

Research Article

The Potential of Mobile Network Big Data as a Tool in Colombo's Transportation and Urban Planning

Sriganesh Lokanathan

LIRNEasia, Sri Lanka

Gabriel E. Kreindler

LIRNEasia, Sri Lanka

MIT, USA

University of Moratuwa, Sri Lanka

N. H. Nisansa de Silva

LIRNEasia, Sri Lanka

Yuhei Miyauchi

LIRNEasia, Sri Lanka

MIT, USA

**Dedunu Dhananjaya
Rohan Samarajiva**

LIRNEasia, Sri Lanka

Abstract

Rapid urban population growth is straining transportation systems. A big data–centric approach to transportation management is already a reality in many developed economies, with transportation systems being fed a large quantity of sensor data. Developing countries, by contrast, rely heavily on infrequent and expensive surveys. With mobile phone use becoming ubiquitous, even in developing countries, there is potential to leverage data from citizens' mobile phone use for transportation planning. Such data can allow planners to produce insights quickly, without waiting for the proliferation of sensors. Using mobile network big data (MNBD) from Sri Lanka, our article explores this potential, producing mobility-related insights for the capital city of Colombo. MNBD-based insights cannot produce all the insights needed, but the high frequency and spatial resolution of the insights that they do provide can complement existing infrequent surveys. For resource-constrained developing economies, even an incremental advance in their ability to produce timely and actionable knowledge can improve existing transportation and urban planning. However, more research will be required before such techniques can be mainstreamed.

To cite this article: Lokanathan, S., Kreindler, G. E., de Silva, N. H. N., Miyauchi, Y., Dhananjaya, D., & Samarajiva, R. (2016). The potential of mobile network big data as a tool in Colombo's transportation and urban planning. *Information Technologies & International Development* [Special Issue], 12(2), 63–73.

1. Introduction

Transportation systems in developing economies are straining under rapid urban population growth. Road congestion is a frequent reality in many cities, particularly on the main thoroughfares that feed the city's day-time working population. Data are needed to identify the choke points and prioritize additions and enhancements. A big data–centric approach to transportation management, based on sensor data, is already a reality in many developed economies, with transportation systems being fed by large quantities of sensor data such as loop detectors, axle counters, parking occupancy monitors, CCTV, integrated public transport card readers, and GPS data from phones as well as from public and private transport systems (Amini, Bouillet, Calabrese, Gasparini, & Verscheure, 2011).

Developing countries, however, rely more on traditional forms of data collection such as questionnaires. Such survey-based methods, administered at peak hours, can be costly in terms of personnel, processing, and traffic disruption. This rather clumsy questionnaire method was used in the Sri Lankan capital of Colombo in 2013. Other less intrusive methods (e.g., automatic traffic recorders) are unable to yield important information such as routes taken and parking.

Mobile network big data (MNBD) have enormous potential to contribute to traffic planning and complement these infrequent surveys. Because the data streams are continuously flowing, the effects of changes in traffic channels—e.g., one-way traffic plans and new roads—can potentially be easily tracked. Although additional costs for data storage may be involved, base transceiver station (BTS) hand-off data can even serve as trackers of traffic speed and disruptions. Moreover, as the proportion of GPS-enabled smartphones increases, it may be possible to achieve the same traffic-tracking objectives with smaller samples, i.e., without collecting masses of BTS hand-off data.

The primary question addressed by the research outlined in this article was: To what degree can mobile network big data inform transportation planning for the city of Colombo? To address this question, we attempted to use MNBD from Sri Lanka to understand where the daytime commuting population of Colombo comes from, that is, to create origin–destination (OD) matrices that explicate the flow of commuters between different geographic areas.

2. Literature Review: Data-Driven Transportation Management

Depending on the network's sophistication, mobile network data can capture a range of location variables. The data can be broadly classified into two types: passive positioning data and active positioning data.

Passive positioning data are automatically generated by the mobile network and captured in the network's logs for billing purposes and also for network management (Ahas, Aasa, Silm, & Tiru, 2010). Every time a subscriber uses her or his mobile phone to make or receive a call, to send or receive an SMS/MMS, or to access the Internet, a BTS generates a record of that event. These records are collectively termed *call detail records* (CDRs). In addition to the identifiers of the parties involved in the event, the event's date, time, and duration are stored. Each record also includes the cell ID (the ID of the antenna), which in turn has geolocation and antenna orientation information (i.e., an azimuth). These passive positioning data based on cell IDs are inexpensive when compared to active positioning data, but with the tradeoff that because they are at the level of network cells, they are less precise than active positioning data.

