Detecting Emerging Space-Time Crime Patterns by Prospective STSS

Tao Cheng Monsuru Adepeju

{tao.cheng@ucl.ac.uk; monsuru.adepeju.11@ucl.ac.uk}
SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineering,
University College London,
Gower Street, WC1E 6BT London

1. Introduction

Detecting crime patterns as they emerge in both space and time can enhance situational awareness amongst security agents and prevent epidemics of crimes in potential problematic areas (Neill and Gorr, 2007). Amongst others, space-time scan statistics (STSS) (Kulldorff et al. 2005) and space -time kernel (Nakaya and Yano, 2010) have been widely used in crime analysis. Stemmed on strong statistical theories, the STSS could provide the significance of the purported crime clusters, and this continues to gain huge popularities for crime hotspot analysis (LeBeau, 2000; Neill and Gorr, 2007; Uittenbogaard and Ceccato , 2011; Cheng and Williams, 2012; Gao et al. 2012). All these works applied STSS to crime clusters detection in a retrospective manner where all clusters within certain frame of time are detected. The approach was found to be very effective for historic analysis of crime outbreaks and near-repeat victimization. However, most STSS-based hotspot analyses were conducted at either region-wide and/or at monthly temporal granularity. This is not appropriate for city-based policing, which requires detailed spatial (local or micro) and temporal (daily) analysis. Few studies have actually attempted prospective detection of clusters with the aim of capturing their growth (emergence) in both space and time simultaneously so as to facilitate early prevention of the phenomenon in question. This was however seen only in epidemiology where outbreaks of diseases were detected employing this approach using the over-the-counter drug sales in Allegheny County from 2/13/04 - 2/12/05 (Neill et al., 2005). However, there is no quantitative evaluation of the significance of the emerging patterns and the rapidness of their emergence.

Therefore, the aim of this research is to explore a prospective detection of emerging crime patterns at detailed spatial and temporal scales so as to facilitate proactive policing. In particular, we use the permutation STSS for the detailed crime emerging pattern detection and evaluate their significance as well as the rapidness of detection, by comparing the results with that of retrospective analysis.

2. Prospective Permutation STSS

Generally, space-time scan statistics (STSS) work on the basic idea of scanning through a geographic region with a large collection overlapping geographic windows moving across space and time (Kulldorff et al. 2005). A scanning window (e.g. a cylinder) moves across the entire area counting the number of geographic events within the window and evaluating the expected count. Considering all the cylinders in a neighborhood, the window with most ‘unusual’ number of observed cases as compared to the expected value, having taken care of multiple hypotheses testing, is noted and reported as the 'most likely cluster'. In other words, a true hotspot or cluster is not only the cylinder in which number of observed crimes is relatively larger than its expected, but also its likelihood ratio is exceedingly large depending on the number of replications generated in a case where Monte Carlo simulation is adopted for multiple testing. The model used for evaluating the expected value and the likelihood ratio of a space-time window can depend largely on the nature of the data as well as the domain in question. Among others, a Poisson and space-time permutation models are generally used for count datasets such as the disease and crime data. The Poisson model, assumes a Poisson distribution for the data, and requires population-at-risk information to accurately evaluate the significant clusters across the region. The population-at-risk information is basically derived from census population data. However, in the context of victim-offender interactions where the population is viewed as a very dynamic phenomenon, hence population-at-risk could be difficult to estimate especially for certain crime types, therefore, the use of census data is seen as inappropriate and alternative model would be required. In the light of this, the space-time permutation model is employed for our case study.

The STSS can be carried out in either a single retrospective manner or a time-periodic prospective surveillance where the analysis is repeated at a regular time interval such as daily, weekly, monthly and so on. The former involves scanning through datasets and evaluate all 'historical' clusters i.e. all clusters that started at any time within the study period and ended before or on the set study period (end date). The prospective time-periodic cluster surveillance on the other hand, involves proactive detection of only 'alive' clusters i.e. clusters that started on or/and ended on the specified surveillance date. The prospective surveillance is mostly used to monitor incoming space-time datasets to proactively monitor outbreaks of geographic events (e.g. disease outbreak) in across a region. It generally stems on an expectation that at a certain moment, a localized cluster will begin to emerge, increasing in intensity as the geographic events continue to occur in the area. In this study we used both retrospective and prospective methods and assess the effectiveness of the latter by checking its results against the former (Fig. 1).

