ASEAN-Japan Seminar on Utilization of Mobile Phone Big-Data(MBD), Jan. 16, 2019

Utilization of MBD analysis for transport policy in Asia

Yoshihide SEKIMOTO, Associate Prof.,
Institute of Industrial Science, University of Tokyo

About me

- 2002
  - PhD, University of Tokyo
- 2002-2007
  - National Institute for Land, Infrastructure and Management (NILIM), MLIT
- 2007-
  - Univ. of Tokyo

People flow project (PFLOW)

http://pflow.csis.u-tokyo.ac.jp (2008-)

What's People Flow Project

Recently, monitoring dynamic changes in people flow has become necessary, in order to mitigate secondary disasters following earthquakes, fires or other major events, as well as to mitigate congestion at nodes in terminal stations. For example, 247 people were killed and injured while rushing to a pedestrian bridge at a 2001 fireworks event in Akashi, Japan, where spectators numbered about 150,000. Moreover, daily ridership comes to about 4 million people in Shinjuku Station, one of the most crowded stations in Tokyo. From the point of view of public facility managers, it is necessary to grasp the people flow comprehensively, for instance, in order to design safe and comfortable spaces, and appropriate urban transport policies. In commercial fields of outdoor advertisement, price systems, which support an effective advertising activity, depend on the traffic volume of people for each location.

In technical terms, tracking mobile objects by GPS or PHS, tracking the number of people who are stationary by CCTV camera, tracking the number of passengers getting on and off according to the number of IC (integrated circuit) tickets through the automatic ticket gates, tracking the number of people who are stationary by the number of registered mobile phones at each base station, and tracking the hourly number of visitors to department stores enables us to measure people flow according to various dimensions.

However, the scope of many of these goes no further than data acquisition technology. Such research cannot be seen as infrastructure data that can aggregate the acquired data and provide an overview of the mass flow. This is true in terms of the comprehensive qualities including spatial/temporal accuracy, acquisition/process cost and value to the user as a service.

Therefore, we start "People Flow Project (PFLOW)" which overviews data process technology, data quality and its common infrastructure for people flow on a large scale. Moreover, we provide spatio-temporal data processing service for all researchers through our platform named "People Flow Analysis Platform (PFLOW-AP)".

- Top
- Data provision service
  - Data provision service
  - Bundled download service
- Data cleaning service
  - Spatial interpolated service
  - GPS data cleaning service
- People Flow Analysis Platform
- Application cases
  - Sample data
  - Examples of joint research
  - Visualization examples
  - Published papers
- Registration for use
- Inquiries

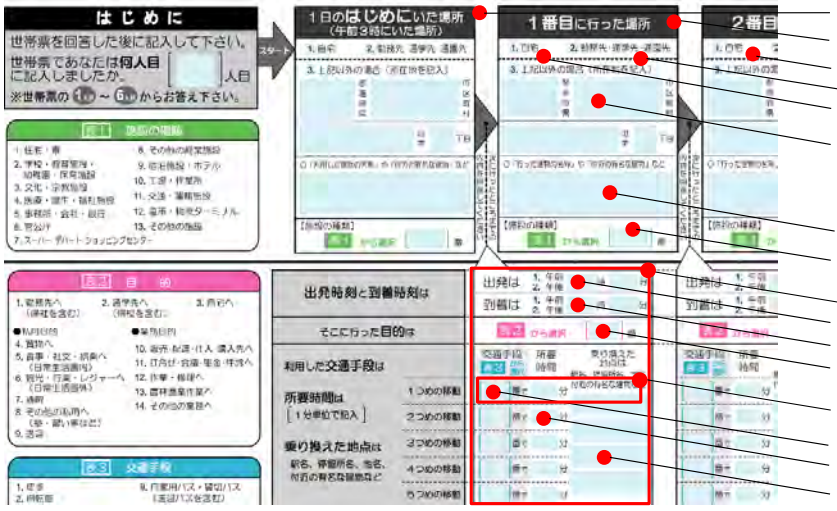
Key factor is rough understanding in whole target area !

2. PFLOW from PT (2008～)

At the first stage…

PFLOW from PT（2008～）

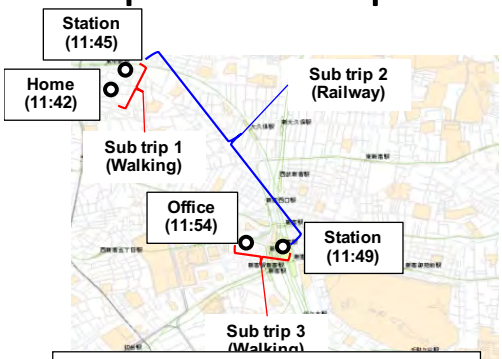
Person trip survey ??



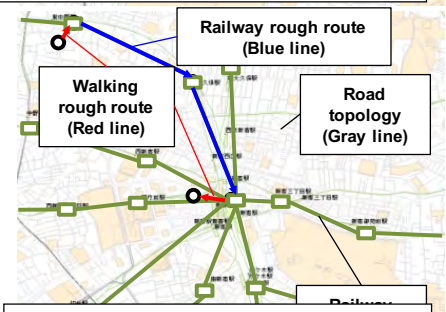
Place staying at 3:00 in the morning
1st place to go
2nd place to go
Home　Office or School
Other places (Rough address)
Other places (Place name)
Kind of place
Trip
Departure time
Arrival time
Purpose
Sub-trip
Transportation mode
Travel time
Transfer point

Main part of PT survey sheet (from the Tokyo Metropolitan Region Transportation Planning Commission web site "http://www.tokyo-pt.jp/data/file/tebiki.pdf")

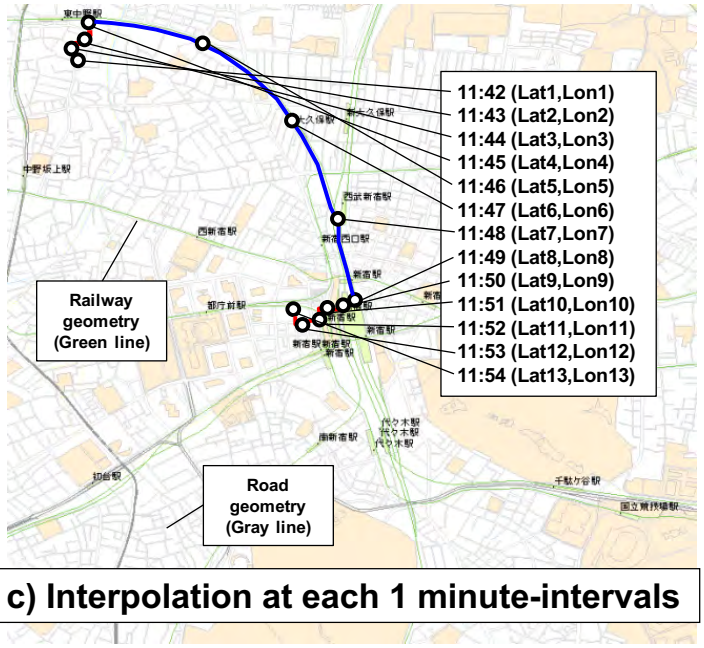
Spatio-temporal interpolation from OD data



a) Geocoded OD of each sub trip



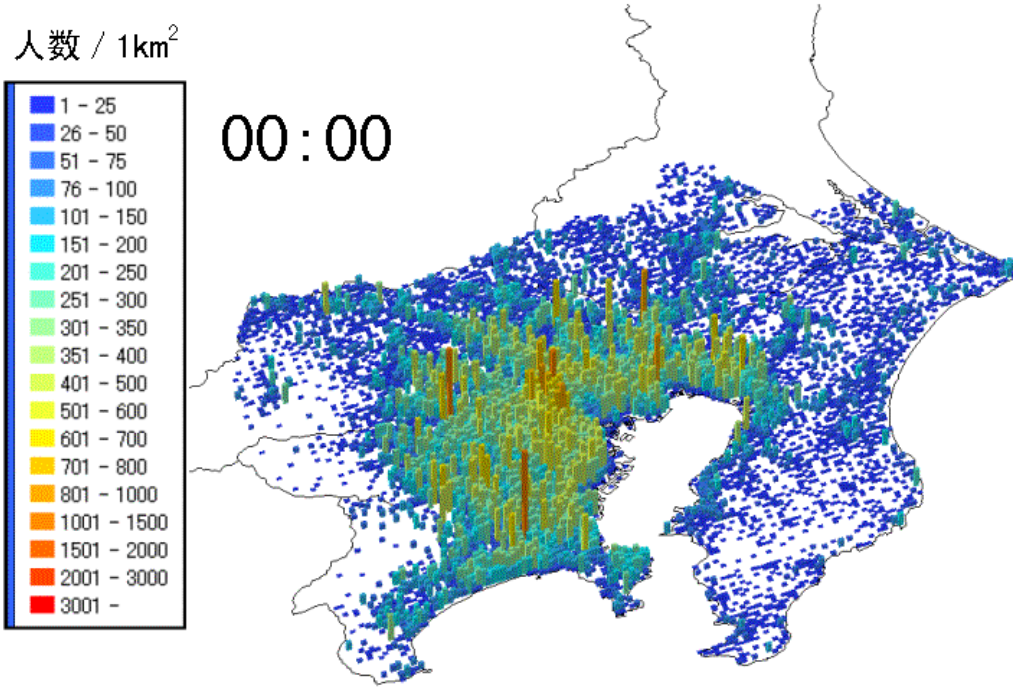
b) Route choice along road/railway topology



