increased for predictive crime works. Examples include studies analyzing repeat and near-repeat patterns of crime that have been observed for a range of crime types, such as burglary (e.g., Johnson et al., 2007; Johnson & Bowers, 2004; Townsley et al., 2003), shooting events (e.g., Haberman & Ratcliffe, 2012; Ratcliffe & Rengert, 2008; Wells et al., 2012; Wyant et al., 2012), vehicle crime (e.g., Johnson, Bowers, Birks, & Pease, 2009), robbery, and aggravated assaults (e.g., Grubesic & Mack, 2008).

In this article, we investigate crime events and develop a crime prevention system for Turkey. Crime data analysis and visualizing processes may be easier in high technology and well-developed countries due to readily available crime databases. However, in Turkey, crime data collection in GIS environment is a new approach. During this study, raw data sets were enhanced in the way of making this a usable crime geodatabase. We think that our approaches and methods may be a suitable model for other developing countries. This article also focuses on burglary crime victimization and prevention methods with new statistical and analytical methods. All the burglary crime events in past 5 years were investigated and solutions were shared.

**Method**

In brief definition, this article focuses on analyzing crime events in spatio-temporal principles and also moving barriers behind lack of ready crime data sets. Raw data set taken from police departments were converted to usable GIS-based crime data set after some standardization and matching operations. This crime data set contains a significant amount of knowledge and statistics, such as time of crime, location of crime, and also gender, age, past crime, and definition of suspect and victim. Furthermore, base layers including road networks, buildings, and city districts were gathered in the same geodatabase after processing in GIS. Thereafter, repeat and near-repeat analysis was applied to these databases and result tables and maps were generated. This has enabled to reveal a recommender crime prevention system for use by police departments. The approaches, methods, and process stages are given in detail in Figure 1 below.

**Repeat and Near-Repeat Analysis**

A large body of research demonstrates that crime is unevenly distributed among offenders, places, and victims, respectively (Blumstein, Cohen, Roth, & Visher, 1986; Johnson et al., 2007; Pease, 1998; Sherman et al., 1989). In addition, crimes, especially in some places, occur in succession. These understandings have been enhanced by the discovery of a near-repeat phenomenon.
Near-repeat crimes refer to the spatiotemporal proximity of crime events, in other words, clustering in both time and place (Sturup, Rostami, Gerell, & Sandholm, 2017; Youstin, Nobles, Ward, & Cook, 2011). One approach to crime analysis suggests that the best way to predict future crime occurrence is to use past behavior, such as actual incidents or collections of incidents, as indicators of future behavior (Caplan, Kennedy, & Piza, 2013).

The empirical findings from research into the patterns of repeats and near-repeats has led some commentators to suggest that recent events provide a powerful indicator for predicting where and when crime is likely to take
place (Bowers, Johnson, & Pease, 2004; Chainey & da Silva, 2016; Johnson & Bowers, 2004; Pease, 1998; Skogan, 1996). In near-repeat phenomenon, not only are locations at risk of repeat victimization, but nearby locations are also at increased risk of crime up to a certain distance and for a certain time (Bowers & Johnson, 2004; Haberman & Ratcliffe, 2012; Ratcliffe & Rengert, 2008; Townsley et al., 2003).

Researchers have identified the repeat and near-repeat phenomenon first among burglaries. This phenomenon indicates that shortly after being burglarized, a location and locations within its immediate proximity are more susceptible to experiencing the same event (Johnson & Bowers, 2004; Wu et al., 2015). Exact-repeat events are defined as consecutive burglaries occurring at the same location, separated by a time interval of any duration; near-repeat burglary events are instead classified as taking place within a set spatial neighborhood of a focal burglary point (Short, D’Orsogna, Brantingham, & Tita, 2009).


“Flag” thesis indicates that some innate features make a location more attractive than others and therefore the location is likely to be repeatedly victimized; in contrast, “boost” thesis suggests that additional victimizations are dependent upon the initial crime (Wu et al., 2015).

