Using Mobile Phone Traces to Understand Activity and Mobility in Dakar, Senegal

Ed Manley, Adam Dennett and Michael Batty

Centre for Advanced Spatial Analysis (CASA), University College London
Gower Street, London, United Kingdom

Summary

With the emergence of mobile phone trace datasets, new opportunities have arisen for improving the understanding large-scale mobility behaviours. The potential impact of these insight derived from these data is no more significant than in the developing country context, where existing data collection infrastructure is limited or non-existent. In this research, mobile phone data for Dakar, Senegal is used to better understand urban activity and mobility dynamics. To achieve this, a clustering method is introduced that extracts the spatial distribution, and the temporal characteristics, of the activities of individual mobile phone users. With this classification of individual locations of activity, citywide trends in activity and mobility over time are derived. The paper concludes in discussing the potential and limitations of this approach, and the outlook for associated analyses that employ mobile phone trace data.

Keywords Big Data; Mobile Phone Data; Mobility; Activity; Developing Countries.

1 Introduction

With mobile phone use nearly ubiquitous in both parts of the world, interest is gathering around the potential for using derivative trace data to better understand human behaviour. The potential impact of these datasets is no greater than in the developing world, where datasets considered standard in many countries are unavailable due to a lack of resources. One significant potential avenue of research relates to establishing patterns of activity and mobility in developing countries.

The increasing use of mobile phone traces for the exploration of activity and mobility patterns has led to the development of a range of novel methodologies. In one study located in Tallinn, Estonia, strong similarities were observed in the temporal dynamics in the activities of different individuals (Ahas et al., 2010). Other applications have involved the automatic classification of land use from mobile phone activity patterns (Toole et al. 2012) and identification of trip purpose (Phithakkitnukoon et al, 2010). Network analyses applied to communication interactions derived from mobile phone in identifying the location of ethnic boundaries in Ivory Coast (Amini et al, 2014).

In this paper, urban activity and mobility dynamics are derived through exploration of individual behaviours, capture within mobile phone traces recorded in Dakar, Senegal. The paper introduces a method for the identification of locations of individual activity over different times of day. On an aggregate basis, the identification of activity clusters enables an analysis population density, possible land use, and mobility within between different areas of the city. The paper first outlines the data used in this study, and the method used for the identification of activity locations, and then moves on to exploring citywide trends in activity and mobility.

2 Study Area and Data Description

The study is located in Dakar, the capital city of Senegal, West Africa. The majority of the city is located on a peninsula on the Pacific Ocean coastline, and contains a population of 1.08 million inhabitants. Like many cities in developing countries, data collection facilities are not as established nor comprehensive as elsewhere. Although the last census was carried out in 2013, only limited data has since been released.

The study utilises mobile phone usage data provided as part of the 2014 Orange Data for Development challenge. The scheme made multiple mobile phone datasets available for research purposes, recorded for selected Orange mobile phone users in Senegal during the course of 2013.

The dataset used for this study describes the mobile phone usage patterns of individual users during three-week periods over the course of the year. Within this dataset, whenever individuals used their mobile phone, the cell tower through which the call or text message was sent was recorded. The cell tower used for transmission is usually that nearest to the individual, unless they are crossing between zones covered by two towers. Within the area of Dakar, indicated in Figure 1, there are 435 cell towers, with an average minimum distance to the nearest next cell tower 404 metres.

