“WHERE” is a Big Question.

Geospatial Big Data

ผศ.ดร. วีระ เหมืองสิน
Asst. Prof. Veera Muangsin
Mobile Application and System Services (MASS) Research Group
Department of Computer Engineering
Chulalongkorn University
veera.m@chula.ac.th

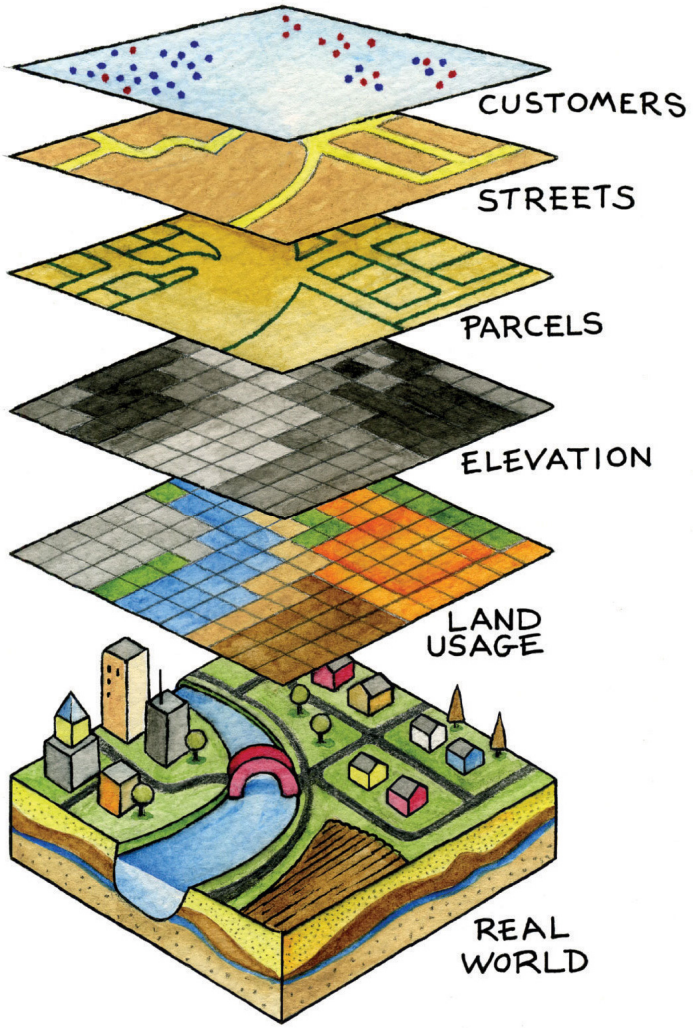
Chula DataScience

CHULA ENGINEERING
Foundation toward Innovation


Ruan Pharot Racha
Demonstration School
โรงเรียนสาธิต
มหาวิทยาลัย...
Exhibition Hall
Chamchuri 6
Chamchuri 4
Chamchuri 5
Indoor Stadium 2
Phayathai Rd
Maha Chakri
Sirindhorn Building
The Royal Bangkok
Sports Club (RBSC)
สนามม้าราชกรีฑา
สโมสร (อาร์บีเอลซี)
Chulalongkorn
University
จุฬาลงกรณ์
มหาวิทยาลัย
CU Auditorium
Faculty of Arts
คณะอักษรศาสตร์
ถนน พญาไท
ถนน อังรีดูนังต์
โรงเรียนสาธิต
จุฬาลงกรณ์...
Stadium
Faculty of Education
Chulalongkorn...
ATM
Chakrabongse Building
Faculty of Engineering
คณะวิศวกรรมศาสตร์
Fallabella, 100 Rajdamri
YOU
ARE
HERE
International School
of Engineering
International School
of Engineering...
Chulalongkorn
University...
โรงเรียนสาธิต
จุฬาลงกรณ์...
Faculty of Law
คณะนิติศาสตร์...
Sala Phra Kieo
Soi Chulalongkorn 42
Gem and
Jewelry Museum
Prem Purachatra Building
ศูนย์วิจัยโรค
เอดส์ สภากาชาดไทย
ซอย จุฬาลงกรณ์ 50
จุฬาฯ 50
Chulalongkorn 15 Alley
Faculty of Economics
คณะเศรษฐศาสตร์
Mahittalathibet Building
Visid Prachuabmoh
Building
The Thai Red Cross
สภากาชาดไทย
ถนน พระรามที่ 4
ศูนย์บริการ
โลหิตแห่งชาติ
Alley
Ratchadamri Rd
Faculty of Medicine
Mandarin Hotel
หัวลำโพง 568/18
Chula DataScience
CHULA ENGINEERING
Foundation toward Innovation

Geospatial Data Layers

- Navigation
- Urban planning
- Land use management
- Natural resource management
- Disaster management



Geospatial Data has always been big.
So, what's new?

- Variety, abundant → Volume
- Velocity → Volume
- Applications → Location-based services

We need big data to answer big questions.

Big Questions:

What do customers/people want?

What does a particular customer/person want?

A person's behavior can be inferred by

Interest:	**What** does she do/have/buy/like?
Social Interaction	**Who** does she contact with?
Mobility	**Where** does she live/go?

Most human activities associate to locations.

Most people have characteristic mobility patterns.

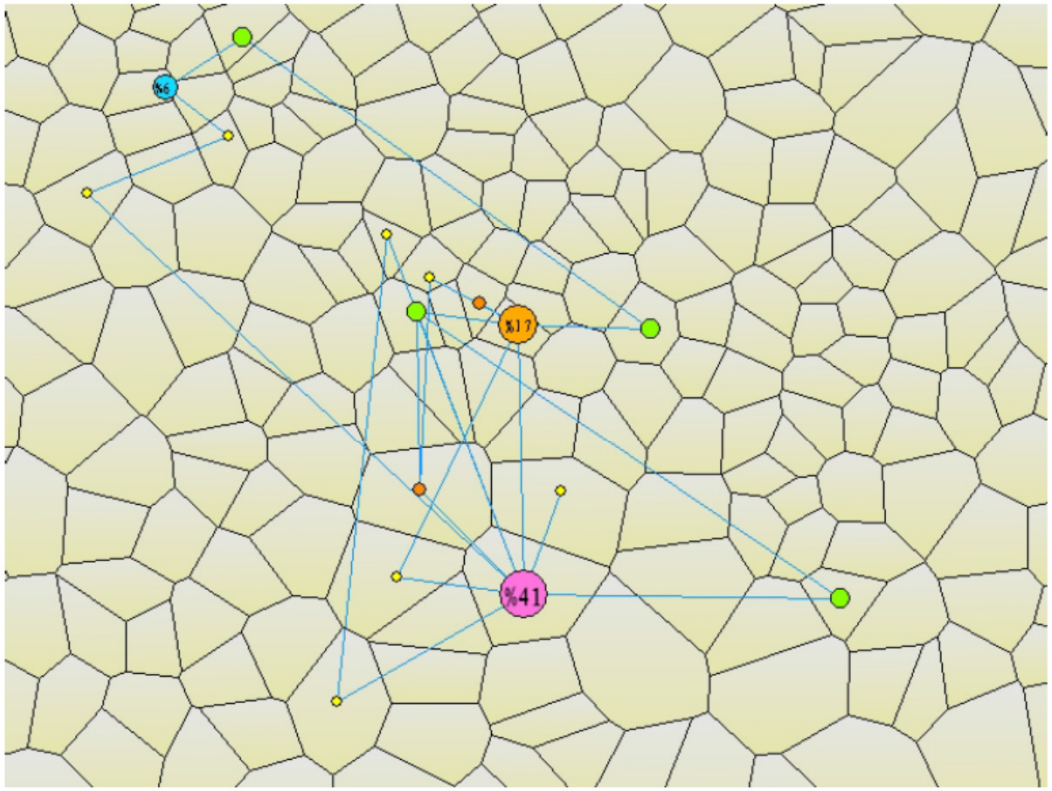


Figure 4. Mobility network associated with a mobile phone user who usually appears at fixed locations.

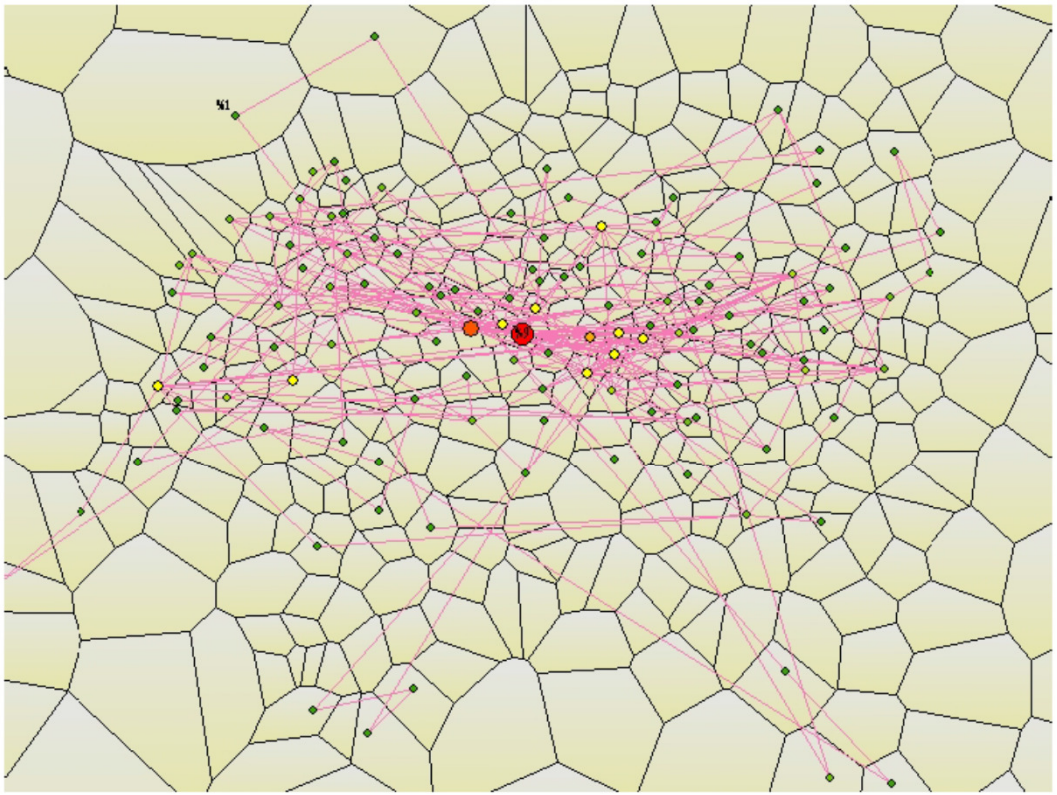


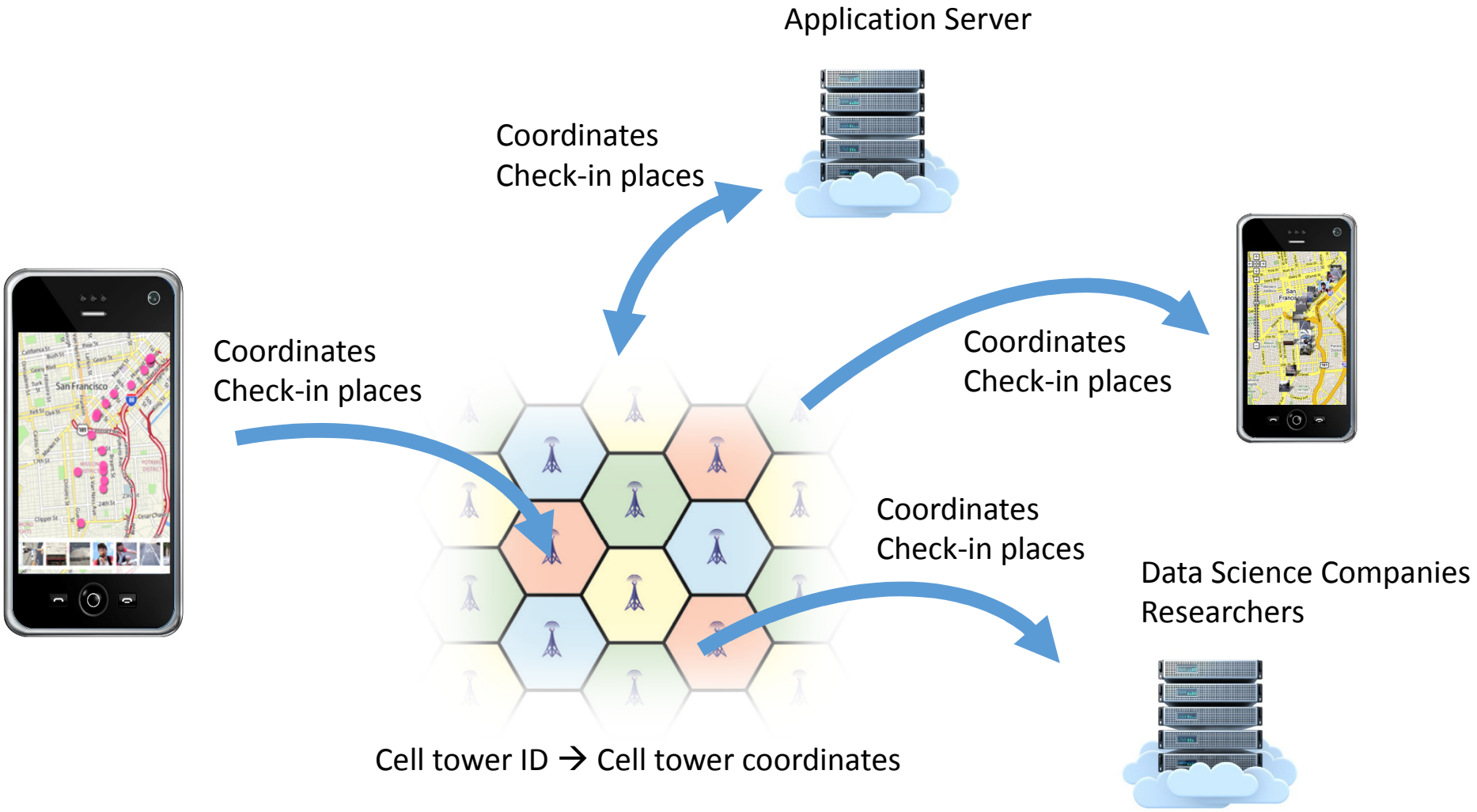
Figure 6. Mobility network associated with a mobile phone user who usually moves across a large set of locations and seldom visits the same place repeatedly.

"Analyzing and Geo-visualizing Individual Human Mobility Patterns Using Mobile Call Records", Chaogui Kang

Most data involving human activities are somehow related to locations and exhibit some patterns.