Analytical Strategy and Formula of Repeat, Near-Repeat

There is limited research regarding the extent to which repeat victims are likely to be repeat offenders, and few studies have assessed whether predictors of repeat victimization and repeat offending are similar (Fagan & Mazerolle, 2011). “Repeat and Near-Repeat Analysis” toolset, added as an extension in ArcGIS, is used for predictive mapping of burglaries. Repeat and Near-Repeat Analysis is used to identify near-repeat victimization patterns and create near-repeat prediction zones. This toolset contains three tools (Export Event Comma-Separated Values [CSV], Repeat and Near-Repeat Classification, and Calculate Prediction Zones) for identifying and analyzing patterns in event data. First, “Export Event CSV” that exports an event layer to a CSV file in the format required by the Near-Repeat Calculator
(NRC) and output fields include $X$, $Y$ coordinates and date of the crime event. To investigate the repeat and near-repeat phenomenon in burglary patterns, NRC is used. The NRC is developed by Jerry Ratcliffe (2009). It is a public-domain software program and can be distributed freely for educational or research purposes. Whether or not there is a statistically significant near-repeat pattern in a crime data set is recognized by this program. The algorithm of NRC is based on the revised Knox close-pair test. The Knox test can be utilized to observe and measure whether crime events are significantly clustered both spatially and temporally, given a threshold of closeness in both space and time (Wu et al., 2015). Then, statistical significance is found through Monte Carlo simulation. This statistical significance is assessed with a tabular file including crime events (originators of near-repeats and near-repeats of originators) generated by the program. Second, the “Repeat and Near-Repeat Classification” classifies events as originators, repeats, or near-repeats according to user-specified distance and time parameters. In addition, it creates a line feature class to link related events; connections help visualize the spatial and temporal relationships between events. Third, “Calculate Prediction Zones” identifies areas at risk of repeat and near-repeat events by specifying the space and time range of influence of past events. As a result, these prediction zones help to identify high-risk areas and support predictive policing efforts.[AQ: 11]

Case Study

Study Area

Trabzon city in Turkey was chosen as the study area. The reasons for choosing this region are that crime data required for the analysis were available in raw format and the authors’ knowledge about the city environment. Trabzon city is located between $38^\circ$ 30’–$40^\circ$ 30’ east longitude and $40^\circ$ 30’–$41^\circ$ 30’ north latitude and the city’s population is about 779,000. The position of the study area in Turkey and Eurasia can be seen clearly in Figure 2.

Geodatabase Design

In this study, first, spatial data that may be directly or indirectly related to the crime were collected. The main parts of this data set were obtained from the police department and local government services in raw formats (Paper, Excel, and CAD). The main types of this data set are address data, street-road networks, buildings, and crime events in Excel format and on papers.
Recorded crime data between 2010 and 2014 were provided by the police department for the city of Trabzon. These crime data include crime types, date, location of crime, and the age and sex of the offender. The burglary data used in this study were between 2010 and 2014; there were 2,238 recorded burglaries. The collected data were arranged, standardized, and made usable. The process of assigning geographic location information to address data, which is known as geocoding, was done. Normally, geocoding operations are being used for address match operations, but in this study, geocoding was used for identifying the location of each crime event. As most of crime data set contained street information, these were solved by linking geocodes of street information. Some of remaining crime data were based on the location and name of buildings. These were coordinated by the use of center coordinates of buildings. Geocoding “hit-rate” was in excess of 98% and this geocoding accuracy of the burglary data is sufficient in quality for the purposes of the research.

**Overview of the Crime Events at Study Area**

Figure 3 shows the number and percentage of crime events for each crime type between 2010 and 2014 in Trabzon, Turkey. According to the recorded events, the most intensive crime type is violent crime, constituting 43% of all
crime events. The remaining crime types are burglary at 13%, irregularity against person and property at 12%, and threat events at 8%. The reason why burglary crimes were selected for analysis is that the rate of the burglary crimes is quite high compared with other crime types and has continued to increase every year between 2010 and 2014. In addition, burglary events are meaningful for statistics and sometimes have a repeating structure in terms of time and location; besides, violent events have random structure.

