ESTIMATION OF TRIP GENERATION IN YANGON CITY BY USING CDRS DATA

Kyaing[a], KoKoLwin[b], Yoshihide Sekimoto[c]

[a]Yangon Technological University, Insein Road, Gyogone, Insein 11011, Yangon, Myanmar.
[a]Email: kyaingkyaing63@gmail.com
[b,c]Institute of Industrial Science, The University of Tokyo, Komaba 4-6-1, Meguro Ku, Japan.
[b]Email: kokolwin@iis.u-tokyo.ac.jp
[c]Email: sekimoto@ii.u-tokyo.ac.jp

Abstract: This paper focuses on estimating the number of person-trips that will produce from or attract to an area in Yangon city by using Call Detail Records (CDRs) data. Multiple linear regression analysis is used in prediction of trip generation. This analysis establishes the relationship between the number of trips generated by township and some socioeconomic attributes. The numbers of trip generated are demonstrated using CDRs over seven days of December 1th to 7th, 2015. CDRs are used to generate trip numbers of townships within certain time frame and time windows. The variables that mostly affect trip production are found to be the car ownership. The number of office and industry are found to be the most effective variables in trip attraction prediction with significant level of 5%. This information is very useful for urban transport planning and especially for public transport planning to improve the deteriorating of urban transport situation.

Keywords: Multiple linear regressions, trip generation, Call Detail Records (CDRs), socio-economic attributes, urban transport planning

1. INTRODUCTION

Yangon is not only the former capital city but also the largest economic centre in Myanmar with about 7.36 millions in the area of 10276.7km^2 (Wikipedia, 2014). After 1988, the government started to launch the market-oriented economy that put into practice trade opportunity to encourage international trade investment. Various governmental departments of Yangon city have been cooperated in the development of the industrial zones. Therefore many people also migrate from rural areas to the city for their better life and income. This occurred because they are pushed by poverty and less chance of opportunities for their better lives or they may be pulled by the attraction of the city.

Numbers of car registration are increased after 2011 due to car import policy changes. In 1990s, the number of vehicle registration is about 74,000 but in 2013, the number rose to about 260,000. According to this economic policy and car import policy, Yangon city becomes rapid urbanization and motorization. In order to satisfy the needs for rapid urbanization and motorization, the city should be well-planned with a variety of facilities. Among the needs, transportation infrastructure plays a vital role in economic growth of country.

The current urbanization and motorization put worse and worse day by day on the existing transport infrastructure in Yangon city. The deteriorating of urban transport situation has become a serious concern socially, politically and environmentally. In order to provide a well transportation planning, it is necessary to understand the human mobility pattern and to predict future traffic demand.

To analyse and predict the traffic demand, many data sources and information are required. Today, especially in developing countries, the process of obtaining relevant real-world indicators is a hard task that needs a lot of expertise and resources. Household travel interviews, census survey data, or traffic counts represent traditional tools to forecast travel demand. Interview and surveys, however, are costly and conducted rarely especially in Myanmar.

The roadside and household surveys involve expensive data collection and thereby limiting sample sizes and lower update frequencies. Moreover, they are prone to sampling biases and reporting (Hajek, 1977).

The census data is also hard to obtain as the first census data was collected in 1973 and the second time was collected in 1983. After 30 years, the third census data was collected in 2014.

In the prediction of the human mobility, one of the largest fields of application is using Call Detail Records (CDRs) for understanding which is the busiest areas and times during the day, to predict the movement of the people and promote a better transportation service (Berlingerio et al., 2013). In Calabrese et al. (2011) mobile data are used for monitoring in real time the traffic conditions and pedestrian distribution. The CDR trace has become a powerful tool to analyse human behaviour patterns and an increased interest towards making use of CDRs to analyse the human mobility cheaply, frequently and especially at a very large scale that has been recorded lately. Mobile phones are becoming pervasive in both developed and developing countries and they can be a precious source of data and information, with a significant impact on research in behavioural science (Berry, 2011; Lazer et al., 2009).

Mobile phone datasets have shown to be particularly useful and popular as they provide temporal and spatial information on a scale and granularity that was not available to researchers. Although CDRs may have some bias on Human mobility studies, but up to now, they have been providing the best data sets to study the human mobility. Therefore, this paper focuses on estimating the number of person-trips that will produce from or attract to an area in Yangon city by using CDRs data.

The structure of this paper is organized as follows: previous research works are presented in Section 2. Section 3 stated that why the study area is selected as a case study. Section 4 briefly describes that data to be used in this study and how to analysis the trip production and attraction from O-D flow matrix. The relationship between trip generation and socio-economic attribute are discussed in Section 5. Finally, conclusion and future work are presented in Section 6.

2. LITERATURE REVIEW

There exists a large body of work on qualitative studies analysing the relationship between socio-economic factors and cell phone usage. Donner *et al.* (2007) presented a survey of 277 micro entrepreneurs and mobile phone users in Ki-gali, Rwanda, to understand the types of relationships with family, friends and clients, and its evolution over time. Among other findings, the author discovered an inverse correlation between the age of the user and the probability of adding new contacts to its mobile-based social network. The author also claims that users with higher educational levels were also more prone to add new contacts to their social net-works.

Similar qualitative studies were carried out by Kwon *et al.* (2000). The authors conducted a study to understand the impact of demographics and socio-economic factors on the technology acceptance of mobile phones. The analyses presented in this paper offer the ability to expand these studies to millions of users by correlating their cell phone behavioural usage to public datasets with socio-economic information.