c) Interpolation at each 1 minute-intervals

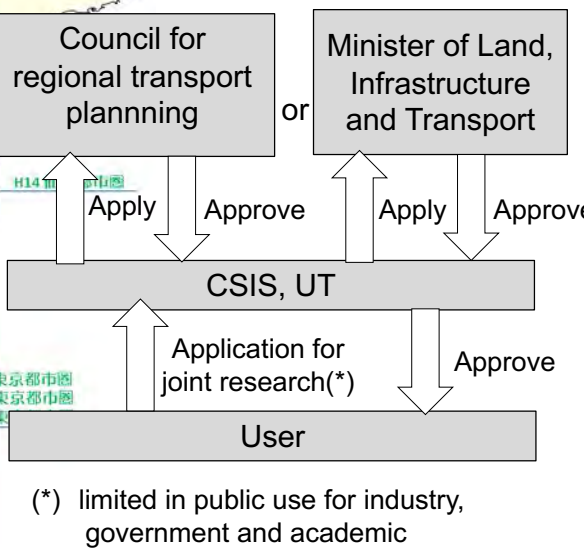
* Y. Sekimoto et al. PFLOW: Reconstruction of people flow by recycling large-scale fragmentary social survey data, *IEEE Pervasive Computing*, Vol.10(4) pp.27-35, 2011.

3D visualization



3D visualization with 1-km2 mesh

Archiving of PFLOW
(14 metropolitan areas and totally 3.5 million)



Many joint researches through "People Flow Data Set"

【Transportation】

Research on improving the efficiency of urban transport systems using portable personal mobility.(iTransport Lab, Ltd.)

A simulation of tourist flow patterns in the Sendai metropolitan area using the People Flow Analysis Platform. Masayoshi Tanishita (Chuo University)

Utilization of statistical data in urban transport planning. (Ritsumeikan Asia Pacific University Department of Asian Pacific Studies)

【Spatio-temporal analysis】

Detection of patterns in travel routes using position information and travel times (Kobe University Graduate School of Engineering)

Development of a spatio-temporal data model for analysis of spatio-temporal behavior using GIS. (Tokyo Metropolitan University)

【Risk analysis】

A model for the transmission of novel infectious diseases.

(University of Tokyo Institute of Industrial Science)

An investigation of of disaster risk using GIS. (Aichi Institute of Technology Department of Environmental Engineering)

【Personal information and security】

On the anonymization of personal information and its two-dimensional use (Information Grand Voyage Project). (Mitsubishi Research Institute, Inc.)

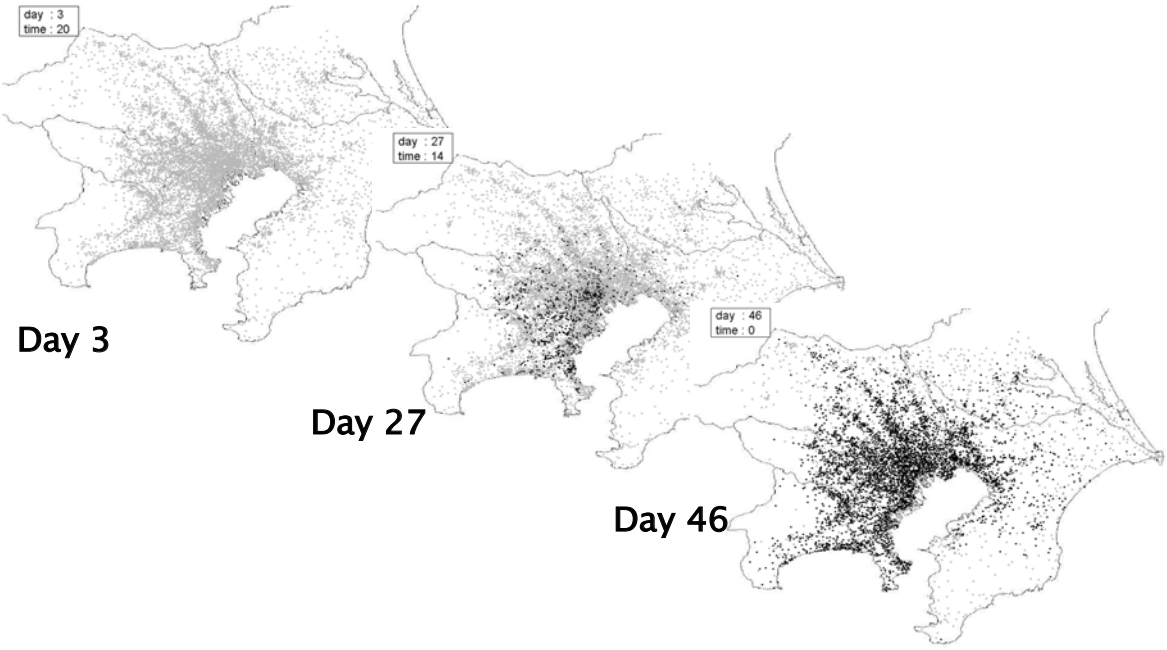
【Environment】

Development of a scenario for fine spatial output and changes in land use via unified system analysis. (National Institute for Environmental Studies)

【Marketing】

A study of consumer respiration models using person-trip data. (Fine Analysis, LLC)