Active positioning data are location data captured via specifically initiated network queries to locate handsets, using either network-based and/or handset-based positioning methods. Location data from GPS or GPS-augmented network triangulation can also be considered active positioning data (Ahas et al., 2010). Use of these methods has come about due to national regulations (e.g., regulations requiring operators to capture high-precision location data for security reasons) or due to the existence of a business case for providing location-based services. However, not all mobile operators generate continuous active positioning data for all their subscribers, and even fewer operators store the data.

Active and passive positioning data have been used with great effectiveness in transportation, helping to measure and model people's movements in both developed and (to a much lesser extent) developing economies. Trip-based OD matrices (traditionally derived through infrequent surveys) have been created using mobile

network big data in South Korea (Yoo, Chon, Kang, & Kim, 2005), Spain (Caceres, Wideberg, & Benitez, 2007), and the United States (Calabrese, Di Lorenzo, Liu, & Ratti, 2011), among others. Researchers have also used MNBD to conduct activity-based modeling of people’s movements (Isaacman et al., 2011), to infer transportation modes (Wang, Calabrese, Di Lorenzo, & Ratti, 2010), and to map traffic flow (Wu et al., 2013).

Even the least-developed mobile network infrastructure generates passive positioning data in the operator logs. Passive positioning cell ID data provide the least spatial resolution. Despite this there has been much recent work using such data in transportation planning. For example, IBM researchers have used passive positioning CDR data from the operator Orange to map citizens’ travel routes and show how data-driven insights could be used to improve planning and management of transportation services in Abidjan, the largest city in Côte d’Ivoire (Berlingiero et al., 2013). The IBM study suggests that overall travel time could be reduced by 10% through network optimization (in this case by extending one bus route and adding four new ones), thus offering a partial solution to the city’s congestion problems. Similar work using passive mobile positioning data to inform transportation planning and management is being conducted in other countries, both developed and developing.

3. The Dataset

Our study used four months of national passive positioning CDR data of voice calls for between 5–10 million SIMs from a Sri Lankan mobile operator.¹ The data were completely pseudonymized by the operator—i.e., a unique computer-generated identifier replaced each phone number—and we were not provided with any mapping information linking the phone numbers and the identifiers.

Each CDR corresponded to a particular subscriber of the operator and a CDR was created each time a subscriber originated or received a call. In the case of an *on-network call* (when both parties to the call were subscribers of the same mobile network), two records were generated, one for each party.

Each record contained the following data:

- call direction: a code to denote if the record was an incoming or outgoing call;
- subscriber identifier: pseudonymized identifier for the subscriber in question;
- identifier of the other party: pseudonymized identifier for the other party to the call;
- cell ID: identification of the antenna the subscriber was connected to at the time of the call;
- date and time the call was initiated; and
- duration of the call.

4. Data Analysis and Findings

Data analyses were done by assigning each SIM in the dataset a unique home and work location at the level of the divisional secretarial division (DSD), which is Sri Lanka’s third subnational administrative level (after provinces and districts). Based on the home and work assignments, it was possible to find regular inter-DSD mobility patterns, which form the basis of the analyses conducted in this article.

4.1 Identifying Home and Work Locations

Two of the themes that characterize human mobility are “home to work” and “work to home.” While there are many other themes (e.g., “home to school”), our study used only these two basic themes because they are the most common and the ones with the greatest relevance to transportation planning.

We devised a methodology to find the home and work locations (in this case at the DSD level) for each SIM. The first step in finding the home and work DSD locations was to find the time bracket within which an average individual would spend at either location. It was assumed that the time brackets falling outside the home and work brackets would be time that individuals were spending commuting between home and work locations.

1. As per our agreements with the operator, we do not name the operator or give a precise figure for the number of SIMs analyzed.

THE POTENTIAL OF MOBILE NETWORK BIG DATA

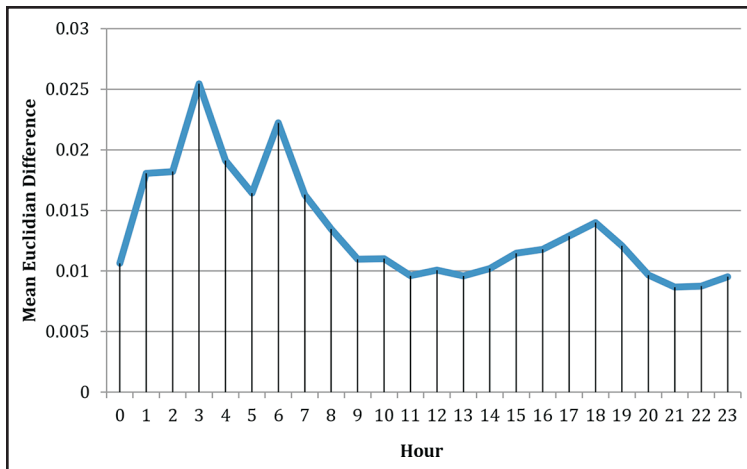


Figure 1. Average SIM mobility during hourly time brackets in a single day.