Location Data

- Coordinates (latitude, longitude)
- Place name, street address
- Proximity of reference (cell tower id)


Application Server
Coordinates
Check-in places
Coordinates
Check-in places
Coordinates
Check-in places
Coordinates
Check-in places
Data Science Companies
Researchers
Cell tower ID → Cell tower coordinates

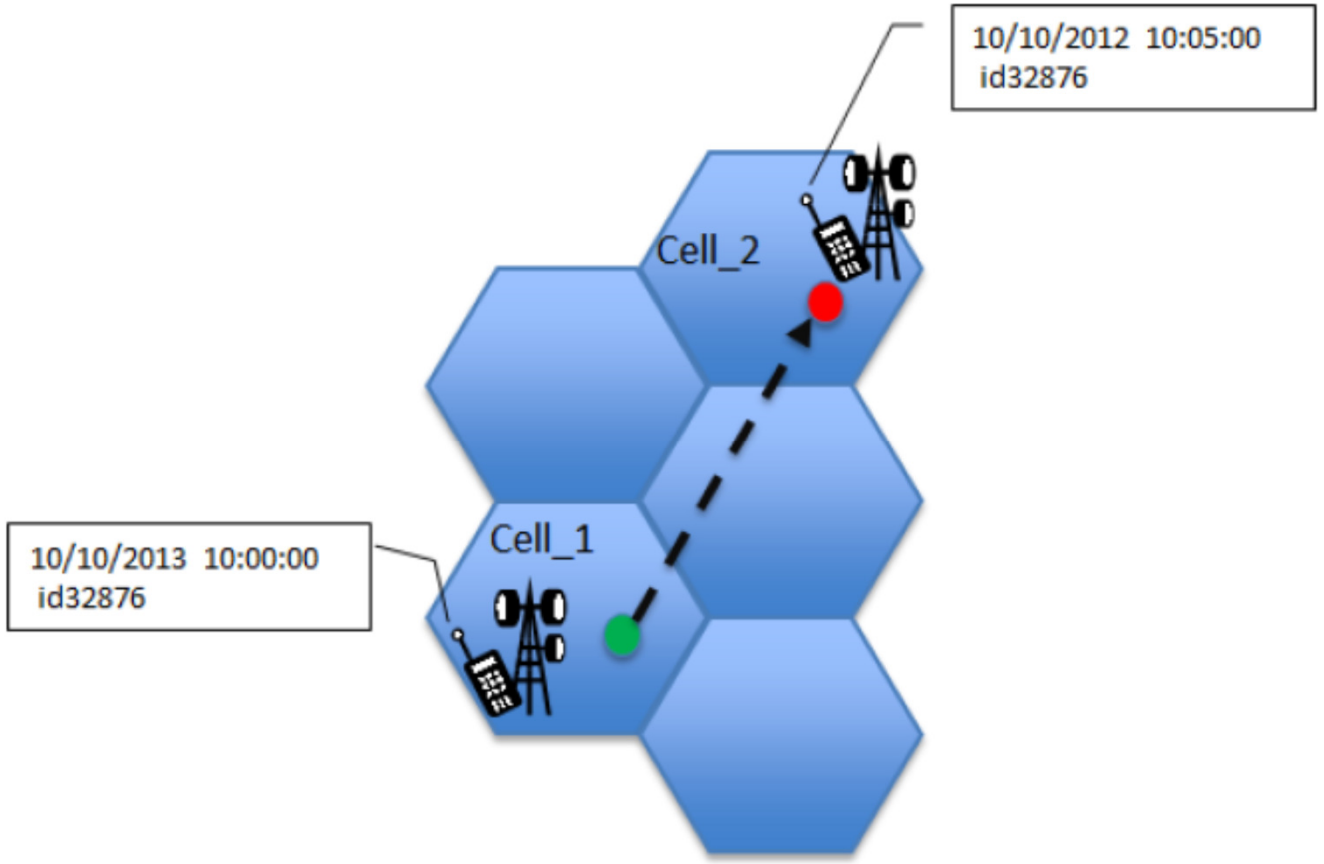
Category of Spatial Analysis

- Area or place categorization
- Area-based population demography or customer segment categorization
- Travelling patterns between areas
- Location analysis
- Route analysis

Research based on Geopatial Big Data

- Mobile network data
- GPS traces
- Social network data
- Crowdsourcing data

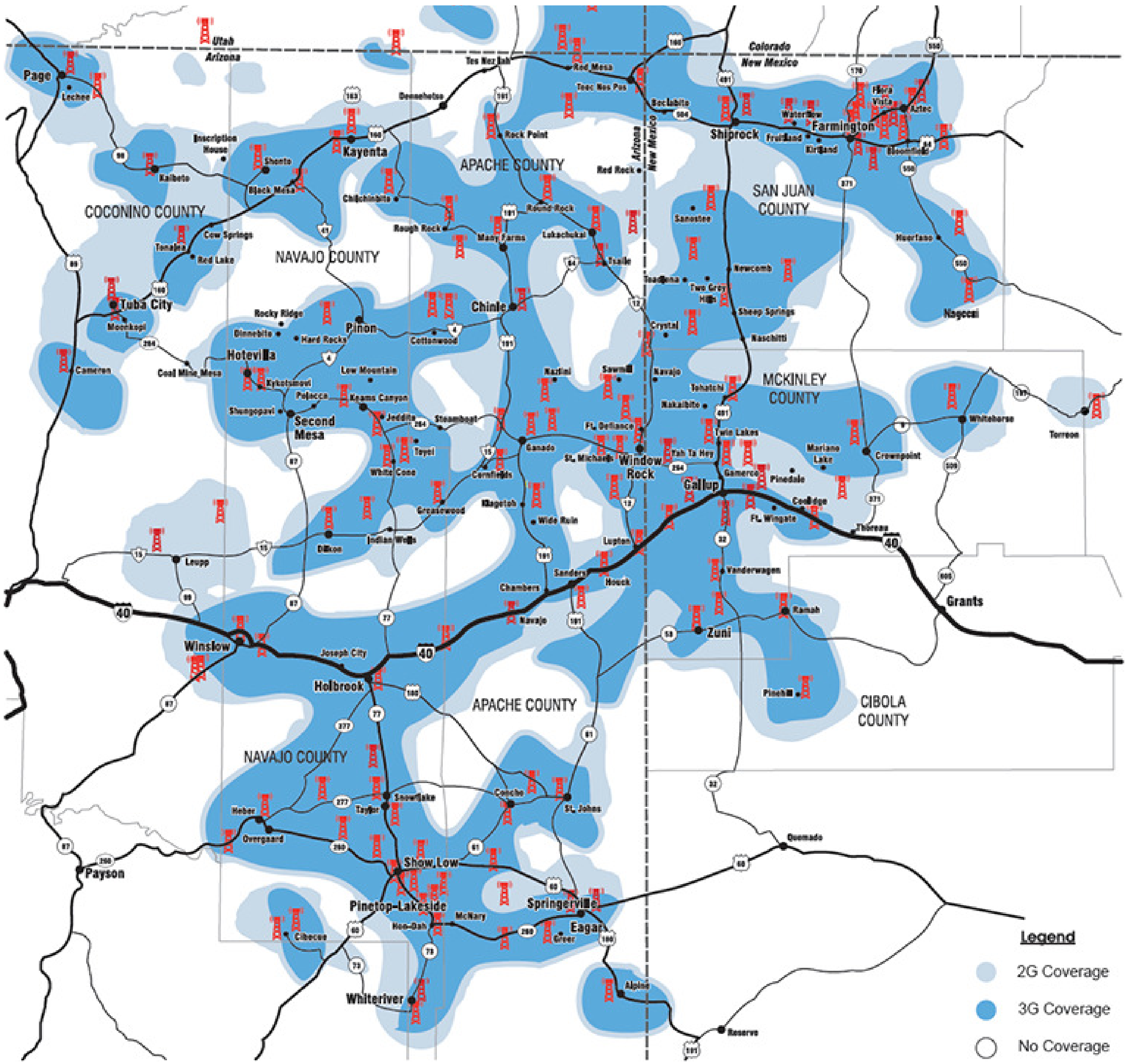
Mobile Phone Call Detail Records are geospatial-temporal data



"Use of mobile phone data to estimate visitors mobility flows", Lorenzo Gabrielli, et al.

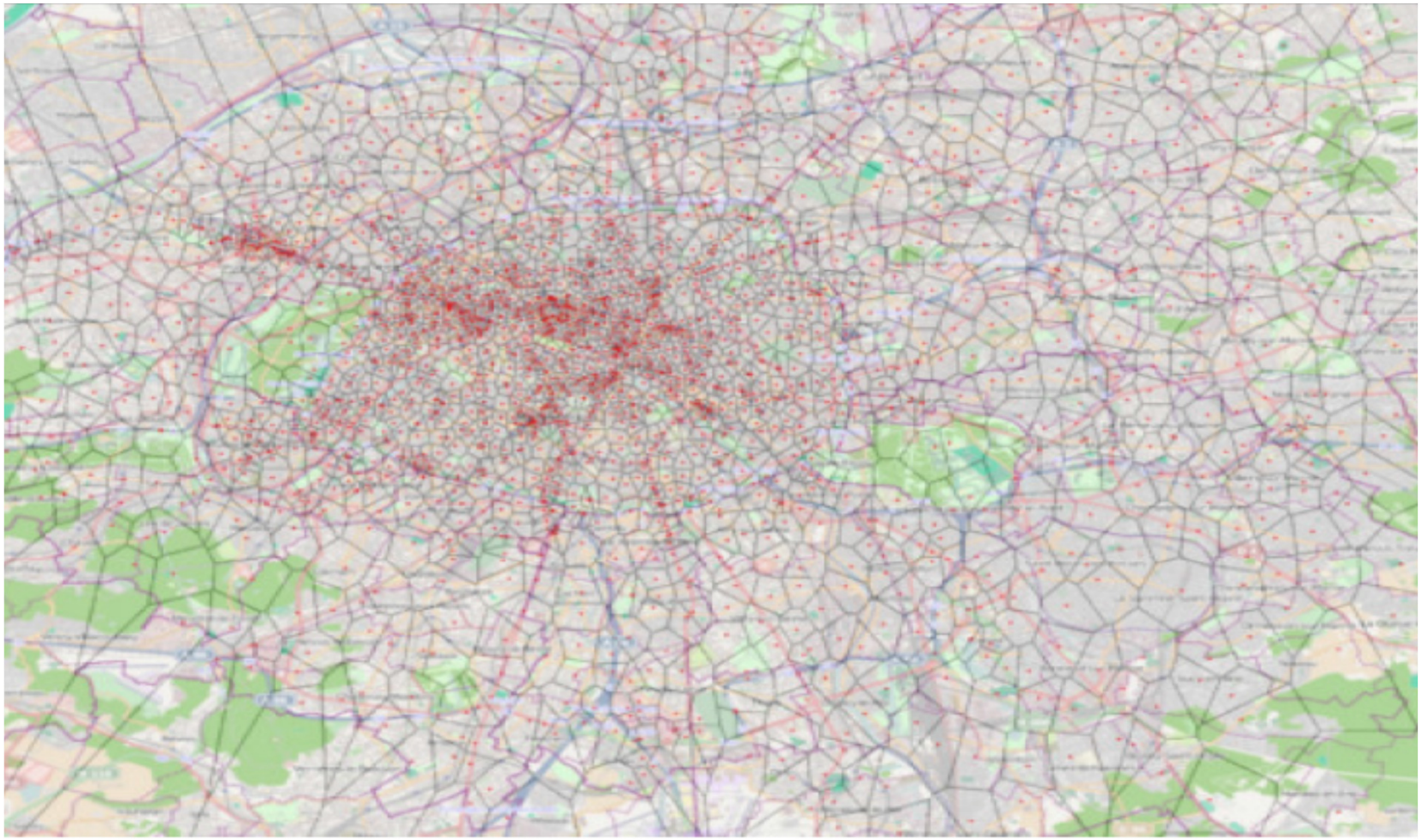
Chula DataScience

CHULA ENGINEERING
Foundation toward Innovation



Utah
Arizona
Colorado
New Mexico
Page
Lechee
Kaibeto
Inscription House
Shonto
Kayenta
Black Mesa
COCONINO COUNTY
NAVAJO COUNTY
APACHE COUNTY
SAN JUAN COUNTY
MCKINLEY COUNTY
CIBOLA COUNTY
Tonalea
Cow Springs
Red Lake
Tuba City
Moenkopi
Cameron
Coal Mine Mesa
Dennehotso
Tes Nez Iah
Red Mesa
Teec Nos Pos
Beclabito
Shiprock
Rock Point
Red Rock
Round Rock
Rough Rock
Chilchinbito
Many Farms
Lukachukai
Tsaile
Chinle
Rocky Ridge
Dinnebito
Hard Rocks
Pinon
Cottonwood
Hotevilla
Kykotsmovi
Polacca
Shungopavi
Second Mesa
Low Mountain
Keams Canyon
Jeddito
Steamboat
Tayei
White Cone
Greasewood
Indian Wells
Dilkon
Leupp
Ganado
Cornfields
Klagetoh
Wide Ruin
Nazlini
Sawmill
Navajo
Ft. Defiance
St. Michaels
Window Rock
Lupton
Sanders
Houck
Chambers
Navajo
Winslow
Joseph City
Holbrook
Farmington
Waterflow
Fruitland
Kirtland
Flora Vista
Aztec
Bloomfield
Sanostee
Newcomb
Toadlena
Two Grey Hills
Sheep Springs
Naschitti
Crystal
Tohatchi
Nakaibito
Twin Lakes
Yah Ta Hey
Gamerco
Gallup
Pinedale
Mariano Lake
Crownpoint
Coolidge
Ft. Wingate
Thoreau
Whitehorse
Torreon
Huerfano
Nageezi
Vanderwagen
Ramah
Zuni
Pinehill
Grants
Heber
Overgaard
Taylor
Snowflake
Concho
St. Johns
Show Low
Pinetop-Lakeside
McNary
Hon-Dah
Cibecue
Whiteriver
Springerville
Eagar
Greer
Alpine
Quemado
Reserve
Payson
Legend
2G Coverage
3G Coverage
No Coverage

Mobile base station distribution in Paris



“Content Consumption Cartography of the Paris Urban Region using Cellular Probe Data”, Stefato Sahar Hoteit, et al. (Orange)

Mobility pattern from mobile phone records

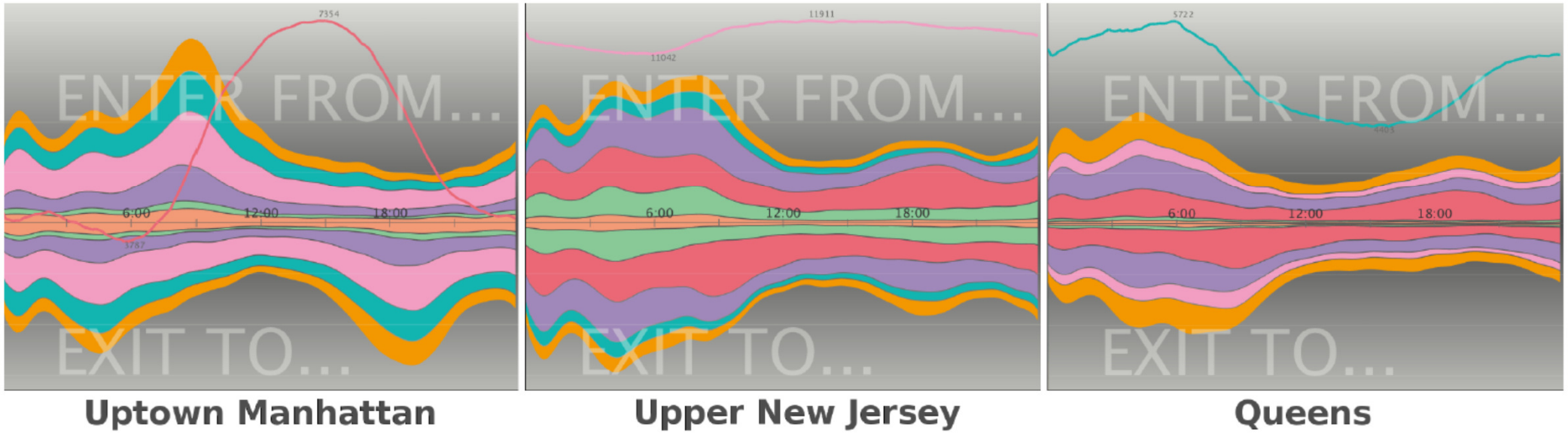


Figure 7: Stack views of Uptown Manhattan, Upper New Jersey, and Queens. Each stack view is augmented with a line chart showing the fluctuation of overall population in the region. Each stack view shows the region's characteristic based on the mobility patterns of mobile users.

