

RESEARCH

Open Access

Spatio-temporal crime hotspots and the ambient population

Nick Malleson^{1*} and Martin A Andresen²

Abstract

It is well known that, due to that inherent differences in their underlying causal mechanisms, different types of crime will have variable impacts on different groups of people. Furthermore, the locations of vulnerable groups of people are highly temporally dynamic. Hence an accurate estimate of the true population at risk in a given place and time is vital for reliable crime rate calculation and hotspot generation. However, the choice of denominator is fraught with difficulty because data describing popular movements, rather than simply residential location, are limited. This research will make use of new 'crowd-sourced' data in an attempt to create more accurate estimates of the population at risk for mobile crimes such as street robbery. Importantly, these data are both spatially and temporally referenced and can therefore be used to estimate crime rate significance in both space and time. Spatio-temporal cluster hunting techniques will be used to identify crime hotspots that are significant given the size of the ambient population in the area at the time.

Keywords: Crime analysis and mapping; Population at risk; Clustering; Big data; Twitter; SatScan

Introduction

The *crime rate* is a common statistic that is often used to summarise the quantity and extent of criminal events. Using crime rates helps to reveal clusters in space and/or time in which the volume of crime is significantly different to that which would be expected given the underlying demographics or physical environment. However, choosing an appropriate denominator is non trivial and studies are often restricted to using residential population data that do not adequately describe the true ambient population. This can lead to the calculation of misleadingly high or low crime rates.

In an attempt to alleviate some of these drawbacks, this research will utilise novel 'crowd-sourced' data to measure the ambient population. In particular, it will use messages generated on mobile devices (such as smart phones) and posted to the Twitter social media service. As this article will discuss, this source of data has the potential to represent the ambient population at much higher spatial and temporal resolutions than by those used previously. However, there are important drawbacks to using these data,

particularly regarding noise and population bias, that temper any conclusions drawn from the results. The overall aim of the research is to identify spatio-temporal clusters of crime that are significant even after taking account of the ambient population, and explore their shifting spatio-temporal dynamics. However, whilst questions remain as to the degree to which social media data truly represent the population under study, the results presented here must be considered preliminary and future work will subject the data to greater scrutiny.

This article is structured as follows: Section 'Background' outlines the background to the study and reviews relevant literature; Section 'Methods' details the methods used and the area under study; Section 'Results and discussion' outlines the results and discussion; and Section 'Conclusions' draws conclusions.

Background

The population at risk in crime analysis

It is well known that inherent differences in their underlying causal mechanisms mean that different types of crime must be analysed distinctly. Hence it can be argued that the denominator used in crime rate calculations should be given the same consideration. It has been recognised that: "a valid rate ... should form a probability statement,

*Correspondence: n.s.malleison@leeds.ac.uk

¹School of Geography, University of Leeds, Woodhouse Lane, LS2 9JT Leeds, UK

Full list of author information is available at the end of the article

and therefore should be based on the risk or target group appropriate for each specific crime category” (Boggs 1965, pg 900). Most research uses the size of the residential population as the population denominator. However, recent studies suggest that the residential population is unsuitable as a measure of the population at risk for crimes that involve mobile victims such as assaults (Boivin 2013), robbery (Zhang et al. 2012) and violent crime (Andresen 2011). That is not to say that residential population is an unsuitable denominator for *all* crime types – there might be little difference between ‘traditional’ denominators and other measures for residential burglary and car theft (Cohen et al. 1985) – but its applicability to crimes that do not rely on the number of residents in a neighbourhood is highly questionable.

However, there is no scientific consensus on the appropriate ways to measure the population risk (Andresen and Jenion 2010). In effect, the problem faced by researchers is that very few non-residential population measures exist. Nevertheless, some research has attempted to move beyond simple residential risk measures. Boggs 1965 presents one of the earliest examples of this approach by using different denominators for the different types of crime under examination. These included the business-residential land use ratio for business crime, parking space for vehicle theft and road area (as a proxy for pedestrians) for street robbery. Some research has made use of pedestrian models to compare rates of street crime using pedestrian volume estimates and traditional residential data, finding significant differences between the two (Chainey and Desyllas 2008). More recently, Andresen and colleagues have made use of the LandScan Global Population Database as a denominator (Andresen 2006; Andresen and Jenion 2010; Andresen 2011; Andresen et al. 2012). LandScan data provides an average estimate of the global ambient population at a spatial resolution of approximately 1km. However, this spatial resolution is relatively poor for crime analysis – research has shown that analysing crime at scales greater than the street level can hide important lower-level patterns (e.g. Andresen and Malleon 2011). Finally, mobile phone data are becoming a popular means of estimating pedestrian flows, and some preliminary research has attempted to use these data to improve crime estimates (Bogomolov et al. 2014). However, as mobile telephone data are privately owned they can be extremely difficult to access for research purposes and pose a number of ethical questions, particularly around privacy and informed consent.