The literature covering large-scale quantitative analyses of the relationship between cell phone usage and human factors is very limited, given the recent availability of large datasets with cell phone call records. Eagle *et al.* (2008) studied the correlation between communication diversity and its index of deprivation in the UK. The communication diversity was derived from the number of different contacts that users of a UK cell phone network had with other users. Eagle combined two datasets:

(i) a behavioural dataset with over 250 million cell phone users whose geographical location within a region in the UK was known, and

(ii) a dataset with socio-economic metrics for each region in the UK as compiled by the UK Civil Service.

The author found that regions with higher communication diversity were correlated with lower deprivation indexes. Although this result represents an important first step towards understanding the impact of socio-economic parameters on mobile use at a regional level, it is necessary to seek to elaborate more fine-grained impact analyses that can draw correlations between human factors and cell phone usage. Blumenstock *et al.* (2010) analysed the impact that factors like gender or socio-economic status have on cell phone use in Rwanda.

Similarly to Eagle *et al.* (2008), the authors combined two datasets, one containing call detail records from a telecom company in Rwanda and the other one containing socio-economic variables computed from personal interviews with the company's subscribers. Their main findings revealed gender-based differences in the use of cell phones and large statistically significant differences across socio-economic levels with higher levels showing larger social networks and larger number of calls among other factors. This approach succeeds to reveal findings at an individual level; however, it limits the scalability of the results to the availability of the subscribers and to the amount of time and money available to carry out personal phone interviews to hundreds of users. To overcome these problems, an approach that combines two large-scale datasets to understand the relationship between cell phone use and specific human factors is proposed.

Because CDRs data are obtained continuously, much analysis related with movement pattern has been performed. UN Global Pulse utilizes such data for social issues, such as disaster management, spread of infectious disaster management, and spread of infectious disease, socio-economics, and transportation. In addition, CDRs data have been used for behaviorioral pattern predictions by M. Ozer *et al.* (2013, 2014), extraction of people's stay by M. Kurokawa (2013), route estimation, and road traffic volume estimation by Y. Hasegawa (2014). Recently, origin-destination trips have also been detected using CDRs by L. Alexandera (2105).

A few studies have measured strong relationships between socio-economic levels and human mobility at specific scenarios such as access to hospital by Propper, C. (2007), travel patters by Carlsson- Kanyama (1999).

Japan International Cooperation Agency (JICA) collected person trips data to forecast trip production and attraction of Yangon in 2013.

3. STUDY AREA

There are 45 townships in Yangon region which are classified as central business district (CBD), inner city, outer city, and new suburban, old suburban and periphery area. Yangon is not only the regional hub for rail and ground transport but also regional and international hub for air and principal seaport. Among these townships, thirty three townships (except periphery area) have selected as the study area. The selected study area is more urbanized and motorized at the infrastructure in Yangon area and also more developed at the public transportation. The current situation of urban transport is deteriorated and it is necessary to be control and manage at the mobility and accessibility to urban service for an efficient and sustainable public transport system and road network. The location of the study area based on district is shown in Figure 1.

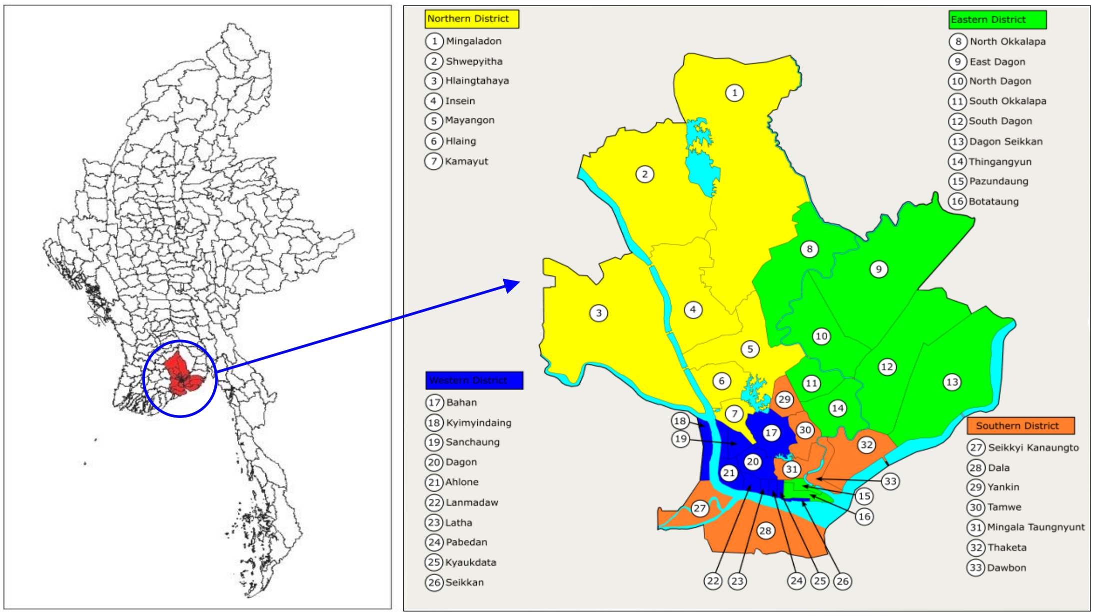

Figure 1. Location of the study area

4. METHODOLOGY

4.1 Data to be used in Study

Two types of secondary datasets were used in this analysis. The first dataset is Call Detail Records (CDRs) data and the second dataset is Socio-economic data.

4.1.1 Call Detail Records (CDRs) data

The CDRs data is used in the estimation of person trip because of its huge volume, wide coverage, real-time production, automated collection, and low cost. Due to the limited roadside surveys and infrequent data collection of census data, CDRs data become a powerful tool to analyze human behavior patterns and are useful for transportation planning. CDRs data will be collected from Myanmar Post and Telecommunications (MPT), which is one of the biggest mobile operators. Although there are 3 telecommunication networks in Myanmar, namely MPT, Telenor and Ooredoo, only the CDR data from MPT is available. The CDRs Dataset contains 7 days of voice calls and data from base transceiver stations (BTS) towers located in the study area during 1st December 2015 to 7th December 2015 across in Yangon.

