MAPPING THE INFORMAL PUBLIC TRANSPORT NETWORK IN KAMPALA WITH SMARTPHONES: MAKING SENSE OF AN ORGANICALLY EVOLVED CHAOTIC SYSTEM IN AN EMERGING CITY IN SUB-SAHARAN AFRICA

I NDIBATYA*, J COETZEE# AND MJ BOOYSEN*

*Department of E&E Engineering Stellenbosch University, South Africa
#GoMetro, South Africa
indibatya@gmail.com, justin@gometro.co.za, mjbooysen@sun.ac.za

ABSTRACT

In cities in the emerging world, public transport networks are governed by a large number of agents, each with their own agendas, priorities, incentives and resources, interacting nonlinearly through complex feedback loops. The transport system in these cities have developed into a semi-chaotic self-organizing structure with seemingly unpredictable behaviour to an outside observer. This is due to user agent actions by passengers, independent determination of operating plans and practices by transport operators, and a managing authority exhibiting a lack of will (both political and institutional), to implement adequate control measures to provide regulation and management of these systems.

Based on the problems that face transport systems in developing cities and public transport in particular, this paper reports on an attempt to understand the supply of public transport in the Kampala area in a novel manner. We describe a mapping approach using a custom-developed smartphone application which was used to quickly and accurately capture informal transport systems for analysis and study of urban mobility where no dependable data was currently available. Secondly, based on the data created by the study project, to provide insights into the routes, operations and characteristics of the minibus taxi network which convey the majority of Kampala's travelers. Our hypothesis is that by understanding the network in geospatial terms, we will be able to create benefits for all role-players and improve the efficiency and effectiveness of the supply of public transport to more closely match the demand for public transport in an emerging world city.

1 INTRODUCTION

Kampala, Uganda, experiences transport system challenges similar to other developing cities in Africa, with three identified main contributing factors to the transport supply not matching the needs of the citizen demand. The first issue is rapid urbanization. The greater Kampala area was home to 3.15 million residents in

2012, and its population is expected to double by 2020 and quadruple by 2040 (World Bank, 2015). This increase in population already places and will continue to place severe pressure on the severely-constrained available space, utilities, roadways and public transport supply of Kampala. The second relevant factor is that there is a low percentage of people with access to private transport by motor vehicle. It is estimated that only 9% of the passengers on Kampala's roads use private transport, while the rest (91%) use public transport (KCCA, 2014). Third, is the lack of a well-regulated and organised public transport service as an alternative to private car usage. In Kampala, public transport has not been provided by the City Council, and consequently public transport supply has been provided by private individuals who own and operate mostly 14-seater minibuses on unregulated routes and schedules, and operate as the mode of transport literature refers to as "paratransit". Fourth, the public transport that is provided by private operators is run at an operating profit and not as a social service or a social good, and so the very poor sometimes cannot afford the high cost of public transport and therefore walk for 50% of their trips which tremendously reduces their mobility and reduces their access to opportunities (KCCA, 2014).
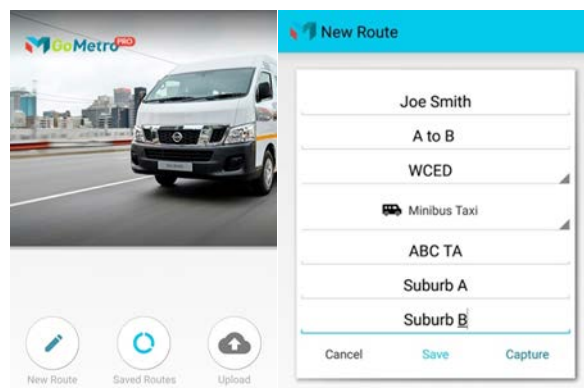

Figure 1: The GoMetro Pro mobile application used to map Kampala paratransit

Based on the problems that face transport systems in developing cities and public transport in particular, this paper reports on an attempt to understand the supply of public transport in the Kampala area in a novel manner, namely smartphone mapping. We describe a mapping approach using a smartphone application, GoMetro Pro, which was used to quickly and accurately capture informal transport systems for analysis and study of urban mobility where no dependable data was currently available, at an acceptable level of accuracy and precision. The resultant dataset was then evaluated and analysed.

