Consumer Data Research Centre
An ESRC Data Investment
LDC

Understanding Footfall using Wi-Fi sensor data in the Smart City

Karlo Lugomer, Balamurugan Soundararaj and Roberto Murcio
University College London

Impacts

- Presents the Smart Street Sensor project, a novel case of producing footfall data using Wi-Fi signals from mobile devices.
- Outlines methodologies for estimating footfall and pedestrian flows across retail centres in the UK.
- Addresses significant gaps in data availability concerning human activities.

Project Background

The accurate measurement of human activities and their spatial and temporal distribution is a fundamental first step towards their understanding and management. Such distributions are highly granular and accurate estimation is crucial for decision-making processes across numerous fields, including urban management, retail, transport planning, and emergency services. In the retail sector, in this case, the inference of footfall aims to identify the number of individuals passing by the shop during a given period and can provide a crucial indicator of store performance.

Current estimates of urban activity and footfall have largely been derived from Census data, sample surveys or traffic counts. Census data, whilst being comprehensive, is only representative of a night-time (residential) population and has a very low frequency of update. Alternately, sample surveys and traffic counts offer day-time estimates with a higher temporal resolution, yet coherent collection of these data suffer from many practical issues. However, recent innovations in technology have provided novel means of mining the small-scale spatio-temporal dynamics of human activity. For example, Wi-Fi signals from mobile devices can be used as a proxy for activity patterns at different times and locations, offering promising opportunity for the estimation of footfall in retail areas.

In partnership with the Local Data Company (LDC) this research presents the Smart Street Sensor project, a novel case of producing footfall data using Wi-Fi signals from mobile devices. The data presented is being derived from Wi-Fi sensors that have been installed in a number of retail establishments across the UK, using proprietary hardware and software developed by the LDC. This case study describes the methods of creating, capturing, and processing this novel data and outlines the methods adopted for initial data cleaning practices. The project aims to use this data to accurately estimate footfall and pedestrian flows across retail centres in the UK whilst also providing insights for the wider research community.

Data and Methods

As of January 2017, approximately 500 sensors have been installed across the UK (see Figure 1). These cover a large number or retail centres throughout England, Scotland and Wales.

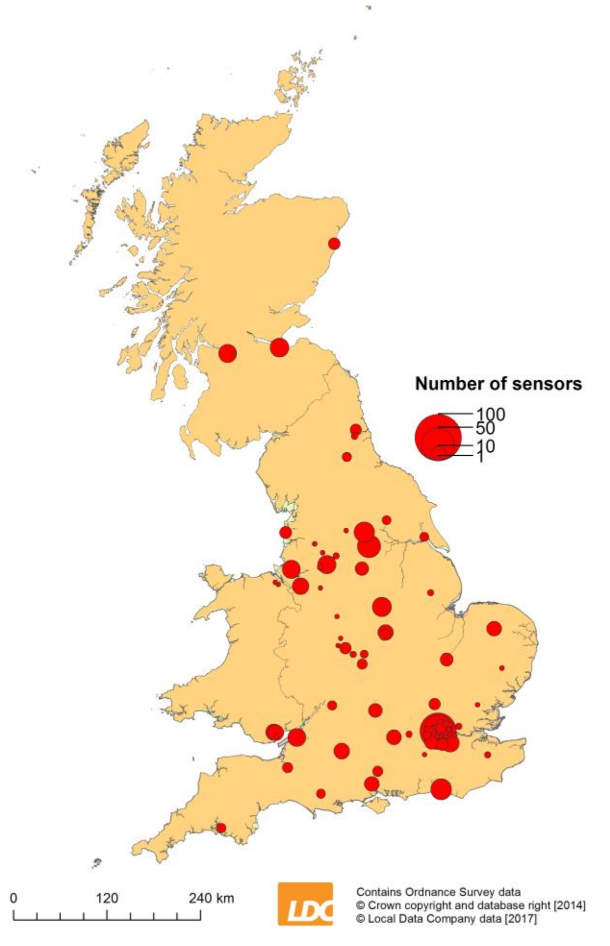

Figure 1. The distribution of sensors installed across Great Britain.