There are 656 BTS towers of MPT service in 33 townships of Yangon and about 60 dead BTS towers as shown in Figure 2(a) and 2(b). The following Figure 3 shows the comparison of total number of BTS and dead BTS in traffic analysis townships.

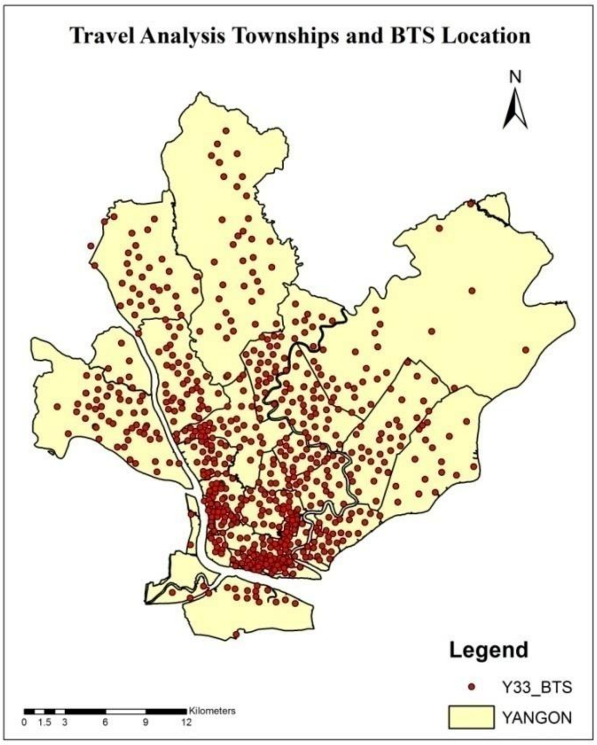

Figure 2(a). Distribution of BTS in traffic analysis townships

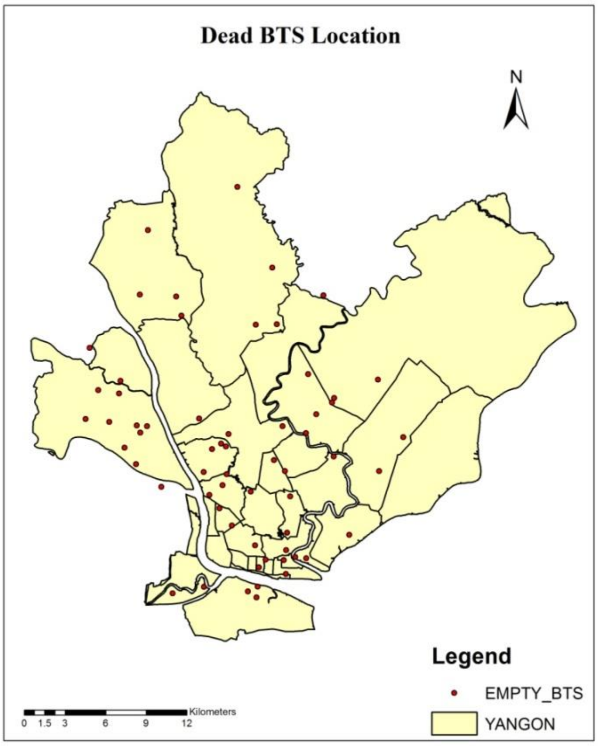

Figure 2(b). Distribution of dead BTS

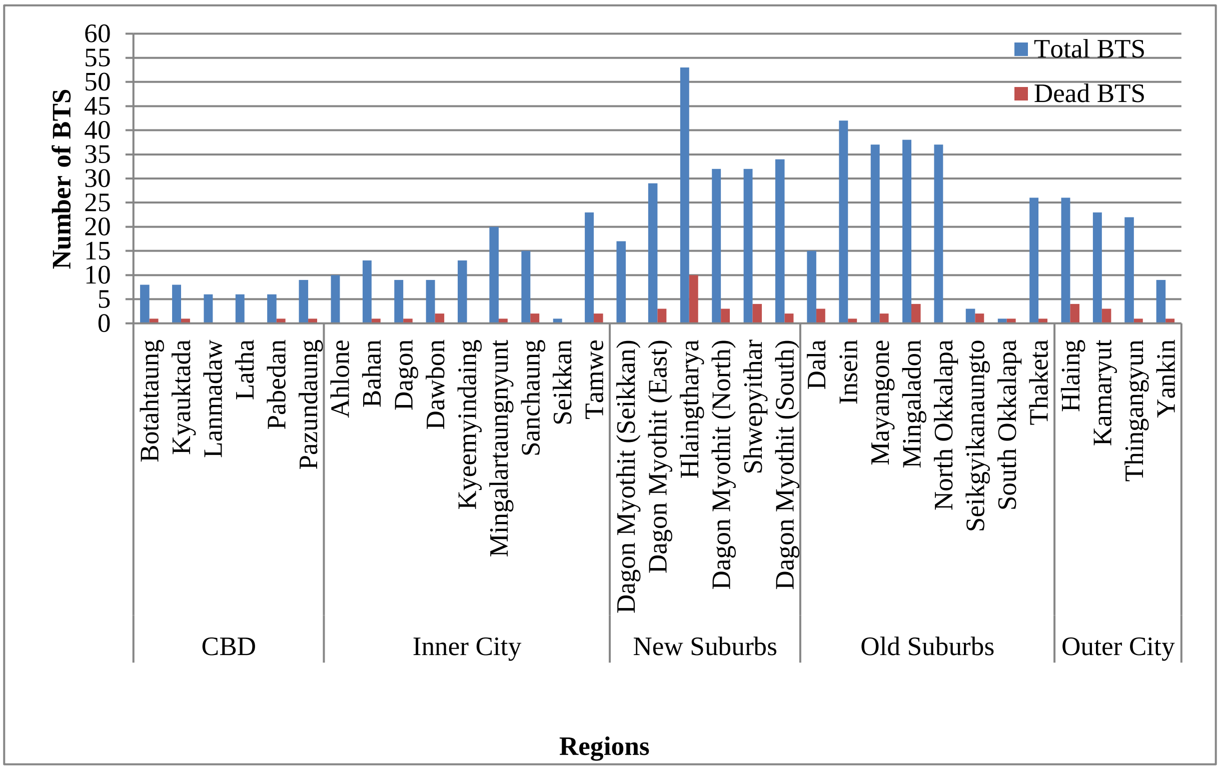

Figure 3. Comparison of number of total BTS and dead BTS in traffic analysis townships

To safeguard personal privacy, individual phone numbers were anonymzed by the operator before leaving their storage facilities, and were identified with a security ID (hash code). Each entry in the dataset has a CDRs comprising of Timestamp, Caller's ID, Call duration in second and Caller's connected cell tower ID. The following table shows the sample of encrypted CDRs data and abbreviation of this information.

Table 1. Encrypted CDRs data

PID	Event ID	DTIME	Duration	Upload	Download	Cell ID
26D8405F8E	1048	20151201160554	64	194	610	1414011000907400
A8D8575F87	982	20151201160556	25478	5624	11687	1414011000606820
3CD94C5F8C	982	20151201160542	77	45462	31939	414010801320823

Abbreviation: PID = Person Identification (Encrypted mobile SIM card number);
EVENT = Call events such as incoming call, outgoing call, etc.;
DTIME = Start call date and time (yy:mm:dd:hr:min:sec);
DURATION = Call duration (second);
CELLID = Cell Identification Number;
UP = Upload data size;
DOWN = Download data size

4.1.2 Socio-economic data

Socio-economic data were used to identify user impact of human mobility pattern and was collected from Ministry of General Administration Department, and Ministry of National Planning and Economic Development (2014 Census data), YUTRA and Yangon Directory. The data collected for socioeconomic data are population, age group, average income of township, car ownership (total number of car in township), number of household, population density, and number of student, number of worker, number of school, number of university, number of office, number of hotel, number of recreation, number of religious, number of industry, number of hospital, number of shopping center, and number of market.

4.2 Method of Analysis

In the prediction of the human mobility, one of the largest fields of processing is trip production and attraction of person or vehicle. In this study, CDRs data was used for analysis of production and attraction of person-trip based on socio-economic data. Figure 4 shows the preprocessing of CDRs data, processing of origin and destination (O-D) matrix to extract trip generation.