In the morning an individual typically leaves the home location and commutes to the work location. Once the individual is at the work location, there is typically less drastic movement. At the end of the workday, the individual leaves the work location and commutes to the home location. Once the individual is at the home location, there typically is less drastic movement until the departure to work the next day. Thus, there should be two peaks of human movement, one in the morning when people commute from their homes to their workplaces and another in the evening when people commute from their work-

places to their homes. The statistical valley in between the two peaks would, accordingly, be the work hours, and the times before the first peak and after the second peak would be home hours.

Based on the assumptions described above, we followed a three-step process to extract mobility graphs using data for a normal working day from the dataset:

1. We calculated the average position (latitude, longitude) for each individual SIM for each hour of the day (24 points in all).
2. We obtained the distance traveled by each individual SIM during each hour bracket by calculating the Euclidian (i.e., straightline) distance between two consecutive hour-wise average locations.
3. We obtained an average hour-wise distance measure for all the SIMs that showed movement during a specific hour and plotted the results on a graph.

Ideally this process should have been done for each working day in the dataset and then combined to obtain an average working day profile, but due to time limitations this was not done.

This three-step process revealed the mobility patterns illustrated in Figure 1.

Figure 1 shows, as predicted, two peaks of mobility: a morning peak at 06h00 and an evening peak at 18h00. There is also another peak in the middle of the night at 03h00. We determined this spike was caused by a small number of individuals moving quickly over long distances, presumably mostly truck drivers. (The reason this minority dominated the values during that period was that the majority of individuals were sleeping then, thus, not making or taking calls, resulting in null CDR values for most SIMs.)

To fix the anomaly of the 03h00 peak (and other problems arising from the paucity of data in the middle of the night), it was decided to interpolate the position data for every individual SIM in the dataset. Thus, if an individual SIM had at least two real location values, the missing values between the real values (i.e., the null values generated by hours when there was no voice telephony activity via the SIM) were interpolated using the real values. Using the interpolated individual data and following the same three-step process described above, the average mobility of a SIM during each hour was calculated, as illustrated in Figure 2.

Based on this second mobility graph, we concluded that the period 10h00–15h00 should be considered the work hours and the period 21h00–05h00 the home hours. This decision was slightly subjective, but was based on finding these to be the two most significant periods with minimal variation in average mobility.

Our next step was to identify home and work DSDs. To do this we employed a variation of the “tower days” concept (i.e., the sum of the number of days a tower was contacted) proposed by Isaacman et al. (2011), that is, we used the logic behind the tower days concept to generate a “DSD count” concept. Accordingly, assignment of home and work DSDs was carried out via a three-step process, as follows.

1. A DSD was assigned for each individual SIM for each home or work slot for each by choosing the DSD used the most for each of the time slots. Weekends and national holidays were left out in assigning work DSDs. For example, in one 31-day month, 10 days fell on weekends and the second Tuesday was a national holiday, so for that month there were 31 home DSD allocations and 20 work DSD allocations for each SIM.
2. For each SIM, all the potential DSDs obtained for the home time slot over the four-month period were listed along with its associated frequency. This was also done for the SIM for the work time slot.
3. An overall home location for each SIM was assigned to the DSD that occurred most in the frequency table generated in the previous step for the home time slot. The same logic was used with the frequency table associated with the work time slot to assign an overall work location for the SIM.

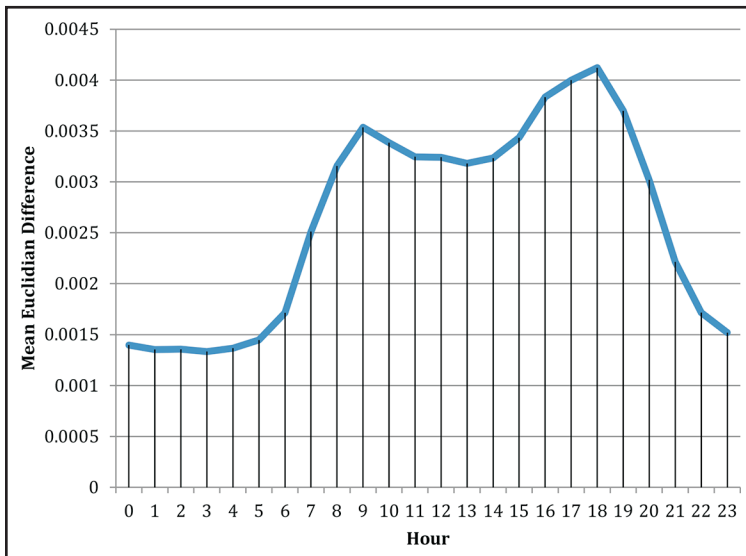


Figure 2. Average SIM mobility during hourly time brackets in a single day using interpolation.