Our hypothesis is that by understanding the network in geospatial terms, we will be able to create benefits for all role-players and improve the efficiency and effectiveness of the supply of public transport to more closely match the demand for public transport in an emerging world city.

2 PUBLIC TRANSPORT IN EMERGING CITIES

Public Transport (which is mostly informal in many developing cities) plays a very important role in the development of a city. It provides an entry point to urban employment for low-income citizens (Ken, 2002), it provides a cheap mode of mobility to the remote locations in the cities (Mutiso, 2011), and considerably improves access to opportunities and services. The chaos that characterises the unregulated informal public transport reduces the positive impacts that public transport would play in the lives of urban citizens. One such negative impact is an inability to easily plan a transport route within the undocumented and dynamic public transport network.

Historically, local or regional governments in Africa have serviced public transport needs through the tendering and management of a monopolised large conventional bus service (Ousmane, 2008). However, due to a lack of proper management, undue political influence, and the inability to adapt to changing customer demands, these bus systems usually incur large losses, are unable to maintain operations and eventually shut down. Consequently, small privately-owned minibuses have spontaneously arisen and filled the public transport supply gap left by large bus system failures. These are known by different names in different regions, such as mini-bus taxi, matatu or trotro.

These 14-seater minibuses are responsible for the majority of daily commutes in the public transport sector in Kampala, transporting 83% of the passengers in the city and contribute 36% of the total vehicles on Kampala's roads (ITP, 2010). To most low-income earners in the city, they provide an easy means to get near to their major destinations in the city, from where they can either walk or board a motorcycle taxi (boda-boda). There is no formal ticketing system, with all transactions being conducted in cash. There is also no centralised planning or pricing function, with an association of loosely-connected owners and drivers largely setting their rules.

However, the minibus taxi operations in Kampala have grown to a scale that it has replaced the need for a large metropolitan bus service, even though it operates much less efficiently and with individual vehicle optimisation in terms of profit and operations. Without a system-wide planning authority with a mandate to provide network-level operational optimisation, the paratransit system operates individualistically with some chaotic outcomes: with no fixed fare-structure, fixed routes, route schedules and designated stops, there is very little information on their operations available to city planners aiming to understand how the City currently moves.

There is very little information available to commuters trying to find directions from one point to another. In Kampala City, it is common to wait for a taxi from a particular location and fail to get any for over two hours. The commuter experience does not promote public transport use. In 2006-8, the UN-Habitat through the Institute for Transportation & Development Policy (ITDP) attempted to develop the network by using human observers at strategic points in the city (Ousmane, 2008). This method,

known as corridor or cordon counts, is prone to errors, and its shortcomings are well-documented (Cameron, 2005; Moodley, 2005).

However, the emergence of large-scale urban sensing data such as mobile phone traces (from smart phone GPS sensors), and the considerable penetration of mobile phones in developing cities present researchers with an opportunity to collect high update frequency data over a long period of time at low costs, and therefore study urban mobility at a much higher resolution than before (Francesco, 2013). We believe that many developing cities can make use of high smartphone penetration rates to collect public transportation network data using GPS trajectories, process it and return indexed and searchable transit information to the citizenry in multiple formats including text message, USSD string, HTML and electronic signage.

3 THE TRADITIONAL APPROACH TO PUBLIC TRANSPORT DATA COLLECTION

Gaibe (2010), in a paper proposing the use of GPS data in order to improve the data generated from the collection of the Current Public Transport Record (CPTR) of a city referenced Moodley (2005) and Cameron (2005) in outlining the key challenges and shortfalls in CPTR data collection efforts.

Some of these were:
- Operations encountered in the field differed significantly from the information sourced from operators and officials.
- Survey data capturing and cleaning was a major task requiring long hours and dedication from the capturing staff. Errors in the datasets were common, diluting the value of the data collected, and the predictability of the performance of models on the transport network data collected.
- There was a requirement in the design of these studies for a uniform interpretation of terminology and methodologies before data collection was executed in the field.
- Little data was collected between ranks, with the focus of data collection on taxi rank counts. Intra-rank demand was therefore not observed and counted in the model. Gaibe (2010) found that in Stellenbosch, up to 57% of total passengers were not accounted for using traditional rank surveys only. For shorter, local routes that figure went up to 70% of passengers not counted at the ranks.