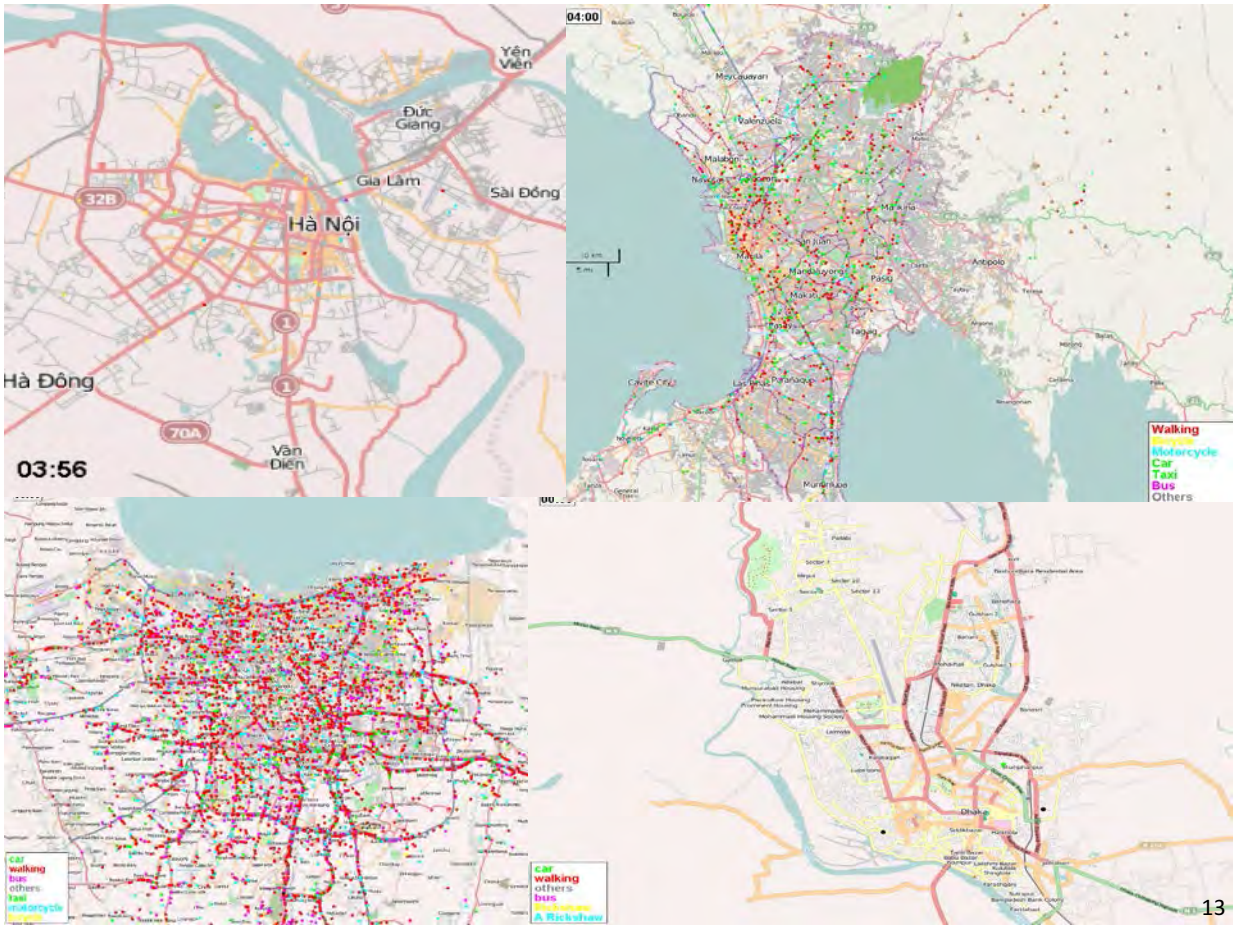
Pandemic simulation based on PFLOW



Aihara & Suzuki lab in IIS, Univ. of Tokyo

JICA-PT data

City	Population (million)	Survey year	Sample size	Number of trips	Ratios of various modes of main transportation (2-wheeler/car/taxi/bus/rail)
Manila	9.45	1996	231,889	471,035	2%/10%/25%/58%/4%
Kuala Lumpur	1.39	1997	80,560	218,460	29%/44%/2%/23%/2%
Damascus	3.08	1998	38,490	81,698	4%/25%/15%/56%/0%
Managua	1.20	1998	24,854	54,138	2%/25%/4%/69%/0%
Bucharest	2.15	1998	67,509	143,311	0%/19%/0%/27%/54%
Phnom Penh	1.15	2000	18,664	40,369	89%/11%*1/-/-
Chengdu	3.09	2000	31,188	70,199	81%/10%/4%/4%/0%
Belem	1.78	2000	24,043	59,529	15%/13%/2%/70%/0%
Jakarta	2.10	2000	423,237	1,083,280	2%/0%/42%/56%/0%
Tripoli	0.33	2001	3,608	7,615	35%/29%/17%/19%/0%
Cairo	14.4	2001	136,070	268,360	2%/13%/46%/29%/10%
Ho Chi Minh City	3.18	2002	27,412	71,890	96%/2.4%*1/1.8%/-
Hanoi	7.16	2004	63,716	188,949	89%/3.6%*1/6.7%/-
Nairobi	4.04	2004	20,980	46,828	2%/29%*1/68%/1%
Lima	8.04	2004	115,728	270,384	17.2%*2/13.7%/69.1%/-

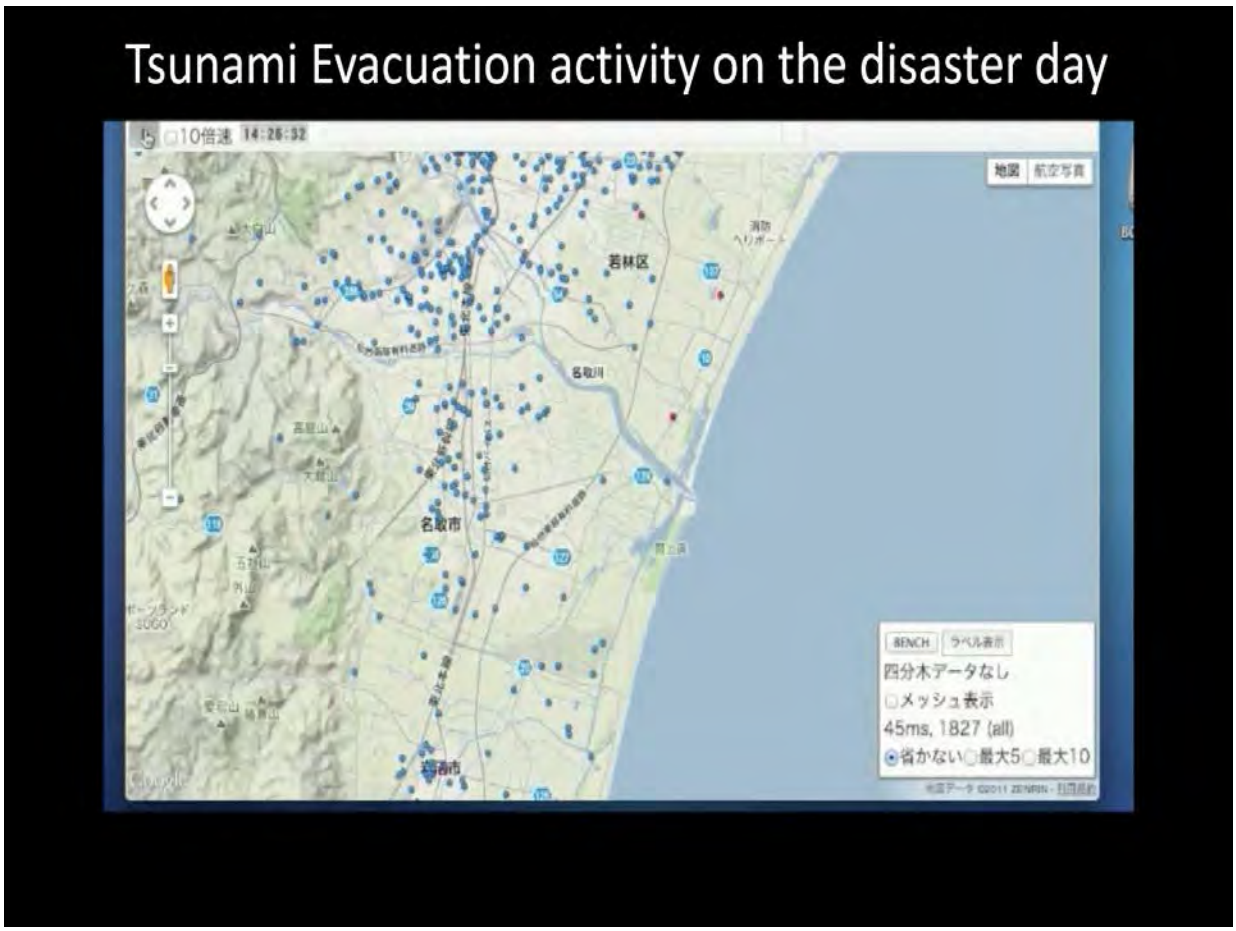


For the data set preparation…

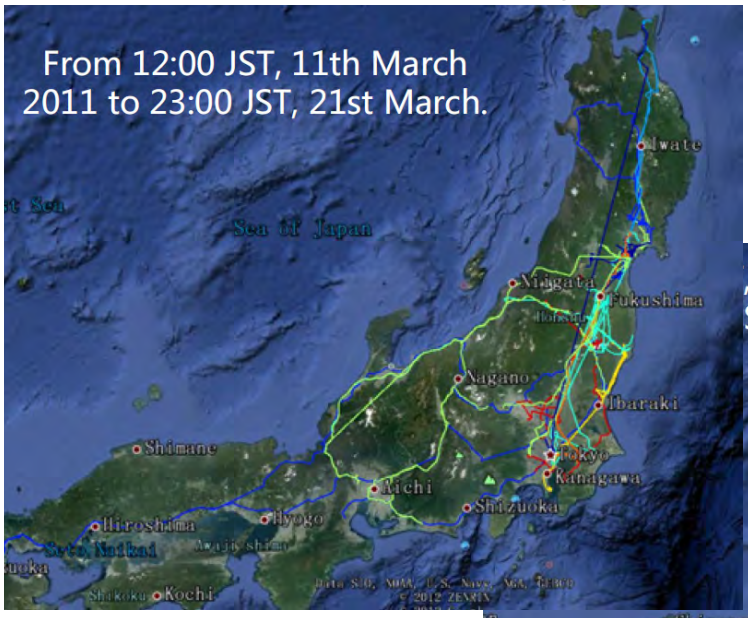
In case of Tokyo metropolitan area

Processing items	Calculation time
Data cleaning (partly manual)	Several days
Geo-coding from zone to lat/lon	Several days
Spatio-temporal interpolation	84 days
Creation of query tables for DB	14 days