To measure the soundness of our methodology, we compared our results with population data from the 2011 national census (Department of Census and Statistics, 2012). For each DSD we compared the total number of SIMs assigned to that particular DSD as a home location (a proxy for resident subscriber population in that DSD) against the DSD population from the national census. Figure 3 shows a log-log graph in which the x-axis is the log of a DSD’s census population, and the y-axis is the log of the DSD home location count.

The adjusted r-squared value of 0.59, while not low, is not very high. This is explained by the fact that the operator does not have consistent subscriber penetration rates in all the DSDs.

The operator’s mobile telephony base station density and penetration were both found to be highest in the Western Province, the province that includes Colombo. Given our desire to focus this research exercise on Colombo, we decided to calculate the relationship between our home DSD counts and the census DSD data for only the Western Province (Figure 4).

The adjusted r-squared value of 0.79 is much better when considering only the DSDs in the Western Province rather than the DSDs in the whole country (0.59 in the latter case). This was hypothesized to be a reflection of the fact that the Western Province has a more uniform mobile penetration for the operator than the rest of the country.

4.2 Colombo City Findings

A finding from an analysis of our Western Province home DSD counts was that the largest contributor to each DSD working population came from among the same DSD’s inhabitant population, that is, that most people were living and working within the same DSD. Accordingly, to better understand the importance of Colombo city (comprising the Colombo and Thimbirigasyaya DSDs) as a work destination, we examined:

THE POTENTIAL OF MOBILE NETWORK BIG DATA

- the extent of each home DSD's contribution to Colombo's working population (i.e., each home DSD's contribution to work DSD counts in Colombo and Thimbirigasyaya DSDs); and
- the relative rank of each home DSD's contribution to Colombo's working population.

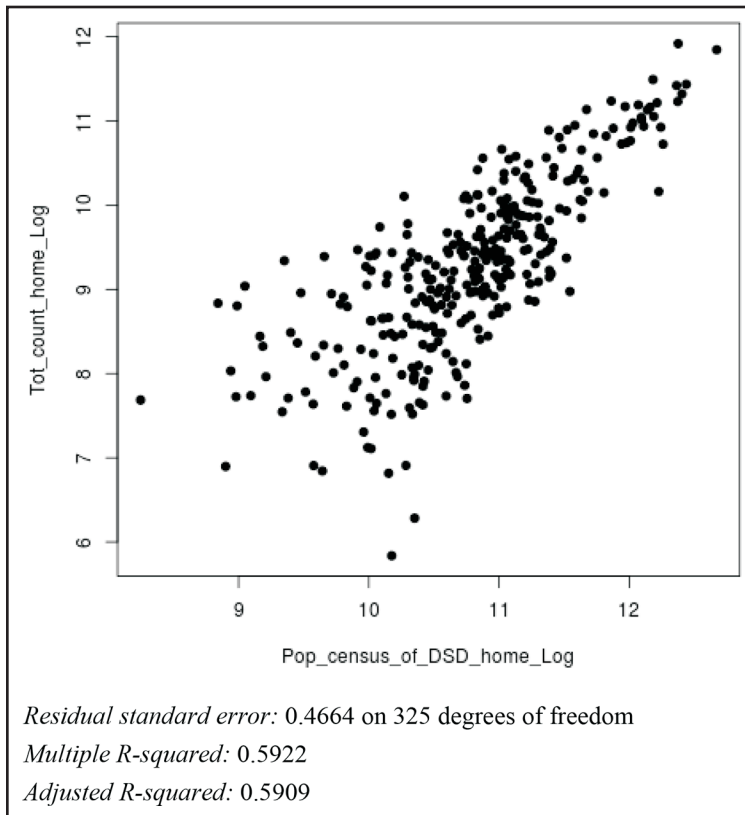


Figure 3. Correlation between derived home DSD population counts and DSD census population in Sri Lanka.

important working location. These source DSDs are generally the DSDs north and northeast of Colombo DSD, not only from Colombo district but also from Gampaha district, which shares its southern boundary with Colombo district.