"Inferring Human Mobility Patterns from Anonymized Mobile Communication Usage", Y. Tanahashi, et al. (AT&T)

Mobility pattern from mobile phone records

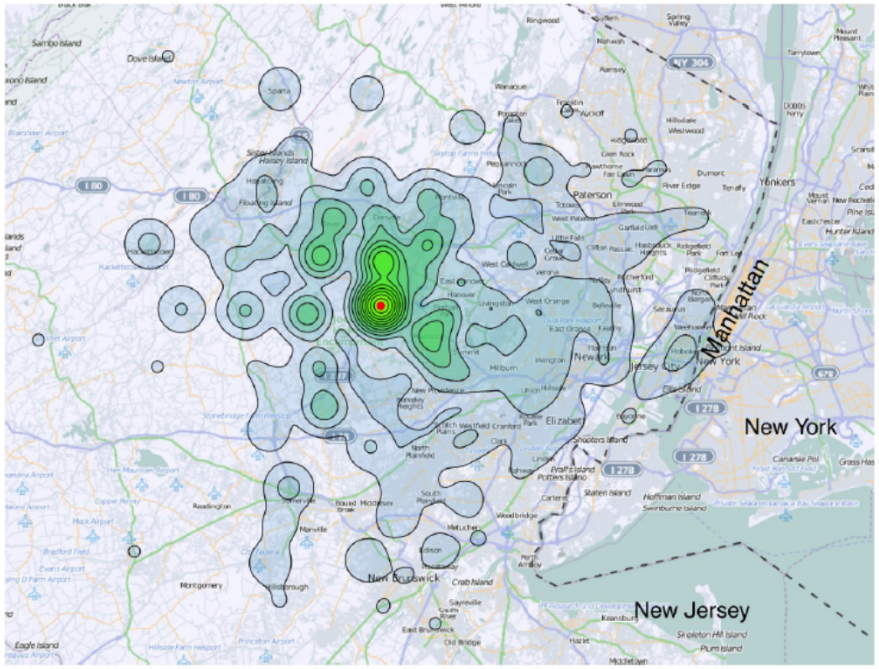


Figure 3: Laborshed of Morristown. The red dot is at the city center. Contour lines divide regions of different concentrations of workers' homes. Workers are identified as those who use their cellphones in Morristown during weekday business hours. Most workers come from nearby areas, but some come from as far as Manhattan.

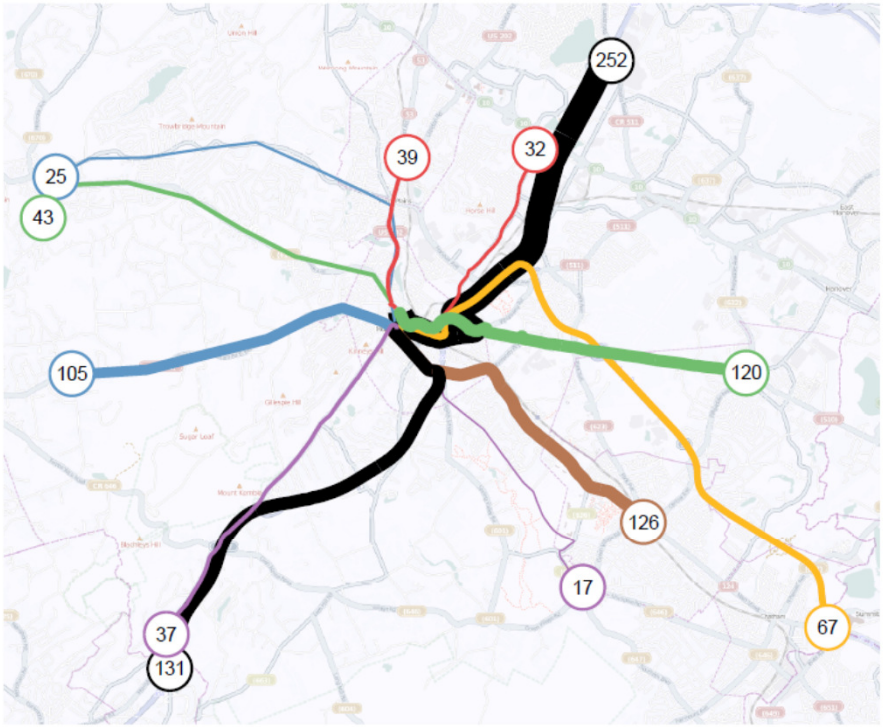


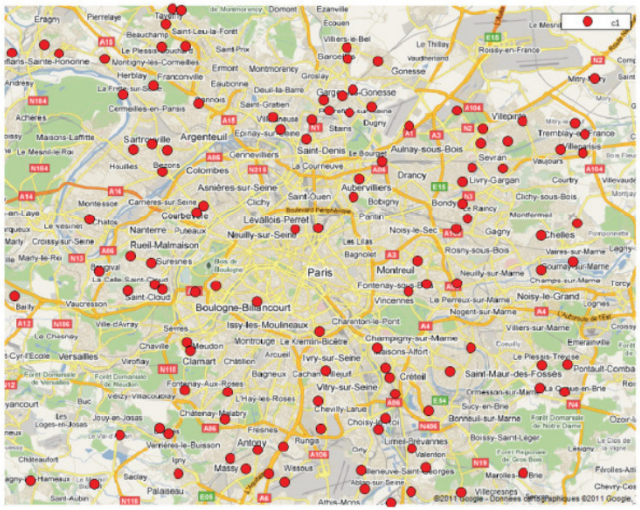
Figure 5: Relative traffic volumes on twelve commuting routes into the center of Morristown as assigned by our route classification algorithms. Line widths are proportional to the estimated volumes. Counts shown at the beginning of each route are normalized to 1,000 moving cellphones.

"Human Mobility Characterization from Cellular Network Data", R. Becker, et al. (AT&T)

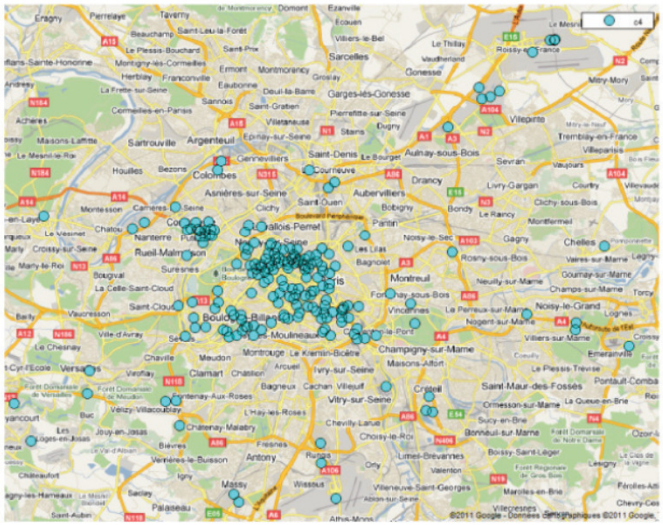
Mobility pattern from mobile phone records

Category A:The "hypermobility" antennas cluster

Category B: Antennas having both mobility and stationarity events



Category C: Antennas having mainly stationarity events



"Urban mobility: velocity and uncertainty in mobile phone data ", Thomas Couronné, et al. (Orange)

Chula DataScience

CHULA ENGINEERING
Foundation toward Innovation

Mobility pattern from mobile phone records

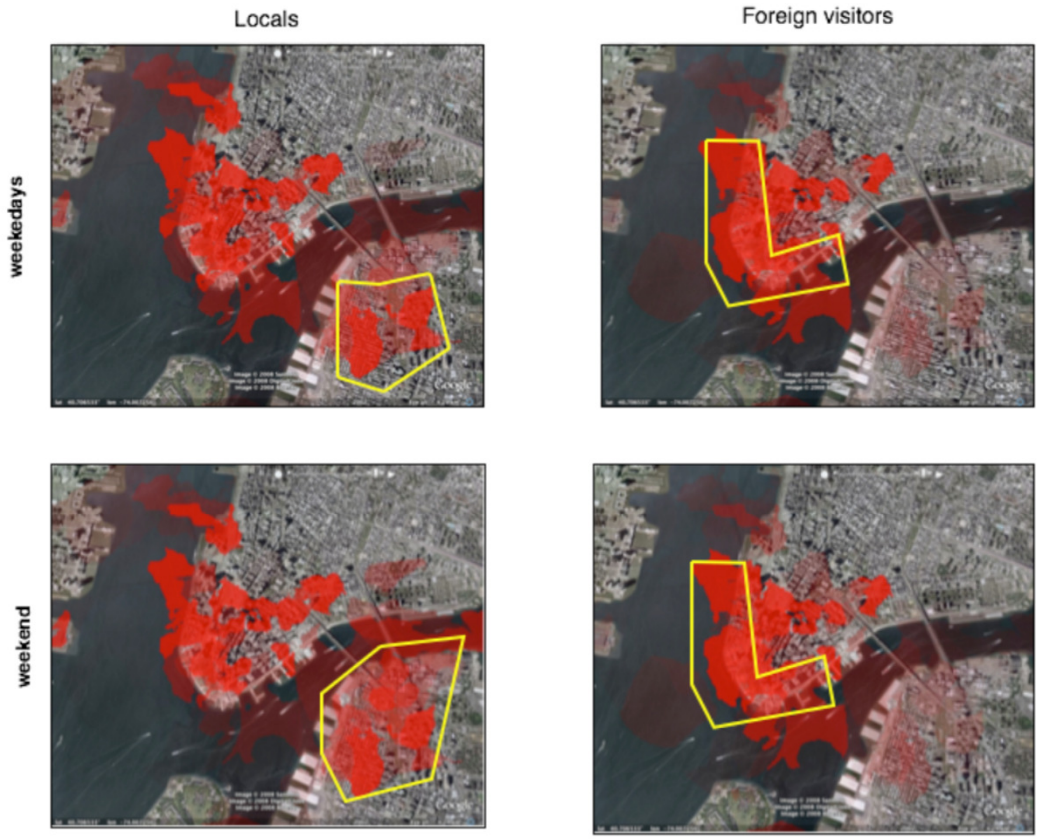


Figure 3. Spatial distribution of locals (New Yorkers) and foreign visitors in the neighborhood of the Waterfall. New Yorkers generate network traffic activity in the financial district and in Brooklyn, a neighborhood not attracting many foreigners compared to the waterfront of lower Manhattan.

"Towards estimating the presence of vistors from the aggregate mobile phone network activity they generate", Fabien Girardin, et al. (AT&T)

Area classification using mobile phone records

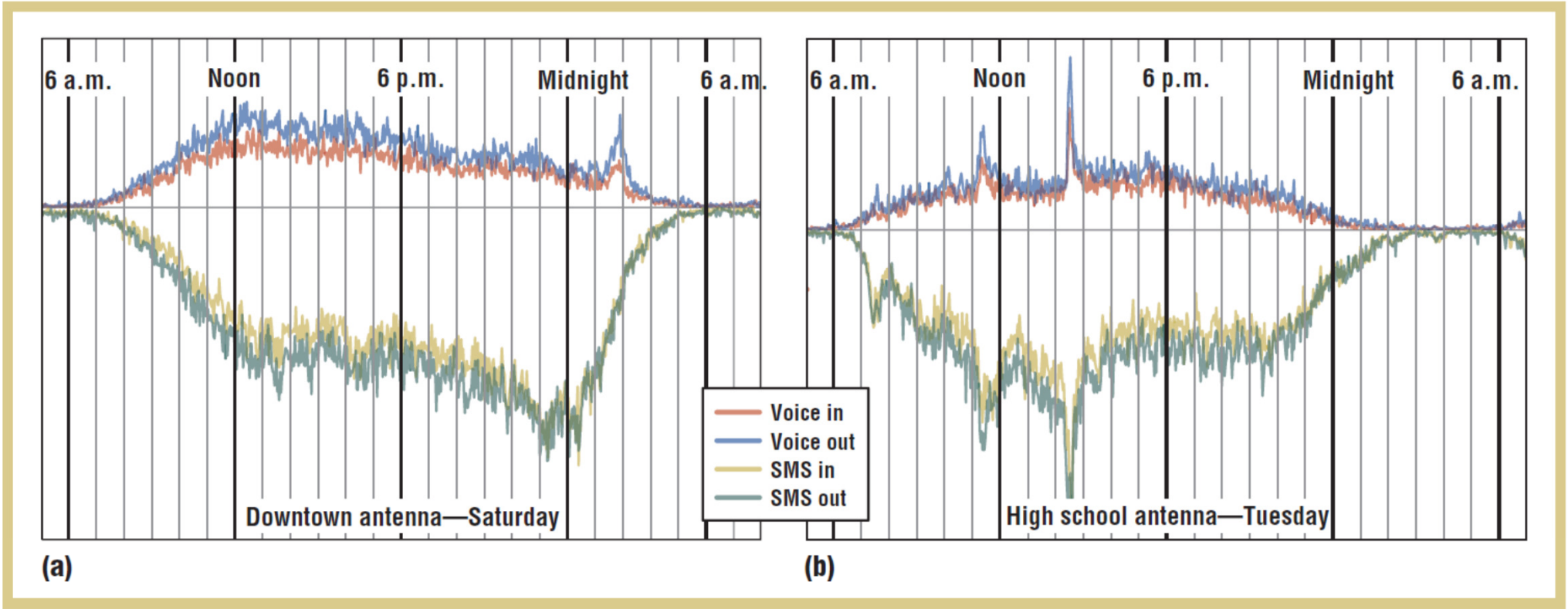
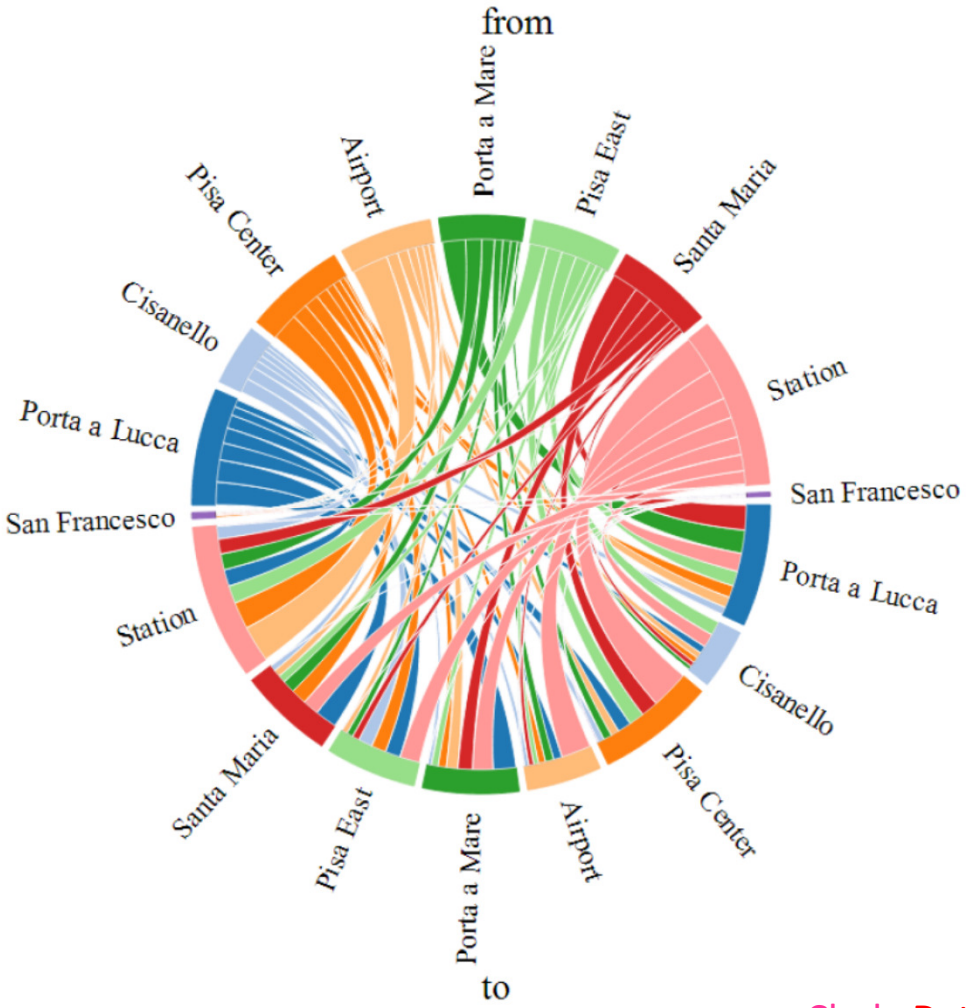