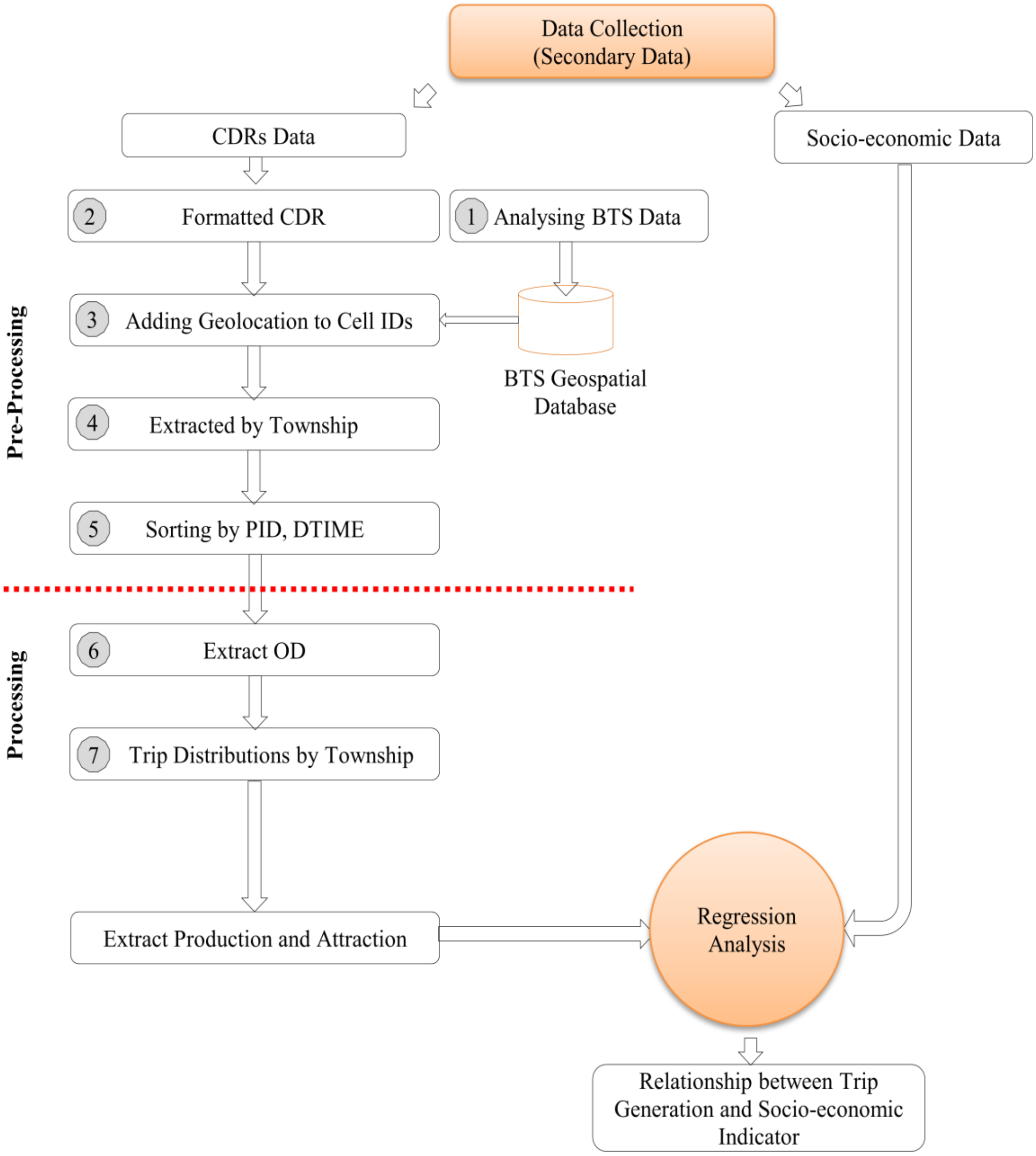

Figure 4. Implementation program

4.2.1 Preprocessing of CDRs data

A call detail record (CDRs) is a data record produced by a telephone exchange or other telecommunications equipment that documents the details of a telephone call or other telecommunications transaction (e.g., text message) that passes through that facility or device (wikiperia).

In preprocessing CDRs data, some error in BTS file, missing BTS name, wrong coordinates and township names were corrected in analyzing BTS data step. And BTS to geographical locations were converted by using a map from UN (United Nations) project for reference purpose.

CDRs data have to format as correct shift digits, adding string to convert number to string. Formatted CDR voice data and formatted CDR data are shown in Table 2 and Table 3.

After formatting, geographic locations (Longitude, Latitude) should be added to each cell IDs as shown in Table 4.

Table 2. Formatted CDRs voice data

PID	EVENT	DTIME	DURATION	CELLID
PID_C2DBCF5F7E00	1103	20151206000752	63	CID_414010035717871
PID_E7D7075F8200	1097	20151206000836	41	CID_414010060520072
PID_31DBE86F0600	1103	20151206000846	9	CID_414011000401661

Table 3. Formatted CDRs data

PID	EVENT ID	DTIME	DUR ATION	UP LOAD	DOWN LOAD	CELL ID
PID_26D8405F8E00	1048	20151201155914	446	440	1394	CID_414010803330561
PID_3CD94C5F8C00	1048	20151201160554	64	194	610	CID_414011002401315
PID_3CD94C5F8C00	982	20151201160556	63	5624	11687	CID_414011002401315

Table 4. BTS geolocation

SITEID	SITENAME	CELLID	LON	LAT	TOWNSHI P	STDIV	SYSTE M
YGN00265	UTYGN00265B1	CID_414010035617552	96.16028	16.7779	Dagon Seikkan	Yangon	GSM
YGN00349	UTYGN00349A1	CID_414010035318151	96.20084	16.84618	South Okkalapa	Yangon	GSM
YGN00263	UTYGN00263A1	CID_414010035317521	96.18746	16.83289	Thingang kyun	Yangon	GSM

SITEID = BTS site identification number; STDIV = State and division

In CDRs data, only 3 data set: PID, DTIME and CELLID data from both voice call and data file are used in this analysis. These data are extracted by township because 33 townships are considered as the origin and destination datasets to extract trip production and attraction.

Table 5. Extracted by township (Yangon)

PID	DTIME	CELLID	CELL_TSMAP
PID_0016A1566800	20151203083404	CID_414010035618021	Lanmadaw
PID_0016A0566100	20151203181200	CID_414010030113432	Mingaladon
PID_0016A1566900	20151203124553	CID_414010035615138	Dagon

Table 6. Sorting by PID & DTIME

PID	DTIME	CELLID	CELL_TSMAP