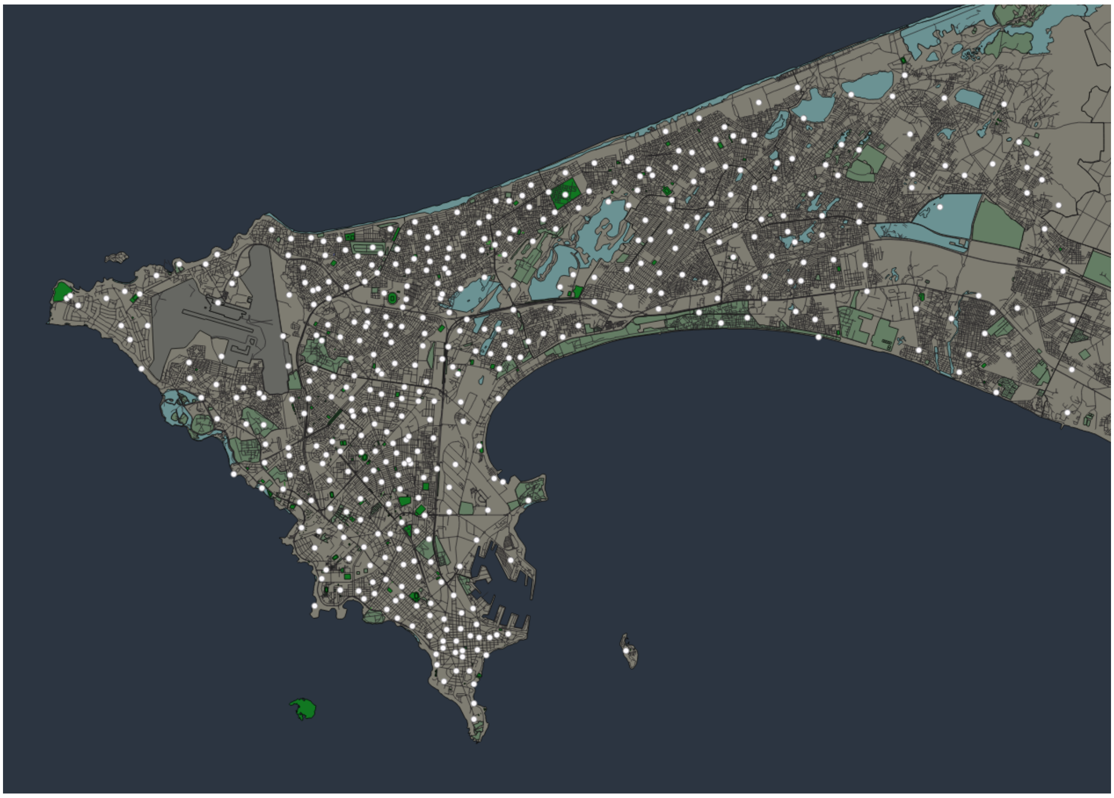

Figure 1: The Dakar study area, showing the location of the 435 cell towers.

3 Identifying User Activity Locations

An analysis of individual mobile phone usage behaviour was undertaken in order to identify areas of *regular activity*. The location – a cell tower in this case – at which an individual is observed on a regular basis can be assumed to hold a strong significance to that individual. The timeframe within which these activities take place may be indicative of the type of activity being undertaken by the individual.

The definition of regularity in this case is defined as the demonstrable presence of an individual at a location, within a given timeframe, on at least 30% of the days they are observed. The 30% threshold ensures the location is visited relatively often and across multiple days, and accounts for the possibility that individuals will not use their device at that location each time they visit.

Locations of activity are identified through clustering the time points at which the individual appears at each unique location. For this purpose, DBSCAN is used, ensuring flexibility in cluster size. The DBSCAN algorithm is specified to only cluster points that fall within 60 minutes of an existing cluster, with all other instances classed as outliers. The result is a set of locations for each individual at which they are observed at around the same time on a regular basis. Each location cluster is defined with a minimum, maximum and mean time.

4 Spatial and Spatiotemporal Variation in Activity

Through identification of individual activity locations over time, one can aggregate to draw out population-level indicators of spatial and spatiotemporal variation in activity across Dakar. This nature of these trends may provide some indication of the types of activity being undertaken at each location.

The first, simplest approach is to extract the sum number of people shown to be visiting each location. This demonstrates the volume of individuals dwelling in each area across the city. This distribution is shown in Figure 2, and indicates a higher proportion of this behaviour around the southern and western areas of the city.