Crime Density Analysis

Kernel density estimation involves placing a symmetrical surface over each point and then evaluating the distance from the point to a reference location based on a mathematical function and then summing the value for all the surfaces for that reference location (Anderson, 2009). Four different methods were tested for creating crime density maps for the zone. These methods are Kriging, Inverse Distance Weighting (IDW), Hot-Spot, and Kernel. The Kernel density analysis method was selected for determining and mapping geometric distribution analysis of burglary events because it gave efficient results (Figure 4). This map helps us to identify where the burglary events are of higher or lower density, after calculating Kernel’s interpolations.

According to Figure 4, the highest rate of burglary crime is at Kunduracilar street in Kemerkaya district. In the second place with high crime rates are Kahramanmaraş and Uzun streets due to crowd and mobility arising from near located work centers. Figure 4 also shows that the lowest crime rate zones are Besirli 1 and Kanuni districts in the city. The least burglary rate is found in nearby zones of Besirli 1 and Kanuni districts.

Figure 3. Distribution of crime types (2010-2014).
Applying Repeat and Near-Repeat Analysis

“The Repeat and Near-Repeat Analysis Tools” tools were used for identifying, visualizing, and predicting repeat and near-repeat patterns for burglary events. In the study area, 2,238 burglary events were used for analysis, recorded between 2010 and 2014. Spatial bandwidth was determined as 200 m and four bands of 200 m, 400 m, 600 m, and 800 m were applied; temporal bandwidth was determined as 7 days and four bands of 7, 14, 21, and 28 were applied.

In ArcGIS 10.4, with the repeat and near-repeat toolset, burglary events were converted to a CSV file from shapefile (Esri) and this file was used in NRC. The date of burglary crimes, coordinates $X$ and $Y$ of burglary points were enough for using NRC. A total of 2,227 records of burglary events were successfully recorded by NRC but for 11 rows in the file due to missing attribute data. We can use the statistically significant space and time band values to populate parameters of Repeat and Near-Repeat Analysis in ArcGIS.

A statistically significant repeat and near-repeat of burglaries pattern was identified with NRC. Table 1 presents spatial and temporal bands for repeats and near-repeats of burglaries were statistically significant in Trabzon and also shows significance levels and risk levels across all spatial and temporal bands. According to Table 1, repeats and near-repeats pattern
in Trabzon was statistically significant ($p \leq .05$) for eight of the 30 different spatial and temporal bands. After an event, there is evidence of an overrepresentation event at the same place up to 14 days after an initial event. The most overrepresented repeat victimization range is significant in the zone from 0 to 7 days from an initial event. The value of 2.24 is interpreted as that the same location is very likely to witness another burglary within the next 7 days, and the chance of another burglary is about 124% greater than if there was no repeat victimization pattern. Within same temporal band, when we expand the spatial band up to 200 m (not including the same location), the chance of another burglary is immediately dropped to 14%, and also when we expand the spatial band for 400 m to 600 m within 7 days, the chance of another burglary is dropped to 13%. Finally, when we expand the spatial band up to 800 m (not including the same location), a significant and meaningful repeat victimization pattern for burglary does not appear. According to this result, the increased chance for a near-repeat burglary to occur after an originating burglary diminishes as the distance and time from an originating event increases. With classification analysis, each burglary event was classified as originator, repeat, near-repeat, and other (not in a pattern). Table 2 shows the number of repeat and near-repeat burglary events per spatial and temporal band and percentage of all burglary events classified as repeat or near-repeat and appearing in each band. These values help us to understand the proportionality of repeat and near-repeat burglary events.

Table 1. Observed Over Mean Expected Frequencies and Significance Levels of Burglary Risk.

