ISSN: 2513-2512
R4TLI Conference Proceedings 2016
Paper Reference: *R4TLI-A23*

An Assessment of Mobile Network Big Data-based Insights for Transport Planning in Sri Lanka

M. K. D. T. Maldeniya
LIRNE*asia*

Sriganesh Lokanathan
LIRNE*asia*

Amal S. Kumarage
University of Moratuwa, Sri Lanka

1. Introduction

With increasing urbanisation and the accompanying trends in human mobility and traffic patterns, Sri Lanka is entering an era that demands greater flexibility and responsiveness from the local transport planning apparatus. The recent global data revolution has seen large transaction-generated datasets being leveraged to understand large-scale patterns in human behaviour. Transaction-generated mobile phone network data is one of the few forms of data with sufficient coverage of the population of Sri Lanka that provides an opportunity to apply similar techniques. The paper assesses the potential of this data for delivering insights that can enhance the traditional transport planning process in Sri Lanka through a series of analyses that quantifies and interprets human movement in urban regions.

2. Literature Survey

CDRs have been used in numerous analyses that can provide insights fitting in to different aspects of the existing transport planning process in Sri Lanka.

These include techniques for identifying home and work locations for subscribers which can be used to understand commuting behaviour [2] as well as more advanced techniques that derive estimations of more general mobility behaviour as Origin-Destination (OD) matrices [1]. Similar techniques have been adapted and evaluated in the Sri Lankan context as well [3], [4].

3. Data

The analyses were based on anonymised CDRs for thirteen contiguous months of activity (2012 - 2013) for nearly 10 million SIMs from Sri Lanka. This represents about 50% -60% of the mobile phone subscriber base of Sri Lanka. Sri Lanka had a

mobile penetration level above 90% at the time. The data provides greater spatial resolution in urban regions and lower resolution in rural regions as a result of the level of infrastructure operators have in place to serve regions based on demand. In addition, data from transportation surveys were used for validation.

4. Methods

4.1 Estimating Commuting Behaviour

Each CDR has location and time attributes. These attributes are leveraged to identify the most frequent locations that a subscriber was observed at during home (9pm – 5am) and work hours (10am – 3pm). For a given subscriber, commuting behaviour is defined as travel from the home location to the work location. Individual commuting trips are aggregated to derive high resolution commuting OD matrices.

4.2 Estimating Different Forms of Aggregate Travel Behaviour

The spatio-temporal information present in mobile network data can be used to extract different forms of travel, based on interpretation. Three such perspectives that define a trip (the basic unit of travel) differently were considered. The individual trips extracted based on these three approaches were aggregated initially at the mobile network base station level to estimate O-D matrices.

1. Stay-based approach: Locations where a subscriber has stayed for more than a minimum amount of time (10mins) are identified. These locations and the time period are considered as 'stays'. A trip is recorded for each pair of consecutive stays.
2. Transient trip approach: Each pair of consecutive non-identical locations in a subscriber's CDR for a day, is considered as the origin and destination of a trip. Trips identified in this manner are likely to be transient in that they may represent intermediate steps in actual trips. This approach is sensitive to noise in the data in the form of 'false' displacements observed due to expansion and shrinkage of base station coverage regions for load balancing purposes.
3. Regular travel approach: First sequences of two locations occurring more often than a threshold frequency (10%) during each day a person is observed in the data are identified. These pairs of locations are considered origins and destinations of frequent trips. The likelihood of making each of the frequent trips during different periods of each day of week is then estimated. Therefore, OD estimates that are derived from this approach are probabilistic.

Considering a of an individual for a where
for a day

(1)

5. Results and Discussion

The OD estimates derived based on different approaches were compared with traditional survey-based forecast at the DSD level for the Western province

Table 1 - Summary of Linear Models for MNBD OD Matrices with Traditional Forecast

Method	Intercept	Estimated trips	R^2
Stay based	35,516***	76.41***	0.819
Transient trip	25,460**	2.66***	0.903
Regular travel	14,770·	1.16***	0.909

6. Conclusions

MNBD based techniques can provide insights with wide ranging applications in the traditional transport planning process in Sri Lanka.

Acknowledgement

This work was carried out with the aid of a grant from the International Development Research Centre, Canada and the Department for International Development UK

References

1. Calabrese, F. et al. Estimating Origin-Destination flows using opportunistically collected mobile phone location data from one million users in Boston Metropolitan Area.
2. Isaacman, S. et al. 2011. Identifying important places in people's lives from cellular network data. *Pervasive …*. (2011), 1–18.
3. Lokanathan, S. et al. 2014. *Using Mobile Network Big Data for Informing Transportation and Urban Planning in Colombo*.
4. Maldeniya, D. et al. 2015. Where did you come from ? Where did you go ? Robust policy relevant evidence from mobile network big data.

Keywords*: Urban transport planning, Big Data, Call Detail Record*