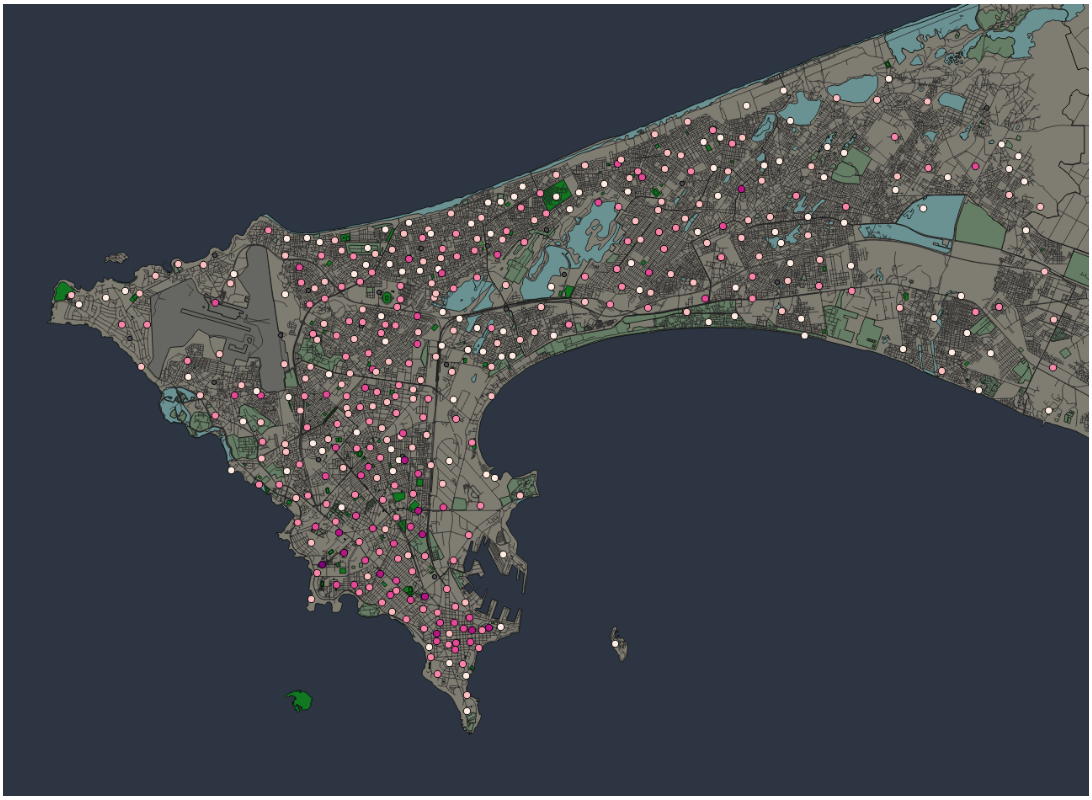

Figure 2: Number of individuals visiting each location on regular basis.

Moving onto temporal variation, a useful initial indicator is to extract the mean time at which individuals, observed regularly, are active at each location. Across all cell towers, the mean regular activity time is 16:13, indicating an evening bias towards phone activity. Spatial variation in mean regular activity time, as shown in Figure 3, should be judged relative to the overall mean.