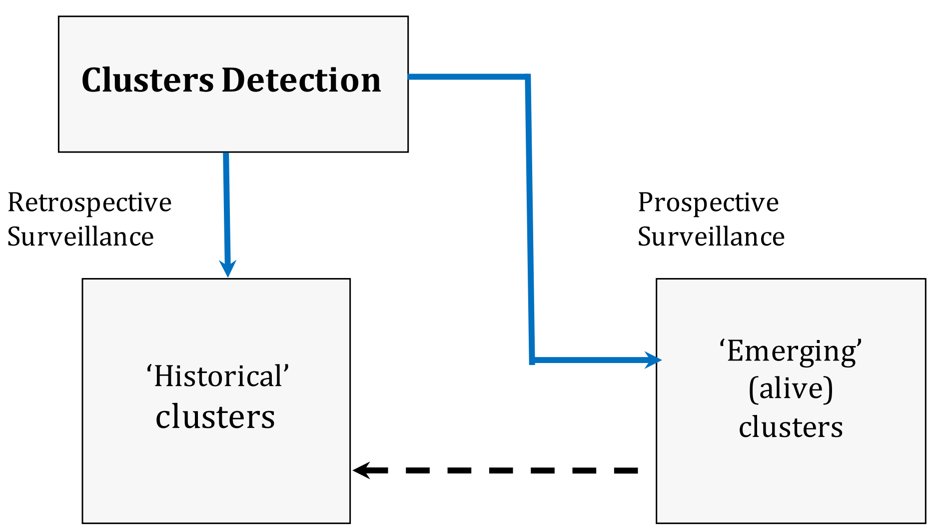

Fig. 1: Space-time cluster surveillance

3. Case study –Camden Borough of London

The Metropolitan Police Service Computer Aided Dispatch (CAD) system is the repository of all crime incident (999) calls within the City of London. The datasets used in this study however, is an extract from this database comprising of 28,686 geocoded crime records of the borough of Camden between 1st March 2011 and 31stMarch 2012 (one year and a month data). Each record consists of relevant information that fully describes a crime incident that occurred inside the borough at certain period of time. The datasets features the most detailed spatial and temporal granularity available with each incident point aggregated to a 250m by 250m grid while the incident time are recorded to the nearestseconds.

The space-time permutation scan statistical model implemented in SatScan™ software (Kulldorff and Information Management Service Inc. 2009) is used in this study for the detection of space-time clusters in each crime type. The model works by scanning through all possible grid points in the datasets and simultaneously iterate over all possible space and time divisions to report significant clusters of varying sizes. The upper limit for the geographic extent was set as 750 meter radius to be able to capture clusters covering considerably large geographical extent. The temporal duration of a certain date ensures that the detection of clusters between the starting date and that particular date is feasible. To evaluate the clusters' significance, the Monte Carlo replication of 999 was used to compute the p-values.

3.1 The retrospective surveillance

A total of 20 significant clusters were reported in the retrospective surveillance, based on the p-value threshold of 0.05 adopted, and each cluster is named in accordance to the administrative wards it covers or intersects. A cluster of significant value of p = 0.05 would mean that possibility of such cluster to have occurred by chance is once in every 20 days. Figure 2 shows the 3D representation of the clusters in ArcGIS 10 environment.