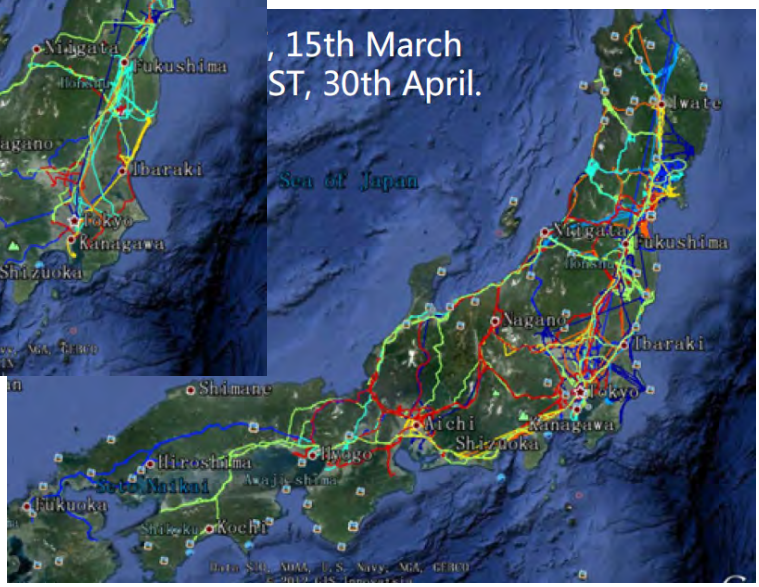
3. PFLOW from mobile phone GPS（2011～）



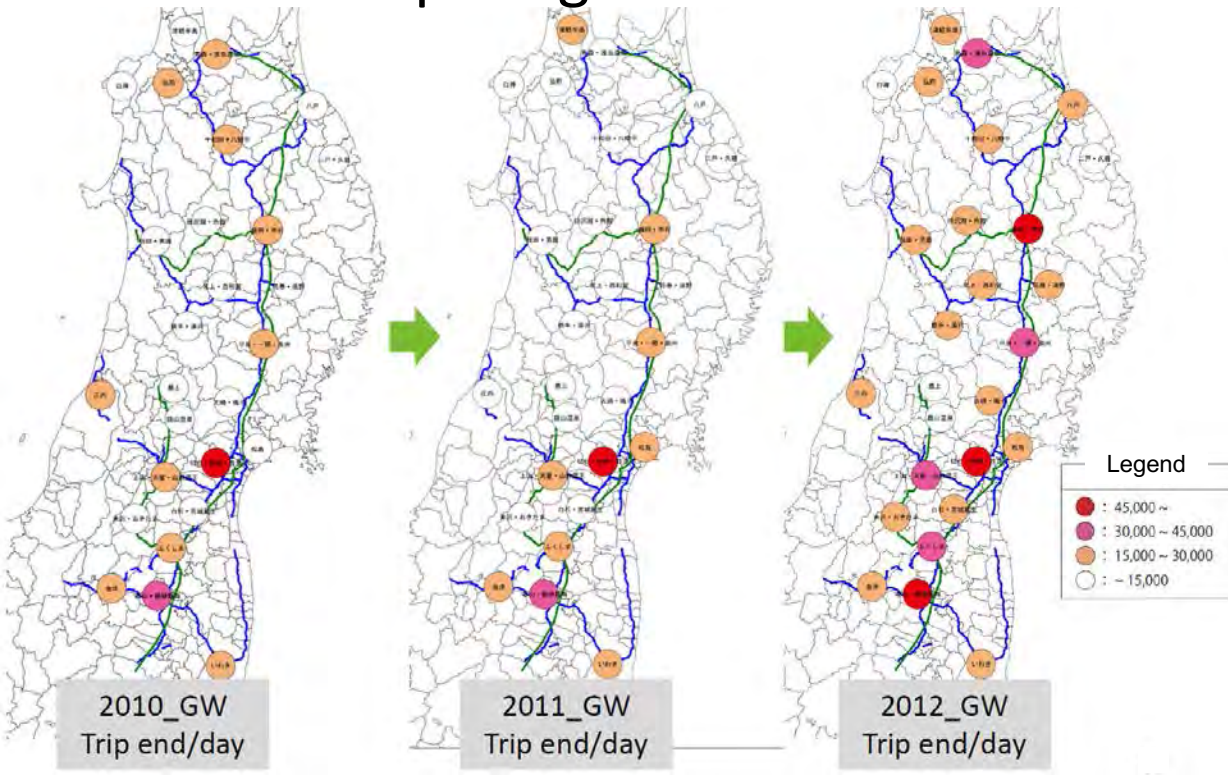
Evacuation activity after the earthquake



From 15th March to 30th April



Understanding of recovery status comparing tourist flow

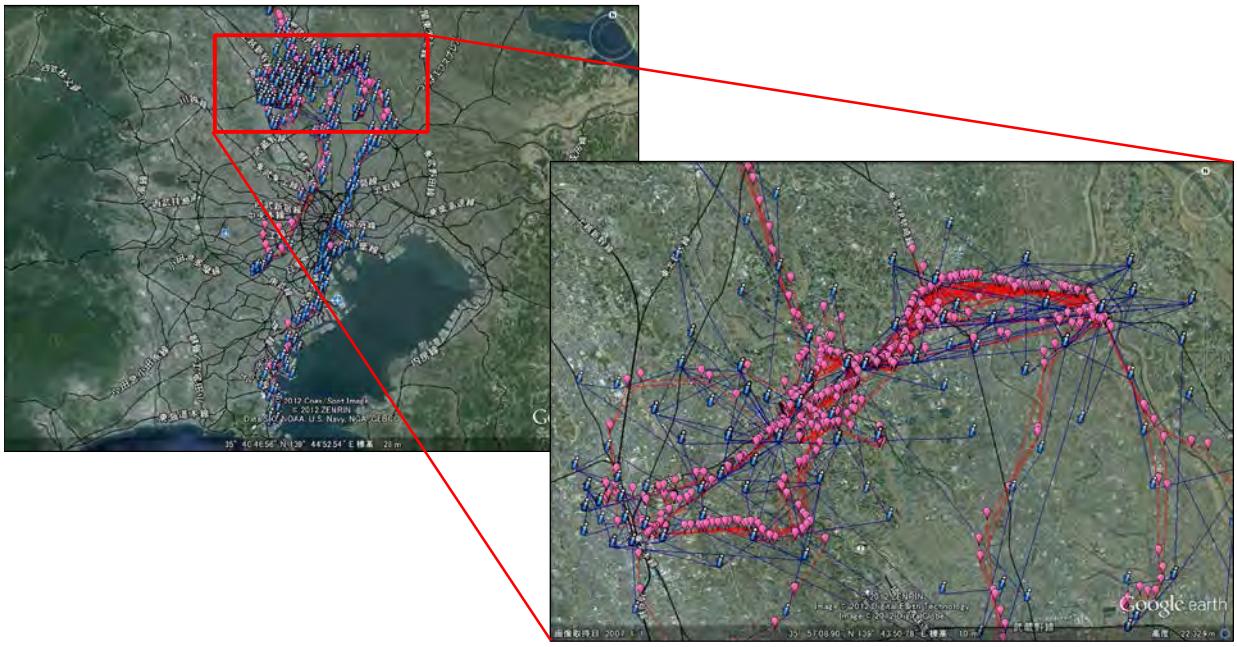


“Density map” from Auto-GPS data

Density map from Auto-GPS data
（ZENRIN DataCom CO.,LTD. http://lab.its-mo.com/densitymap/）

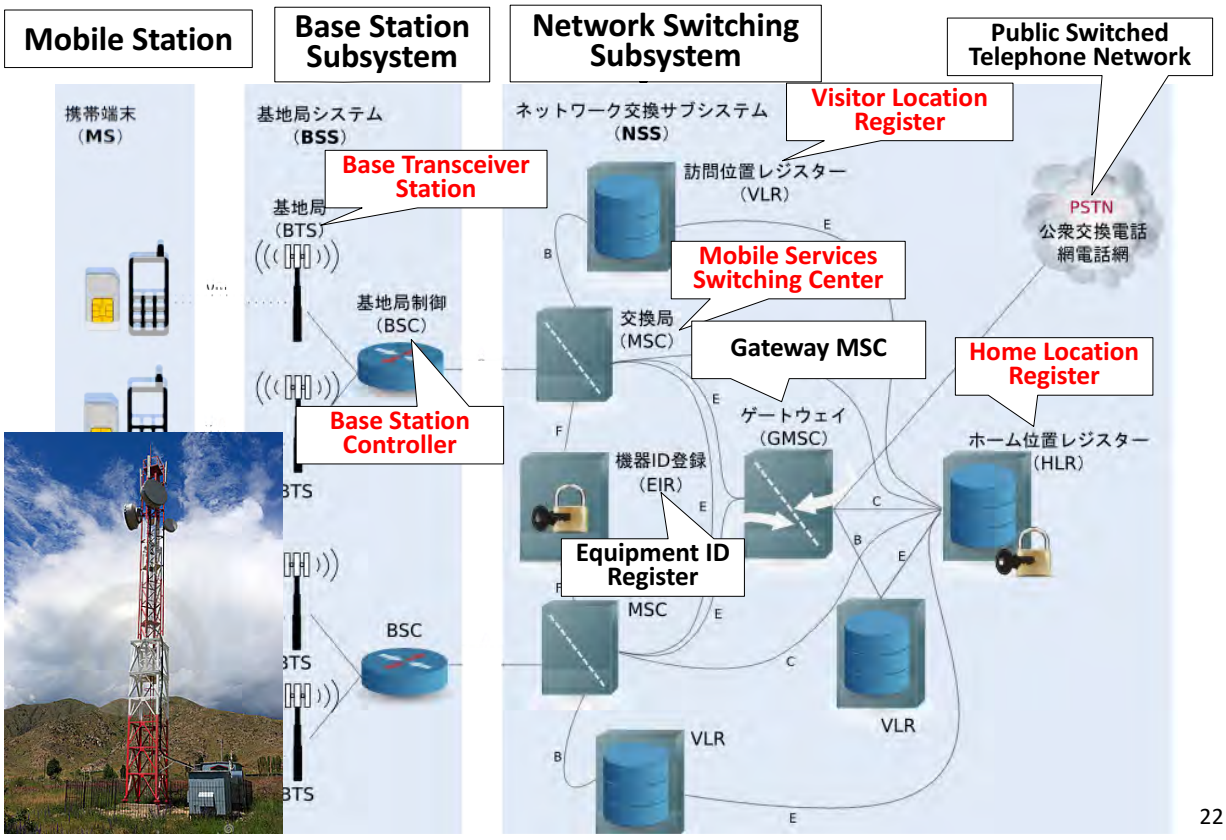
4. PFLOW from mobile phone CDR（2012～）

Call detail record (CDR) from mobile phone base station



Comparison of GPS (pink) and CDR (blue)