Figure 4. Lip plots of voice call and SMS volumes show unusual spikes highlighting local patterns or events in Morristown, New Jersey. Call volume (plotted upward: inbound, red; outbound, blue) and SMS volume (plotted downward: inbound, light green; outbound, dark green) on two antennas are shown. The antenna in (a) points towards the commercial and restaurant district and the antenna in (b) points toward the high school. A voice peak occurs Saturday at 2 a.m. when the bars close. Both voice and SMS peaks occur Tuesday when the school lets out.

"A Tale of One City Using Cellular Network Data for Urban Planning", Richard Becker, et al. (AT&T)

Origin-Destination



"Use of mobile phone data to estimate visitors mobility flows", Lorenzo Gabrielli, et al.

Chula DataScience

CHULA ENGINEERING
Foundation toward Innovation

Human Mobility from Different Sources of Data

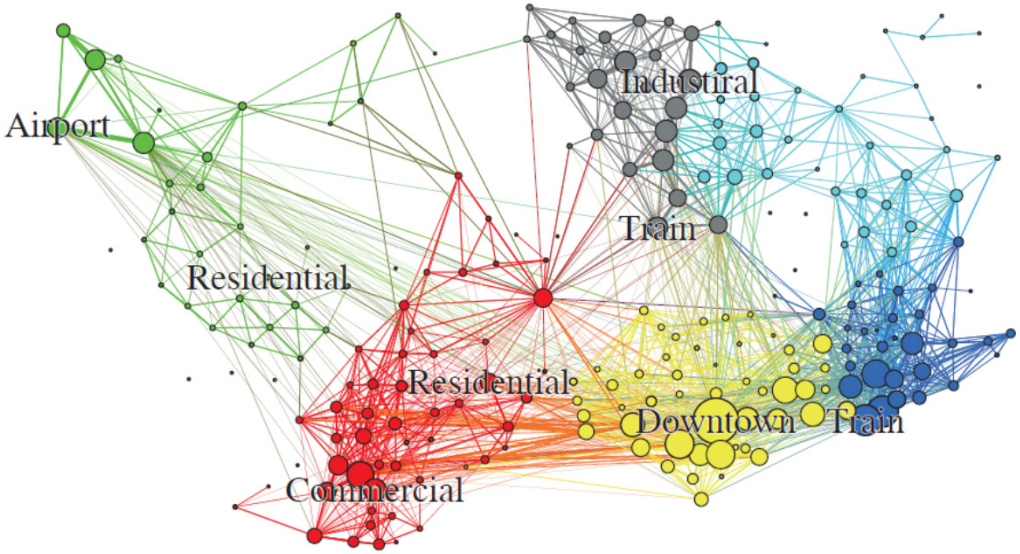


Fig 2: Human Mobility from Cellphone Data

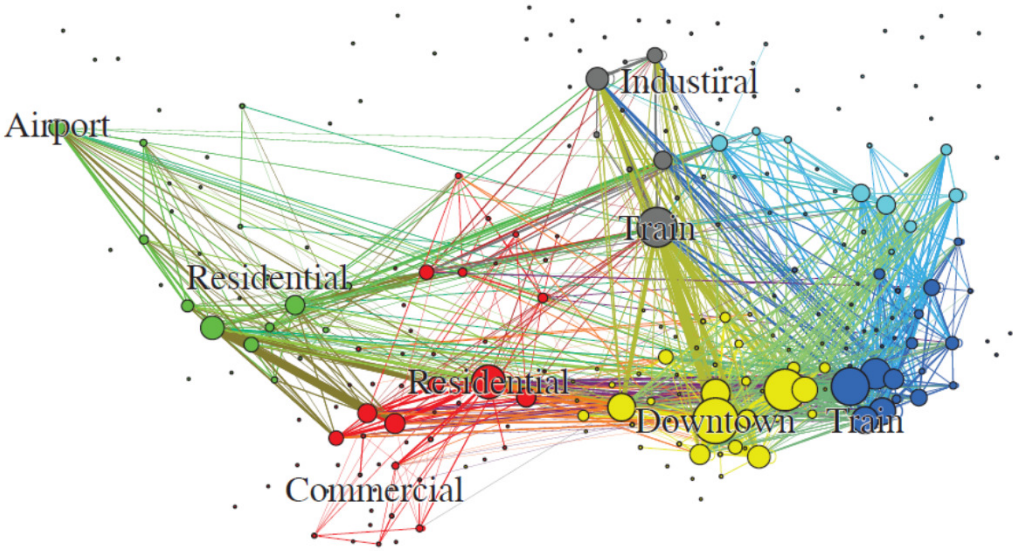


Fig 3: Human Mobility from Transit Data

"Exploring Human Mobility with Multi-Source Data at Extremely Large Metropolitan Scales", Desheng Zhang, et al.

Real-time Event Detection by Social Sensors

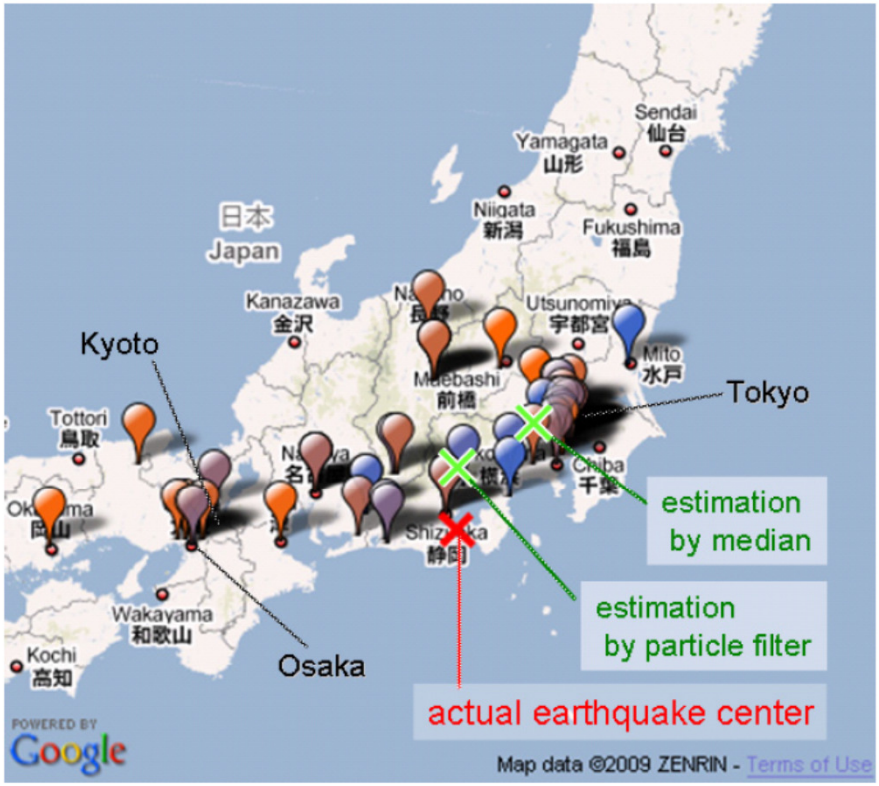


Figure 9: Earthquake location estimation based on tweets. Balloons show the tweets on the earthquake. The cross shows the earthquake center. Red represents early tweets; blue represents later tweets.

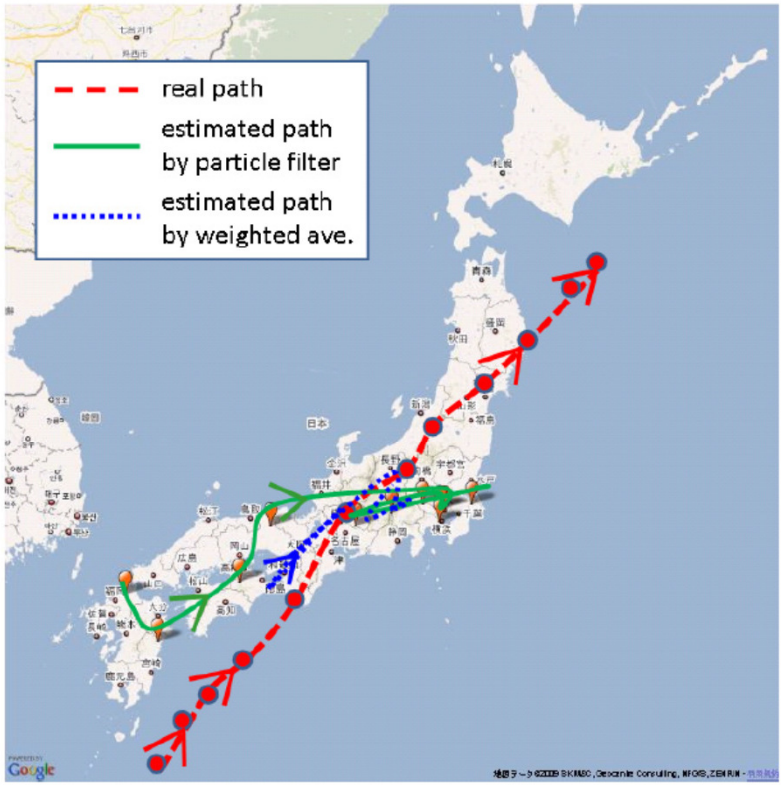


Figure 10: Typhoon trajectory estimation based on tweets.

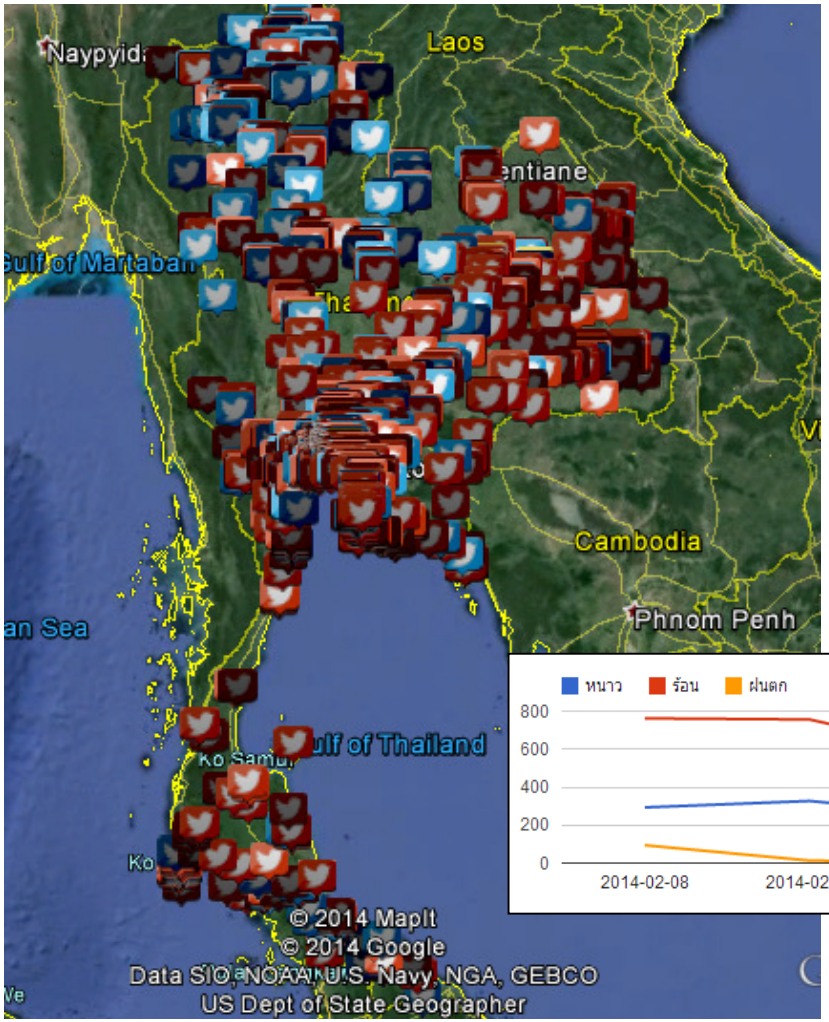
"Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors", Takeshi Sakaki, et al.

