Predicting Future Shooting Crime Locations Using Principles of Data Analytics (SHOPS)

M. Said Varlioglu
MS in School of Information Technology
Research Asst. in School of Criminal Justice
University of Cincinnati
Cincinnati, OH
varlioms@mail.uc.edu

Murat Ozer, Ph.D.
Assistant Professor
School of Information Technology
University of Cincinnati
Cincinnati, OH
m.ozer@uc.edu

ABSTRACT

Shooting crime is a serious public health problem in the US. The analysis of any historical crime data reveals that crime is non-randomly distributed in time and space. Based on this notion, hot spots policing has gained its momentum to effectively predict future crime locations. Recent studies; however, pointed out that traditional hot spots policing occasionally predict rare crimes such as homicides and shootings due to their less frequent recurring counts in a given place and time (specifically for shorter time periods such as weeks and months). Given this context, we developed a new shooting prediction system (SHOPS) to explore whether recent dynamic/mobility activity patterns of known violent individuals increase the prediction of short-term fatal and non-fatal shootings compared to the traditional hot spots policing. Findings suggest that SHOPS predicts fatal and non-fatal shooting locations more precisely by identifying fewer hotspot locations. Policy implications of the study were discussed in the conclusion section.

KEYWORDS

Crime prediction; prevention; crime patterns, crime analysis, data analytics

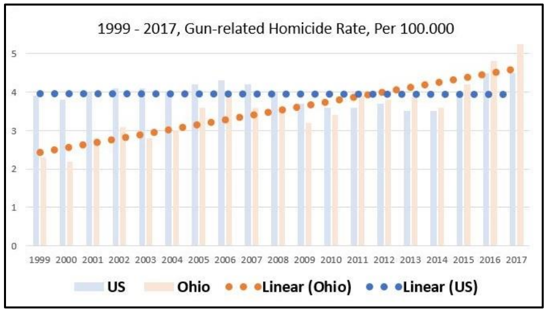

Figure 1: Gun-related homicide rates, OH & US

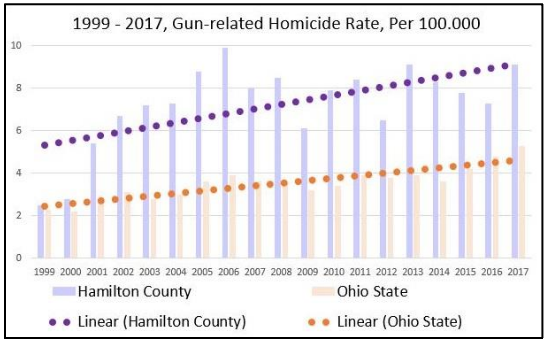

Figure 2: Gun-related homicide rates, OH & Hamilton County

INTRODUCTION

It is the fact that crime affects all aspects of human life. Gun violence – as a serious kind of crime - has long been a major problem of the US cities. [1, 2, 3, 4, 5]. Likewise, Ohio and its major cities have been suffering from gun violence for many years. For instance, Fig.1 and Fig.2 show that gun-related homicide rates in Ohio and Hamilton County in which Cincinnati locates have been increasing above the national rates since 2011. Overall national statistics of 2017 ranks Cleveland as the 5th, Cincinnati as the 9th, and Columbus as the 21st deadliest cities of the nation [6, 7]. Given the serious public health problem of the Ohio State localities, we wanted to study fatal and non-fatal shootings with the City of Cincinnati data. In this context, we will first review the literature to explore the existing studies that focus on the prediction of fatal and non-fatal shootings. Next, we will introduce a new shooting prediction system to proactively identify places for likely interventions. Finally, we will compare the predictive power of the newly proposed shooting prediction system with the predictive power of traditional systems.

