Trends in urban flows: from Wi-Fi data to pedestrians' route choices

Roberto Murcio[1], *Karlo Lugomer*[2], *Balamurujan Soundararaj*[2].

(1) Centre for Advanced Spatial Analysis, UCL. London, UK, r.murcio@ucl.ac.uk.

(2) Consumer Data Research Centre, Geography, UCL, London, UK.

The accurate estimation of human activity in cities is one of the first steps towards understanding the structure of the urban environment and estimating them with confidence is crucial for decision-making in numerous applications. Using data generated by a network of sensors (Wi-Fi enabled devices) installed at around 1000 locations across UK that capture a series of public signals - known as probe request frames, we constructed a daily footfall (FF) signal by location (Figure 1.a) and derived a series of results to characterise each area in terms of two features: Core activities and Pedestrian flows / route complexity between two locations located at 5 min. walking distance from each other.

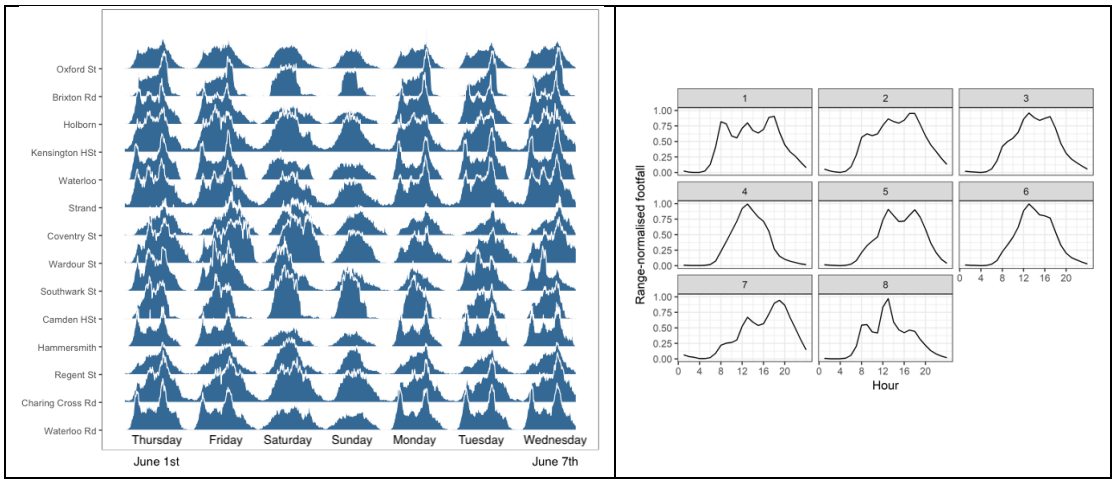

Figure 1. a) A week of measures at fourteen London locations where it is clear that different locations produce different FF signals; b) Weekday clustering of FF signals. Again, there's a clear distinction between locations, depending on their particular core activities.

Core Activities. We produce the typical weekday/weekend signal for each location and then cluster them (using a Dynamic Time Warping approach) and found a core of eight canonical FF signals that tell us a story about the activities around these locations (Figure 1.b). For example, cluster 1 is a typical job location (FF peaks at morning/lunch/afternoon). **Pedestrians Flows and route complexity**. Without tracking people, we provided a measure (based on the concept of Local Transfer Entropy) to determinate the most likely direction for the flow of people between locations, relating the number of pedestrians reaching location S1from S2 (and vice versa) by measuring the uncertainty between both signals. We noticed that for ~40% of the pair's locations, the flow profiles were the same between S1↔S2 or meaningless (LTE≤0). To complete our analysis, we performed a semantic exploration over the walking directions (from the Google API) from S1 to S2 and then derived a route complexity score S, based on an assigned weight to each common word W_i and its associated distance d_i: S = $\sum_{i=1..12} W_i/d_i$. This S, in conjunction with the LTE flows, explains the different local dynamics observed in areas with the same FF profiles and with the same walking distance between them. In conclusion, in these nearby locations, street configuration and core activities (and not distance), are the drivers behind the FF footprints observed.