

National Research Council of Italy



Engineering

Mobile Traffic Data Analytics

M. Fiore *CNR-IEIIT, Italy*

AlgoTel

Bayonne, France

May 25, 2016

Contributors



Razvan Stanica



Vincent Gauthier



Mounim El Yacoubi



Ghazaleh Khodabandelou Angelo Furno



Marco Gramaglia



Diala Naboulsi



Outline

\rightarrow Embarrassingly quick intro to mobile traffic data

The why and what of mobile traffic data

→ Some (contextualized) research examples

- Classification of a multi-disciplinary literature
- Overview through case studies
 - 1 *land use* detection
 - 2 dynamic population estimation
 - 3 mobile demand profiling
 - (4) *privacy of subscriber trajectory* databases
- References



r

[Part 1]

[Part 2]



Introducing mobile traffic data Context and collection

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016



Context (I)

- The surge of mobile activity...
 - Who doesn't own a smartphone today?

Mobile devices popularity rises them to status symbols
 Mobile apps and services are driving societal changes

- ...and some collateral effects of our interest
 - Oigital activities are easily tracked and recorded

Such rich datasets
 represents a unique
 opportunity to *monitor human behavior* at
 huge spatiotemporal





scales



Instagram

Context (II)

- Relevance of mobile traffic data
 - Mobile service providers
 - new service models, e.g., LBS
 - tailoring of services (and ads) to users
 - reselling of (aggregate?) data to third parties
 - Mobile operators
 - huge challenge in accommodating the demand
 - understanding subscribers can entail better network design and resource management
 - Research community C-RAN, SDN, NFV 🔀
 - multiple disciplines: studies can be run at scale on very large populations of individuals (impossible with traditional approaches)







~



M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016

Mobile traffic data collection (II)



RNC probes

Monitor signaling related to Radio Resource Control (RRC) operation. Fine-grained state changes at each mobile device, e.g.

- attach and detach operations
- start and conclusion of sessions
- all HO and LU events
 Plus, record key performance
 indicators on data transmission,
 such as the uplink and downlink
 throughput experienced by devices.
- imaximum positioning precision

not supported by all equipment large number of probes



Mobile traffic data collection (III)



(**(**p) femtocells





Cell

sector

MSC probes

Similar to RNC probes, but only in the circuit-switched (CS) domain. Same type of information as RNC probes, but

- limited to voice/texting
- missing information on events managed locally by RNC, e.g., intra-RNC HO

🙂 fair positioning precision fair number of probes

🔆 only voice/texting information







Location

boundary

Area

AlgoTel | May 25, 2016

Mobile traffic data collection (IV)

GGSN/PGW probes

Monitor data transiting by packet-switched (PS) domain gateways in order to reach the Internet. For each data traffic session of every device, record Packet Data Protocol (PDP) Context information:

- device identifier
- session start and end time
- traffic volume

 type of service (at different levels of detail, from transport protocol, to service type, up to precise application)
 These probes can also store device locations, corresponding to the cell where the device last updated its PDP Context from. They monitor Authentication, Authorization and Accounting (AAA)
 procedures triggered by PDP Context updates passing by GGSN/PGW.

🙂 detailed information on IP traffic

very large areas (e.g., countries) covered by a few probes

🙁 no voice/texting information

stale view of device location (infrequent PDP Context updates)







Mobile traffic data collection (V)









connected devices



CGF probes

0

Monitor information stored for billing purposes, i.e., so-called Call Detail Records (CDR). These include, for each device activity:

- device identifier
- start (and possibly end) time
- cell where the activity starts (and possibly ends)
- type of activity: voice, texting, data (possibly including traffic volume)
 In some rare cases, these probes can also record HO information.
- typically one single probe readily available, well formatted acceptable positioning precision

🕑 no precise traffic information





M. Fiore – Mobile Traffic Data Analytics

Mobile traffic data collection (VI)

Takeaway messages

- Not all mobile traffic datasets are created equal!
- The *quality* of mobile traffic datasets is increasing
 - new probes, more extensive deployments in both RAN and CN
- The *availability* of mobile traffic datasets is growing
 - operators become more open-minded
 - challenges are great occasions to access data!