4 PUBLIC TRANSPORT DATA COLLECTION BY MEANS OF MOBILE PHONES

A public transport mapping project using smartphones was implemented by the University of Stellenbosch Department of Electrical and Electronic Engineering, and funded by GoMetro in November to December 2015. The data-warehouse for the project runs on GoMetro Pro and is hosted by GoMetro on its data servers. The

authors would like to make this data available to any urban researcher looking to replicate the methods used, but perhaps in another city.

Public transport network mapping was done in three phases. Phase one involved the identification of all gazetted and ungazetted taxi ranks, the routes that originate from the respective ranks and documenting the major stops along the routes (informally known to the taxi drivers and conductors). Phase two consisted of the actual tracing of selected routes by volunteers using the GoMetro Pro application, as well as constructing a description of operations for each route in terms of trip frequency. During this phase a team of three volunteers activated the GoMetro Pro mobile application (a free mobile app from GoMetro that is also called GoMapp in some regions) and rode in the available public service vehicles as normal passengers. Route traces of individual riders were collected by the mobile app as a series of GPS coordinates along the routes for two weeks. Furthermore, the volunteers tagged the locations of all the vehicle stops during the trips. This data was used to characterise the informal public transport network and its dynamics on a route-by-route basis. Phase three involved data processing and plotting, and map-matching algorithmic analysis to match the GPS data with the road infrastructure to correctly locate the roads used by the minibus taxis.

4.1 Phase 1: Documenting Taxi Ranks and Route Origin and Destination pairs

Volunteers traveled around the five major divisions that make up Kampala (Central, Kawempe, Makindye, Nakawa & Lubaga) to locate the formally gazetted Taxi ranks and other gazetted taxi ranks. At each rank, the GPS coordinates of the taxi rank were recorded, and the major routes that originate from the taxi ranks were registered in a central database. In some cases, the routes were seen written on small signage which were carried from one minibus taxi to the other in order of departure sequence. The Table 1 shows the number of taxi ranks in each division that make up Kampala City, the number of routes originating from the taxi ranks, and the average taxi fare per division. Taxi ranks and routes were categorized as either major or minor, classed as such based on whether the frequencies of departure were high or low. Major taxi ranks are the ranks officially gazetted by the authorities and minor ranks organically evolved (in most cases along the routes).

Table 1: Showing the number of taxi ranks, number routes originating from taxi ranks, and the average   route costs.

S/N	KCCA Division	Taxi Ranks			Routes				Average taxi fare per route	
		Major	Minor	Total	Major	Medium	Minor	Total	UGX	ZAR
1	Central Business District	3	6	9	55	17	50	122	2732	11
2	Nakawa	0	1	1	3	3	6	12	6750	26
3	Makindye	0	4	4	4	5	6	15	3646	14
4	Rubaga	0	2	2	14	12	4	30	2983	12
5	Kawempe	0	2	2	6	4	2	12	991	4