Out of the various signals emitted by wireless mobile devices, the sensors focus on capturing the Wi-Fi probe request frames, which are sent from mobile devices to search for new Wi-Fi networks. Since Wi-Fi capability to connect to access points is almost ubiquitous on mobile devices, this is a particularly advantageous approach. For instance, it is not only non-intrusive and passive (thus improving the

Consumer Data Research Centre
An ESRC Data Investment
LDC

Understanding Footfall using Wi-Fi sensor data in the Smart City

Karlo Lugomer, Balamurugan Soundararaj and Roberto Murcio
University College London

participation rate), but also contains a MAC address of the device. This, when hashed, can act as a unique identifier without compromising participants' privacy.

Initial methodologies aimed to perform exploratory analyses on the data and identify uncertainties in sensor counts by drawing comparisons with manually detected footfall measured on site. Both simple data mining algorithms and more laborious ground-truthing were adopted in order to understand and address ambiguities further.

Key Findings

Figure 2 shows the distribution of sensors ranked by the total number of probe counts captured. The lack of normal distribution may suggest that the probe counts are not random and are a function of the natural processes in urban areas.

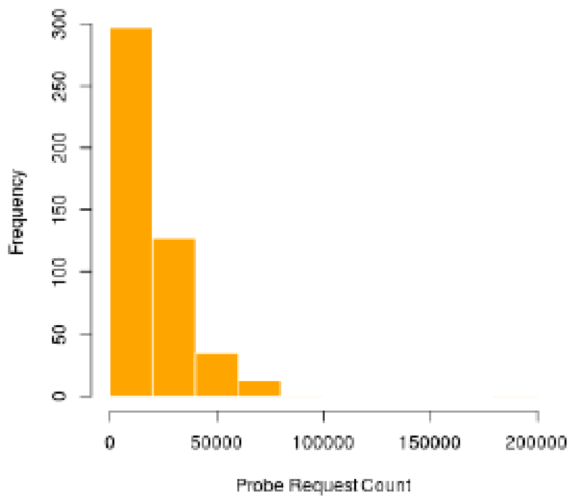

Figure 2. The distribution of sensors ranked based on the total number of counts captured.

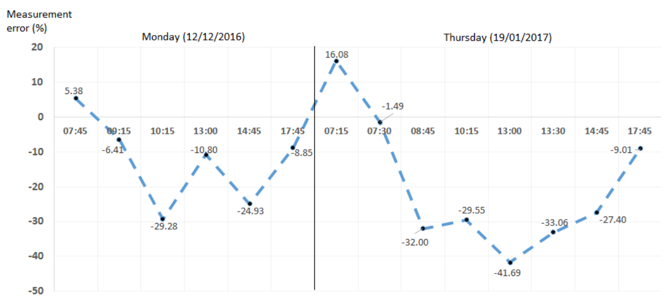

Figure 3. Temporal variation of residual error after initial data cleaning recorded by the sample sensor in London.

Results demonstrated a discrepancy between the number of mobile devices detected by the sensor and manual footfall counts. The researchers suggest that these errors may be divided into two groups: those leading to over-counting and those leading to under-counting. Causes of over-counting ranged from the presence of recurring smartphones, non-smartphone devices with Wi-Fi capability such as printers, or passing traffic. On the other hand, under-counting factors could be attributed the presence of specific physical obstacles which distract the signal or the proportion of people who turn off the Wi-Fi on their phone.

Figure 3 shows an example of measurement errors for one sensor located in London, illustrating that the magnitude of error fluctuated significantly throughout the day. In addition, variations appeared inconsistent between different sensors under consideration. The researchers suggest that further calibration of the data will involve utilising manual counts on multiple occasions and applying average adjustment factors to minimise the observed measurement errors.

Future Directions

The Wi-Fi sensors are currently producing a rich, temporally and spatially granular source of data, which offer promising potential for future applications in human activity modelling. On-going research aims to formulate a more comprehensive and robust calibration method that accounts for both the over and under-counting factors. This will enable reliable modeling of the spatio-temporal variation of pedestrian flows. In addition, understanding the flows in terms of their geodemographic characteristics hopes to aid exploration of more complex classifications of retail areas. This may be achieved through the integration of the temporal distribution of footfall and demographic indicators.

This research hopes to address a significant gap in current data availability regarding footfall, which is anticipated to be of importance for a wide range of applications both commercially and academically.