RELATED WORKS

Crime prevention studies suggest that crime frequently occurs in relatively small locations [8, 9] known as crime hotspots. For this reason, many law enforcement agencies use GIS technologies to identify spatial patterns [10, 11,], thereby increasing their effectiveness by directing scarce resources to the places where the need is greatest [12, 13]. This approach is known as hots spots policing in the criminology field. Recent studies suggest that hots spots policing rarely work to successfully/accurately predict future shooting locations for shorter time periods (e.g., week, month) due to the small number of recurring events (shootings and homicides) [14]. There are some studies in the literature that attempted to add a dynamic feature to the traditional hot spots policing formula to increase short-term crime predictions. In this context, Wang et al. [15] developed a machine learning algorithm named “Series Finder” to automatically detect crime patterns that were committed by repeat offenders. Similarly, certain researchers examined spatiotemporal patterns of human spatial behavior to improve short-term crime prediction [16, 17, 18]. These efforts increased the precision of general crime prediction (i.e., combined crimes like violent and property crimes); however, their predictive power for homicides and shooting crimes still remained the same compared to traditional hot spot policing. For this reason, Mohler [14] developed a different system to specifically predict homicides and shooting crimes by addressing the shortcomings (e.g., relying on a small number of cases in the case of rare events) of traditional hot spots policing methodology. In his model, Mohler [14] gives a higher weight to the recent shooting locations if they overlap with the chronic places. This dynamic model better predicted short term homicides using the recent homicide locations. Indeed, Mohler’s model indirectly incorporates dynamic human activities to the formula by giving higher weights to the recent fatal and non-fatal shootings.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2008	28	10	28	39	35	45	55	31	39	51	28	29
2009	34	20	36	35	39	31	45	44	34	26	32	40
2010	21	23	49	29	35	42	22	49	46	47	34	33
2011	28	17	30	39	47	40	55	31	41	38	29	32
2012	38	16	28	32	34	28	37	38	36	30	31	26
2013	29	28	27	38	47	50	45	35	39	43	23	23
2014	27	22	26	26	39	36	32	32	49	25	27	34
2015	20	21	32	46	55	53	44	47	42	55	25	39
2016	22	24	35	26	32	49	62	33	44	42	23	34
2017	35	35	42	20	38	27	49	39	39	31	33	23
2018	25	10	19	13	46	46	32	25	31	27	29	27
2019	27	12	--	--	--	--	--	--	--	--	--	--

Table 1: Jan ‘15 to Feb ‘19 Shoooting Data

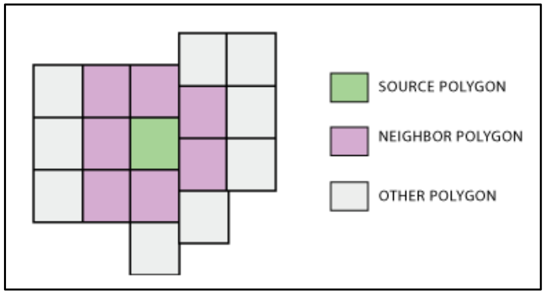

Figure 5: Identifying Neighbor Polygon

We wanted to replicate and develop Mohler’s dynamic model by merging the two widely known thoughts of criminology and crime prevention theories: Criminological theories suggest that a small number of individuals (known as chronic offenders) commit the majority of crime. And crime prevention theories stress that crime is non-randomly distributed in time and space. We integrated these two school of thoughts and hypothesized that if a study merges dynamic human patterns of chronic offenders with static locations of crime prevention theory, the prediction of short term crime locations can be more precise. Therefore, *the main research inquiry of the present study is to explore whether dynamic human activity patterns of known violent individuals increase the short-term shooting predictions over and beyond the prediction of traditional hot spots policing.*

METHODOLOGY

Data

The current study employs four different data sources from the Cincinnati Police Department. These data sources include police-reported data, arrest data, suspect-victim data, and field interview reports data from January 1, 2015 to February 28, 2019 as demonstrated in Table 1.

Dependent Variable

Th Dependent variable of the study is the locations of fatal and non-fatal shootings that occurred between November 1, 2018 and February 28, 2019. This time period was designated as the future shooting locations to test the predictive power of traditional hot spots policing and the newly proposed system that used the past data (from January 1, 2015 to October 31, 2018). We named our prediction system as Shooting Prediction System (SHOPS); therefore, SHOPS refer to the new shooting system for the rest of the study.