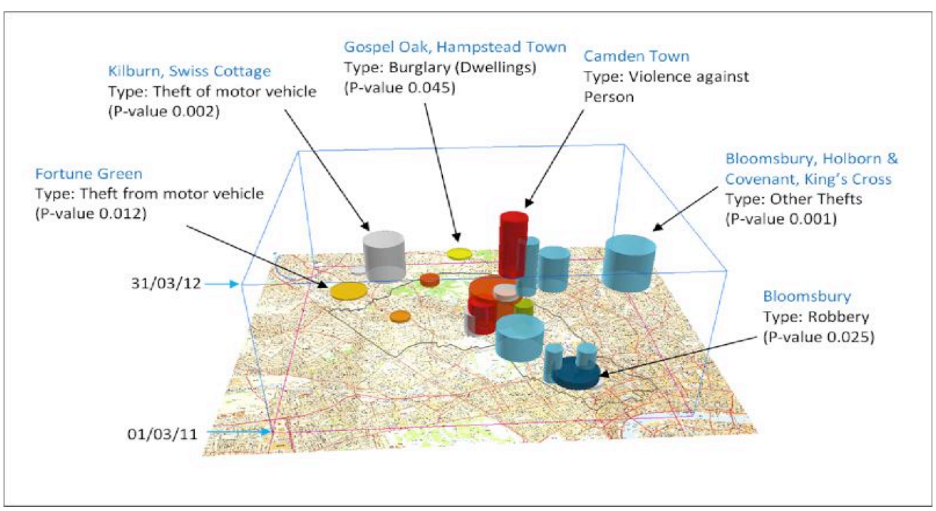

Fig. 2: Space-time display of clusters detected in the retrospective analysis

3.2 The prospective surveillance

The start_date for each cluster detected in the retrospective analysis served as the targets in our prospective analysis. The emerging clusters were captured by watching out for specific regions with low but yet to be fully significant clusters. As soon as such regions are spotted, continuous monitoring is done for subsequent days to ascertain whether the trend would continue or not. If the trend continues i.e. the p-value continues to decrease, such clusters would be considered emerging until the p-value gets to p=0.001 which is the maximum significance value based on 999 Monte Carlo replication specified during the analysis. In our surveillance however, only six (6) clusters were observed to show this pattern and are shown in Table 1 below. However, the King's Cross cluster (the 3rd in the table) failed to reach the maximum p of 0.001 but the pattern showed within that very short period of time warrants its inclusion in the table. More importantly, the percentage (or the density) of crimes inside each emerging clusters is relatively high compared to the occurrences across the entire area.

Table 1: Emerging clusters detected during prospective surveillance

SN	Emerging Cluster	Duration (2011)	Radius (m)	Observed crimes	Total crime	Recurrent Interval (days)	Reported P-value	% of crimes inside Emerging cluster
1	Gospel Oak, Hampstead (Burglary Dwelling)	30/08-02/09	235	5	23	200	0.005	22
		05/09	247	2	12	167	0.006	17
		06/09	247	2	13	1000	0.001	15
2	Fortune Green (Theft from Vehicles)	08/06-13/06	500	5	12	143	0.007	42
		14/06	500	2	5	500	0.002	40
		15/06	500	1	2	1000	0.001	50
3	King's Cross (Theft of motor vehicles)	17/10-25/12	555	23	264	22	0.046	9
		26/12-27/12	555	1	4	36	0.028	25
		27/12-28/12	555	1	6	67	0.015	17
		28/12-29/12	499	1	6	333	0.003	17
		29/12	499	1	3	500	0.002	33
		30/12	499	1	3	250	0.004	33
4	Camden Town (Other thefts)	13/07-15/07	248	7	58	22	0.046	12
		16/07	248	4	34	1000	0.001	12
		18/07	248	3	28	1000	0.001	11
5	Hampstead (Shoplifting-Theft)	23/05	250	3	5	167	0.006	60
		24/05-26/05	250	2	9	1000	0.001	22
6	Haverstock, Belsize, Gospel Oak, Camden Town, Kentish Town (Shoplifting Theft)	23/05-26/05	750	4	13	15	0.067	31
		26/05-29/05	750	4	17	167	0.006	24
		30/05	750	3	6	1000	0.001	50

To assess effectiveness of the prospective surveillance, we compared the retrospective *start_date* and *end_date* of each cluster with their respective prospective *detection_date* to examine how early or late the emergence was detected (refer to table 2). A very effective emerging cluster detection could be described as the one in which the cluster was detected prospectively long before its *end_date* or just a little later than its retrospective *start_date* depending on the overall duration of the cluster as detected in retrospective surveillance. Only the ‘King’s Cross’ (Theft of motor vehicle) cluster was detected two days before its actual retrospective *start_date* while the ‘Hampstead Town’ (Theft) cluster was detected prospectively on exact day as the retrospective *start_date*. Three (3) other clusters were detected few days after their retrospective *start_dates* (Table 2), but long enough before the *end_dates*. The last cluster as shown in the table (Shoplifting –Theft) took more than two months before it was detected prospectively, but still nearly 2 months earlier than the *end_date*. Generally, the spatial extents of the clusters barely changed as the clusters emerge in time (Table 1). Conclusively, we can say that the detection of these 6 emerging clusters were effective.