Social media and the ambient population

The ambient population is highly dynamic and exhibits strong spatial and temporal fluctuations at various scales (e.g. hourly, daily, seasonal). Traditional data, such as censuses or other household-based studies, are temporally

static and do not capture adequate information about activities and behaviour outside of the home. Hence it becomes apparent that a lack of data is the main barrier to further research into the ambient population. Fortunately, the recent proliferation of user-generated content on social-media services has the potential to reveal considerable information about peoples’ daily spatio-temporal movements that might prove invaluable for developing more accurate population at risk estimates. Examples of social-media services include messages posted to Twitter (some of which include accurate GPS coordinates); the Foursquare service (that allows users to publicise their current location); geo-located photos posted on the Flickr website; etc. Some examples that make use of these data include the mathematical analysis of human mobility patterns (Cheng et al. 2011), new neighbourhood boundary definitions based on the characteristics of the people who commonly frequent them (Cranshaw et al. 2012) and the identification of events such as earthquakes (Crooks et al. 2013). Furthermore, a recent special issue of *Cartography and Geographic Information Science* 40(2) 2013 entitled ‘Mapping Cyberspace and Social Media’ has a number of examples, and some preliminary research compares data from the UK Census and Twitter to uncover spatial crime clusters (Malleon and Andresen 2015). However, the authors are unaware of any research that uses social media data directly to better understand the ambient population who are susceptible to crime victimisation.

Drawbacks of using social media data

The main drawback associated with data from social media, that emerges largely due to their novelty, is that they have undergone limited validation. It is therefore difficult to quantify the extent to which they provide a reliable characterisation of the *true* ambient population.

For example, it is difficult to estimate the proportion of the population who use Twitter at all, let alone those who use it sufficiently regularly to contribute to an accurate ambient population measure. Although there is evidence that the use of social media is becoming more prolific – for example “two-thirds of online adults (66 %) use social media platforms ... ” ($N = 2,277$) (Smith 2011) – it is undoubtedly a minority who participate regularly. Furthermore, the data used here only include messages that have been attributed with a GPS location. These are predominantly created on location-aware mobile devices such as smart phones. The proportions of such messages vary, but percentages between 1 and 5 % of the total number created are common. Hence the measure of the ambient population is reduced first to a sample of Twitter users, and then further to those users who report their location accurately.

A further drawback relates to the potential for participation inequality due to the disparity of access to

the Internet and related technologies, termed the ‘digital divide’ (Yu 2006). Although the digital divide has traditionally referred primarily to a disparity in *access* to a computer or the Internet, a more nuanced definition has arisen in recent years to encapsulate the lack of adequate technical skills to fully participate in online culture, as well as the simple availability of hardware (Fuchs 2008; Schradie 2011). Indeed, the digital divide is now seen by some as social problem as much as a technical one (Fuchs 2008; Smith and Brenner 2012) and one that has “emerged along the familiar fault lines of social inequality” (Chen and Wellman 2005). For example, recent research has found that: digital content creators are predominantly from more affluent groups (Brake 2014); young people (and particularly colleague graduates) are more likely to contribute to websites (Brake 2014); and there is a “growing production divide” between the poor, working class and more affluent internet users (Schradie 2011). There is some evidence that these trends are less evident with Twitter use – e.g. higher rates of Twitter use among black internet users (Smith and Brenner 2012) – but, again, it is very likely that inequality in Internet use will distort the ambient population measure used here.

Although there are clear drawbacks to the use of social media data, these should not preclude their use in crime research. Traditional residential-based population measures are likely to be more representative of the underlying population, but they also have drawbacks. These include their inability to measure dynamic, mobile populations (i.e. the daytime rather than nighttime population) and the length of time that can amount between their collection and use (particularly with censuses that are captured once per decade). Therefore research with new social media data is an important area that warrants further investigation, as long as drawbacks with the underlying data are clear, and conclusions are drawn with care.