Mechanism of mobile phone base station



Mobile base station data

Anonymized CDR (Call detail record data) per person

pid	time	lat	lon	err1	err2	method	code
f00d09160971f089c928242e195c1d5c	2009-10-01 03:38:50	-71.7971	42.26691	330	185	2	800
f00d09160971f089c928242e195c1d5c	2009-10-01 03:39:34	-71.7975	42.26697	318	158	2	508
f00d09160971f089c928242e195c1d5c	2009-10-01 03:39:52	-71.7979	42.26625	394	152	2	508
f00d09160971f089c928242e195c1d5c	2009-10-01 03:40:17	-71.7951	42.26833	298	310	1	508
f00d09160971f089c928242e195c1d5c	2009-10-01 03:41:20	-71.7963	42.2668	367	212	2	508
f00d09160971f089c928242e195c1d5c	2009-10-01 05:26:18	-71.8037	42.27411	549	252	3	508
f00d09160971f089c928242e195c1d5c	2009-10-01 05:36:01	-71.8463	42.18137	460	350	1	508
f00d09160971f089c928242e195c1d5c	2009-10-01 05:36:43	-71.8398	42.17625	1144	428	1	508
f00d09160971f089c928242e195c1d5c	2009-10-01 05:48:05	-71.8618	42.05578	2164	417	1	508
f00d09160971f089c928242e195c1d5c	2009-10-01 05:49:25	-71.8772	42.05683	2568	575	1	508

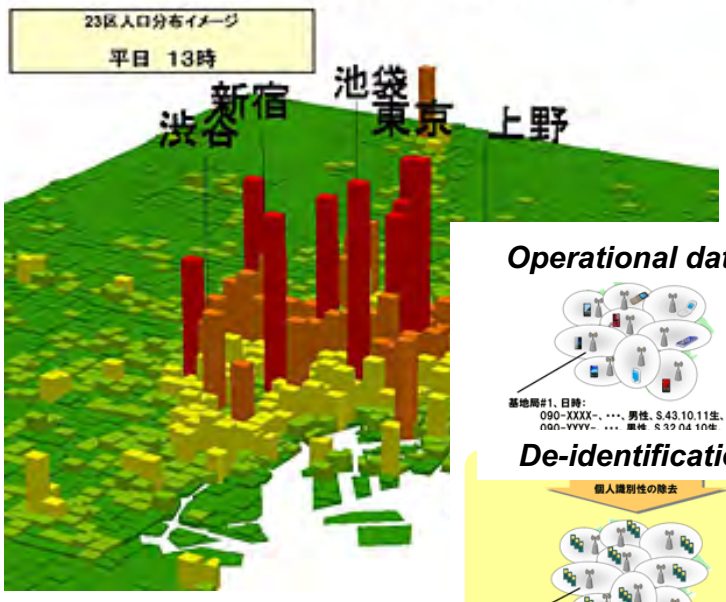
Aggregated as erlang value per each cell station of base station

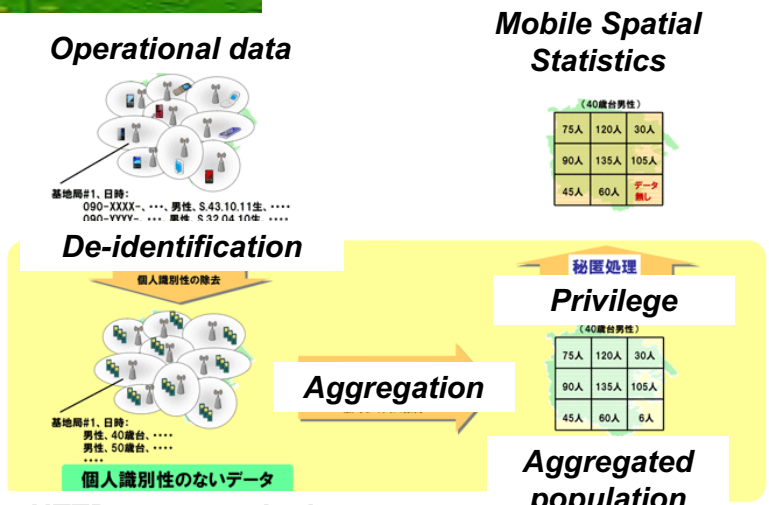
Erlang is the total amount of *resources* used over a period of time

cellid	lat	lon	Start_time	erl
BKKC1	13.75697	100.5594	2008/03/01 9:00	33.98
BKKC2	13.75697	100.5594	2008/03/01 9:00	18.93
BKKC3	13.75697	100.5594	2008/03/01 9:00	33.17
PTWA1	13.75138	100.5402	2008/03/01 9:00	20.75
PTWA2	13.75138	100.5402	2008/03/01 9:00	17.93
PTWA3	13.75138	100.5402	2008/03/01 9:00	33.07

“Mobile Spatial Statistics” from base station data

モバイル空間統計イメージ：東京23区周辺の人口分布





Mobile Spatial Statistics

〈40歳台男性〉

75人	120人	30人
90人	135人	105人
45人	60人	データ無し

Mobile Spatial Statistics（From NTTDocomo web site: http://www.nttdocomo.co.jp/corporate/disclosure/mobile_spatial_statistics/）

Commercial mobility data (in Japan)

- Personal information is not linked under the personal information revised law (2015)
  - Grid-based
    - ZENRIN DataCom (Hourly), NTT Docomo (Hourly), Agoop
  - Link-based
    - Navitime, Pioneer
  - Dot-based
    - Agoop
  - OD-based

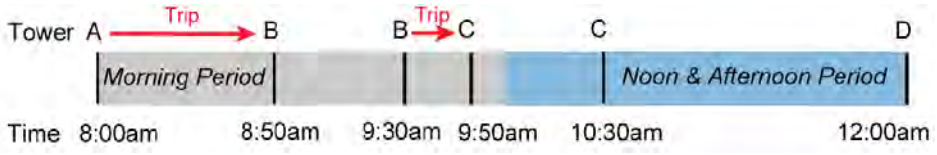
Joint research with Bangladesh

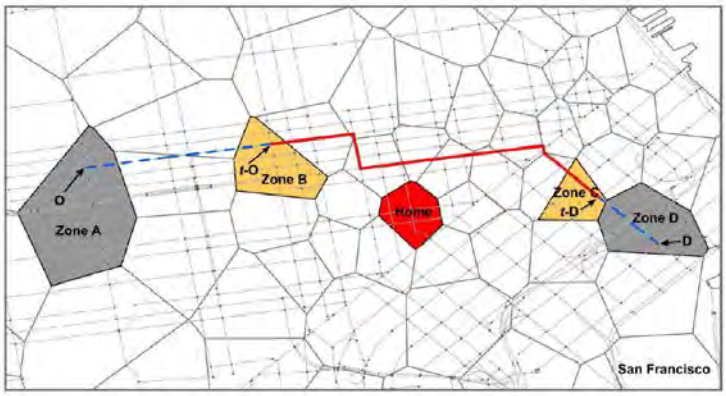
People distribution based on mobile phone base station in Dhaka (provided by GrameenPhone)

Period	2012/6/19～7/18
# of user	6.85 million
# of mobile phone base station	1,362 (1125 in DMA)
Target data	Anonymized individual time and station ID when talk occur
# of average log	**5.8 / 1day** (The # of target is 4 million people whose home can be extracted)

Sparse !!