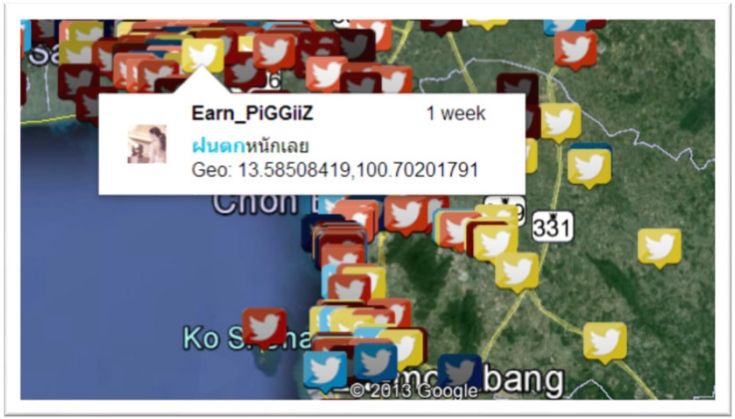
Our Research

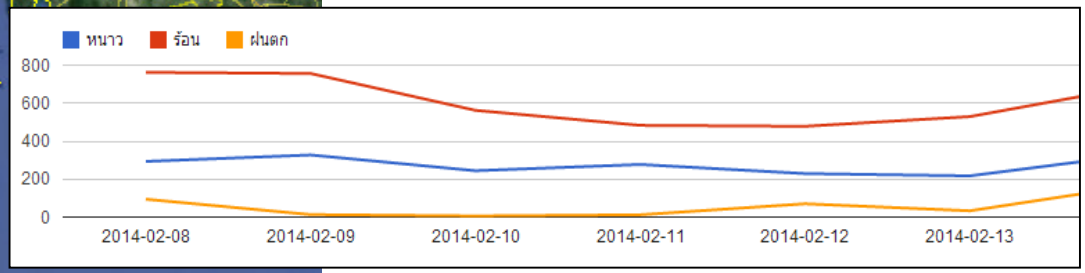
Spatial and mobility analysis using multiple sources of data

- Social network
- GPS traces
- Mobile crowdsourcing
- Mobile network data

Spatial-Temporal Analysis using Social Network Data (Twitter)

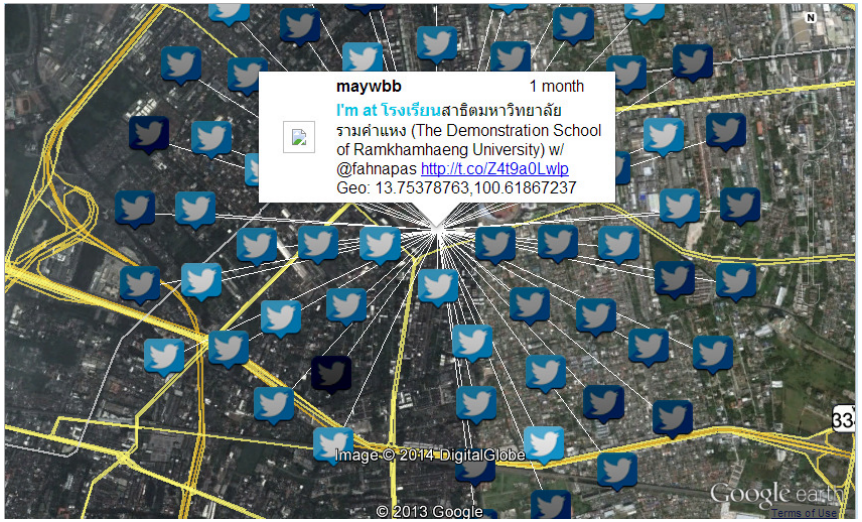


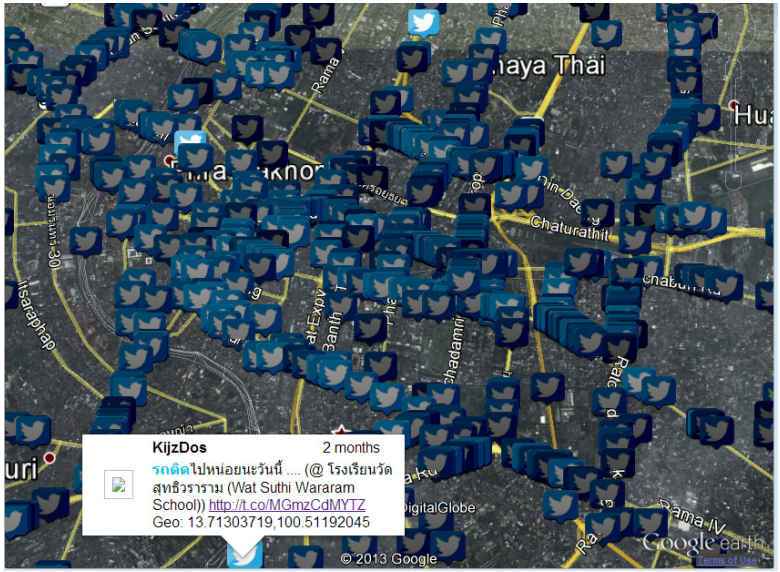


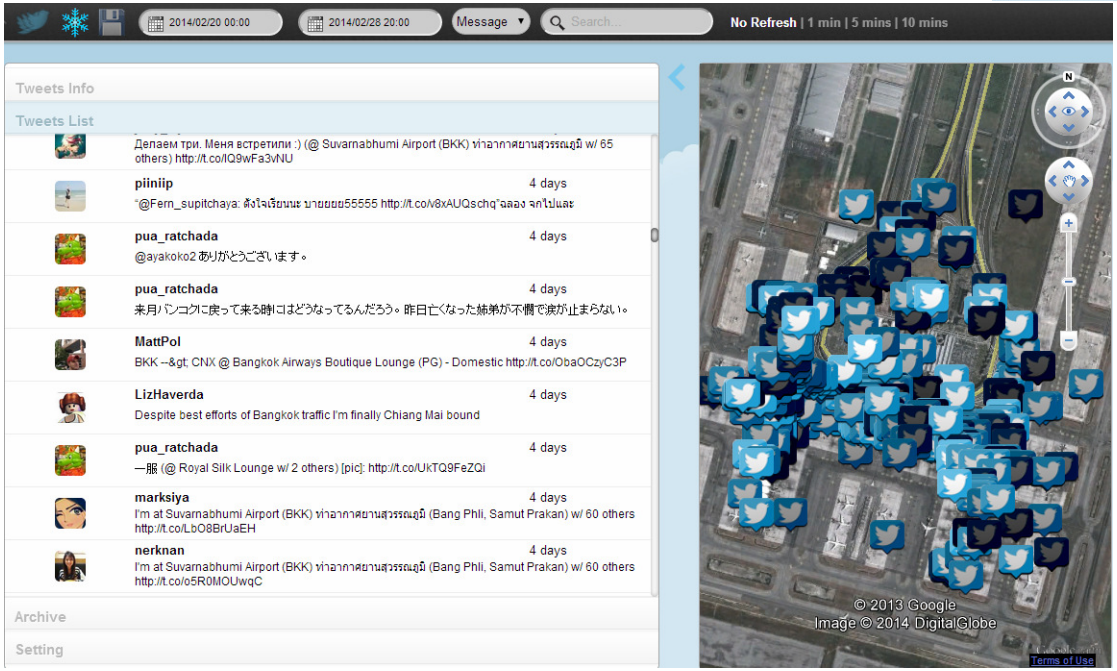


Chula DataScience

CHULA ENGINEERING
Foundation toward Innovation

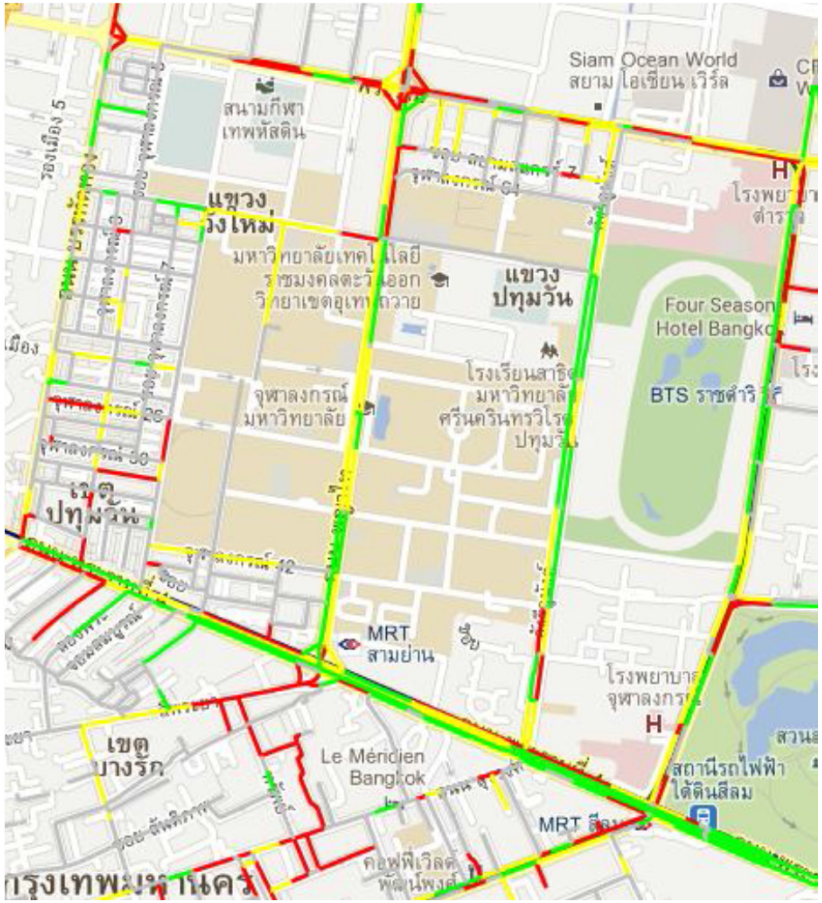

maywbb 1 month
I'm at โรงเรียนสาธิตมหาวิทยาลัยรามคำแหง (The Demonstration School of Ramkhamhaeng University) w/ @fahnapas http://t.co/Z4t9a0LwIp
Geo: 13.75378763,100.61867237
Image © 2014 DigitalGlobe
© 2013 Google
Google earth


KijzDos 2 months
รถติดไปหน่อยนะวันนี้ .... (@ โรงเรียนวัดสุทธิวราราม (Wat Suthi Wararam School)) http://t.co/MGmzCdMYTZ
Geo: 13.71303719,100.51192045
© 2013 Google
Google earth


2014/02/20 00:00
2014/02/28 20:00
Message
Search...
No Refresh | 1 min | 5 mins | 10 mins
Tweets Info
Tweets List
Делаем три. Меня встретили :) (@ Suvarnabhumi Airport (BKK) ท่าอากาศยานสุวรรณภูมิ w/ 65 others) http://t.co/lQ9wFa3vNU
piiniip 4 days
"@Fern_supitchaya: ตั้งใจเรียนนะ บายยยย55555 http://t.co/v8xAUQschq"ฉลอง จกไปและ
pua_ratchada 4 days
@ayakoko2 ありがとうございます。
pua_ratchada 4 days
来月バンコクに戻って来る時にはどうなってるんだろう。昨日亡くなった姉弟が不憫で涙が止まらない。
MattPol 4 days
BKK --> CNX @ Bangkok Airways Boutique Lounge (PG) - Domestic http://t.co/ObaOCzyC3P
LizHaverda 4 days
Despite best efforts of Bangkok traffic I'm finally Chiang Mai bound
pua_ratchada 4 days
一服 (@ Royal Silk Lounge w/ 2 others) [pic]: http://t.co/UkTQ9FeZQi
marksiya 4 days
I'm at Suvarnabhumi Airport (BKK) ท่าอากาศยานสุวรรณภูมิ (Bang Phli, Samut Prakan) w/ 60 others http://t.co/LbO8BrUaEH
nerknan 4 days
I'm at Suvarnabhumi Airport (BKK) ท่าอากาศยานสุวรรณภูมิ (Bang Phli, Samut Prakan) w/ 60 others http://t.co/o5R0MOUwqC
Archive
Setting
© 2013 Google
Image © 2014 DigitalGlobe
Terms of Use

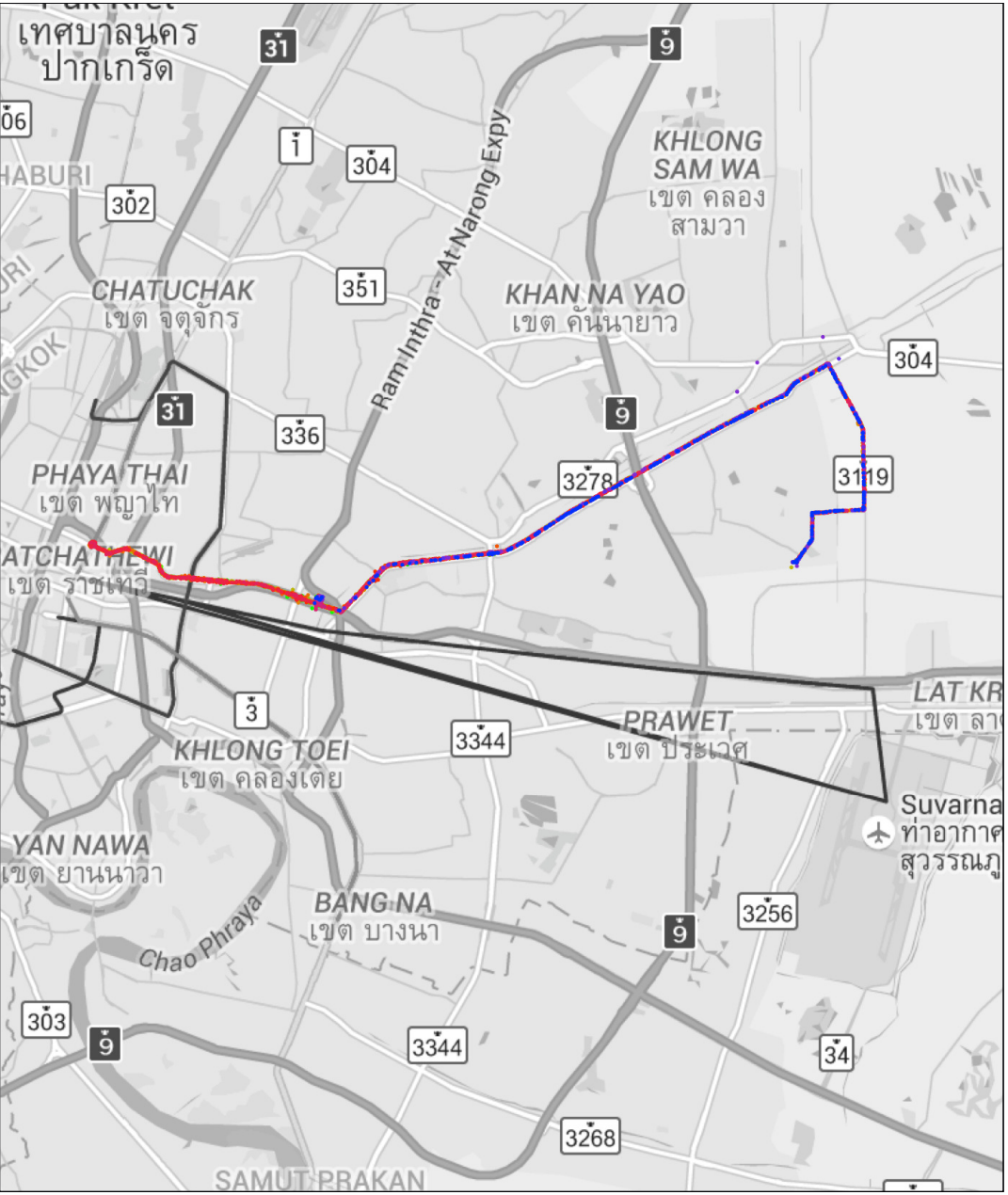

Samui
Ko Samui
Image © 2014 DigitalGlobe
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
© 2013 Google