Methods

Data and the study area

The chosen study area is Leeds, UK. The Leeds local authority district is the third largest in the UK with a population estimated to be 757,655 in 2012. As illustrated by Fig. 1, the geography of Leeds is typical of many British cities. The city has a central business and retailing area with high concentrations of shops, businesses and entertainment facilities that, along with the busy

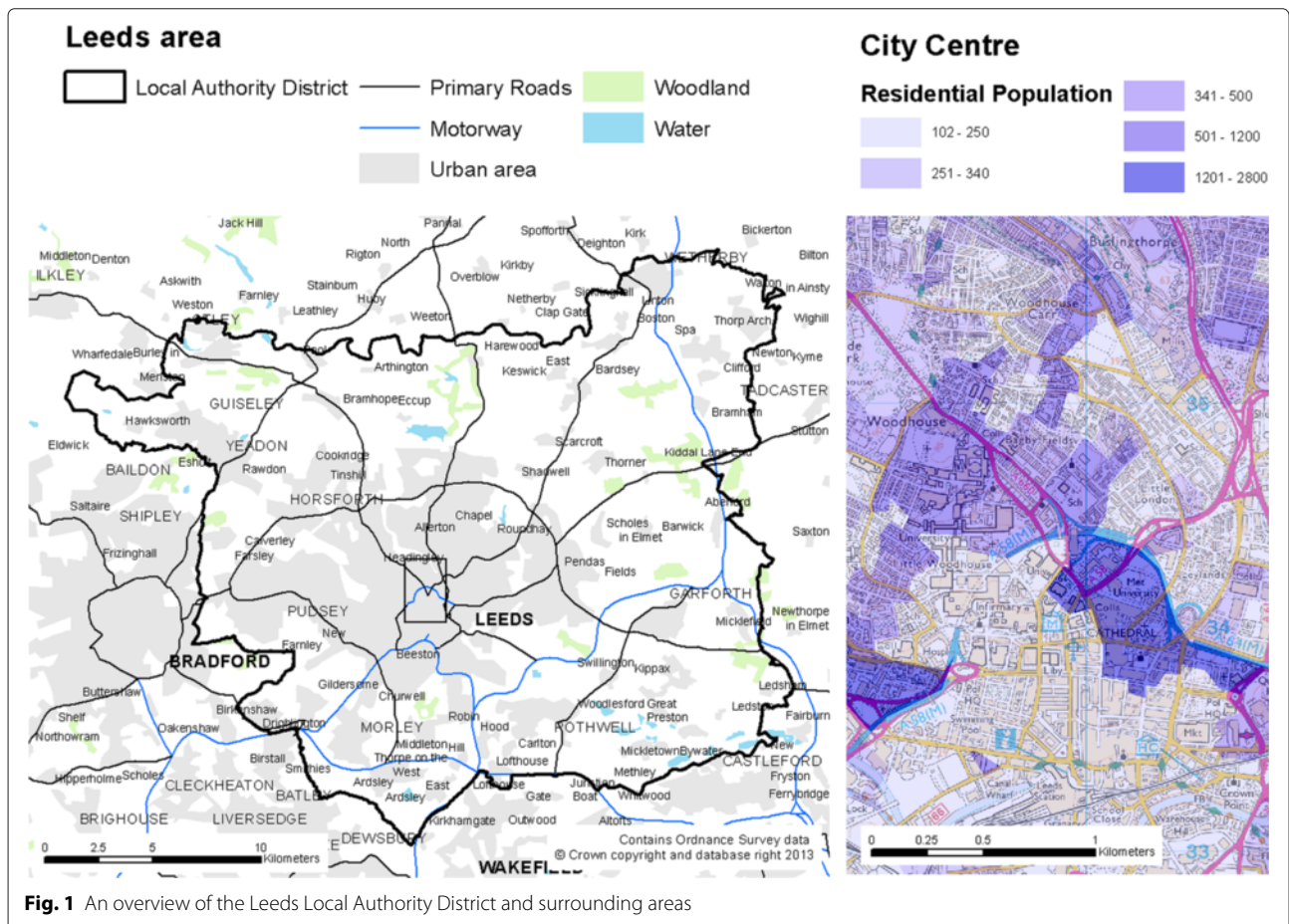


Fig. 1 An overview of the Leeds Local Authority District and surrounding areas

railway station, attracts a large number of visitors to the centre. These facilities act as both crime generators and crime attractors (Brantingham and Brantingham 1995) that naturally cause an increase in certain types of crime. Importantly, relatively few people live in the city centre. This makes the residential density an inappropriate denominator for some crime rate calculations.