As Table 3 also shows, the workforce of the other Colombo city DSD, Thimbirigasyaya, also gets a large contribution of its working population (47.1%) from within the same DSD, while the top outside contributors are the neighboring Colombo district DSDs (to the areas south and east of Thimbirigasyaya DSD).

Another interesting picture emerges when we consider the interpolated individual data. Figure 4 compares results from our analysis of mobile network interpolated individual data (Figure 5a) with the results from a costly survey conducted in the Western Province (Figure 5b) that was used to understand mobility and transportation patterns.² The findings are similar.

Figure 5a, based on our analysis of mobile telephony data, depicts the relative population density in Colombo city and its surrounding regions at 13h00 compared to the previous midnight on a weekday in 2013. Thus, the image shows the population movement from the city's outskirts to its center during the day. The yellow-to-red colors show areas where density at 13h00 has increased relative to midnight (i.e., the more red, the

Our analysis found that nearly 47% of Colombo city's working population came from outside the city, that is, just over 53% of SIMs had *both* a home DSD count and a work DSD count in Colombo city (i.e., in Colombo and/or Thimbirigasyaya DSDs). Table 1 shows the top home DSDs for the working population of Colombo city.

When we drilled down further and looked at the two constituent DSDs that make up the city of Colombo, namely Colombo and Thimbirigasyaya, an even clearer mobility picture emerged.

Colombo DSD

For Colombo DSD, our data analysis found that nearly half (49.5%) of its working population also lived in that DSD, with 50.5% living outside Colombo DSD (see Table 2).

Table 2 also shows that among the home DSDs from which Colombo DSD's working population is drawn, the Colombo DSD is either the second or third most

2. The 2013 OMTRANS study was done between 2012–2013 in the Colombo and greater Colombo areas and comprised data from a 40,000-household survey (Ministry of Transport, 2013).

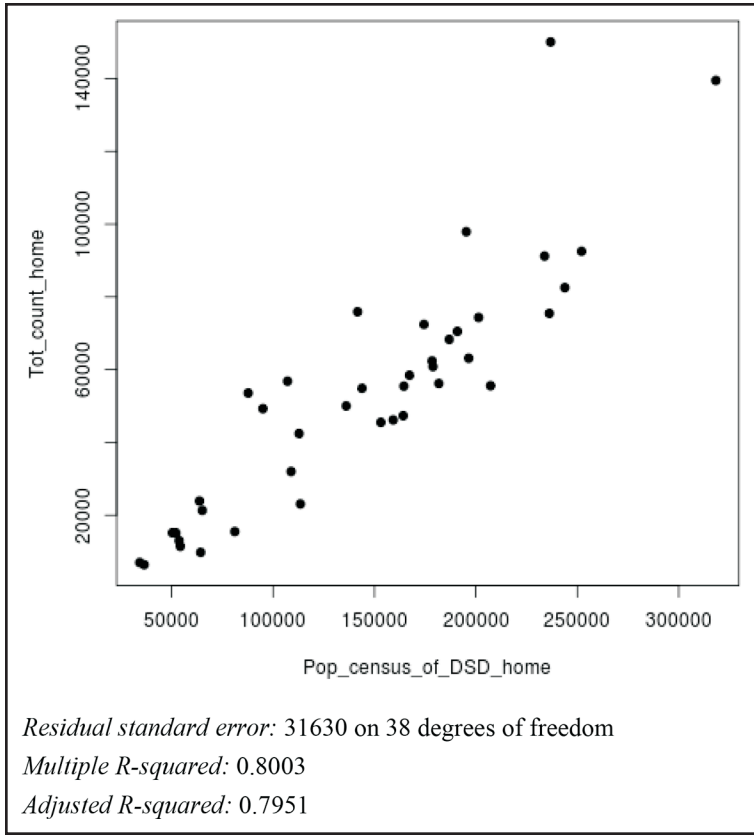


Figure 4. Correlation between derived home DSD population counts and DSD census population in the Western Province.

greater the rise in density), and the blue color depicts areas where density at 13h00 has decreased relative to midnight (i.e., the darker the blue, the greater the loss in density). The clear, noncolored areas are those where the overall density at 13h00 is unchanged relative to midnight.

Figure 5b depicts the major transportation transit points (the darker the location, the more people it attracts, with town centers being circled) as identified by the 2013 COMTRANS study. The figure shows a quasi-identical finding to the image in Figure 5a generated by our analysis of mobile big data. The green arrows running from Figure 5b to Figure 5a show the key points of similarity between the two.

Both images suggest how Colombo city acts a sink, pulling in people from the surrounding regions during the workday. Understanding these movements of people and their sources on a near-real-time basis via mobile network big data could greatly enhance transportation planning.

Table 1. Sources of Colombo City's Working Population.

