Enumerating the Ambient Population in the Context of Crime: A Comparative Study of ‘Contemporaneous’ and ‘Traditional’ Data Sets

Emily Sheard, Nick Malleson, and Mark Birkin

What is the ambient population and why does it matter?

The term ‘ambient population’ refers to the actual number of persons who are present within a particular area at any given time. Population counts have traditionally been derived from decennial censuses but these have tended to be limited in their ability to capture daily fluctuations in population size that can occur when people move between different activity types and locations. For example, ONS (2013, p. 1) note that the City of London typically experiences a 56 fold increase from usually resident to workday population! Since reliable population figures are fundamental to the efficient allocation of public resources and service provision, there is growing interest in the use of alternative data sources as enumerators of the population.

The ambient population in the context of crime

Accurate representations of the ambient population are extremely relevant for the field of crime analysis because they can facilitate the identification of areas where the risk of victimisation is comparatively high, or low, relative to the size of the underlying population. To illustrate, a UK city centre people-based offence rate based on a ‘traditional’ denominator, such as a usually resident population count, is likely to overstate the risk of victimisation because these areas tend to be less densely populated i.e. there might appear to be more crime per head of population than is actually the case. An ambient population figure is likely to generate a more representative crime rate because it includes additional constructs of the population (e.g. workers and visitors) and also because it can be aligned more closely with the timing of offences.
Data and preparation

A key aim of the dissertation research was to ascertain if the risk of victimisation varies depending on the type of background population data set that is analysed. A 2011 census-based usually resident count represented ‘traditional’ enumeration and anonymised and spatially aggregated mobile telecoms data served as the ‘contemporaneous’ source. The latter was for a one week period in May 2013 and included hourly population estimates for a number of mobile network grid cells in south east England.

Recognising that ‘contemporaneous’ data sets are not universally accessible, a 2011 census-based workday population count was also included to see if it could be utilised as an intermediary source in the absence of other data. Spatially anonymised and temporally aggregated open source data relating to ‘theft from the person’ offences (THEP) in May 2013 was downloaded from data.police.uk. The crime and census data was then assigned to the mobile network geography for the purposes of analysis.

Case study – City of Westminster borough, London

The City of Westminster was chosen as a case study area because it contained the mobile network grid cell with the highest number of THEPs, together with a wide range of ‘routine activity’ nodes, including retail premises and tube stations. Because the THEPs were grouped by month, an assumption was made, based on the literature, that some of the offences probably occurred during the peak time period 16:00-19:59 (Newton et al, 2014; Smith et al, 2006); an average hourly mobile population estimate was calculated for every grid cell to reflect this. The Geographical Analysis Machine (Openshaw et al, 1987) was then used to test each of the three data sets to see if there were any statistically significant clusters of THEPs after each of the background populations had been taken into account.

In other words, were there any locations where the number of offences was higher than expected relative to the number of potential victims?
Findings and practical application

As might theoretically be expected, the residential population count generated the highest level of excess incidence whereas the mobile population data resulted in the lowest maximum cluster value. It should be noted, however, that significant clustering was still evident even when the mobile phone data was analysed (see Figure 1.). This type of information should be of interest to crime practitioners because it indicates where the risk of victimisation is greatest, as well as providing a means by which the success of crime reduction initiatives can be measured.

Future research will continue this line of work, but broaden the analysis to include other factors that might explain and predict crime rates.

References