The research makes use of two data sources: reported crime data collected by West Yorkshire Police, and messages posted to the social media service Twitter. The Twitter data consist of messages posted from within Leeds during the period 22nd June 2011 to 14th April 2013. Only messages with associated GPS coordinates have been included; these are commonly created using mobile devices by users who have explicitly opted to publish their present location. After removing messages from business accounts (e.g. weather forecasts, car advertisements, etc.) the number of messages, $N = 1,955,655$.

The crime data consist of crimes recorded by West Yorkshire Police in Leeds that occurred in the period April 2001 – March 2004. The data were filtered such that only ‘street’ crimes were included, as it is these crimes that are most likely to be influenced by the ambient population. The crime types used were ‘Theft from person’ and ‘Robbery’ (similar types to (Chainey and Desyllas 2008)). Although in some instances these crimes will occur within buildings, in most cases they occur outside. In some cases the exact time of the occurrence was not known and hence the average time point between start and end periods was used.

A drawback with the two sources used here is that the dates of the occurrences do not align. Although more recent crime data are available publicly from the [police.uk](http://www.police.uk) service (<http://www.police.uk>), those data have been temporally aggregated to the nearest month, which makes them unusable here. However, it is inevitable that many crime studies must use data that originate from different time periods, particularly as censuses occur very infrequently. Furthermore, it is reasonable to assume that the underlying structure of the ambient population has not changed substantially in Leeds in the 10 years prior.

Data preparation

As the behaviour of the ambient population in an urban area like Leeds is largely regular in space and time, both the crime and social media data were temporally aggregated such that they describe an ‘average’ week on an hour-by-hour basis. Hence all datum were grouped into one of $24 * 7 = 168$ distinct time periods, starting at 4am on Monday morning (i.e. all events that occurred between 4am and 5am on Monday morning were assigned to the first temporal group). Although this will mask any seasonal trends and those that occur with a periodicity of less

than one hour, it will highlight the interactions between crime events and typical daily urban flows such as people commuting, partaking in leisure activities, shopping, etc.

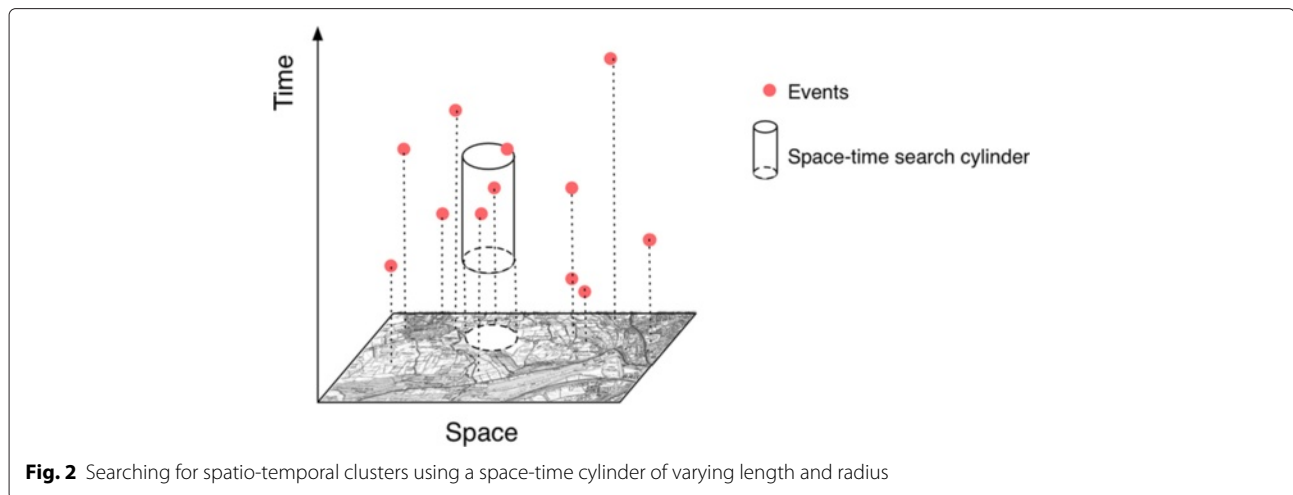
Furthermore, a requirement of the clustering algorithm chosen is that both data sets share a common geography. Hence it was necessary to perform some spatial aggregation. The police.uk service has created a set of *anonymous map points* that are used to spatially anonymise individual crimes but maintain the overall spatial structure in the data. Therefore each crime and social media message were snapped to the nearest anonymous map point.