Rank	Where the working population of Colombo city lives (home DSDs)	% of Colombo's working population
1	Colombo city (Colombo + Thimbirigasyaya DSDs)	53.07%
2	Maharagama	3.66%
3	Kolonnawa	3.52%
4	Kaduwela	3.32%
5	Sri Jayawardanapura Kotte	2.92%
6	Dehiwala	2.60%
7	Kesbewa	2.54%
8	Wattala	2.47%
9	Kelaniya	2.08%
10	Ratmalana	1.95%
11	Moratuwa	1.82%

THE POTENTIAL OF MOBILE NETWORK BIG DATA

Table 2. Top Sources for the Working Population of Colombo DSD.

Source DSD (district)	Source (home) DSD's rank (as source of Colombo DSD workforce)	% of Colombo DSD's working population that comes from Source DSD	Colombo DSD's rank among the source DSD workforce's destinations for work
Colombo (Colombo district)	1	49.5%	1
Thimbirigasyaya (Colombo district)	2	6.0%	2
Kolonnawa (Colombo district)	3	4.0%	2
Wattala (Gampaha district)	4	3.5%	2
Maharagama (Colombo district)	5	2.8%	3
Kaduwela (Colombo district)	6	2.7%	3
Kelaniya (Gampaha district)	7	2.7%	2
Sri Jayawardanapura Kotte (Colombo district)	8	1.9%	3
Biyagama (Gampaha district)	9	1.9%	2
Kesbewa (Colombo district)	10	1.9%	3
Ja-Ela (Gampaha district)	11	1.7%	2

Table 3. Top Sources for the Working Population of Thimbirigasyaya DSD.

Source DSD (district)	Source (home) DSD's rank (as source of DSD Thimbirigasyaya workforce)	% of Thimbirigasyaya DSD's working population that comes from source DSD	Thimbirigasyaya DSD's rank among the source DSD workforce's destinations for work
Thimbirigasyaya (Colombo district)	1	47.1%	1
Maharagama (Colombo district)	2	4.5%	2
Kaduwela (Colombo district)	3	3.9%	2
Sri Jayawardanapura Kotte (Colombo district)	4	3.8%	2
Colombo (Colombo district)	5	3.7%	2
Dehiwala (Colombo district)	6	3.4%	2
Kesbewa (Colombo district)	7	3.2%	2
Kolonnawa (Colombo district)	8	3.1%	3
Ratmalana (Colombo district)	9	2.5%	2
Moratuwa (Colombo district)	10	2.2%	2
Homagama (Colombo district)	11	2.0%	2

5. Challenges and Future Work

We pose several challenges and offer potential future research areas by using mobile network big data for transportation planning in Sri Lanka.

5.1 Travel Motifs

One potential difficulty with taking the mobile big data approach to transportation is its reliance on a generalization of human behavioral patterns to a human behavior theme, holding that people work in the daytime at