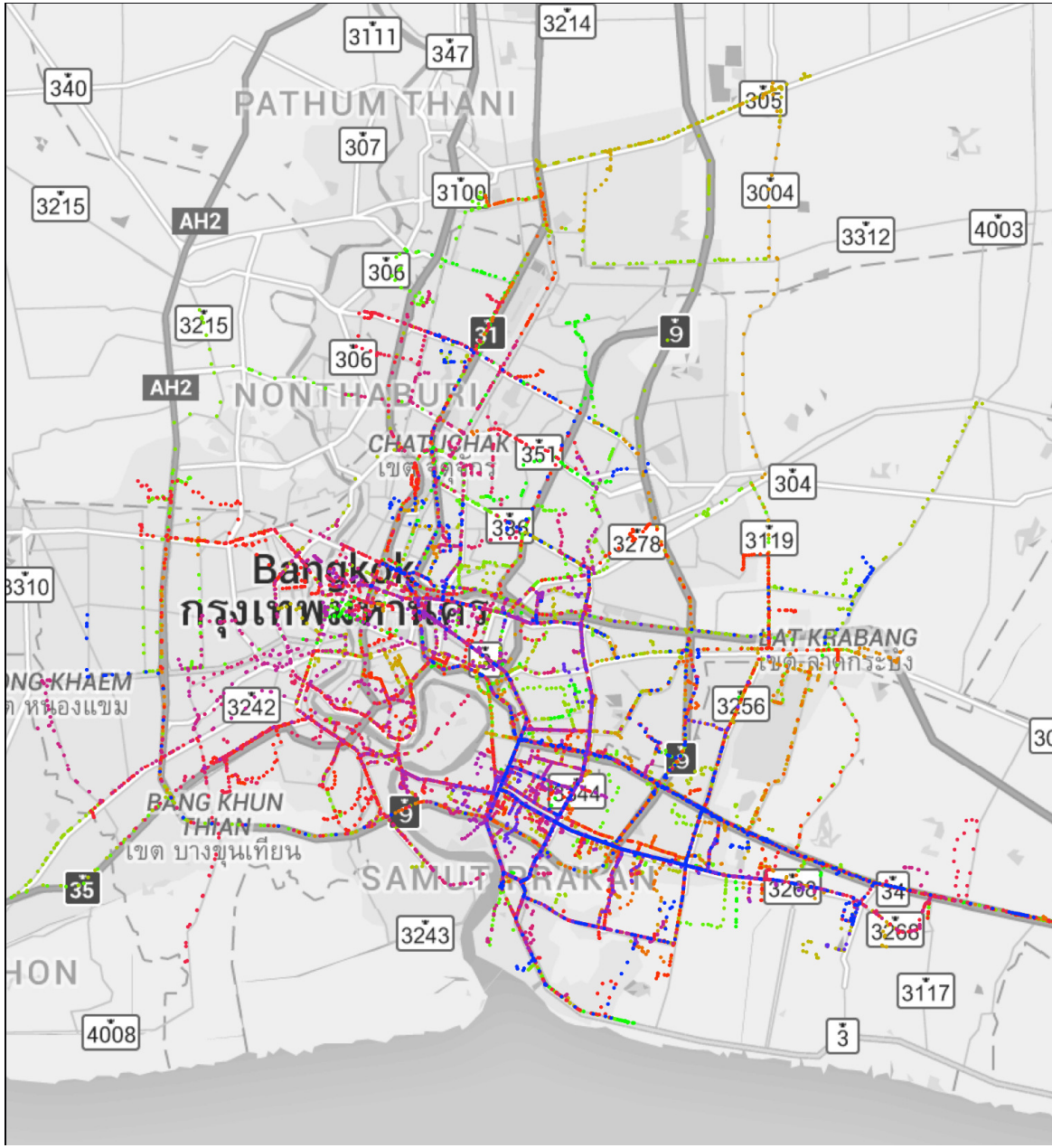
Traffic Analysis using GPS Traces



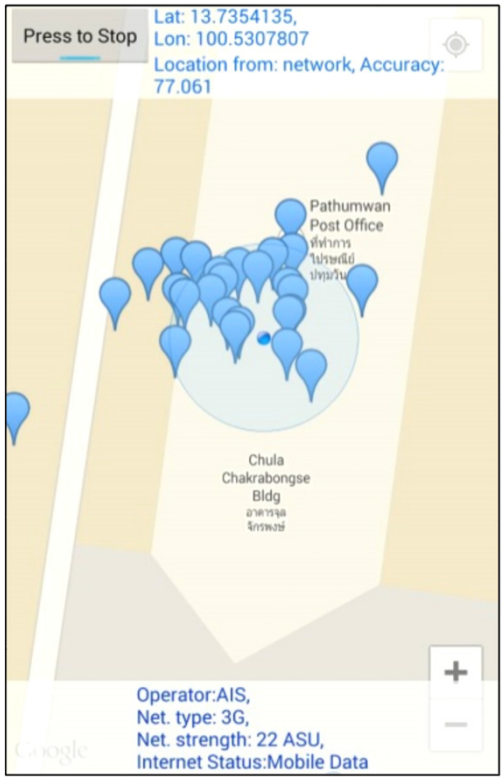
Chula DataScience

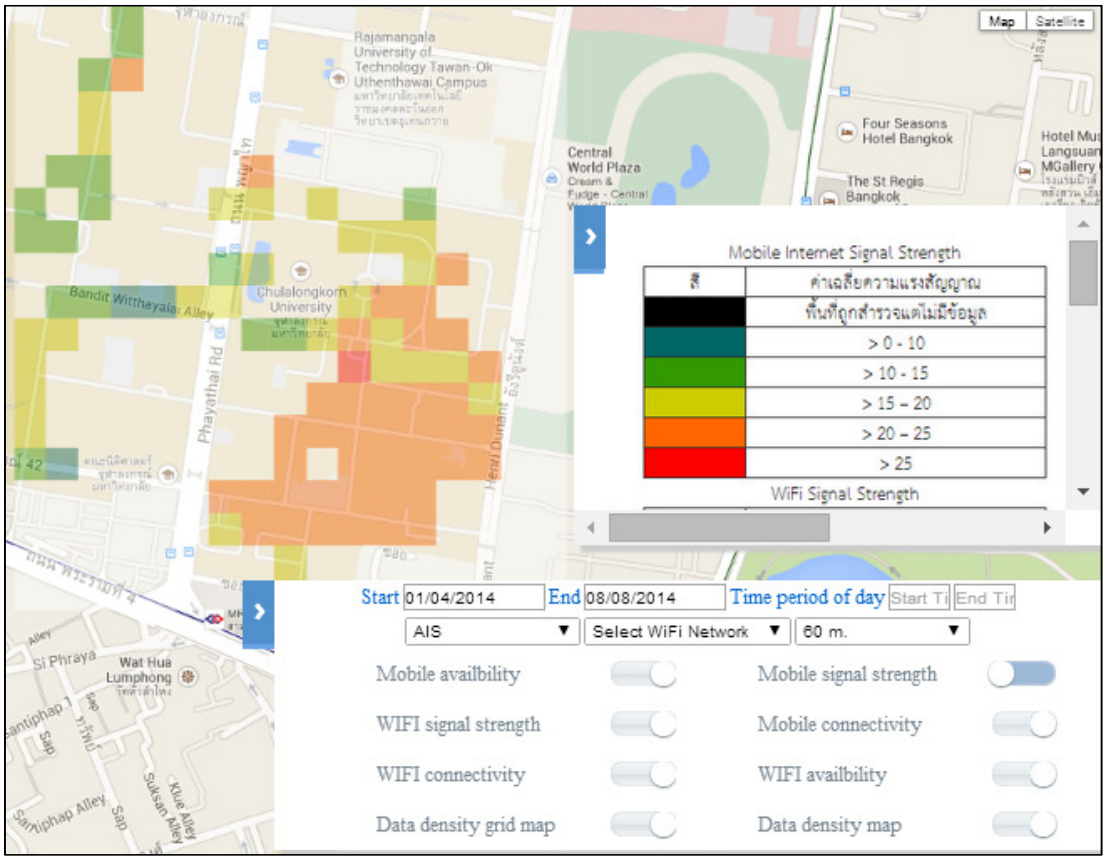
CHULA ENGINEERING
Foundation toward Innovation


เทศบาลนคร
ปากเกร็ด
31
9
1
304
302
Ram Inthra - At Narong Expy
KHLONG
SAM WA
เขต คลอง
สามวา
CHATUCHAK
เขต จตุจักร
351
KHAN NA YAO
เขต คันนายาว
304
31
336
9
3278
3119
PHAYA THAI
เขต พญาไท
ATCHATHEWI
เขต ราชเทวี
3
3344
PRAWET
เขต ประเวศ
LAT KR
เขต ลา
KHLONG TOEI
เขต คลองเตย
Suvarna
ท่าอากาศ
สุวรรณภู
YAN NAWA
เขต ยานนาวา
BANG NA
เขต บางนา
Chao Phraya
3256
9
303
9
3344
34
3268
SAMUT PRAKAN


3111
347
3214
340
PATHUM THANI
305
307
3100
3004
3215
AH2
3312
4003
306
3215
306
31
9
AH2
NONTHABURI
CHATUCHAK
เขต จตุจักร
351
304
336
3278
3119
3310
Bangkok
กรุงเทพมหานคร
LAT KRABANG
เขต ลาดกระบัง
NG KHAEM
หนองแขม
3242
3256
9
3244
BANG KHUN
THIAN
เขต บางขุนเทียน
9
35
SAMUT PRAKAN
3268
34
3243
3268
HON
3117
4008
3

Network Coverage Analysis using Crowdsourced Signal Data

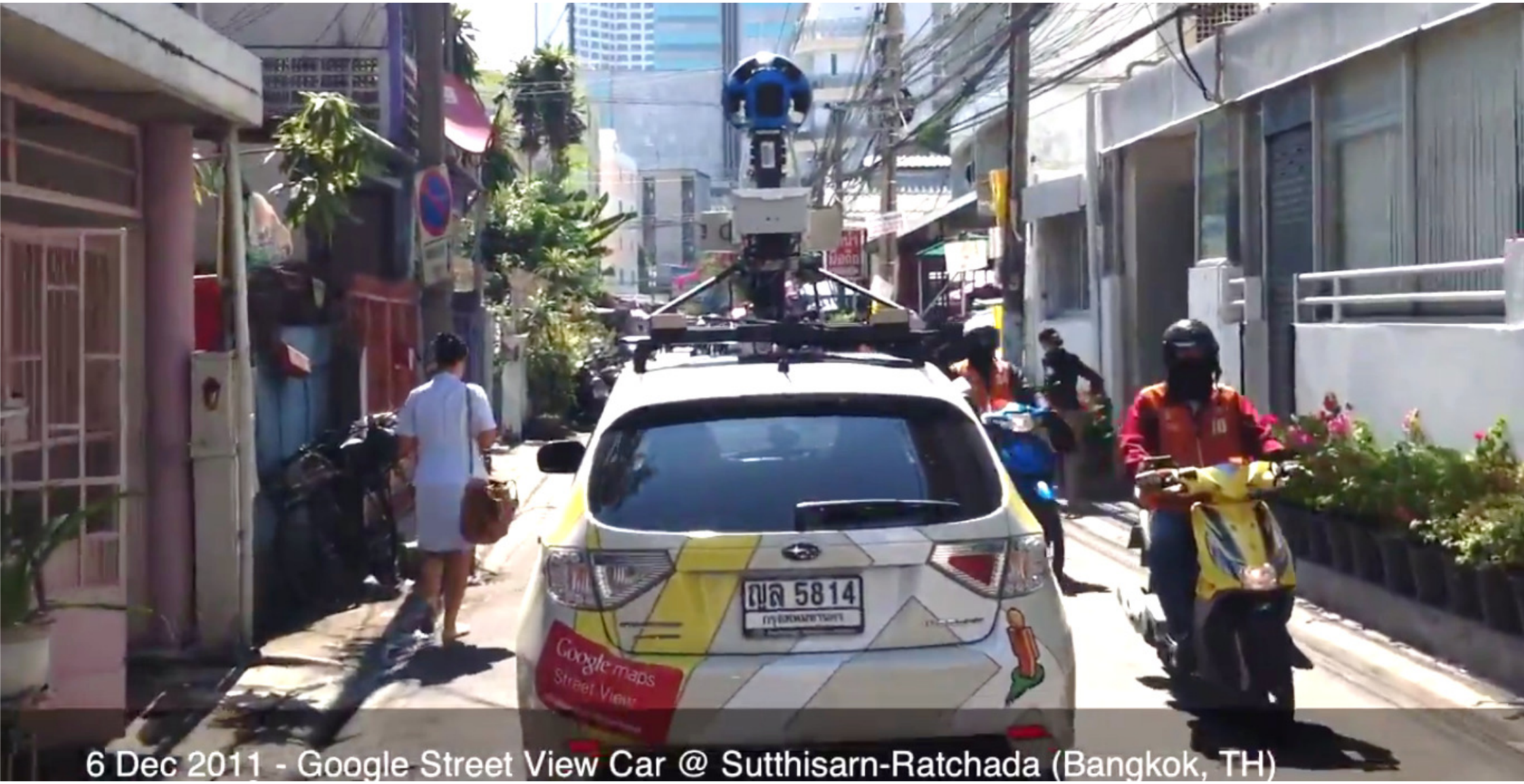




Chula DataScience

CHULA ENGINEERING
Foundation toward Innovation

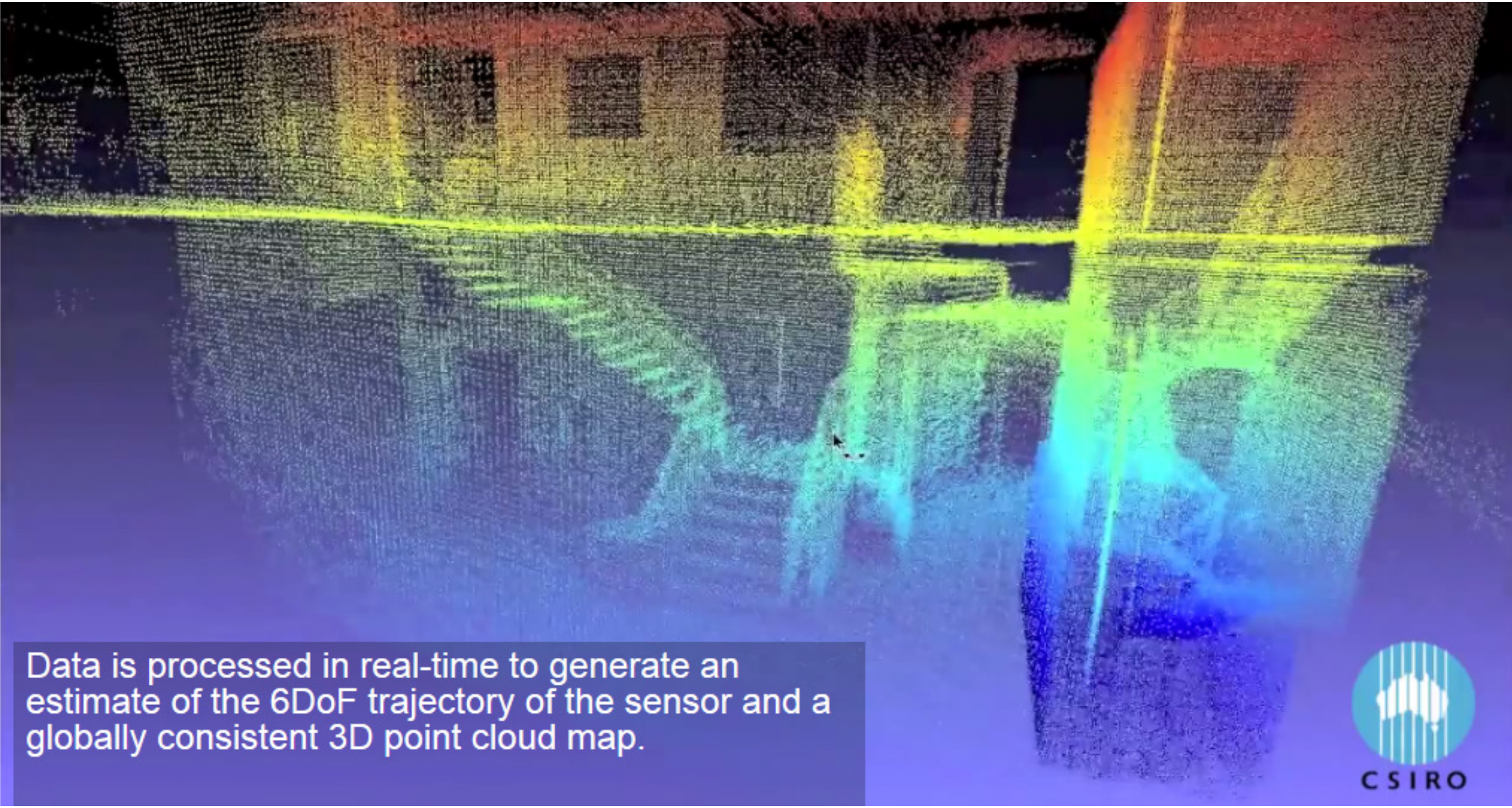
With modern mapping techniques,
everywhere will be on high-resolution map.