Spatio-temporal clustering

Spatial scan statistics are designed to detect significant clusters of events by moving a circular window of varying radius across a dataset and recording those circles that exhibit statistically significant numbers of events, optionally taking some population at risk into account (Kulldorff 1997; Openshaw 1987). Such methods have proven popular in a crime context (Corcoran et al. 2003; Malleon and Andresen 2015), but are limited because they do not take account of temporal variations. Therefore, Space-Time Scan Statistics (e.g. Kulldorff et al. (2005)) are also becoming popular (Cheng and Adepeju 2013; Gao et al. 2013; Leitner and Helbich 2011; Nakaya and Yano 2010). They employ a similar method, but use a scanning window that covers a temporal dimension as well as the spatial dimension. Conceptually, this can be perceived as a *cylinder* of varying radius and length that scans a through three dimensional data – space on the horizontal axis and time on the vertical axis. For example, see Fig. 2. Following Nakaya and Yano (2010), this conceptual model will also be used to visualise the resulting clusters.

However, most space-time research projects do not take account of the population at risk in the identification of clusters. Here, a Discrete Poisson model (Kulldorff 1997) that takes the population at risk into account was employed. The model assumes that the number of cases (crimes) in a given space-time search cylinder follows a Poisson distribution. The null hypothesis assumes that the expected number of cases at each point will be proportional to the population at risk at the same time and space. The algorithm has been implemented in SaTScan (<http://www.satscan.org/>).

Cluster sizes were limited to 1km in space and 8 units (hours) in time. If a cluster extends beyond these limits then it is more likely that the method has merged a number of distinct clusters, such as a hotspot that emerges around school closing as children leave school and travel through the city centre, and another that emerges a few hours later as adults begin to visit bars or pubs (for example).



Results and discussion

Figure 3 presents the significant clusters returned by the Discrete Poisson model, as well as the overall (a-temporal) density of Twitter messages and crimes. There is considerable insight that could be gained from a closer inspection of each of the clusters. However, exploring each instance in detail is beyond the scope of this paper. Instead, the analysis will focus on two substantial clusters that emerge in the centre of the city. Table 1 defines the temporal shape of the two clusters.

Like many other cities, Leeds city centre exhibits a relatively large volume of crime. However, prior a-temporal research has argued that after taking the equally large ambient population into account, the city centre hotspot loses significance (Malleon and Andresen 2015). Here, we extend the analysis by taking account of the time of offence as well as the location. The Discrete Poisson model identified two clusters in the city centre area. The discussion will focus first on an explanation for their significance, followed by an analysis of the differences in crime type and/or motivation that they might represent.

The first cluster, A, covers a relatively small area (radius approximately 400 m) in the city centre, characterised by bars, shops, restaurants, etc. The cluster extends from 10:00 to 17:00 on Saturday. This cluster is particularly pertinent because it has a large overall volume of crime *and* a large volume of social media messages. Hence with an a-temporal cluster analysis (as in Malleon and Andresen (2015)) the cluster would not be statistically significant. However, the volume of crime during the day on Saturday is substantial enough to be significant even given the large ambient population. It is, of course, possible that this cluster is an artefact of fewer people using Twitter at the time and hence due to a mis-representation of the true size of the ambient population. We argue that this is unlikely however; there is nothing to suggest that visitors to the

city centre on a Saturday are less likely to participate in social media than other groups.

The second cluster, B, is larger (radius approximately 1 km) and hence less homogeneous with respect to the physical environment. However, it is notable that a large portion of the cluster is part of the University of Leeds campus and the surrounding predominantly student accommodation. The cluster extends from 21:00 on Saturday evening until 02:00 on Sunday. Unlike the first cluster, it does not cover an area with a consistently large volume of crime or social media messages. Again, therefore, using an a-temporal analysis this cluster would likely not have been statistically significant.

There are some particular differences in the two clusters that might highlight variations in the underlying causal mechanisms that lead to their emergence. Cluster A occurs during the daytime on a Saturday, in an area that is well known for its retail offering and will be very busy. Hence it is very likely that the hotspot is a consequence of thefts from individual people (e.g. pickpocketing) or from shops (unfortunately the crime classification provided in the data is not detailed enough to distinguish between these different types). Given the area and time, the victims are most likely to be shops or shoppers. The second cluster, however, occurs later in the day, and covers an area that is largely dominated by the University and its students. Hence the victims are much more likely to be students enjoying activities in the evening. Therefore the clusters probably represent crimes against very different victim groups and will be committed by different offenders using very different crime templates. Whilst it is too early to provide any concrete recommendations from these preliminary results, particularly given the questionable reliability of the social media data, the emergence of these diverse clusters might begin to shed light on the shifting spatio-temporal distributions of crime and their potential victims.