Table 2: Assessment of the prospective surveillance

SN	Crime Type	Cluster	trospective Start Date	Retrospective End Date	Prospective Detection Date	Radius (m) (Prosp)	Radius (m) (Retro)	P–Value (Prosp)
1	Burglary (Dwelling)	Gospel Oak, Hampstead Town	30/08/11	08/09/11	02/09/11	234	303	0.005
2	Theft from vehicles	Fortune Green	08/06/11	15/06/11	13/06/11	500	464	0.007
3	Theft of motor vehicles	King's Cross	26/12/11	10/01/12	24/12/11	555	267	0.046
4	Other thefts	Camden Town	14/07/11	17/01/12	15/07/11	248	249	0.046
5	Shoplifting (Theft)	Hampstead Town	23/05/11	14/06/11	23/05/11	250	250	0.006
6	Shoplifting (Theft)	Haverstock, Belsize, Gospel Oak, Camden Town, Kentish Town	18/03/11	13/07/11	25/05/11	235	668	0.067

4 Conclusions and future research

The space-time permutation scan statistic can serve as an important tool in prospective systematic time-periodic geographical surveillance for early detection of rapidly emerging crime outbreaks. Our study was able to demonstrate the possibility of detecting rapidly evolving space-time crime clusters within a geographical area. This analysis could serve a great asset towards early crime intervention in potential crime outbreak areas. Thus, daily prospective surveillance of crimes will facilitate situational awareness among the security agents and enable them to carry out crime intervention more intelligently. Detection of crime emergence using different time aggregates (weekly, monthly, seasonally) could as well be investigated. This will help in adjusting for certain temporal trends in the datasets (e.g. daily cyclic patterns) so as to investigate how crime varies in absence of those trends.

Acknowledgements

This research is part of the CPC (Crime, Policing and Citizenship) project supported by UK EPSRC (EP/J004197/1), in collaboration with the London Metropolitan Police.

References

Cheng, T. and Williams, D. (2012). Space Time Analysis of Crime Patterns in Central London. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci* ., Vol. XXXIX-B2, paves 47-52, 2012 Accessed on 20th Nov. 2012 from http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXIX-B2/47/2012/isprsarchives-XXXIX-B2-47-2012.pdf

Gao, P., Guo, D., Liao, K., Webb, J. J., and Cutter, S. L. (2012). Early Detection of Terrorism Outbreaks Using Prospective Space–Time Scan Statistics. *The Professional Geographer*, (ahead-of-print).

Kulldorff, M. and Information Management Services (IMS), Inc. (2009). *SaTScan version 9.1.1: Software for the Spatial and Space-Time Scan Statistics*. http://www.satscan.org. Accessed on 10th November, 2012.

Kulldorff, M., Heffernan, R., Hartman, J., Assuncao, R., and Mostashari, F. (2005). A Space-Time Permutation Scan Statistic for Disease Outbreak Detection. *PLoS Medicine Vol.* 2, No. 59

LeBeau J. L. (2000). Demonstrating the Analytical Utility of GIS for Police Operations: A Final Report. National Criminal Justice Ref. Service. NCJ 18710. Source: https://www.ncjrs.gov/pdffiles 1/nij/187104.pdf. Accessed on: 20th November 2012

Nakaya, T., and Yano, K., (2010). Visualising Crime Clusters in a Space-Time Cube: An Exploratory Data-Analysis approach using Space-Time Kernel Density Estimation and Scan Statistics. Transactions in GIS, Vol. 14, Issue 4, Pages 223-239.

Neill, D. B. and Gorr, W. L. (2007). Detecting and Preventing emerging Epidemics of Crime. Heinz Schl. of Public Policy, Carnegie Mellon University, Pittsburgh. Available at: http://www.isdsjournal.org/articles/1945.pdf

Neill, D. B., Moore, A. W., Sabhnani, M., and Daniel, K. (2005). Detection of emerging space-time clusters. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining* (pp. 218-227). ACM.

Uittenbogaard, A. C. and Ceccato, V. (2011). Space-Time Clusters of Crime in Stockholm, Sweden. Conference paper at Spatial Statistics 2011- Mapping Global Change. Available at: http://kth.diva-portal.org/smash/record.jsf?pid=diva2:489962 Accessed on 10th November, 2012