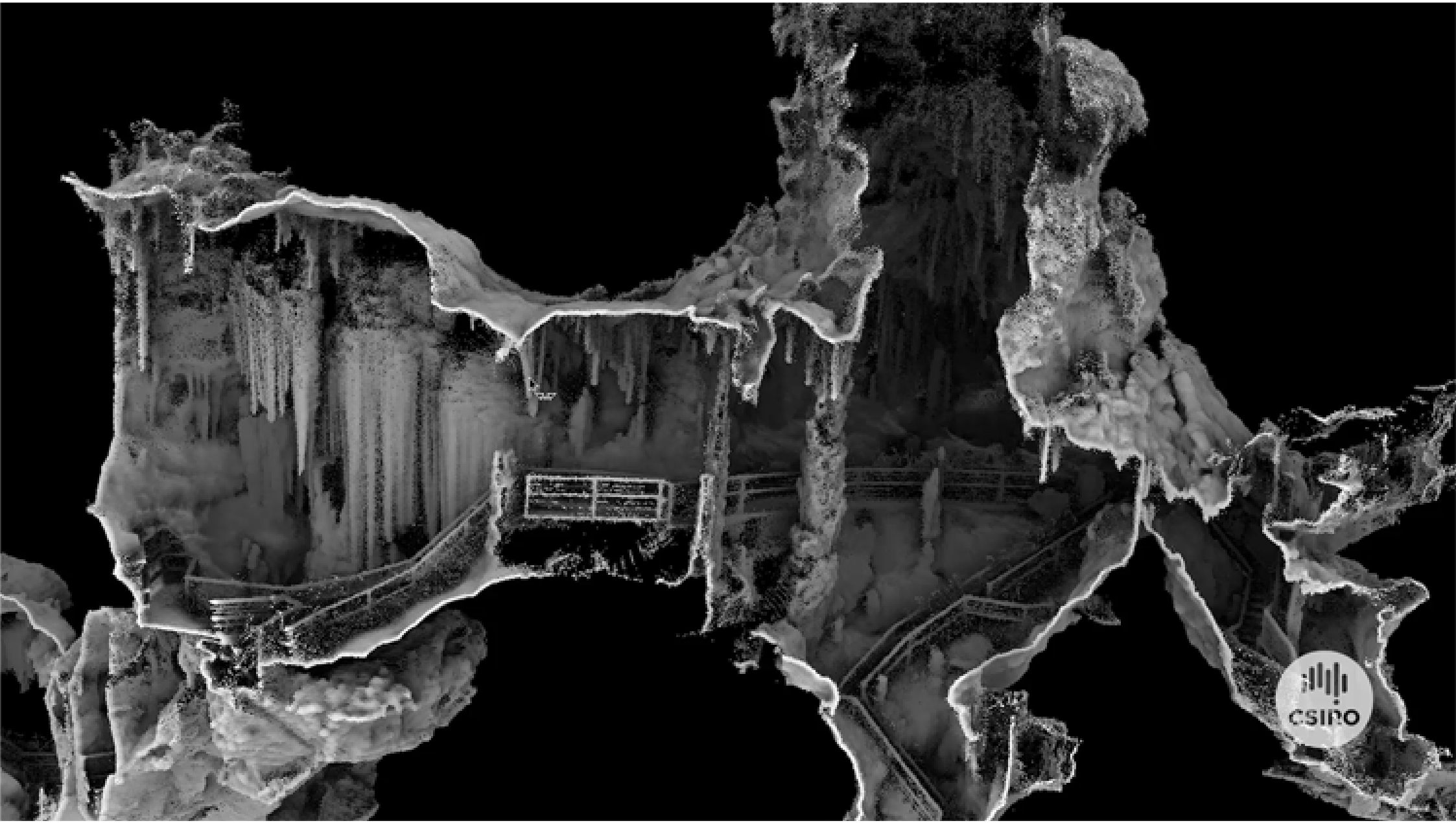


6 Dec 2011 - Google Street View Car @ Sutthisarn-Ratchada (Bangkok, TH)


CATLIN

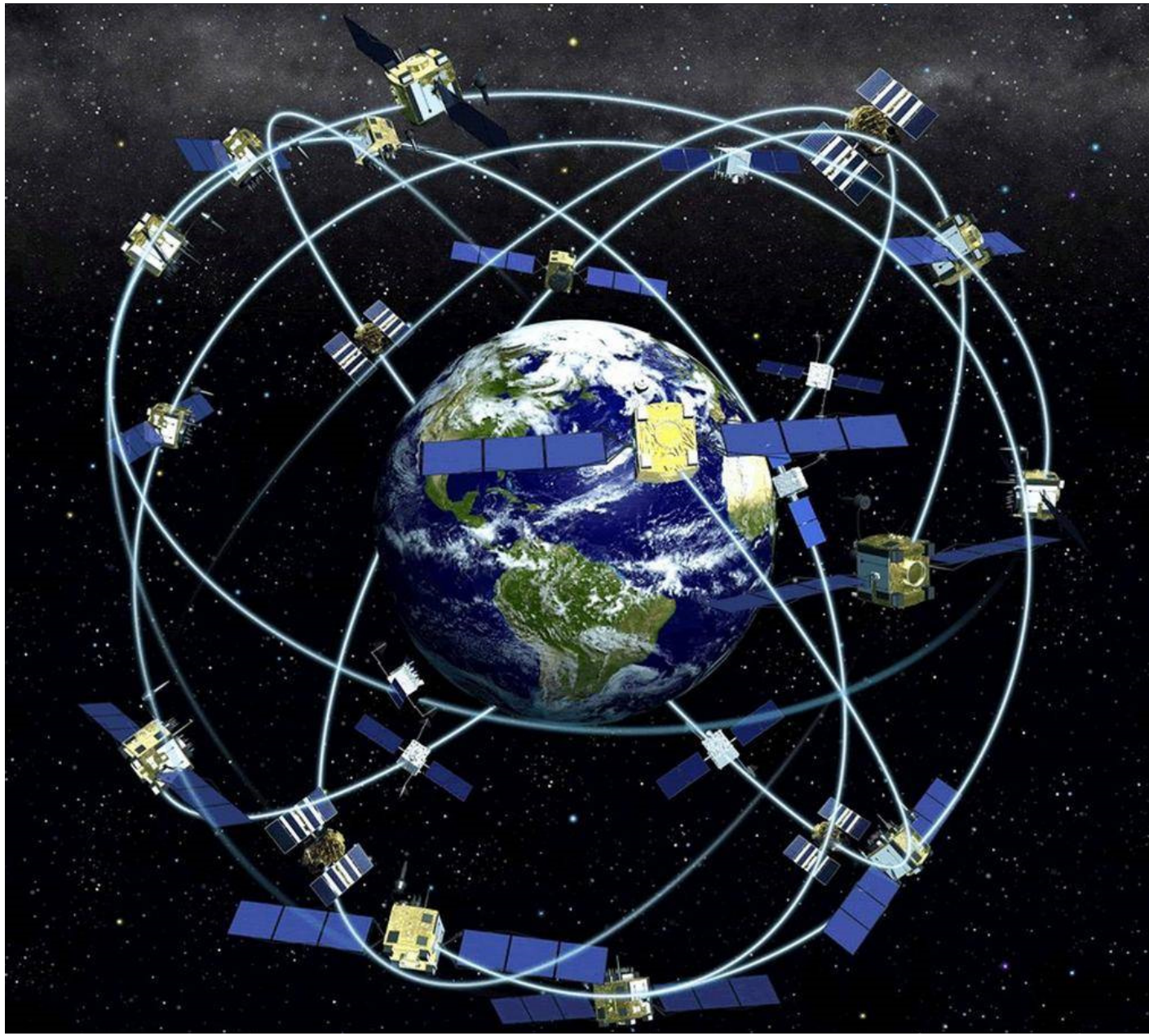

UAV LiDAR solutions


Data is processed in real-time to generate an estimate of the 6DoF trajectory of the sensor and a globally consistent 3D point cloud map.
CSIRO


CSIRO

With modern positioning techniques, everything can be located.

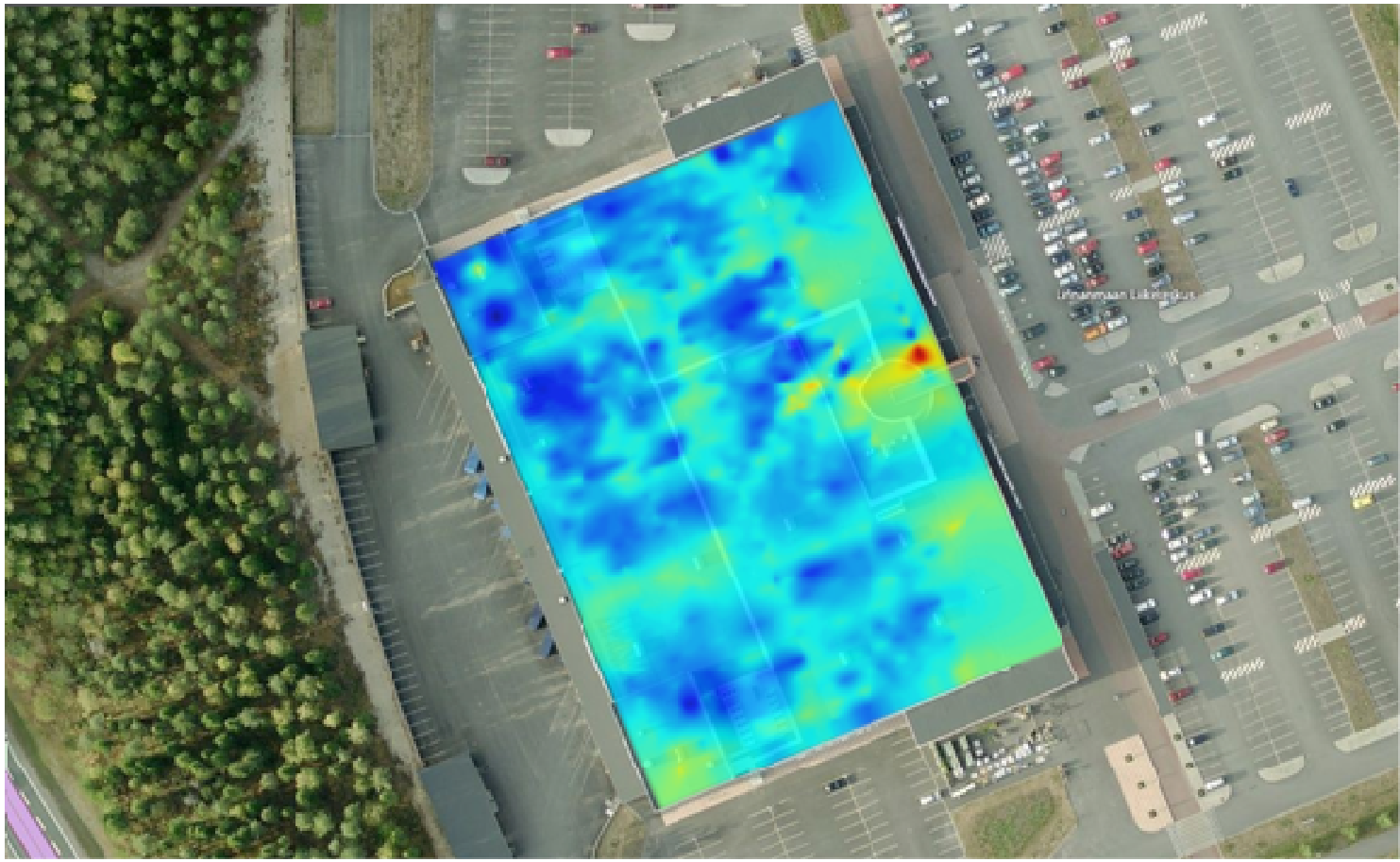
GPS



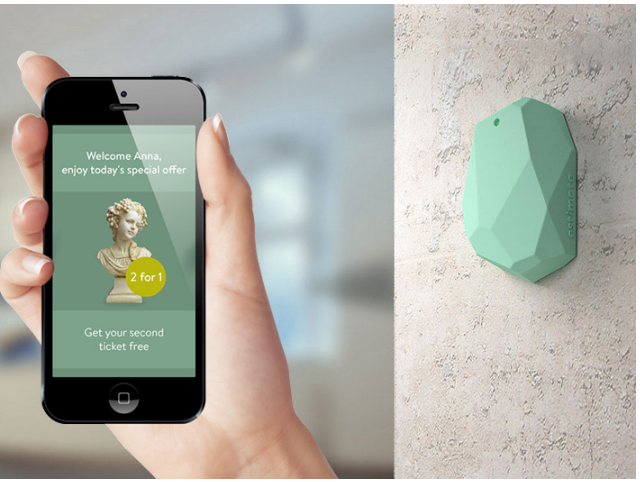
Finger Print

Signal strength of nearby WiFi access points

Magnetic field distortion by steel structure

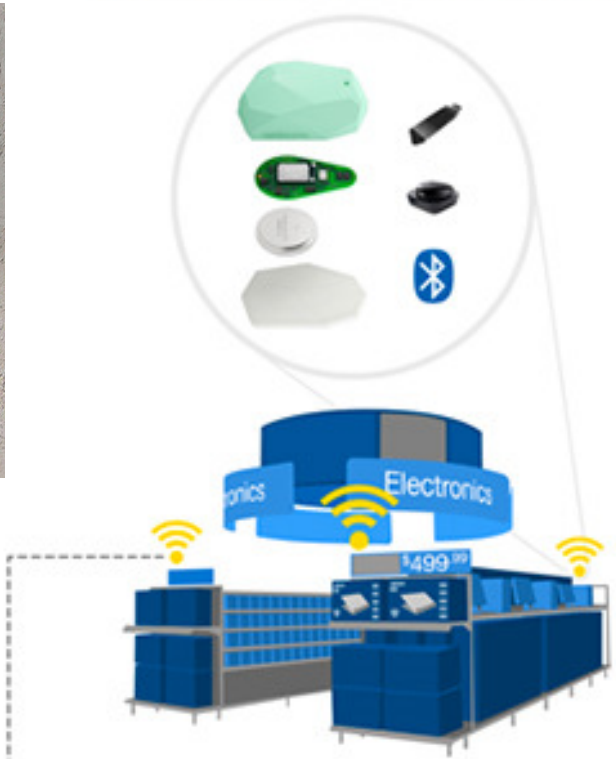


Beacon



HOW iBEACONS WORK

Bluetooth Low Energy (BLE) beacons are small enough to be located anywhere



Each beacon transmits a unique one way ID

THERE ARE MANY PRACTICAL BUSINESS APPLICATIONS OF iBEACON TECHNOLOGY.