4.2   Route Tracing

Volunteers boarded taxis from the taxi ranks following selected routes up to the destinations. The volunteers used the custom-developed GoMetro Pro (formerly GoMapp) mobile application to automatically log the GPS locations of the minibus taxi every 30 seconds during the journey, and to tag each stop along the way by means of the application. Although a user can collect boarding and alighting data using the app, that was not part of this study, and so was not done. The objectives of this phase were to: (1) establish the actual route that taxis follow to their destinations; (2) to tag the informal and formal stops along the route; and (3) to establish the cost of traveling on a given route and the inter-stop (major stops) costs. Figure 2 shows the minibus taxi stops (formal and informal), and the underlying minibus taxi network.
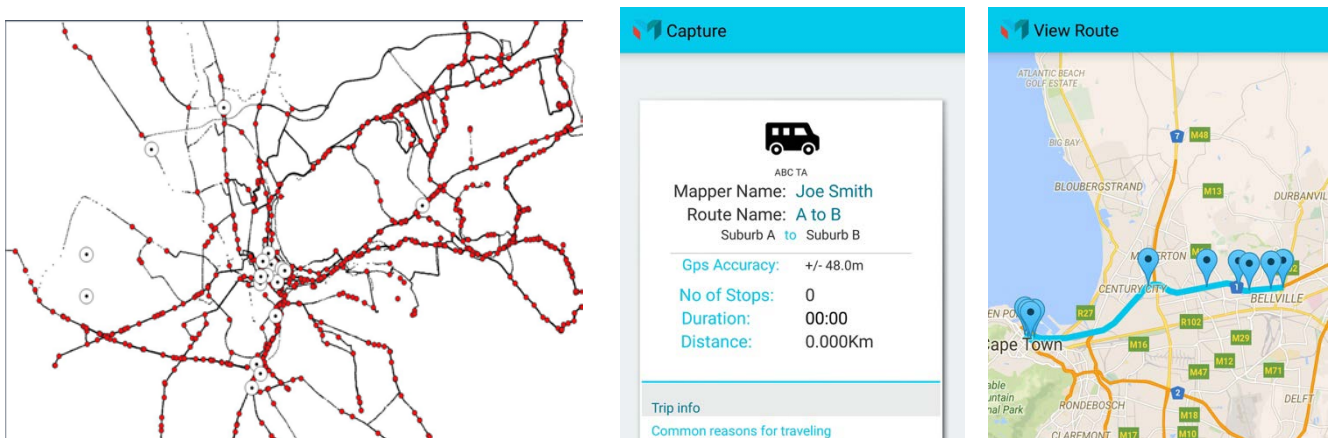

Fig 2: Figure showing the minibus taxi network, the stops and ranks in Kampala

Phase 3 – data processing

Primary data collected by the GoMetro Pro mobile application and transmitted to the GoMetro servers was downloaded and passed through a series of processing steps to ensure consistency and accuracy. Data transfer was done using the ProtoBuf protocol. We used the Python programming language and server-side libraries to prepare, stage and analyse the dataset. After cleaning, the dataset of 81,567 GPS traces (Latitudes and Longitudes) was used in the rest of the analysis. Table 2 shows a sample of the dataset used.

Table 2: Showing the sample data set used

id	roadid	latitude	longitude	route_id	stop_id	stoplat	stoplon
11233	59	0.369890362	32.64934921	11232	11240	0.36989	32.64935
11234	59	0.369768977	32.64928055	11232	11242	0.36989	32.64935
11238	59	0.369327009	32.64917374	11232	11244	0.369653	32.64923
11291	61	0.370179594	32.73626328	11290	11440	0.367664	32.73895
11292	61	0.369937807	32.7364502	11290	11442	0.366804	32.73978
11293	61	0.369733423	32.73667908	11290	11444	0.36606	32.74044

QGIS, a free Geographical Information System (GIS) tool for Linux was then used to plot the data and study the trends. The GPS data was overlaid on the roads network obtained from Open Street Map.

The location accuracy of GPS data depends on many variables. These include the weather (i.e. accuracy is low when it is cloudy compared to when the sky is clear), the strength of the receiver antenna (in our case, the smartphone), the obstructions (e.g. trees, buildings) and other variables. The GPS data we collected had some errors (multi-path errors) and in most cases, the GPS points were off the roads and scattered a few meters away from the road as shown in Figure 3.
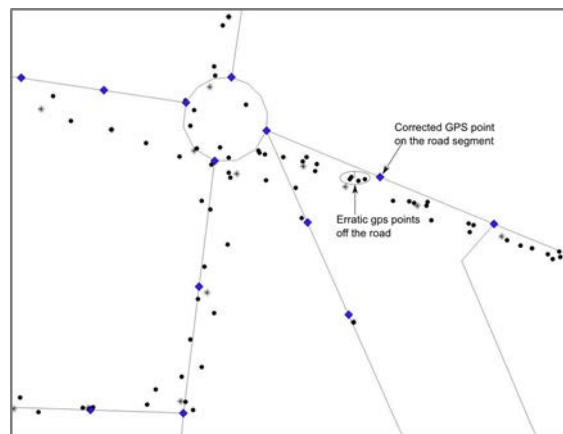

Fig 3: Figure showing the occurrence of multi-path errors and the matched positions by the map-matching algorithm.