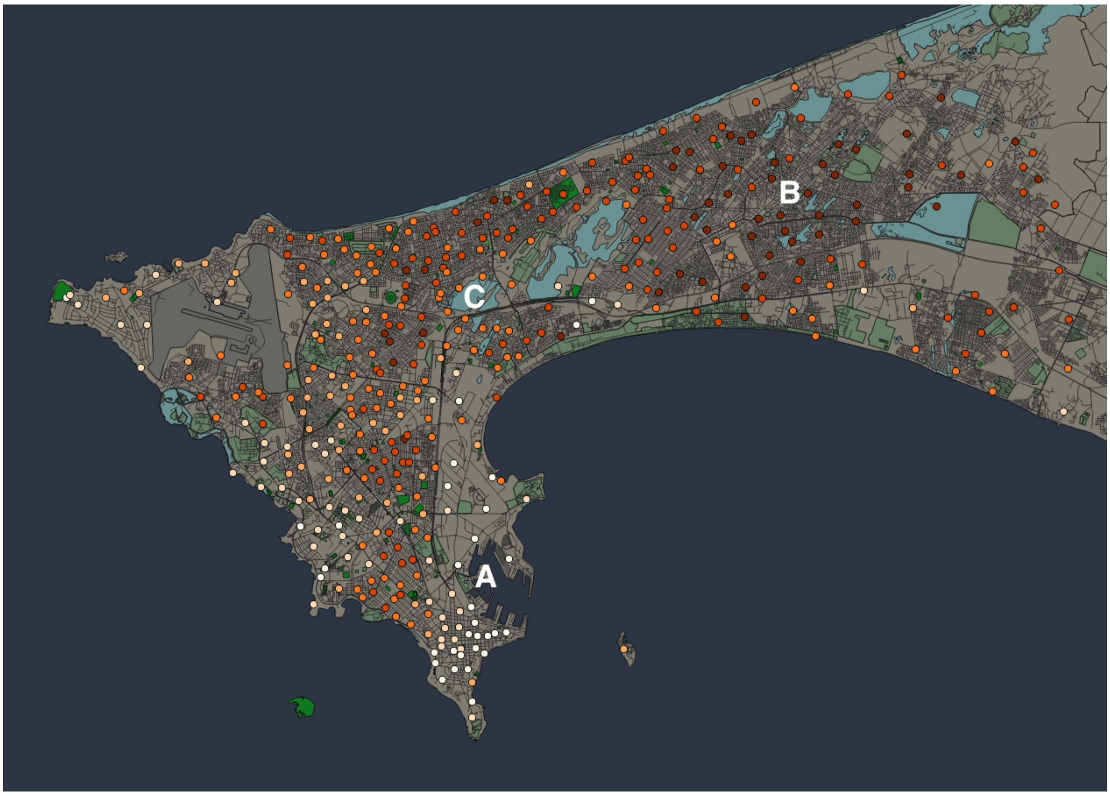

Figure 3: Mean regular activity time for each cell tower in Dakar.

As indicated in Figure 3, regions of earlier and later regular activity are present. Most notably in south-western parts of the city (in the area marked A), the mean time dips to around 2pm, the lowest being cell tower ID 224, with a mean time of 13:32. Conversely, in the north-west mean regular activity times are later (area marked B), with far western areas reaching 17:30, the latest mean is found at cell tower 384, where the mean time is 17:45. Likewise, the area marked C also reflects a later mean activity time than the rest of the city.

Average time, however, may mask the presence of conflicting early and late peaks in activity at each location. An alterative viewpoint is to explore the proportion of regular users present at a location within a specific time period, relative to all other times of day. Two time periods are examined – daytime, defined as 10am to 4pm, and evening, defined here as 6pm to 12am – and are shown in Figures 4 and 5.