Independent Variables

Predicted Shooting Locations Obtained from SHOPS: The main idea of hotspot policing is to increase police visibility in crime dense locations to deter likely criminals for crime commission. Given this context, studies suggest that street segments also known as street blocks are the appropriate unit of analysis regarding its size (average 450 ft) because the entire surface of the street segment will have the equal police visibility during an intervention. Therefore, we aggregated the police contacts (e.g., arrest, suspect, field interview report, or being victim of a crime) of known violent individuals to the street segments and coded a street segment as a hot place/spot if two or more violent individuals contacted with the police within the last 30 days in that street segment. This process gives us repeat police contacts with known violent individuals in a given street segment. In addition to this, near repeat crime concentration as an important factor for future crime predictions. In this context, we employed ArcGIS Polygon Neighbors function (Fig.5) to find first degree neighbors of each street

Street Segment ID	Neighboring Street Segment ID
37	46
37	38
37	47
38	37
38	46
38	47
46	37
46	55
46	56
46	38
46	47
...	...

Table 2: The rows of Street Segment IDs

Address Range	Number of Contacts	Number of High Violence Police Contacts in the First-Degree Street Segments
1721 - 1778 GILSEY AV	2	3
1851 - 1998 SUNSET AV	2	3
3601 - 3698 MCHENRY AV	2	4
1139 - 1230 LINN ST	2	3
701 - 798 CLARK ST	2	3
801 - 898 CLARK ST	3	4
1001 - 1098 MAIN ST	2	3
1001 - 1098 SYCAMORE ST	2	3
909 - 998 SYCAMORE ST	3	5

Table 3: Identification of Future Shooting Locations

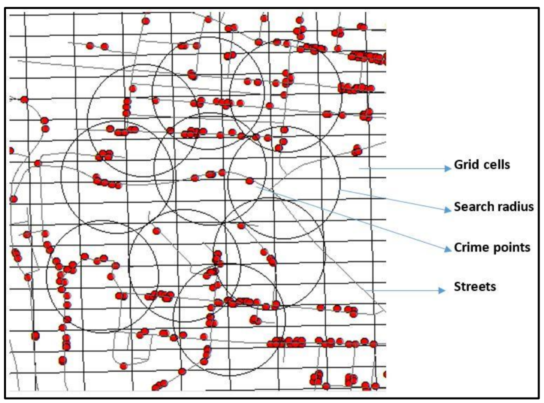

Figure 6. Kernel density map

segment. This method created a table with statistics to show source street segment (with ID numbers) and all neighboring street segments (with ID numbers).

The total row number of the newly generated table is 57,096 which basically displays each street segment's ID and neighboring street segments(s)'s ID as seen in Table 2. This is an important step because the current study wants to increase the precision of crime hot spot locations by considering nearby street segment(s)'s criminal activity as well. Hot spots interventions in the literature suggest more diffusion of benefit effect rather than crime displacement to nearby areas. For this reason, considering nearby street segments' criminal activities while choosing crime hotspots might help researchers to better select the crime locations to maximize diffusion of benefit effect to nearby street segments for optimal crime reduction. This method is different from the kernel density mapping used in the traditional hot spots policing as seen in Fig. 6. Kernel density map first creates a fishnet, then users define the search radius to find the total number of points (shootings in our case) that fall into each cell and the search radius. However, users generally keep the search radius large to find the clusters or crime concentration on the map. This clustering approach makes the map nicer but increases uncertainty because it points a larger location for crime prevention efforts. In contrast to kernel density map grids/cells, we will use street segments to identify hot spot locations because recent hots spots studies suggest using street segments to maximize the influence of any intervention. Likewise, instead of using a search radius of kernel density maps, the current study will employ first-degree neighboring street segments criminal activity patterns. We automated the above explained model in an algorithm (SHOPS) that daily scans the Cincinnati Police Data to find the locations where known violent individuals mostly contacted with the police. The final product of the algorithm is a list of places (Table 3) that are nominated as future shooting locations for the next days.