Multi-path errors in GPS data occur due to Radio Frequency (RF) signals from GPS satellite either being blocked by obstacles such as trees and buildings or the signals reaching the antenna after bouncing off the ground or the obstacles. It causes the receiver to miscalculate the direct distance between itself and the satellite.

We developed a map matching algorithm to match the road segments with the data in our data set within a 20-meter radius. The algorithm makes several reference points along the road segment and then searches for all the GPS traces within the defined radius. Once traces are found, the algorithm returns the current reference point as the point of reference for the taxi position. In addition, the algorithm computes and returns the sum of all traces in the region, and the average of the Latitudes and longitudes in the region. Table 3 shows the sample output from the map matching algorithm we developed.

Table 3: Table showing the results of the map matching Algorithm.

id	road_id	segment_id	mid_idx	mid_lat	mid_lon	countOfGps	avg_lat	avg_lon
1	2	1	6	0.345888	32.65649	2	0.345732	32.65666
2	3	1	0	0.333723	32.61793	31	0.333695	32.6179
3	3	1	1	0.333603	32.61803	2	0.333452	32.61819
4	3	2	1	0.332997	32.61878	3	0.333018	32.61881
5	3	2	3	0.3333	32.61841	2	0.333289	32.61839
6	3	2	6	0.333262	32.61846	1	0.333173	32.6186

5 RESULTS OF THE MAPPING EFFORT USING GOMETRO PRO

The public minibus taxi network is made up of major nodes (Gazetted taxi ranks) and minor nodes (ungazetted taxi ranks) that act as major sources and sinks of travellers in the city. The nodes are interlinked with series of stops (intermediate nodes), some of which are known and gazetted while others are ad hoc and organically develop according to demand. Taxi movements through the network can be described by random Markov walks, they originate and end at specific points while following several routes in the network. Several factors are responsible for the random walks. They include; passenger demand, presence of traffic officers on the main route, the state of the roads, the choice of passengers in the taxi at the time and weather conditions. It is a collection of all these agents that form chaotic tendencies in the public transport network in developing cities.

A sample schematic transit map of roads, taxi ranks, and stops was developed from this research. The map can be enhanced by more data and involving more participants through crowdsourcing methods (Zegras, 2014). Operational schedules for each route mapped and categorised should be compiled in the General Transit Feed Specification format. This is to include pull-away times from every station, rank and terminal – as well as the development of stop-times along each route. The schedules can be fixed or frequency-based. Schedules are to reflect the different operating calendars.

6 CONCLUSION AND RECOMMENDATIONS

The dynamics of public transport networks in developing countries are complex. To fully characterise the entire network, a major amount of data is required from all the actors in the public transportation, i.e. minibus taxis, buses, boda-boda and individual movements of the population. In that case, it is possible to model the complete trips of individuals through the public transport network. In this paper, we sought to study the dynamics of the public transport network in Kampala city using GPS traces from passenger mobile smart phones collected by means of a custom-developed mobile application.

The ultimate objective is to collect two geospatial datasets of Kampala's informal transport network. The first dataset, the one under consideration in this paper, has been developed from a controlled process by this research team. The second dataset is to be developed from a crowdsourcing initiative by Kampala citizens, with a goal to understand if it is possible to use crowdsourcing techniques to create a public transport network dataset that is of a similar level of quality and accuracy as the dataset that a controlled study would produce. This subsequent work will give us two datasets generated in very different processes that we can compare in terms of performance and coverage of a crowdsourcing project.

The dataset generated contains metadata and content to draw up in the General Transit Feed Specification (GTFS) standard of transit data processing. A next step for this work will be to generate a GTFS-compliant operations schedule, service definition and SHP file creation of the current system. This will provide a basis for comparison models with other cities who have developed GTFS-compliant datasets (such as Nairobi and Cape Town) for further analysis and benchmarking of the study area transport performance with other emerging market cities.