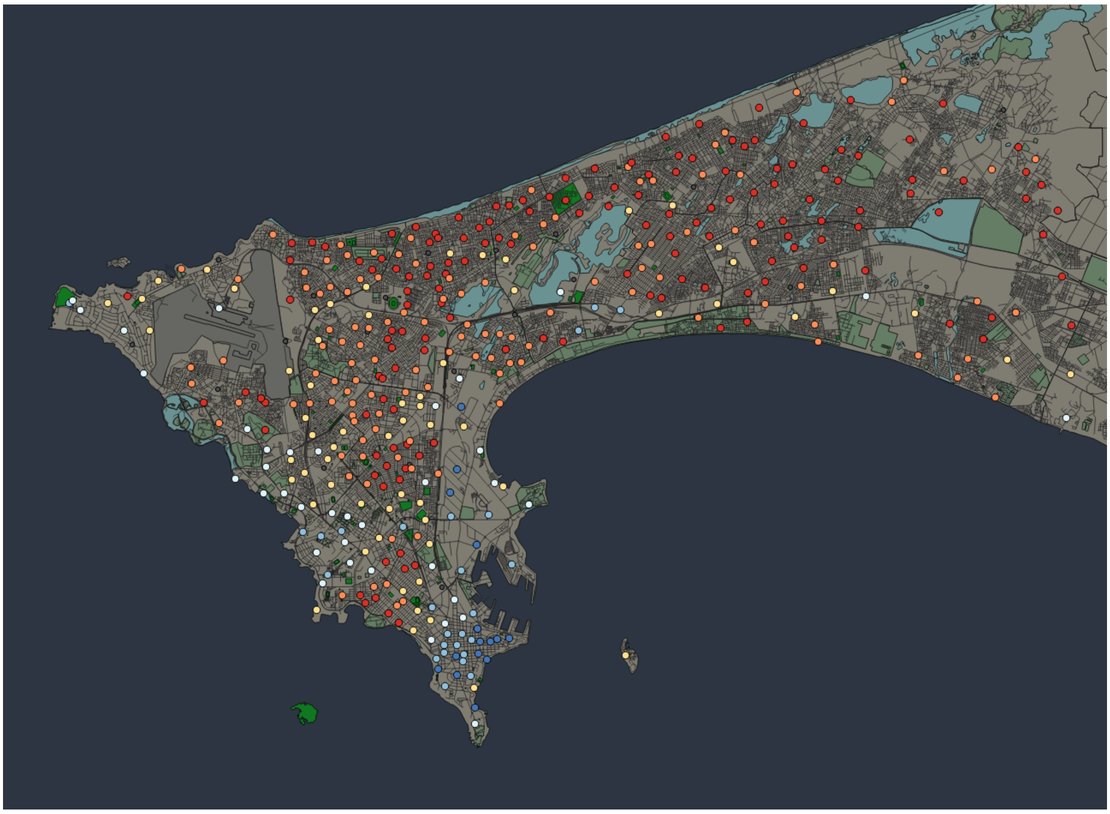

Figure 4: Spatial variation in regular activity during daytime hours (10am to 4pm).

The maps provide a little more insight into the spatiotemporal variation in activity across Dakar. It is clear that the southern tip of Dakar is predominantly active during the daytime, where many of the cell towers see 60-70% of their users during the daytime. During the evening, activity in the southern areas drops significantly, dispersing to the northern and western areas. In these areas, 50-60% of all regular users are observed during the evening hours.