The server uses these microlocation trigger IDs to tell either the mobile device, or another system, to perform a contextual action designed by the business owner.

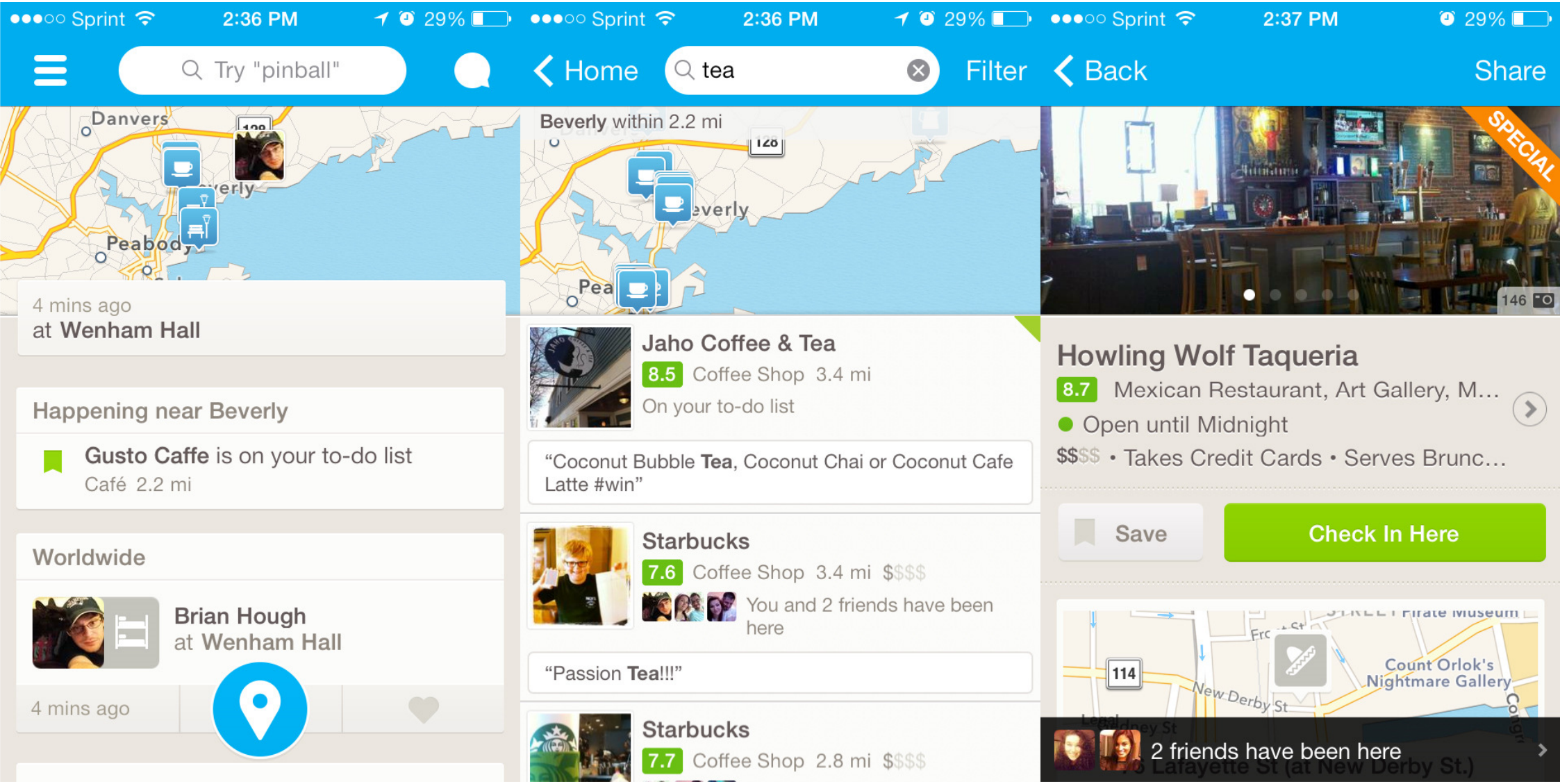
Dynamic in-store displays and contextual information

Light


Indoor Location-Based Services Using LED Lighting
How it Works
1. ByteLight-enabled GE LED fixtures "communicate" a unique light pattern using Visible Light Communication and Bluetooth Low Energy
2. Connected shoppers opt-in to "listen" with retailer's app on any smartphone and tablet with a camera and/or Bluetooth Smart
3. Camera detects unique light pattern and Bluetooth signal emitted by GE Lumination™ LED Luminaires; application notifies ByteLight platform of shopper's position and direction with sub-meter accuracy
4. Platform ties to retailer's digital marketing systems to deliver location-based services and personalized content to each shopper
ByteLight

Any kind of smart devices with known locations can be used as sensors that collect geospatial data.

Consumer Sensors

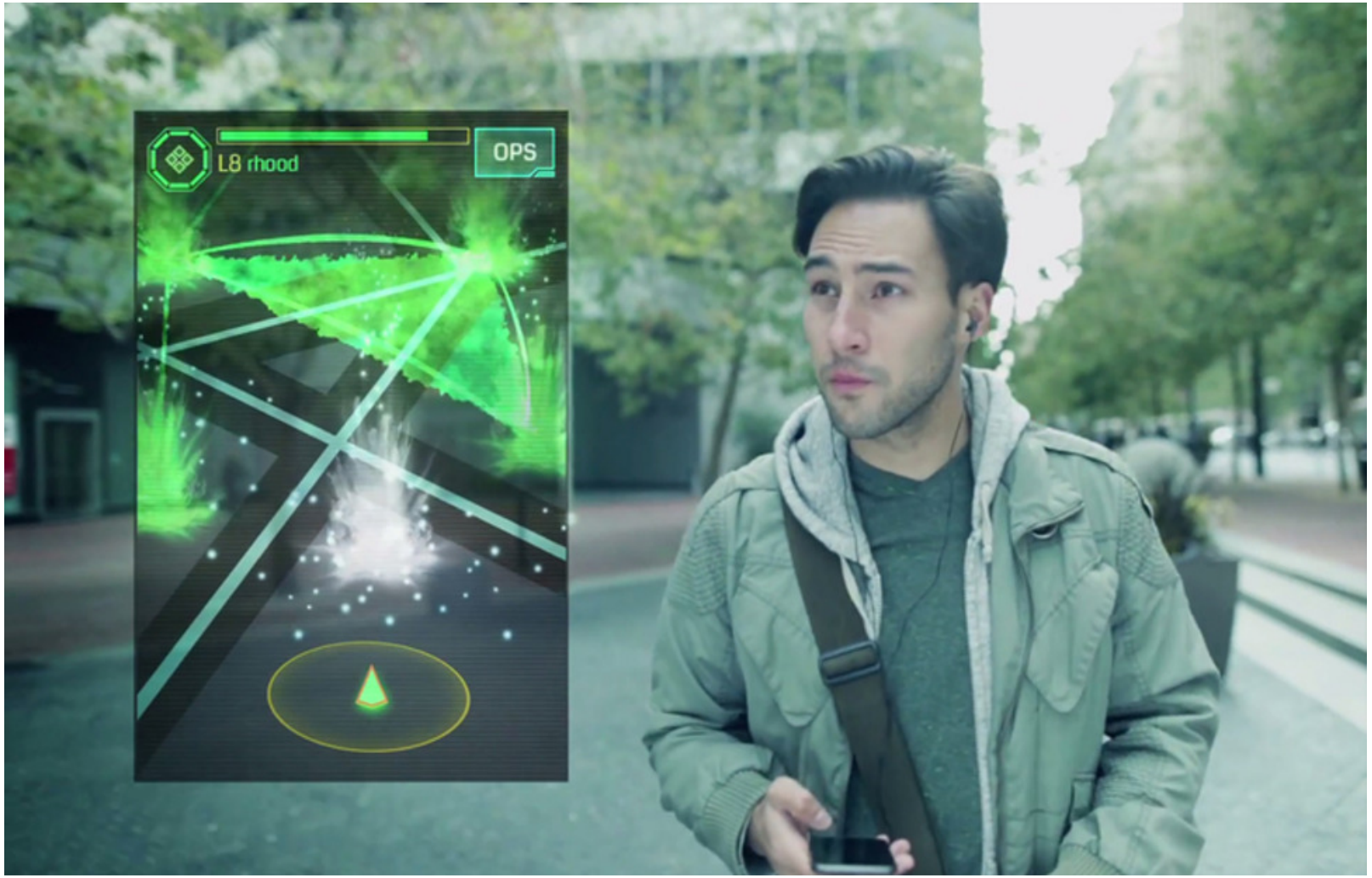


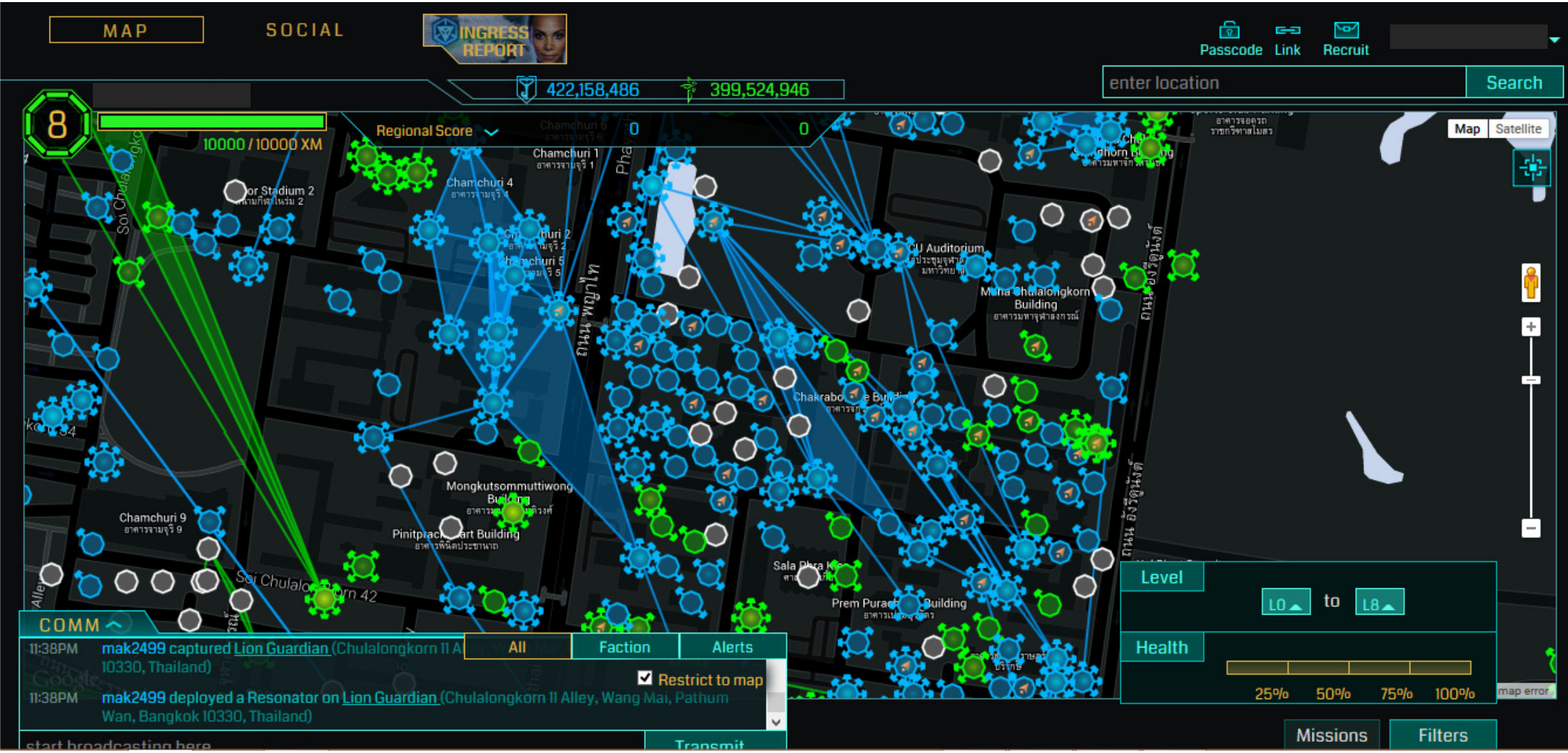
Online
Shopping

Chula DataScience

CHULA ENGINEERING
Foundation toward Innovation

Human Mobility Sensors




MAP
SOCIAL
INGRESS REPORT
Passcode
Link
Recruit
422,158,486
399,524,946
enter location
Search
8
10000 / 10000 XM
Regional Score
0
0
Map
Satellite
Chamchuri 1
Chamchuri 4
Chamchuri 9
Mongkutsommuttiwong Building
Pinitpracha Art Building
CU Auditorium
Maha Chulalongkorn Building
Prem Purachatra Building
Soi Chulalongkorn 42
Level
L0
to
L8
Health
25%
50%
75%
100%
map error
COMM
All
Faction
Alerts
Restrict to map
11:38PM mak2499 captured Lion Guardian (Chulalongkorn 11 Alley, Wang Mai, Pathum Wan, Bangkok 10330, Thailand)
11:38PM mak2499 deployed a Resonator on Lion Guardian (Chulalongkorn 11 Alley, Wang Mai, Pathum Wan, Bangkok 10330, Thailand)
start broadcasting here
Transmit
Missions
Filters


ยินดีต้อนรับ
Welcome
180.00
SAMSUNG
คู่ชีวิต
20:38
AIS


EASY PASS
ช่องอัตโนมัติ
EASY PASS
5
A
6
A
EASY PASS
เฉพาะรถ 4 ล้อ
EASY PASS
เฉพาะรถ 4 ล้อ
EASY PASS
เฉพาะรถ 4 ล้อ
EASY PASS
เฉพาะรถ 4 ล้อ
3.2ม.