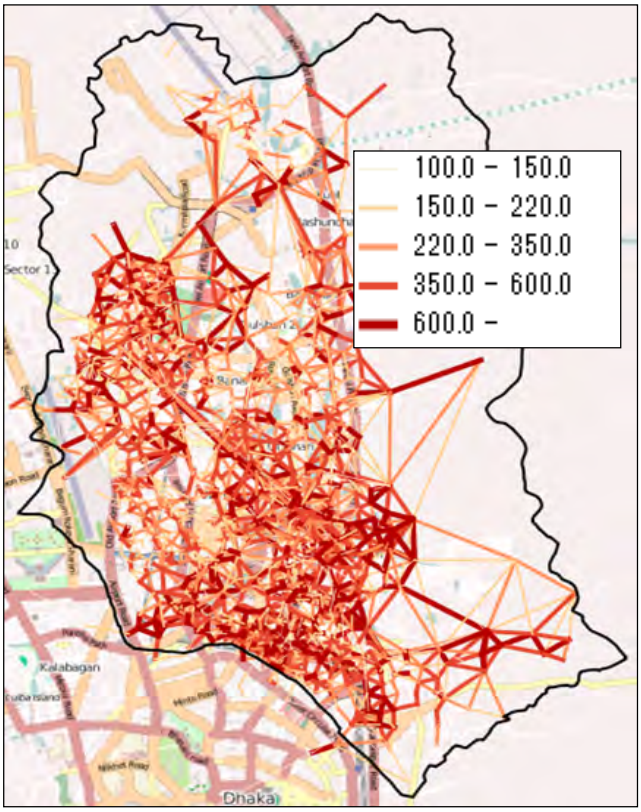
Trips accumulating sparse CDRs data



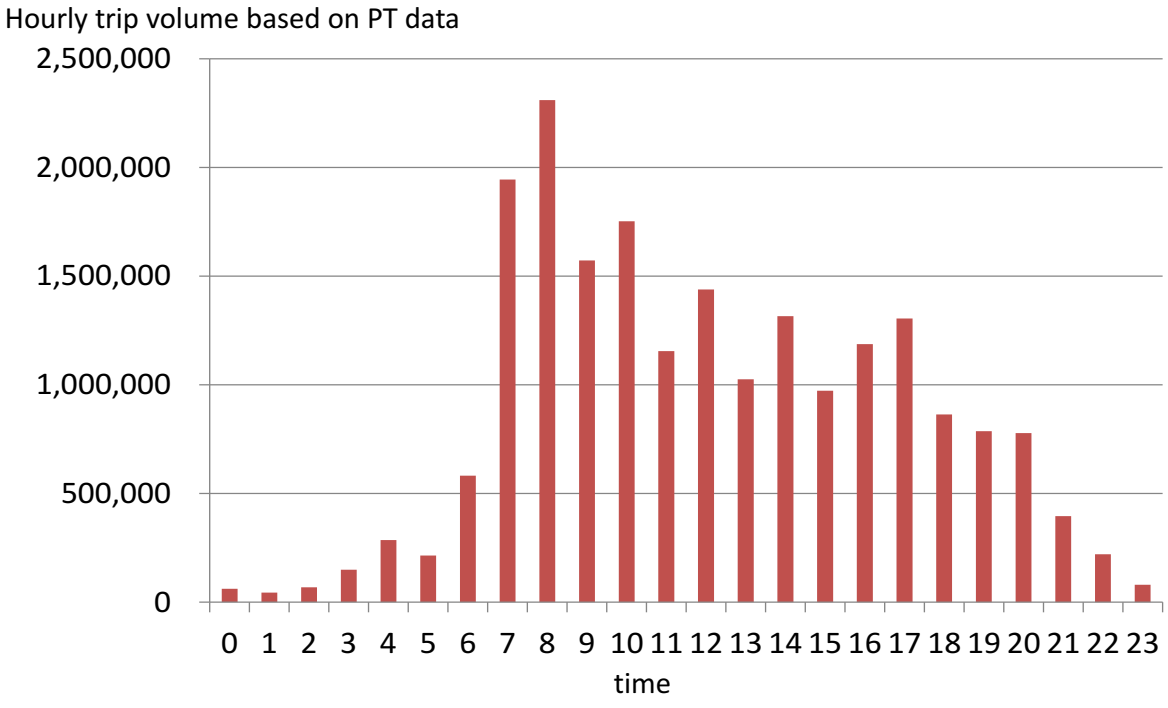


Target (30 days in total)	# of trip pattern	# of trip	# of user	# of user whose trip can be extracted
All users who have CDR	695,637	77,798,677	6,854,189	4,790,888
Users whose home can be extracted	677,670	70,014,818	3,995,353	3,311,932
Users limited in related Voronoi area	647,469	55,383,314	2,877,197	2,430,487
Magnified based on real populations	647,469	172,843,289	7,597,256	-
Vehicle users	619,219	126,921,796	-	-

OD table of trips



Past hourly link traffic volume based on PT survey data (questionnaire data)



Simple route assignment based on "ITA method"（Chen etc., 1991）

- Divide traffic volume into 40%, 30%, 20% and 10% blocks and calculate trip time (t_a) and update trip time

$$t_a = t_f\left(1 + \alpha(VOC)^{\beta}\right) \quad \alpha = 0.15, \beta = 4$$

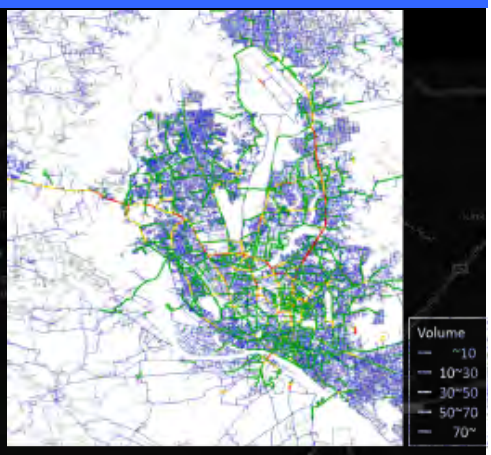
t_f: **Trip time in free flow**

t_a: **Trip time based on congestion (VOC**

VOC: **congestion**: $\frac{\textbf{link traffic volume}}{\textbf{road capacity}}$

Target of our traffic and people flow group

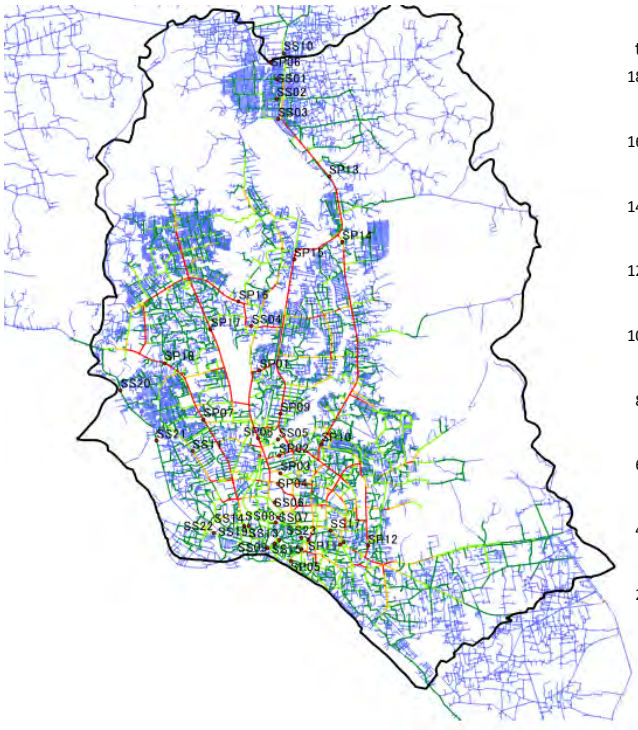
Develop dynamic social observation and assessment system with mobile sensor in order to understand actual condition flow of traffic and crowd