A GTFS-compliant dataset will also allow for the analysis of the transport network performance before the implementation of any interventions, as well as the study area performance after the implementation of proposed transit system improvement interventions. It will now be possible to produce a searchable web-based map that produces a list of every route and stop of every transport mode within a pre-defined radius of a pre-selected co-ordinate. This data collection work can also form the precursor of smartcard or smartphone ticketing systems – which require comprehensive data on a clear operational plan and fare-structure per route or route variation – which this type of dataset starts to collate and improve upon.

This dataset will also now be the basis for the use of web-based tools such as OpenTripPlanner or GoMetro Pro to visualise the level of transport service provision per city block, in order to measure and visualise by reports and heatmaps, the high and low concentrations of service options and service frequencies. New trip-planning software such as OpenTripPlanner or GoMetro Pro can process these datasets in order to generate a travel plan consisting of walking and multiple modes in a trip plan that can be consumed by commuters by means of a website, a USSD string (for non-smartphone users), a mobile app,

We have shown that by understanding an informal public transport network in geospatial terms, we will be able to create benefits for all role-players and improve the efficiency and effectiveness of the supply of public transport to more closely match the demand for public transport in an emerging world city. We recommend similar studies to be undertaken in other developing cities and the development of smart data processing algorithms that can use the routes' data to provide useful information to the travellers. This will reduce the time city travellers try to figure out how to move from one place to the other in the city and give them the ability to manage their daily commutes with some level of certainty. The authors are of the

view that major data collection studies in emerging market cities should be executed with mobile phone tracking creating unique geospatial user traces.

7 BIBLIOGRAPHY

Cameron, J. W., Questions about the Quantitative Basis of Municipal Transport Plans. Presented at South African Transport Conference. Pretoria, 2005.

Francesco Calabrese, Mi Diao, et al. Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. Transportation Research Part C: Emerging Technologies, Volume 26, January 2013, Pages 301-313, ISSN 0968-090X.

Gaibe H and Vanderschuren M. An Investigation into the Methodology of Mini-bus Taxi Data Collection as part of the Current Public Transport Record: A Case Study of Stellenbosch in the Western Cape. Proceedings of the 29th Southern African Transport Conference. Pretoria. 2010

Integrated Transport Planning Ltd. Pre-feasibility Studies for the Development of a Long-term Integrated Bus Rapid Transit System for Greater Kampala Metropolitan Area. World Bank report. May 2010.

Kampala Capital City Authority. Strategic plan 2014 – 2019. Kampala. September 2014

Moodley, G. Y. Developing the Ethekwini OLS: How useful is the CPTR information? Presented at the South African Transport Conference. Pretoria, 2005.

Mutiso W and Behrens R. Boda-Boda Bicycle Taxis and their Role in Urban Transport Systems. Proceedings of the 30th Southern African Transport Conference. Pretoria. 2011.

Naddumba, E K. A Cross-Sectional Retrospective Study of Boda-Boda Injuries at Mulago Hospital in Kampala-Uganda. 2000.

Nazarius T M, Atuyambe L M, Kobusingy O K. Factors Associated with Injuries among Commercial Motorcyclists: Evidence from a Matched Case Control Study in Kampala City, Uganda. PLoS ONE. 2016

Ndibatya I, Booysen M J, Quinn J. An Adaptive Transportation Prediction Model for the Informal Public Transport Sector in Africa, IEEE ITSC, 2014, Qingdao, China.

Ousmane T. Overview of Public Transport in Sub-Saharan Africa. Trans-Africa Consortium. UITP - International Association of Public Transport. Brussels, Belgium. 2008.

Williams, K G. Cities on the Move: A World Bank Urban Transport Strategy Review. World Bank. Washington DC. 2002.

World Bank Group. Promoting Green Urban Development in African Cities, Kampala, Uganda. International Bank for Reconstruction and Development and the World Bank. Washington DC. September 2015.

Zegras C, et al. Tracing a path to Knowledge? Indicative user impacts of introducing a Public Transport map in Dhaka, Bangladesh. Cambridge Journal of Regions, Economy and Society. Cambridge, Massachusetts. August 2014