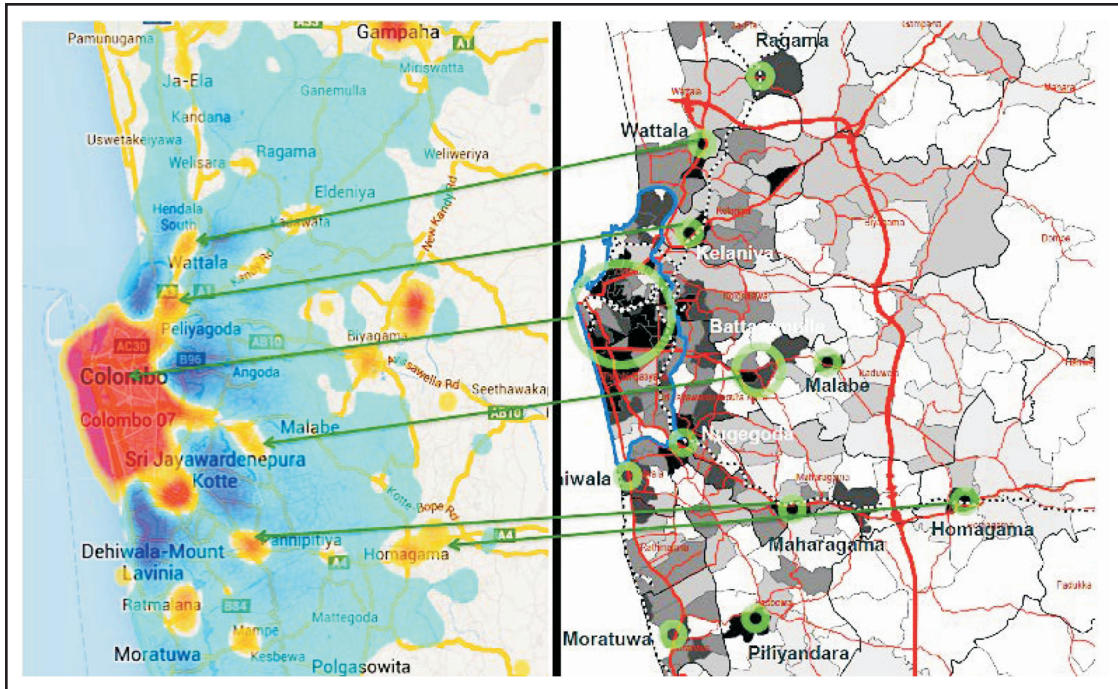


Figure 5a. Result of analysis of mobile network big data.

Figure 5b. Result of analysis of survey data.

an office and sleep at home at night. While our data analysis proved this hypothesis to be true, we do understand there are segments of the working population who do not match this model. The behavioral patterns of two such segments of people would have varying impacts on our intermediate and final results, as we outline in the next two subsections.

5.1.1 Individuals with Inverted Work Shifts

Individuals who stay at home during the day and go to work at night are not captured in our hypothesized theme, but it is interesting to note that this segment would not significantly alter the mathematical model because, just like the individuals who work in the daytime at an office and sleep at home at night, these individuals who would work at nighttime at an office and sleep at home during the daytime will have a night location and day location, and will commute between those two DSDs, thus not altering the overall shape of the mobility graphs. The only potential data problem would be misclassifying these individuals' home locations as work locations and vice versa.

5.1.2 Individuals with Multiple Work Locations

Individuals who have multiple work locations are not captured in our hypothesized theme. It could be proposed that these individuals—if it is assumed they have a uniformly random-distributed number of work locations over uniformly randomly distributed times throughout the day—would generate data points that cause a smoothing effect on the mobility graphs. One possible solution would be to identify these individuals by creating individual mobility graphs and selecting the graphs that have the most deviation from the average mobility graph (provided the stationary individuals were removed as well).

It can be argued that with these multiple-workplace individuals removed, the average mobility graph would have a finer level of detail. However, bear in mind that even without removing such individuals, our average mobility graph displayed the expected shape, suggesting that the proportion of individuals with multiple work locations is comparatively much smaller than the proportion of individuals with a single work location.

5.2 Using DSDs as the Smallest Geographical Unit

Given the large physical size of the DSDs, using them as the smallest geographical unit means that there can be great variation among the mobility patterns of individuals whose home DSD is the same as their work DSD. Some of these individuals will not actually be moving for their jobs, while others will have moderate movement. While this distinction may seem insignificant from a statistical perspective, it can be significant from a policy perspective, since the socioeconomic realities of a nonmoving individual (e.g., a homemaker, a self-employed craftsman) will often be significantly different from those of an individual who moves a moderate distance for work (e.g., a grocery store owner). The way to solve this issue would be to work with geographical units smaller than DSDs so as to generate finer-grained details.

5.3 Possible Extensions and Future Work

Removing individuals who do not match the work-in-the-daytime-at-an-office, sleep-at-home-at-night pattern and analyzing their behavioral patterns would extend our case and provide greater understanding of the socio-economic dynamics of mobility. Then the mobility graph for the remaining work-in-the-daytime-at-an-office, sleep-at-home-at-night people would be more representative of the target population.

Moving to a smaller unit of denomination than the DSDs would be preferable. In this respect we considered working with Grama Niladari Divisions (GNDs), which comprise the DSDs. However, there are many GNDs that lack a BTS within their boundaries. Often a single base station provides coverage for a cluster of adjoining GNDs. And due to the shapes of some GNDs, when a GND does have a base station, that GND only contributes a small portion of the traffic going through the station.

To solve this problem, the use of *voronoi* cells³ to map the base stations was considered; however, *voronoi* cells do not reflect any direct relationship with political or social divisions in the country. We also considered combining the two approaches: using both GNDs and *voronoi* cells to identify base stations. In this proposed solution, first a mapping function would be created between each *voronoi* cell coordinate and the GNDs that the specific *voronoi* cell covers. Then, the *voronoi* cell's mobile traffic would be modeled as a decaying density function that would be allocated to the GNDs based on area overlap.

6. Conclusions

The findings from our research suggest that in developing countries such as Sri Lanka, which have limited sensor networks, active and passive positioning data from mobile network operators (as well as real-time GPS traces from mobile phones) could revolutionize transportation management and improve the efficiency and reliability of transportation systems.