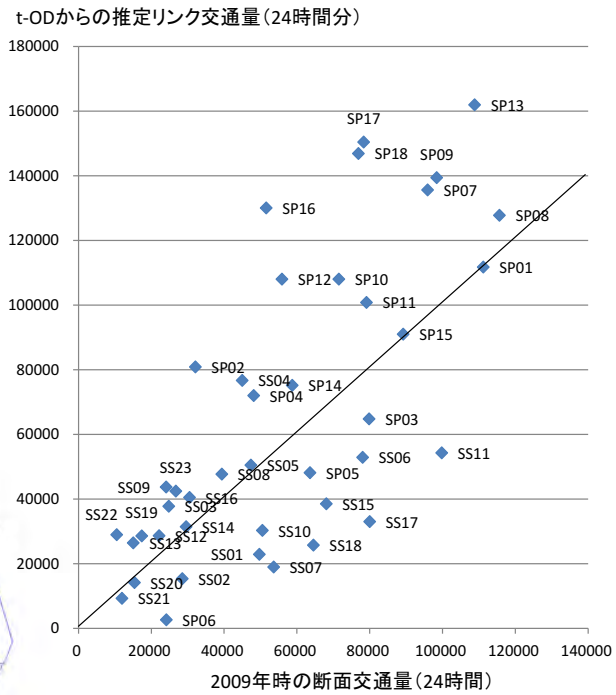


Calculate link traffic volume from cell-phone base station

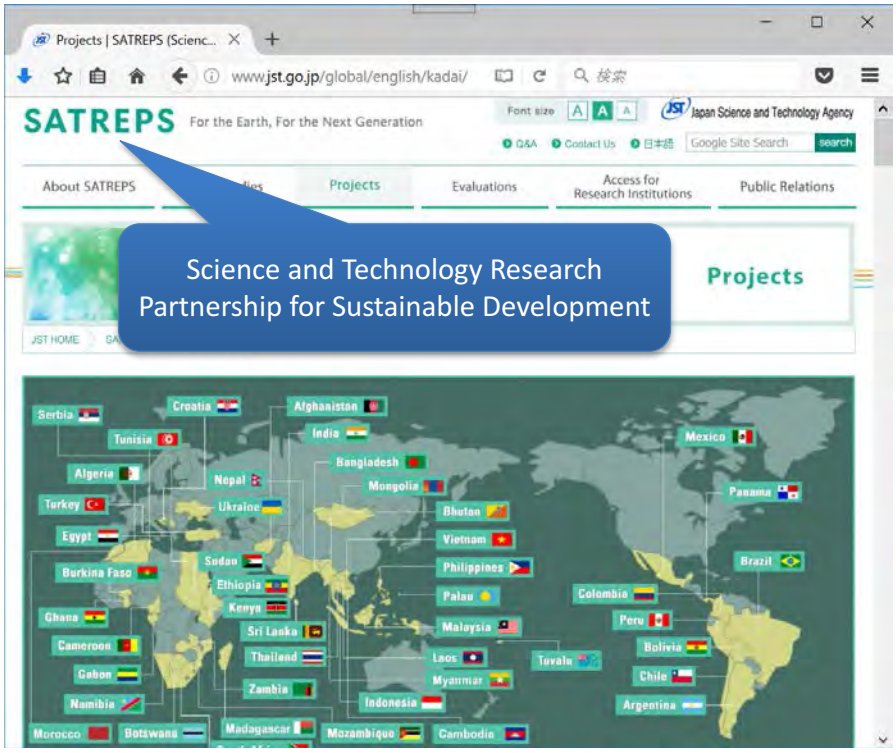
Cell-phone communication information in Dhaka by GrameenPhone (Base station's CDR: Call Detail Record)

Assigned link based traffic volume from OD table



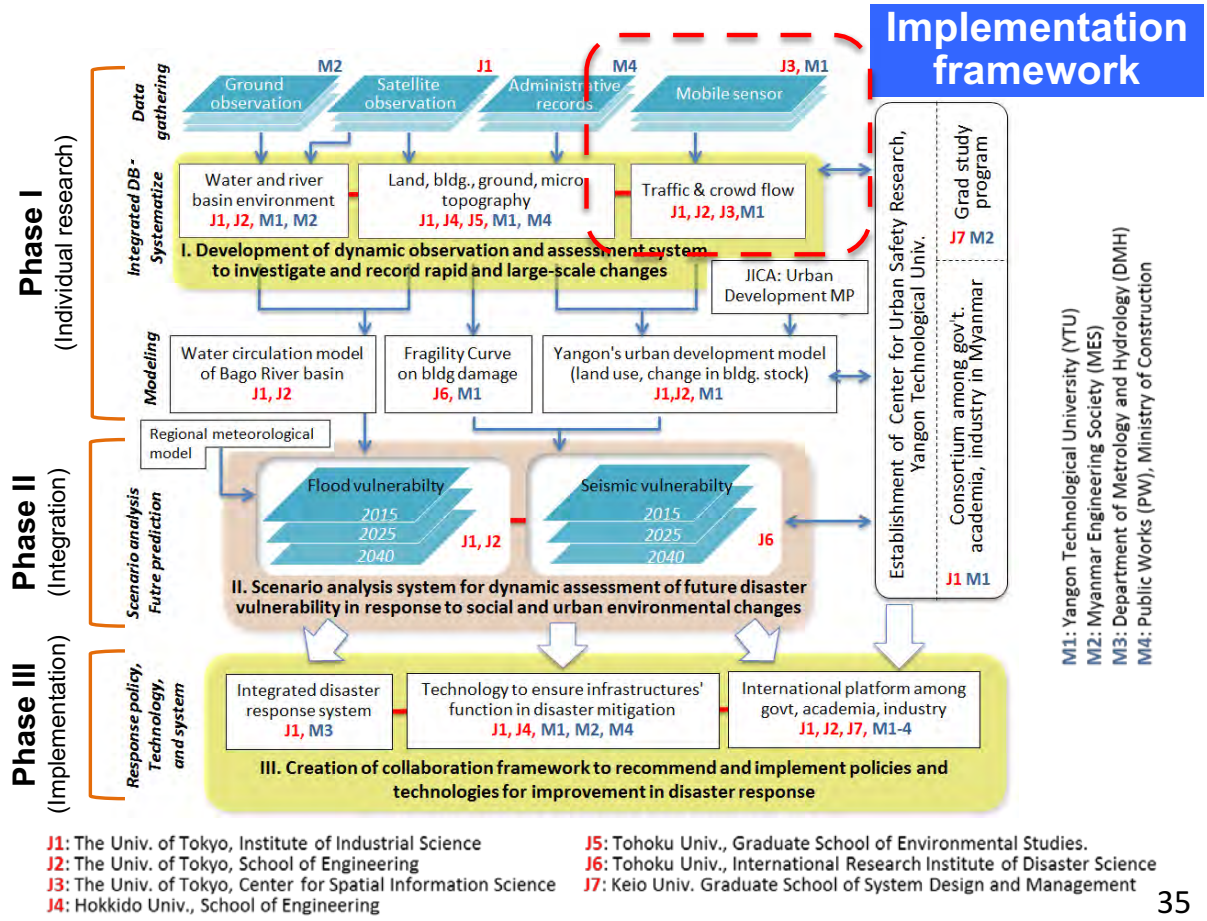


Joint research with Myanmar (JICA/JST SATREPS Project)



"ANZEN" Project in SATREPS (2015-2020)

"Development of a Comprehensive Disaster Resilience and Collaboration Platform in Myanmar"



J1: The Univ. of Tokyo, Institute of Industrial Science
J2: The Univ. of Tokyo, School of Engineering
J3: The Univ. of Tokyo, Center for Spatial Information Science
J4: Hokkido Univ., School of Engineering
J5: Tohoku Univ., Graduate School of Environmental Studies.
J6: Tohoku Univ., International Research Institute of Disaster Science
J7: Keio Univ. Graduate School of System Design and Management

Transport and People Mobility Group

Assoc. Prof. Sekimoto
IIS, UT

Lecturer. Daw Kyaing
Transportation, YTU

Dr. KoKo Lwin
Proj. Assist. Prof. IIS, UT

Prof. Kato
Civil Engineering, UT
(Advisor)

Mr. Fukushima
Researcher, IIS, UT

Prof. Shibasaki
CSIS, UT
(Advisor)

Assoc. Professor. Htay Win
Transportation, YTU

Ms. AyeHninHninNaing
Transportation, YTU

Property of CDR data

- CDR is call detail record of mobile phone

Period	Dec. 1-7, 2015
The number of CELL_ID	14,284
The number of subscriber's ID	4,435,321
The average number of daily records	16,161,366

MSISDN	EVENT ID	DATE&TIME	DURATION(Sec)	Upload(B)	Download(B)	CELL ID
3845230	1048	20151201155914	446	440	1394	00414010803330561
9903911	1048	20151201160658	0	0	0	01414011002401651
3938428	1048	20151201160554	64	194	610	01414011002401315
3938428	982	20151201160556	63	5624	11687	01414011002401315
0501317	1048	20151201090220	25478	30365	45173	01414011000907402
0501317	982	20151201090221	25478	566903	434232	01414011000907402
4016148	982	20151201160007	393	5873889	2136043	01414011000606823
4016148	1048	20151201160009	391	1301	3737	01414011000606823
9776353	1048	20151201153037	2164	51550	102593	00414010801320823
9776353	982	20151201153040	2160	1195309	8558137	00414010801320823
3587200	1048	20151201160457	104	2154	4916	00414010802312061
3908916	982	20151201160542	77	45462	956158	00414010030210102
5765675	982	20151201160406	174	106573	455079	01414010035112128
3961736	982	20151201155642	619	2951466	124273505	01414010335142151

PFLOW from reconstructed trip data

Traffic simulation for the future Yangon

- ✓ Input
  - ✓ OD matrix with each transportation mode: Past person trip survey and CDR data
  - ✓ Road network: JICA survey
- ✓ Output
  - ✓ Main is hourly link-based traffic status (traffic volume, speed and congestion=vol/capacity)
  - ✓ Others are some animations.