Predicted Shooting Locations Obtained from The Traditional Methods: Traditional hot spot policing uses the past/historical data to identify chronic shooting locations to predict future shooting locations. As Fig. 7 displays, we determined the chronic shooting locations as the street segments where at least three shootings occurred between January 2015 and October 31, 2018.

Control Variables

In this study, we will not use any control variables because we assume that everything is the same and the only difference in the study is the identification of the hot spots. We presumed that no omitted variable bias will threat the internal validity of the study since we apply both approaches to the same city under the same conditions. Therefore, no control variables such as poverty level of locations, number of parks, and number of liquor stores, which have proven as an important predictor of crime concentration in the crime prevention literature. [19]

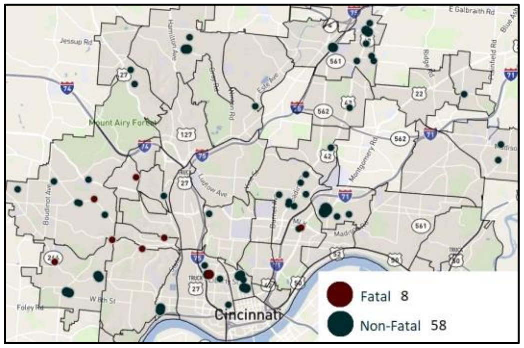

Figure 7. Dec. 2018 to Feb. 2019 Shootings

Analytical Plan

We will use point distance analysis of ArcGIS to measure the distance of predicted shooting locations of SHOPS and the predicted shooting locations of traditional hots spots policing from the locations of actual shootings occurred between Noveber 2018 and February 2019. If the distance of predicted shooting locations is 200 feet or less to the actual shooting location, that will be counted a successful prediction for a future shooting location. Following the distance analysis, we will compare the prediction results of SHOPS with the prediction of traditional hot spots policing using basic descriptive statistics.

RESULTS AND CONCLUSION

Traditional hot spots policing technique identified 103 chronic street segments using the date between January 2015 and October 2018. Initial analysis suggests that traditional hot spot policing successfully predicted 19 out of 85 (22%) fatal and non-fatal future shooting locations. On the other hand, SHOPS identified in average 30 shooting locations for each month and these locations predicted 23 out of 85 (27.1%) future shootings. Even though the prediction looks similar, uncertainty level of traditional hot spot policing is much higher than SHOPS (103 vs. 30); therefore, SHOPS provide a similar prediction to law enforcement with fewer efforts such as conducting fewer directed patrols to gain the similar outcome. Further comparative analysis showed that traditional hot spot policing generally predicts apartment complexes because those places stand as chronic shooting locations as a result of static/historical data examination. For this reason, the prediction of traditional hot spot policing is somehow deceptive because it is very difficult to conduct a treatment/intervention for the entire apartment complexes. On the other hand, SHOPS predicts criminogenic street segments in precision without relying on the static data; therefore, it brings criminogenic street segments (other than apartment complexes) to the attention of decision-makers to conduct effective preventive strategies. In conclusion, even though the current study confirms that hot spot policing has its merit and successfully predicts shooting locations, the uncertainty level (i.e., listing too many places for treatment) of hot spot policing is high which then undermines conducting effective interventions. For instance, a police department may not conduct directed patrols at 103 different places which many of them are apartment complexes. On the other hand, SHOPS predicts future shooting locations in a better precision by only suggesting 30 hot locations to police departments for the same result. The substantial overlap between human activity patterns of known violent individuals and future shooting locations also support the notion that a small number of population (less than 5%) commit the majority of shooting crimes. In this context, our study is in line with Papachristos et al.'s study [20] that concludes: shootings spread like a disease among a network of people with known shooting criminal history. In the lights of current study findings and existing studies, the policy implication of the study is that police departments can maximize their shooting crime prevention strategies by targeting both SHOPS identified shooting places and individuals who drive the shooting crimes for those identified locations.