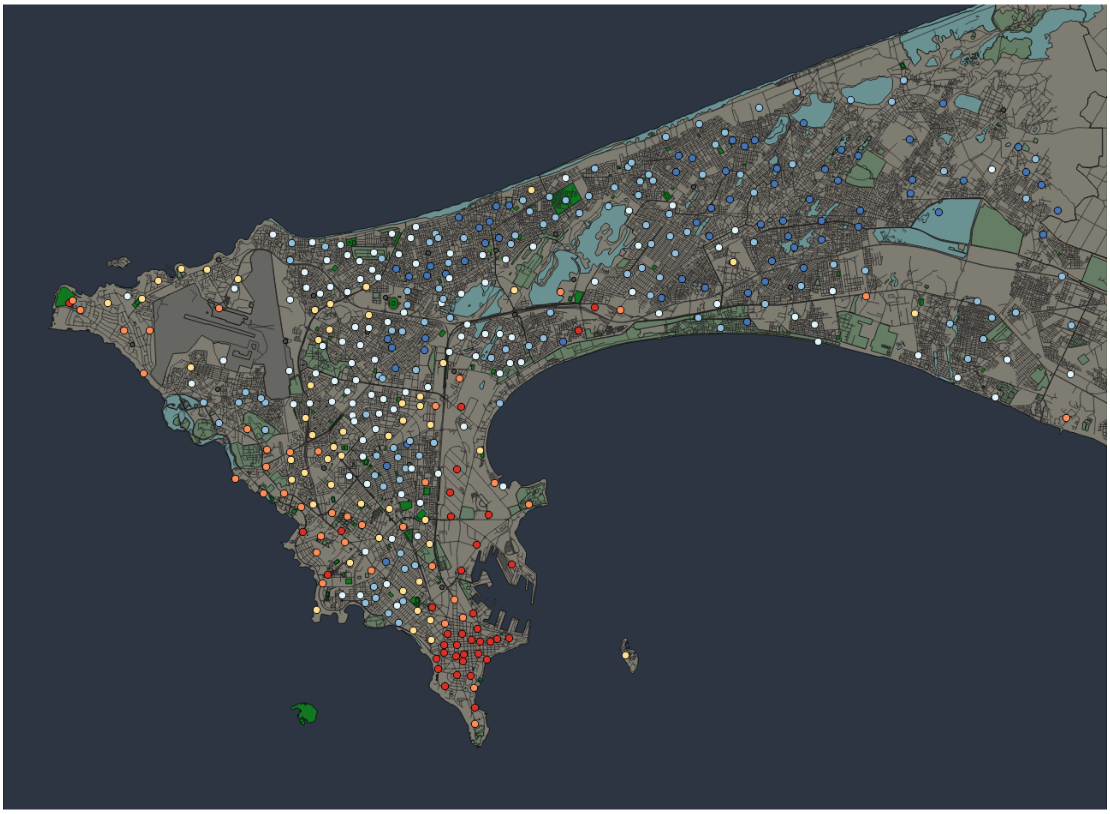

Figure 5: Spatial variation in regular activity during evening hours (6pm to 12am).

5 Mobility Indicators

In addition to identifying location of activity, it is possible to explore the degree of connectivity between cell towers, as demonstrated by the movement of users. This is particularly interesting with respect to identifying movements between ‘home’ regions and other areas of the city.

In exploring this process, a definition of ‘home’ is required. Once more the locations of regular activity are used. In this case, however, the time periods are adjusted to incorporate only individuals observed regularly at locations between 9pm and 5am. These rules identify regular night time locations for 708463 individual users. From this point we assume these are the home locations of these users.

With the definition of night time location assigned to each user, the locations to which these users travel away from this area is identified. To counter the potential that users may be tracked travelling to these locations, only those cell towers at which individuals are shown to dwell (as identified through clustering) are included here. These dwell locations are not required to fit within a certain time period, nor do they need be observed on multiple occasions (as specified in identifying home locations). However, only cell towers outside of a 500m radius of the home location are included, in order to ignore very local movements.

The result is an indication of the mobility of all individuals living at each location, and the locations to which they travel. Through the construction of these networks of mobility for each location, initial insight can be gained by extracting the most popular destinations for all travellers. Using these data, Figure 6 shows the balance between the number of individuals visiting a location and number indicated to be living in that area.