<table>
<thead>
<tr>
<th></th>
<th>0-7 days</th>
<th>8-14 days</th>
<th>15-21 days</th>
<th>22-28 days</th>
<th>More than 28 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same location</td>
<td>2.24*</td>
<td>1.23*</td>
<td>0.85</td>
<td>1.21*</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>$p &lt; .05$</td>
<td>$p &lt; .05$</td>
<td>$p &lt; .05$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-200 m</td>
<td>1.14*</td>
<td>1.00</td>
<td>0.95</td>
<td>0.98</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$p &lt; .05$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>201-400 m</td>
<td>1.04</td>
<td>1.09*</td>
<td>1.02</td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$p &lt; .05$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>401-600 m</td>
<td>1.13*</td>
<td>1.04</td>
<td>1.02</td>
<td>1.09*</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>$p &lt; .05$</td>
<td></td>
<td></td>
<td>$p &lt; .05$</td>
<td></td>
</tr>
<tr>
<td>600-800 m</td>
<td>1.00</td>
<td>0.99</td>
<td>0.97</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>More than 1,200 m</td>
<td>0.98</td>
<td>0.99</td>
<td>1.00</td>
<td>0.99</td>
<td>1.00*</td>
</tr>
</tbody>
</table>

*Statistical probability is chosen as $p = .05$ (20 iterations).
A total of 439 repeat events (those under 1 m in distance) were observed in a 28-day time band, and 1,596 near-repeats occurred within 800 m and 21 days; 19.7% of all events in the data set occurred within 1 m and 28 days of the original events. In addition, 71.7% of all events were a near repeat, occurring within 800 m and 28 days of an originating event. In total, 91.4% of all events were part of a repeat or near-repeat victimization pattern when defined near repeats as within 28 days and 800 m.

According to this high proportion of repeats and near-repeats, an operational strategy based on patrolling in these high-risk areas would be an efficient use of resources.

With classification analysis, a line feature class was created to link related burglary events that visualized the spatial and temporal relationships between burglaries. Figure 5 presents geographic connections of repetitive burglary events in 7, 14, 21, and 28 days in a region of study area. According to these geographic connections between burglary crime points, we can observe at which distance and time interval the other crimes were committed after a crime has been committed in a specific area. For example, it is shown in Figure 5 that a burglary crime was committed on April 9, 2014, and then the other burglary crimes were committed 207.08 m near the first crime on April 12, 2014, and 384.85 m near the first crime on April 15, 2014. This means the determined area is a high-risk area for burglary crime events in April. Similarly, January is a high-risk time because a crime was committed on January 15 and then two burglary crimes were committed 2 days later, on January 17, at 128.15 m and 580.17 m, and also another crime was committed 3 days later, on January 18, at 726.61 m. This information, based on space and time, is crucial in determining possible future crimes.
After repeat and near-repeat burglary patterns were identified with classification analysis, areas of risk for future burglary events were identified by specifying the space and time range of influence of past events between 2010 and 2014. Risk prediction was created based on two risk ranges; 400 m of spatial bands and 14-day temporal bands and prediction zones were mapped based on these values of bands. Risky areas and maps of these areas have been created separately for each spatial and temporal band. According to Figure 6, the highest risky areas are zones nearby Kemerkaya and Iskenderpasa districts. Nearby zones of Bahcecik and Fatih District are also high-risk areas.

**Conclusion**

In this study, we have tested repeat and near-repeat methods for Turkey’s Trabzon city with the burglary type of crime. This methodology not only addresses ready data process but also describes the operations and approaches to make raw data sets useful. Result of repeat and near-repeat analysis stated that the most overrepresented repeat victimization range is significant in the zone from 0 to 7 days from an initial event. The value of 2.24 is interpreted as impactful, meaning the same location is very likely to witness another burglary within the next 7 days, and the chance of another burglary is about 124% greater than if there were no repeat victimization patterns. Besides this, thematic crime density maps were created using Kernel algorithm. Four
different interpolation algorithms (Kernel, Kriging, IDW, and Hot-Spot) were tested for creating crime density maps but Kernel has given the best geographic distribution and visual appearance. Created prediction zones can be used for crime prediction analysis and mitigation of events. This supports predictive policing efforts across an entire community. This system is essential for predictive theories and would be effectively used for a high proportion of repeat and near-repeats events by developing an operational strategy based on patrols in these high-risk areas.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Gamze Bediroglu http://orcid.org/0000-0003-2755-3206

References


**Author Biographies**

[AQ: 22]