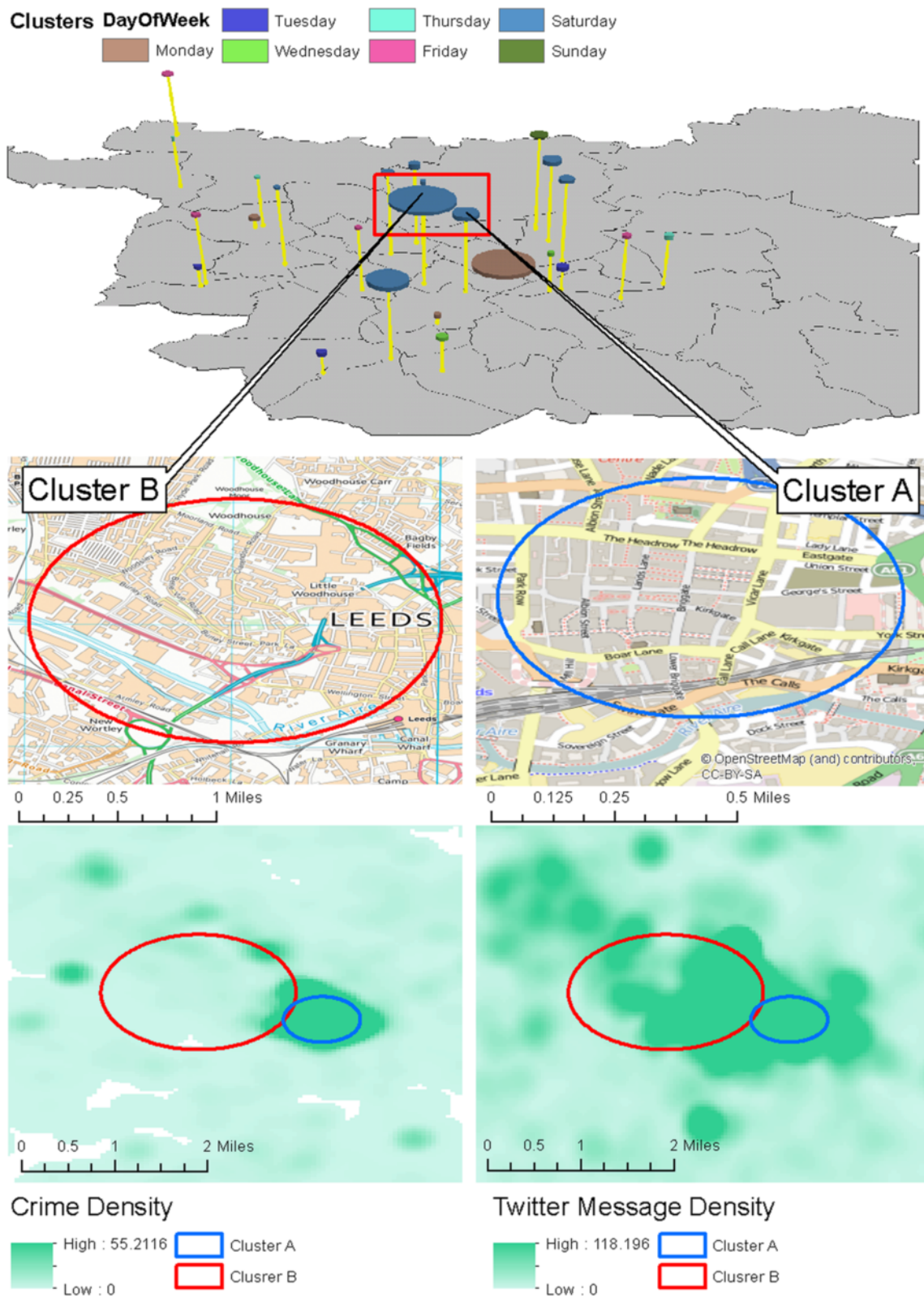


Fig. 3 The significant clusters returned by the Discrete Poisson model Time is represented on the vertical axis. Two city-centre clusters are highlighted. The density of crimes and social media messages (calculated using Kernel Density Estimation (Silverman 1986)) are also illustrated relative to the two central clusters. ©OpenStreetMap contributors, CC-BY-SA ©Crown copyright and database right 2015.