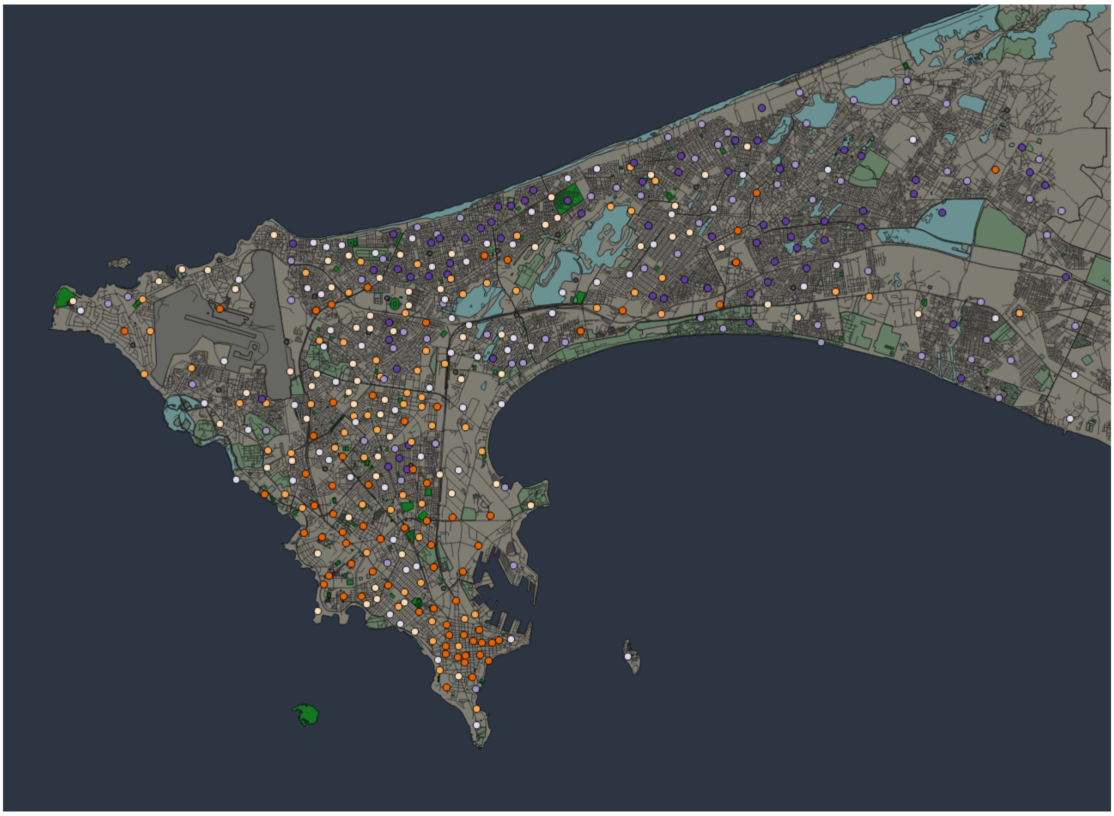

Figure 6: Differences in individual living at each location against the number regularly visiting each location

As can be seen from these results, many of the southern areas of the city are popular destination locations relative to the number of people living in these areas. This trend appears to align with the earlier identified patterns, whereby these same areas were more often visited during daytime than the evening. The highest imbalance between residents to visitors is found at cell tower 77 on the western coast. Maps of Dakar suggest that this tower is positioned on the University of Dakar campus, and so these results would align with expectations.

Another alternative measure than can be drawn from exploring mobility is the mean total mobility of individuals originated at each location. These measures provide an indication of the degree of mobility of individuals living at each location. The results for each location are shown in Figure 7.