PID_0016A0566100	20151203181200	CID_414010030113432	Mingaladon
PID_0016A1566800	20151203083404	CID_414010035618021	Lanmadaw
PID_0016A1566900	20151203124553	CID_414010035615138	Dagon

4.2.2 Processing of CDRs data

Cell phone usage of each region can be used to generate trip numbers of townships within certain time frame and time windows. For each mobile user, it is assumed that a trip is made between two consecutive calls occurring within a study period with different BTS. Trip generation is computed based on CDRs data, there is the limitation that is the trips maker do not use mobile phone so that the trips data do not available. The number of trips (flows) can be analyzed as the movement of people who travel from starting point in the origin region and ending in the destination region. Therefore time window is necessary to be defined for the construction of O-D matrices. Generally, time window intervals are defined according to a specific goal such as flow analysis during peak hours, slack periods, weekday flow or weekend flow analysis. In this study, the results are aggregated into average of seven days in 24:00 hour of time frame for the whole day data. Time frame will be 7:00 to 10:00 AM for morning peak and off peak is 1:00 to 3:00 PM. Every element of O-D matrix represents the number of trips (flows) from one origin area i to one destination j. Although interval of determined time window is 3 hour for the whole day and morning peak, off peak is 2 hour.

It is possible to trace location of people when they make trips through mobile phone data because this data set includes latitude and longitude of each BTS towers. Origin-destination trips can also be constructed by time of day and purpose: home-based work (HBW), home-based other (HBO), and non-home based (NHB).

NHB is defined as to which the user travels the particular places between the time frame of 10 am and 4 pm. The reason of choosing 10 am to 4pm is that it is working period. If the call logs (before 10 am) and the call logs (after 4pm) are observed same cell ID is defined as home location otherwise is work location. O-D matrix represents the mobility flows between the limited time frame and it is square matrices of townships for the study area. Every row and column of the matrices states that the trip production and attraction of the human mobility flows between townships in Yangon.

The following Figure 5 through 7 shows the trip production and attraction of the person trip between townships in study period.