Given that the underlying data constitute the activity of citizens, there are some natural concerns related to privacy. The analysis conducted in this article produces only aggregate insights as the final output, albeit the process of generating those insights starts at an individual level. A discussion of modes of conducting such research while preserving privacy, tradeoffs, and closely related issues such as marginalization and exclusion are beyond the scope of this article.

While we are cognizant of the fact that there are still methodological improvements required, even our preliminary data analysis has proved that mobile network big data show promise as a source of timely and relatively cheap insights for transportation planning. Even after further refinement it is not expected that the insights derived from mobile network big data will cover all the needs of transportation planners in developing economies, nor that they will be perfect, nor that they will remove the need for surveys. For example, the type of MNBD used in this article (i.e., CDR data analyses) will not be well suited to identifying vehicle types or traffic conditions. However, public policy is about the art of the possible, not the art of the perfect. Even an incremental increase in the type and frequency of insights possible in resource-constrained developing economies can potentially facilitate more effective planning processes. ■

3. A *voronoi* cell in this context is the inferred spatial coverage area of a BTS, assuming the BTS is located at the inferred area's center of gravity.

Sriganesh Lokanathan, Senior Research Manager, LIRNEasia. sriganesh@lirneasia.net

Gabriel E. Kreindler, PhD Economics Student, Massachusetts Institute of Technology and Research Fellow, LIRNEasia. gek@mit.edu

N. H. Nisansa de Silva, Researcher, LIRNEasia and Lecturer, Department of Computer Science & Engineering, University of Moratuwa. nisansadds@cse.mrt.ac.lk

Yuhei Miyauchi, PhD Economics Candidate, Massachusetts Institute of Technology and Research Fellow, LIRNEasia. miyauchi@mit.edu

Dedunu Dhananjaya, Software Engineer, LIRNEasia. dedunu@lirneasia.net

Rohan Samarajiva, Founding Chair, LIRNEasia. rohan@lirneasia.net

References

- Ahas, R., Aasa, A., Silm, S., & Tiru, M. (2010). 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), 45–54.
- Amini, L., Bouillet, E., Calabrese, F., Gasparini, L., & Verscheure, O. (2011). Challenges and results in city-scale sensing. *2011 IEEE SENSORS Proceedings*, 59–61. doi:10.1109/ICSENS.2011.6127189
- Berlingerio, M., Calabrese, F., Di Lorenzo, G., Nair, R., Pinelli, F., & Sbodio, M. L. (2013). AllAboard: A system for exploring urban mobility and optimizing public transport using cellphone data. *Mobile Phone Data for Development: Net Mob 2013*. Retrieved from <http://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
- Caceres, N., Wideberg, J. P., & Benitez, F. G. (2007). Deriving origin-destination data from a mobile phone network. *IET Intelligent Transport Systems*, 1(1), 15. doi:10.1049/iet-its:20060020
- Calabrese, F., Di Lorenzo, G., Liu, L., & Ratti, C. (2011). Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing*, 10(4), 36–44. doi:10.1109/MPRV.2011.41
- Department of Census and Statistics—Sri Lanka. (2012). *Census of Population and Housing 2012—Final Report*. Colombo, Sri Lanka: Author
- Isaacman, S., Becker, R., Cáceres, R., Kobourov, S., Martonosi, M., Rowland, J., & Varshavsky, A. (2011). Identifying important places in people's lives from cellular network data. *Pervasive Computing: Lecture Notes in Computer Science* (Vol. 6696, pp. 133–151). Berlin, Heidelberg, Germany: Springer Berlin Heidelberg. doi:10.1007/978-3-642-21726-5_9
- Ministry of Transport—Sri Lanka. (2013). *Urban Transport Master Plan for Colombo Metropolitan Region and Suburbs*. Colombo, Sri Lanka: Author
- Wang, H., Calabrese, F., Di Lorenzo, G., & Ratti, C. (2010). Transportation mode inference from anonymized and aggregated mobile phone call detail records. *13th International IEEE Conference on Intelligent Transportation Systems* 318–323. doi:10.1109/ITSC.2010.5625188
- Wu, W., Cheu, E. Y., Feng, Y., Le, D. N., Yap, G. E., & Li, X. (2013). Studying intercity travels and traffic using cellular network data. *Mobile Phone Data for Development: Net Mob 2013*. Retrieved from <http://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf>
- Yoo, B.-S., Chon, K., Kang, S.-P., & Kim, S.-G. (2005). Origin-destination estimation using cellular phone information. *Journal of the Eastern Asia Society for Transportation Studies*, 6, 2574–2588. Retrieved from http://www.easts.info/on-line/journal_06/2574.pdf