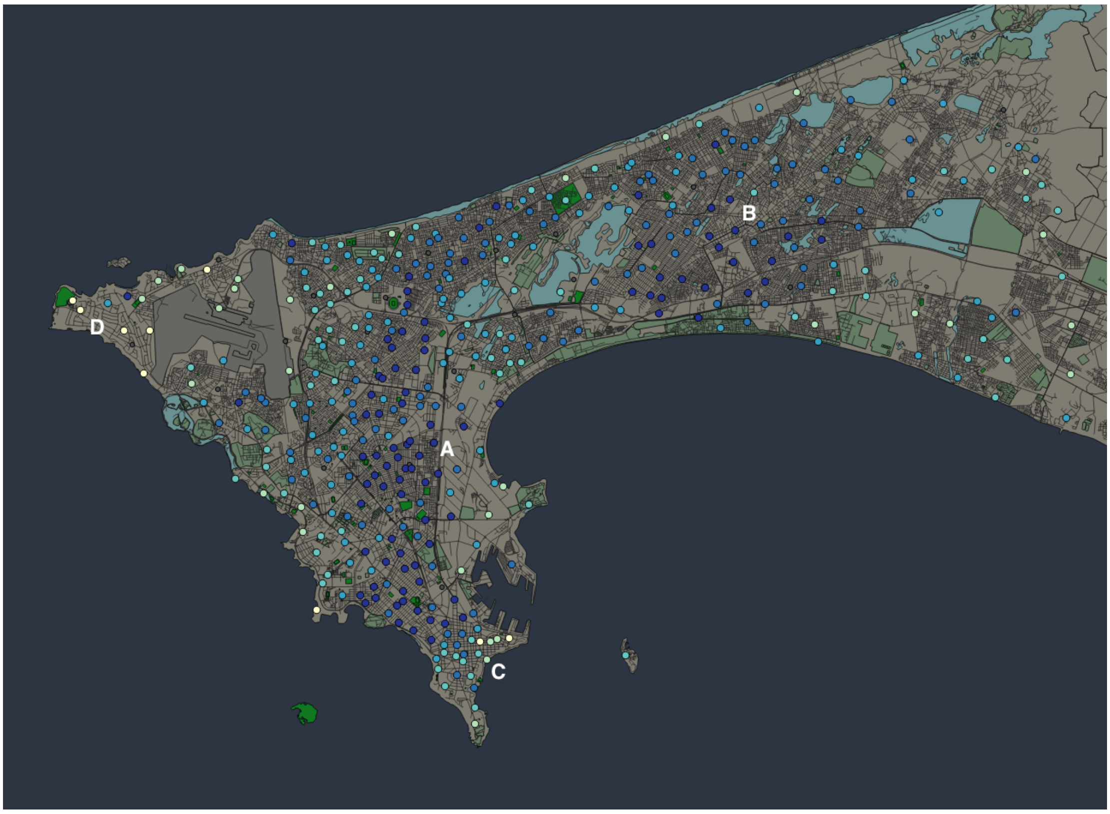

Figure 7: Spatial variation in average total mobility of inhabitants at each cell tower.

The results indicate considerable variation in average total mobility across the city. In areas such as A and B, as indicated in Figure 7, total mobility is relatively low on average, at between 5km and 8km during the time period. This increases to upwards of 20km on average in the regions marked C and D. This latter trend is interesting as these areas are equally those of high attraction to visitors. It may be concluded that the individuals living in these attractive areas are more affluent, and there have more opportunity to travel than others in the outer suburbs.

6 Discussion and Conclusions

This paper has demonstrated how mobile phone usage data can be used to better understand activity and mobility patterns within a developing world context. Through the identification of individual locations of regularity – including those locations that may be assumed to be home locations – it has been possible to generate indications of activity for individuals residing at each cell tower. While these initial explorations currently lack validation, there are some clear, promising trends that should be further investigated.

Despite the promise of these initial findings, some notes of caution should be highlighted. The identification of user home locations is clearly problematic, as the repeated visit of a location at night does not necessarily indicate that that location is that person's home. Furthermore, within this dataset, location is only captured when the phone is used. This biases the location data for an individual towards places where they stop for a reason amount of time. It furthermore means it is not possible to fully capture the complete set of trips undertaken by an individual. As such, the macroscopic picture provided at this stage is likely the most appropriate use for the data. Finally, there are likely to be inherent demographic and usage biases within the dataset too. The requirement that an individual has a phone and uses it with some regularity requires that that person must have some free income. This likely leads to a more affluent study group than may be present within the wider population.

These limitations aside, the initial results described here indicate the potential for further exploration with these datasets. With little quantitative understanding of activity and mobility within developing world cities, mobile phone data may reflect a useful indicator for planners looking to better understand place and transportation needs.

References

Ahas, Rein, et al. "Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: case study with mobile positioning data." *Transportation Research Part C: Emerging Technologies* 18.1 (2010): 45-54.

Amini, Alexander, et al. "The impact of social segregation on human mobility in developing and industrialized regions." EPJ Data Science 3.1 (2014): 6.

Phithakkitnukoon, Santi, et al. "Activity-aware map: Identifying human daily activity pattern using mobile phone data." *Human Behavior Understanding*. Springer Berlin Heidelberg, 2010. 14-25.

Toole, Jameson L., et al. "Inferring land use from mobile phone activity." *Proceedings of the ACM SIGKDD international workshop on urban computing*. ACM, 2012.