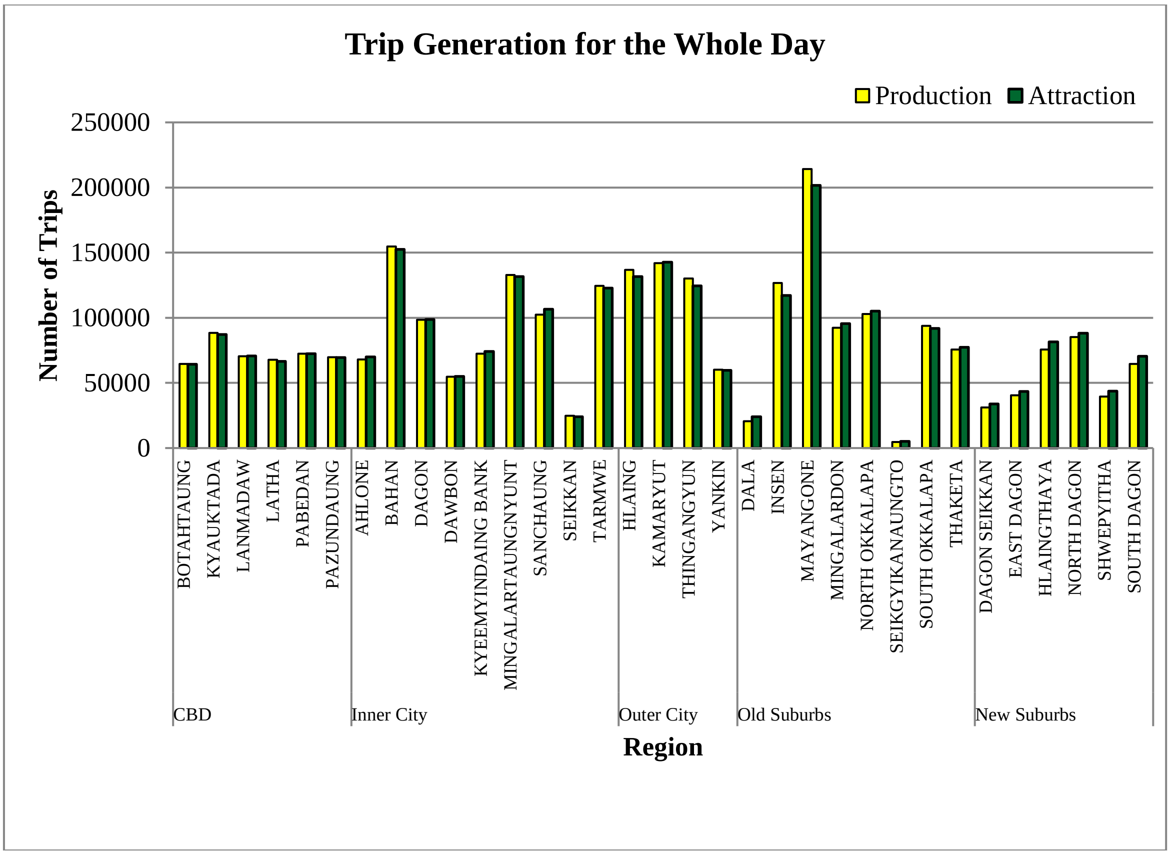

Figure 5. Current trip production and attraction of townships

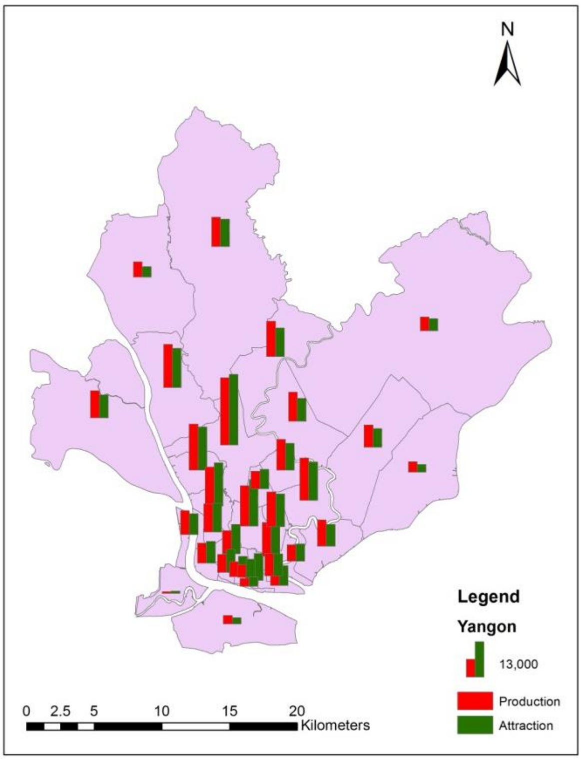

Figure 6. Trip production and attraction of township in 7:00 to 10:00 AM

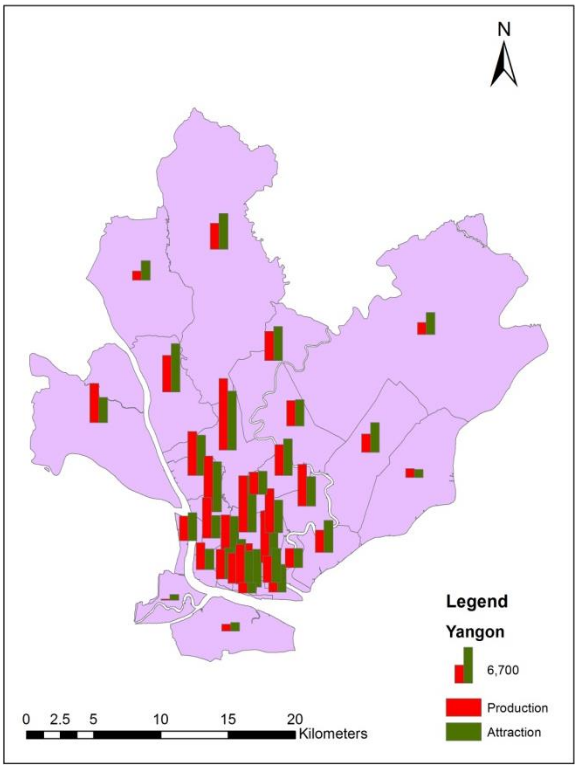

Figure 7. Trip production and attraction of township in 1:00 to 3:00 PM

In order to predict the people movement of the area, the relation between trip generation and socioeconomic factors needs to be analyzed. This analysis is carried out by using the following multiple linear regression models for 33 townships in Yangon City.