Table 1 The temporal definition of the clusters highlighted in Fig. 3

	Day	Time Period
Cluster (A)	Saturday	10:00 – 17:00
Cluster (B)	Saturday	21:00 – 02:00 (Sunday)

Conclusions

This paper has used ‘crowd-sourced’ data to estimate the ambient population and identify spatio-temporal crime clusters that are significant *given the number of potential victims present at the time of the offence*. The marriage of crime data with a temporally dynamic ambient population is, as far as the authors are aware, a novel contribution.

As discussed, data from social media offer advantages over some traditional sources in that they reflect the high spatial and temporal dynamism inherent in the ambient population. However, they also suffer some considerable drawbacks. Traditional data, such as those compiled from surveys, are usually rigidly defined and contain minimal errors or omissions. From these relatively small sets of data, social scientists have developed quantitative tools that are effective at extrapolating to a much wider portion of society (Savage and Burrows 2007). Conversely, social media sources are much messier. Omissions will be numerous, the structure of the data will vary, and it will be difficult to determine which groups of people are over- or under-represented in the data.

The optimist will, however, be confident that the drawback of larger measurement error will be offset by a considerably lower sampling error (Mayer-Schönberger and Cukier 2013). It is also a point that has been made by Savage and Burrows (2007) who foresee a ‘crisis’ in an empirical sociology that fails to embrace these new data and methods, relying instead on small, carefully constructed samples. However, the process of reducing measurement error through larger sample sizes is irrelevant if the samples are being drawn from an inherently biased population (i.e. the group of people who use Twitter). As mentioned previously, the ‘digital divide’ will undoubtedly bias the data used here. But questions remain as to the extent to which the estimate of the ambient would differ were there no inherent structural bias. Future work must attempt to better estimate these errors and biases.

The data used here are a proxy for the ambient population. As ‘Big Data’ and social media become more prevalent and pervasive, the quality of the proxy will undoubtedly increase. Therefore this research illustrates the *potential* that these new forms of data and methods can offer to crime analysis.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

NM contributed to the design of the study, performed the data preparation and analysis work, and drafted the manuscript. MA contributed to the design of the study and contributed to the preparation of the manuscript. All authors read and approved the final manuscript.

Author details

¹School of Geography, University of Leeds, Woodhouse Lane, LS2 9JT Leeds, UK. ²Institute for Canadian Urban Research Studies, School of Criminology, Simon Fraser University, 8888 University Drive, V5A 1S6 Vancouver, Canada.

Received: 1 February 2015 Accepted: 6 May 2015

Published online: 24 May 2015

References

- Andresen, MA (2006). Crime measures and the spatial analysis of criminal activity. *British Journal of Criminology*, 46(2), 258–285. doi:10.1093/bjc/azi054.
- Andresen, MA, & Jenion, GW (2010). Ambient populations and the calculation of crime rates and risk. *Security Journal*, 23(2), 114–133. doi:10.1057/sj.2008.1.
- Andresen, MA (2011). The ambient population and crime analysis. *The Professional Geographer*, 63(2), 193–212. doi:10.1080/00330124.2010.547151.
- Andresen, M, & Malleon, N (2011). Journal of Research in Crime and Delinquency, 48(1), 58–82.
- Andresen, MA, Jenion, GW, Reid, AA (2012). An evaluation of ambient population estimates for use in crime analysis. *Crime Mapping: A Journal of Research and Practice*, 4, 8–31.
- Boggs, SL (1965). Urban crime patterns, 30(6), 899–908. doi:10.2307/2090968.
- Boivin, R (2013). On the use of crime rates. *Canadian Journal of Criminology and Criminal Justice*, 55(2), 263.
- Bogomolov, A, Lepri, B, Staiano, J, Oliver, N, Pianesi, F, Pentland, A (2014). Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data. In *ICMI '14 Proceedings of the 16th International Conference on Multimodal Interaction*. doi:10.1145/2663204.2663254, <http://dl.acm.org/citation.cfm?id=2663254> (pp. 427–434). New York, NY, USA: ACM.
- Brantingham, P, & Brantingham, P (1995). Criminality of place. *European Journal on Criminal Policy and Research*, 3(3), 5–26. doi:10.1007/BF02242925.
- Brake, DR (2014). Are we all online content creators now? web 2.0 and digital divides. *Journal of Computer-Mediated Communication*, 19(3), 591–609. <http://onlinelibrary.wiley.com/doi/10.1111/jcc4.12042/full>.
- Chainey, S, & Desyllas, J (2008). Modelling pedestrian movement to measure on-street crime risk. In L Liu & J Eck (Eds.), *Artificial Crime Analysis Systems: Using Computer Simulations and Geographic Information Systems*. Hershey, PA: Information Science Reference.
- Chen, W, & Wellman, B (2005). Minding the cyber-gap: the Internet and social inequality. In M Romero & E Margolis (Eds.), *The Blackwell Companion to Social Inequalities* (pp. 523–545): Blackwell Publishing Ltd.
- Cheng, Z, Caverlee, J, Lee, K, Sui, DZ (2011). Exploring millions of footprints in location sharing services. In *Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM)*. Menlo Park, California: AAAI Press.
- Cheng, T, & Adepeju, M (2013). Detecting emerging space-time crime patterns by prospective STSS. In *Proceedings of the 12th International Conference on GeoComputation*. <http://www.geocomputation.org/2013/papers/77.pdf>.
- Cohen, LE, Kaufman, RL, Gottfredson, MR (1985). Risk-based crime statistics: A forecasting comparison for burglary and auto theft. *Journal of Criminal Justice*, 13(5), 445–457. doi:10.1016/0047-2352(85)90044-3.
- Corcoran, JJ, Wilson, ID, Ware, JA (2003). Predicting the geo-temporal variations of crime and disorder. *International Journal of Forecasting*, 19(4), 623–634. doi:10.1016/S0169-2070(03)00095-5.
- Cranshaw, J, Schwartz, R, Hong, J, Sadeh, N (2012). The livelihoods project: Utilizing social media to understand the dynamics of a city. In *Sixth International AAAI Conference on Weblogs and Social Media*. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4682>. Palo Alto, California: AAAI Press.
- Crooks, A, Croitoru, A, Stefanidis, A, Radzikowski, J (2013). #earthquake: Twitter as a distributed sensor system. *Transactions in GIS*, 17(1), 124–147. doi:10.1111/j.1467-9671.2012.01359.x.
- Fuchs, C (2008). *Social Science Computer Review*, 27(1), 41–58. doi:10.1177/0894439308321628.