REFERENCES

[1] Pear, V. A., Castillo-Carniglia, A., Kagawa, R. M., Cerdá, M., & Wintemute, G. J. (2018). Firearm mortality in California, 2000–2015: the epidemiologic importance of within-state variation. Annals of epidemiology, 28(5), 309-315.

[2] FBI, U.S. Department of Justice (2018. Crime in the U.S. in 2017. Criminal Justice Information Services Division. Retrieved from: https://ucr.fbi.gov/crime-in-the-u.s/2017/crime-in-the-u.s.-2017

[3] Grinshteyn, E., & Hemenway, D. (2016). Violent death rates: the US compared with other high-income OECD countries, 2010. The American journal of medicine, 129(3), 266-273.

[4] Gabor, T. (2016). Confronting gun violence in America. Springer. DOI 10.1007/978-3-319-33723-4_1

[5] Kerr, S. E. (2018). Gun Violence Prevention?: The Politics Behind Policy Responses to School Shootings in the US. Springer.

[6] Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death 1999-2017 on CDC WONDER Online Database, released December, 2018. Data are from the Multiple Cause of Death Files, 1999-2017, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Accessed at http://wonder.cdc.gov/ucd-icd10.html on Mar 14, 2019 10:01:18 PM

[7] Ferrell, K. (2017). “Cleveland, Cincinnati, Columbus ranked some of the deadliest cities in the U.S.” derived from https://www.nbc4i.com/news/u-s-world/cleveland-ranked-5th-deadliest-city-in-the-u-s/1096237298

[8] Weisburd, D., Maher, L., Sherman, L., Buerger, M., Cohn, E., & Petrisino, A. (1992). Contrasting crime general and crime specific theory: The case of hot spots of crime. Advances in criminological theory, 4(1), 45-69.

[9] Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. Justice quarterly, 31(4), 633-663.

[10] Bowers, K. J., Johnson, S. D., & Pease, K. (2004). PROSPECTIVE HOT-SPOTTING: The future of crime mapping? The British Journal of Criminology, 44(5), 641-658. doi:10.1093/bjc/azh036

[11] Eck, J., Chainey, S., Cameron, J., & Wilson, R. (2005). Mapping crime: Understanding hotspots.

[12] Weisburd, D., & Green, L. (1995). Policing drug hot spots: The Jersey City drug market analysis experiment. Justice Quarterly, 12(4), 711-735.

[13] Weisburd, D., & Telep, C. W. (2011). The efficiency of place-based policing. Evidence Based Policing, 17, 247-262.

[14] Mohler, G. (2014). Marked point process hotspot maps for homicide and gun crime prediction in Chicago. International Journal of Forecasting, 30(3), 491-497.

[15] Wang, T., Rudin, C., Wagner, D., & Sevieri, R. (2013, September). Learning to detect patterns of crime. In Joint European conference on machine learning and knowledge discovery in databases (pp. 515-530). Springer, Berlin, Heidelberg.

[16] Ratcliffe, J. H. (2006). A temporal constraint theory to explain opportunity-based spatial offending patterns. Journal of Research in Crime and Delinquency, 43(3), 261-291.

[17] Toole, J. L., Eagle, N., & Plotkin, J. B. (2011). Spatiotemporal correlations in criminal offense records. ACM Transactions on Intelligent Systems and Technology (TIST), 2(4), 38.

[18] Short, M. B., D'orsogna, M. R., Pasour, V. B., Tita, G. E., Brantingham, P. J., Bertozzi, A. L., & Chayes, L. B. (2008). A statistical model of criminal behavior. Mathematical Models and Methods in Applied Sciences, 18(supp01), 1249-1267.

[19] Willits, D., Broidy, L., Gonzales, A., & Denman, K. (2011). Place and neighborhood crime: Examining the relationship between schools, churches, and alcohol related establishments and crime. University of New Mexico.

[20] Papachristos, A. V., Braga, A. A., & Hureau, D. M. (2012). Social networks and the risk of gunshot injury. Journal of Urban Health, 89(6), 992-1003.