$$Y_i = A + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots + B_i X_i + E_i$$

Where,

Y = The number of person-trips that will begin from or end in each traffic analysis township

$X_{1i}, X_{2i}, \ldots, X_{ni}$ = Socio-economic variables

A, B_j = Parameters determined through calibration

E_i = Error term (in aggregate model depending on size of the zone)

The observed relationship between the socioeconomic factors and the volume of person trips are translated into regression models for future estimate of interaction among the township. This regression model assumed as linear because the number of observations greater than numbers of independent variables. Statistical Package for Social Sciences (SPSS) was used to carry out these analyses.

5. ANALYZING TRIP GENERATION

5.1 Trip Production Model

There are several independent variables which are used in trip production model. The independent variables used in this model are population, age group, car ownership (total number of car per township), and average income, number of household, number of student and number of worker.

Before developing the model, the correlation between the independent variables and the dependent variables are tested. It should have the strong correlation between the independent and dependent variable while the correlation among independent variables are rejected greater than 0.70. The value of VIF is chosen less than 2.0. It can be measured the 95% confidence interval and Durbin-Watson. Therefore, the estimated regression can be identified linear regression for the above solution.

The regression analysis was conducted several times. In each stage, the regression equation was evaluated according to statistical measures. The iteration was made by reducing the independent variables which has the lowest correlation to the dependent variables. The results of the analysis are shown in Table 2.

Table 2. Results of regression analysis for trip production

Variables	Model 1			Model 2			Model 3		
	Coeff:	t	p	Coeff:	t	p	Coeff:	t	p
Constant	46249.27	3.96	0.00*	4211.96	3.54	0.00*	-5299.10	-2.58	0.02*
Independent Variables									
Population		NA			NA			NA	
Age Groups Age 0-14 Age 15-26 Age 27-64 Age > 65+		NA			NA			NA	
Car Ownership of township	10.65	3.94	0.00*	1.32	4.79	0.00*	0.43	2.66	0.01*
Average income			NA			NA	0.04	4.67	0.00*
Household			NA			NA			NA
Student			NA			NA			NA
Worker			NA			NA			NA
R^2	0.33			0.43			0.54		
Adjusted R^2	0.31			0.41			0.51		

Note (*) : significant at 95% level; NA : Not Available

The model 1 is the estimation of trip production per day. Model 2 is taken for the trip production during morning peak period from 7 to 10 am while the model 3 is estimated for the off peak period which is from 1 to 3 pm.

In model 1, the relationship shows the positive relationship between the dependent variable and independent variable. This means that if the car ownership increases, the trip production will also increase. This model has Significant test (t test) of individual variable gives t test value 3.94 (p value = 0.00). These t statistic and p value are significant for 5%

confident level. The coefficient of determination is 0.334 implying that 33.4% of the total variance is explained by the car ownership.

In model 2 and 3, it can be seen that income is significant in off peak period while it is not significant in morning peak. When the income is higher, it produces more trip production because these people can go shopping and recreation centers. But the most effective variable is car ownership in both models. The higher income persons are travelling more trips because the level of car ownership is higher than the other.

5.2 Trip Attraction Model

The number of trip attracted to a township is influenced by several socioeconomics attributes. These attributes are used as the independent variables in establishing trip attraction model as follow;

i. Number of School, which are the total number of school in a township. The schools consist of primary school, middle school and high school.
ii. Number of University
iii. Number of Office shows the number of total offices include government offices in a township.
iv. Number of Hotel
v. Recreation Facilities; consists of parks, sports centres, game centres, restaurants)
vi. Religious Places, which include the pagodas, monasteries, church, mosque, etc.
vii. Industry; It is stated as the total number of primary, secondary and tertiary industries in a township.
viii. Hospital consists of private and public hospitals and other medical clinics.
ix. Shopping Centres, which shows the total number of shopping centres and super markets in a township.
x. Number of markets, shows the total number of markets and bazaars.

As the production model, the correlation test is made between dependent and independent variables. Like in trip production model, several trip attraction models were developed using some combinations of independent variables. The results for the trip attraction models are shown in Table 3.

Table 3. Results of regression analysis for trip attraction

Variables	**Model 1**			**Model 2**			**Model 3**		
	Coeff:	**t**	**p**	**Coeff:**	**t**	**p**	**Coeff:**	**t**	**p**
Constant	40462.93	5.25	0.00*	4143.91	4.48	0.00*	4029.36	9.84	0.00*
Independent Variables									
School			NA			NA			NA
University			NA			NA	466.30	2.09	0.05*
Office	2403.71	7.10	0.00*	290.16	7.13	0.00*	135.62	5.85	0.00*
Hotel			NA			NA			NA
Recreation			NA			NA			NA
Religious			NA			NA			NA
Industry	305.36	3.71	0.00*	28.42	2.88	0.01*			NA
Hospital			NA			NA			NA
Shopping Centre			NA			NA			NA
R^2	0.67			0.65			0.57		
Adjusted R^2	0.59			0.63			0.54		

Note (*) : significant at 95% level; NA : Not Available

In above three models, it can be observed that the number of office is found to be the most significant variable among the independent variables and have positive relationship. In model 1 and 2, the number of industry is found to be second influence factors but in model 3, the influence factor is the number of university. The coefficients of determination for model 1 and 2 are about 70% while in model 3, the coefficient of determination is about 60%.