- Gao, P, Guo, D, Liao, K, Webb, JJ, Cutter, SL (2013). Early detection of terrorism outbreaks using prospective space-time scan statistics. *The Professional Geographer*, 65(4), 676–691. doi:10.1080/00330124.2012.724348.
- Kulldorff, M (1997). A spatial scan statistic. *Communications in Statistics - Theory and Methods*, 26(6), 1481–1496. doi:10.1080/03610929708831995.
- Kulldorff, M, Heffernan, R, Hartman, J, Assunção, R, Mostashari, F (2005). PLoS Med, 2(3), 59. doi:10.1371/journal.pmed.0020059.
- Leitner, M, & Helbich, M (2011). The Impact Of Hurricanes On Crime In The City Of Houston, TX. *Cartography and Geographic Information Science*, 38(2), 214–222. doi:10.1559/15230406382213, <http://www.tandfonline.com/doi/abs/10.1559/15230406382213>.
- Malleon, N, & Andresen, MA (2015). The impact of using social media data in crime rate calculations: shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science*, 42(2), 112–121. doi:10.1080/15230406.2014.905756.
- Mayer-Schönberger, V, & Cukier, K (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think*. London: John Murray.
- Nakaya, T, & 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*, 14(3), 223–239. doi:10.1111/j.1467-9671.2010.01194.x.
- Openshaw, S (1987). An automated geographical analysis system. *Environment and Planning A*, 19(4), 431–436.
- Savage, M, & Burrows, R (2007). The coming crisis of empirical sociology. *Sociology*, 41(5), 885–899. doi:10.1177/0038038507080443.
- Schradie, J (2011). The digital production gap: The digital divide and web 2.0 collide. *Poetics*, 39(2), 145–168. doi:10.1016/j.poetic.2011.02.003.
- Silverman, BW (1986). *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.
- Smith, A (2011). Why americans use social media. Technical report, Pew Research Centre. <http://www.pewinternet.org/Reports/2011/Why-Americans-Use-Social-Media.aspx>.
- Smith, A, & Brenner, J (2012). Twitter use 2012. Technical report, Pew Research Center. <http://pewinternet.org/Reports/2012/Twitter-Use-2012.aspx>.
- Yu, L (2006). Understanding information inequality: Making sense of the literature of the information and digital divides. *Journal of Librarianship and Information Science*, 38(4), 229–252. doi:10.1177/0961000606070600.
- Zhang, H, Suresh, G, Qiu, Y (2012). Issues in the aggregation and spatial analysis of neighborhood crime. *Annals of GIS*, 18(3), 173–183.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