- ✓ First target
  - ✓ See traffic congestion changing traffic demand between 2016-2035 with current infrastructure (road, railway…)
  - ✓ Adding item 1, see the congestion increasing road capacity (e.g. the number of lane)

JICA traffic demand prediction

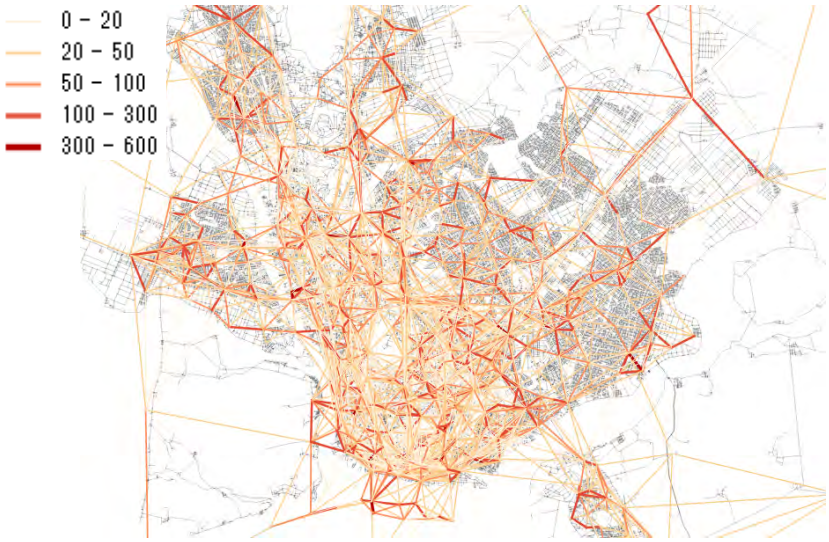
Table 3.1.1 Assumed Socio-economic Framework for Greater Yangon, 2035

Indicator		2016	2035	Ave. Growth Rate 16-35 (%/yr)	2035/2016
Population (000)		6,188	9,495	2.3	1.5
GRDP (Million USD)		14,148	60,990	8.6	4.3
Per capita GRDP (USD)		2,286	6,423	6.2	2.8
Employment	Total	2,339	5,717	4.7	2.4
	% (1/ 2/ 3) [1]	4/ 21/ 75	3/ 25/ 72		
School Enrollment	Total	1,531	2,482	2.6	1.6
	(P/ S/ H/U) [2]	98/ 44/ 67/18	100/ 100/ 90/ 45		
Vehicle Ownership (% to HH)		15	35	4.6	2.3
Trip Rate	Including Walking/Bicycle	2.0	2.3	0.7	1.2
	Excluding Walking/Bicycle	1.1	1.4	1.3	1.3
Average Trip Length (km) excluding Walking		10.1	11.3	0.6	1.1

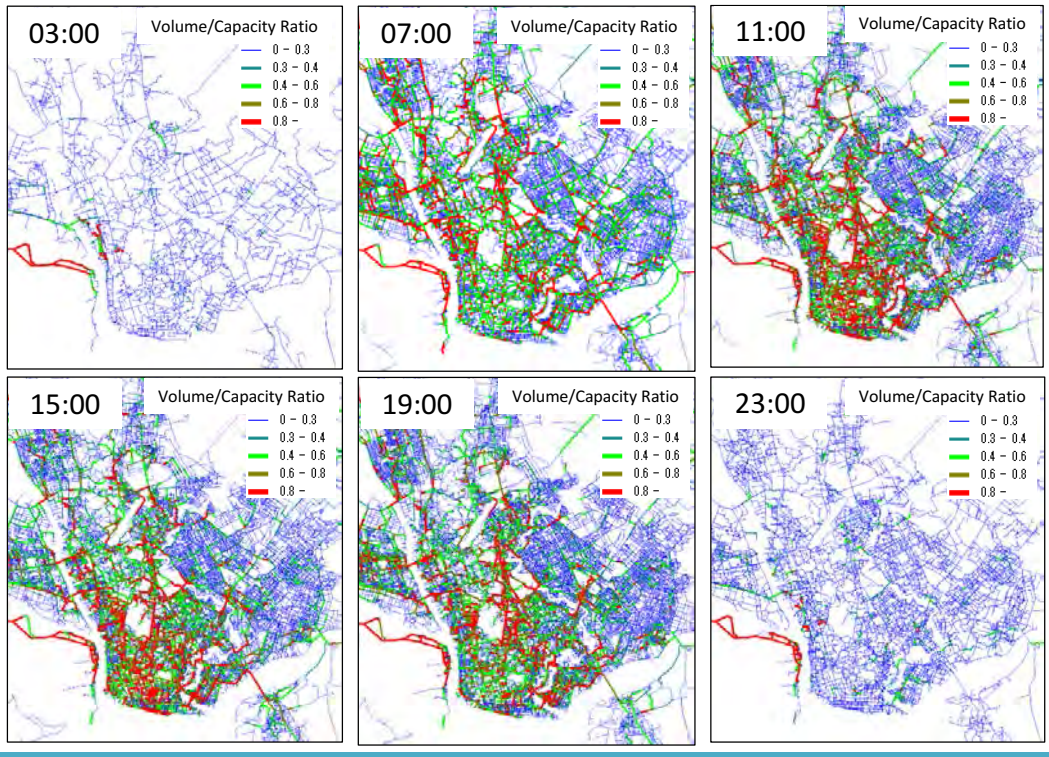
Source: Worked out by JICA Study Team based on various sources.
1) Economic sector: 1 = primary, 2 = secondary, 3= tertiary.
2) Education level: P=primary school (6-10 years old), 2=secondary (11-14), H=high school (15-17), U=university (18-30).

Vehicle OD visualization

✓ Vehicle OD volume at 9 a.m.

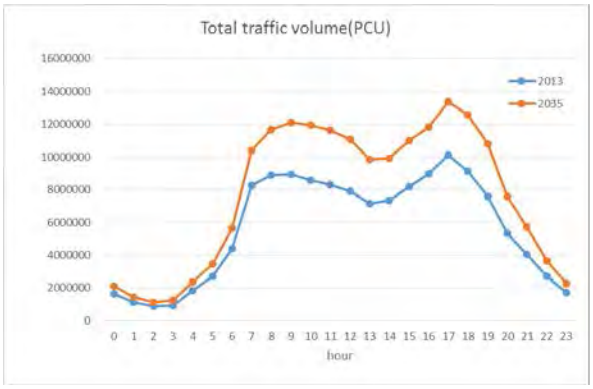


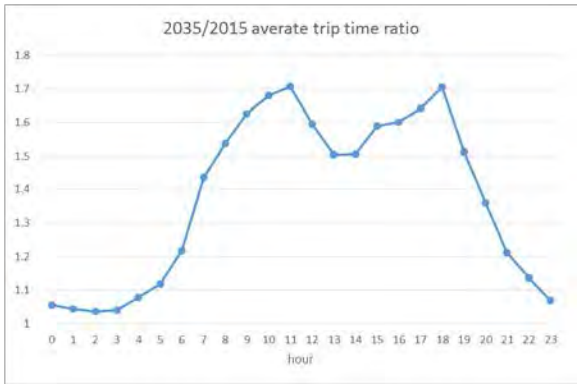
Hourly traffic congestion



Result 1
(Congestion change depending on the future demand)

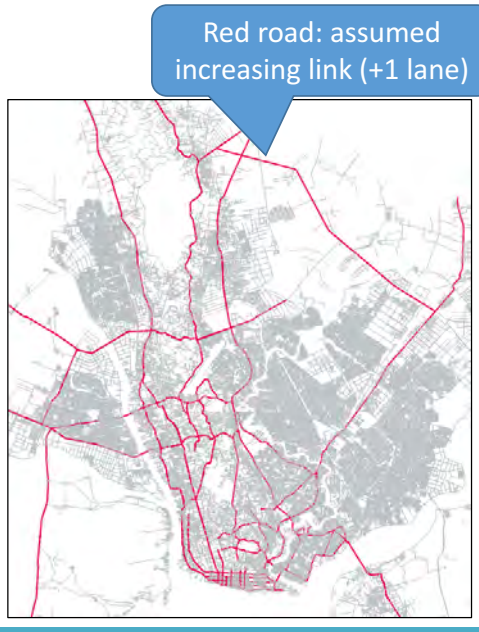
✓ Can confirm increase in trip time by increase of traffic volume in 2035 (1.7 x in some times)





Result 2
(Congestion depending on the demand & road capacity)

✓ Can confirm the save of trip time by increase the number of one lane on major road



MBD trial analysis in Hanoi, Vietnam

Don’t hesitate to contact me !!

http://sekilab.iis.u-tokyo.ac.jp

sekimoto@iis.u-tokyo.ac.jp