6. CONCLUSION

This paper focused on the estimation for the number of person-trips that will produce from or attract to an area in Yangon city by using CDRs data. The relationship between person trip and socioeconomic factors are analysed by using multiple linear regression model. In trip production model, it can be observed that the car ownership is the most significant variable.

In trip attraction model, it is found that the number of office is the most significant and influential variable among other variables. After the number of office, the number of industry is the second most significant variables. These two variables are significant for the whole day model and the peak hour model. But in off peak period, the number of university is found to be significant due to the staggered hour of University time.

This prediction model is based only on the CDR data and magnification factor is not applied to that model. Therefore, it is necessary to validate this model with the real time survey. As for the future studies, actual person trip survey for the selected areas will be collected for validation. The remaining three steps of traffic demand forecasting such as trip distribution, modal split and traffic assignment will be done for the prediction of future traffic.

This information is very useful for urban transport planning and especially for public transport planning to improve the deteriorating of urban transport situation. In urban transport planning, it can know how people come and go and can help to determine where to deploy infrastructure such as public transit stations, improvement of public transit systems, public transport scheduling, and networking. If the person trips from CDRs data can be changed to the vehicle trips, the traffic congestion mitigation plan can be provided in the busy area. CDRs data are used for monitoring in real-time the traffic conditions. Therefore, it enables and to understand how many people are going in which direction at what time. It can improve disaster management and emergency preparedness by estimating the location of vulnerable people against disasters at the occurrence of disastrous events. Moreover, transport planner can manage which route is suitable for detoured person to go back to their home.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to Myanmar Post and Telecommunications (MPT) and Science and Technology Research Partnership for Sustainable Development (SATREPS-JICA) Project for data provided of this research. Special thanks are also dedicated to all my Professors for their invaluable and contributions throughout for this research.

REFERENCES

J, Donner. (2007) The use of mobile phones by micro entrepreneurs in Kigali, Rwanda: Changes to social and business network. *Information Technologies and International Development,* 3(2).

H, Kwon and L, Chidambaram. (2000) A test of the technology acceptance model: The case of cellular telephone adoption. *In Proceedings of the 33rd Hawaii International Conference on System Sciences.*

N, Eagle. (2008) Behavioural inference across cultures: Using telephones as a cultural lens. *IEEE Intelligent Systems,* 23:4:62–64.

J,Blumenstock and N. Eagle.(2010) Mobile divides: Gender,socioeconomic status, and mobile phone use inRwanda. *In Proceedings of the 4th International Conference on Information and Communication Technologies and Development*.

Filippo Maria Bianchi , et al. (2016) Identifying User Habits Through data Mining on Call data records.Department of Information Engineering, University of Rome, Italy.

SanjiScepanvonic, et al. (2015) Mobile phone Call data as a Regional Socio-Economic Proxy Indicator. Department of Computer Science, Alto University , Helsinki, Finland.

Vanessa Frias Martine, et al. (2012) On the Relationship Between Socio-Economic Factors and Cell phone Usage.Telefonia Research Madrid, Spain.

SantiPhithakkitunkoon, et al. (2012) Socio- Geography of Human Mobility: A Study Using Longitudinal Mobile Phone Data. Cultural Lab, School of Computing Science, Newcastle University.

Weifeng Li, et al. (2014) A Framework for Spatial Interaction Analysis Based on Large-Scale Mobile Phone Data. Key Laboratory of Road and Traffic Engineering of the Ministry of Education, 4800 Caoan Road, Tongji University ,Jiading District, Shanghai 201804, China,

L. Alexandera, S. Jiangb, M. Murgaa, M. C. Gonzáleza. (2015). *Origin–destination trips by purpose and time of day inferred from mobile phone data*. Transportation Research Part C: Emerging Technologies.
http://dx.doi.org/10.1016/j.trc.2015.02.018

Y. Hasegawa, Y. Sekimoto, T. Kashiyama, H. Kanasugi. (2014) *Transportation Melting Pot Dhaka:Road-link based traffic volume estimation from sparse CDR data*. The 1st International Conference on IoT in Urban Space (Urb-IoT 2014), Rome.

M. Kurokawa, T. Watanabe, S. Muramatsu, H. Kanasugi, Y. Sekimoto, R. Shibasaki. (2013) *Extracting people's stays from cellular network data*. NetMob 2013, 2013.05.

M. Ozer, I. Keles, I. H. Toroslu, P. Karagoz. (2013) *Predicting the change of location of mobile phone* users. MobilGIS2013, pp. 43-50, 2013.

M. Ozer, I… Keles, I. H. Toroslu, P. Karagoz, S. Ergut. (2014) *Predicting the next location change and time of change for mobile phone users*. MobiGIS2014, pp. 51-59.
UN Global Pulse, Mobile Phone Network Data for Development, http://www.unglobal pulse.org/sites/default/files/Mobile%20Data%20for%20Development%20Primer_Oct 2013.pdf, 2013.10.

Propper, C., Diamiani, M., Leckie & G., Dixon, J. (2007) Impact of patient's socioeconomic status on the status on the distance travelled for hospital admission in the English National Health Service" in Journal Health Serv. Res. Policy; 12: 153-159.

Carlsson- Kanyama, A & Liden, A. (1999) Travel patterns and environmental effects now and in the future: implications of differences in energy consumption among socio-economic groups" in the Economics 30(3), pp 405-417