BIGDATACHALLENGE 2015



- Caveat
 - Here, we are only considering "half of the world"
 - **client-side** mobile traffic data is an equally rich source of knowledge

Nokia Mobile Data Challenge MIT Reality Mining Microsoft GeoLife





A literature classification Considering works with operator-side data

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016

Literature classification

- A highly multidisciplinary research field
 - Mobile traffic data is used in many domains (with overlaps)



3 Social analysis And land use detection

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016

5

Social analysis (I)

- Relating mobile traffic and social features
 - Physics, sociology, epidemiology, networking



Social analysis (II)



- Example: environment
 - Intertwining of *land use* and mobile traffic activity



Social analysis (III)



- Example: environment
 - Case studies: 10 major cities in France and Italy











M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016 🔁 🚺

Social analysis (VI)

• Example: environment



[1] Soto et al., ACM HotPlanet 2011; [2] Cici et al., ACM MobiHoc 2015

Social analysis

Demographics

Mobile call graph models

User interactions

Mobile call graph properties

Epidemics

And population estimation

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016 🖊

Mobility analysis (I)

- Extracting mobility information from mobile traffic
 - Physics, sociology, transportation, networking, epidemiology





Mobility analysis (II)

Example: transportation systems



Estimating citywide *inhabitant density*





Mobility analysis (III)

- **Example: transportation systems**
 - Estimating citywide *inhabitant density*



Mobility analysis

Human mobility

Fundamental laws

Validation

Models

Mobility analysis (IV)

- Example: transportation systems
 - Estimating citywide *population distribution over time*
- 6. extending the model to day hours



these mobile traffic variables enable a day-long estimation

population density becomes a function of **presence** and **average activity**







Validation

Models



Mobility analysis

Human mobility

Fundamental laws







model parameters lpha and eta are linear functions of the users' average activity

consistent through evening-night hours





5 Networking analysis And mobile demand profiling

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016 🖊

Networking analysis (I)

- Adapting network design to the mobile demand
 - Networking, sociology



Networking analysis (II)

• Example: mobile demand

Aggregate access network *traffic profiling*



Networking analysis (III)

- **Example: mobile demand**
- **1.** Mobile traffic *snapshot* at one time instant
- metric
- selection of call, text, data traffic volume
- temporal resolution
- one hour
- regions from spatial spatial profiling aggregation







Number of clusters

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016



Networking analysis (IV)

- Example: mobile demand
 - Case studies: major cities in France, Italy and Ivory Coast
 - Shared (expected) profiles...
 - **1.** working hours
 - evening/weekend
 - 3. night (diverse)
 - ... but unexpected profiles exist!



Aggregate access network traffic

nightlife profile





Date

AlgoTel | May 25, 2016

commuting profile

Network analysis

Networking solutions

Technologies

Privacy solutions

Marketing solutions

Mobile demand

Individual access network traffic

Networking analysis (V)

- Example: mobile demand
 - A *limited set* of profiles captures all major dynamics in network usage...
 - ...yet there are outliers!
- Training (sub-)dataset
 Test (sub-)dataset

typical week profiles k-means of each snapshot on *k* identified profiles

Test snapshots are *outliers* if they are distant from the expected profile for that hour of the week [3,4]

Halloween night, Milan patron saint day, **opening of opera season** at "La Scala" theater, Christmas holidays and events, several public holidays, several football matches of AC Milan and Inter Milan, one collection **probe crash**, etc.



[3] Keralapura et al., ACM MobiCom 2010; [4] Shafiq et al., ACM SIGMETRICS 2011



Networking analysis (again) But a different case study: privacy

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016



Networking analysis (VI)

Example: technologies

enhancing user privacy in *trajectory datasets*





Networking analysis (VIII)

• Example: technologies

measuring the *anonymizability* of trajectories in a dataset



AlgoTel | May 25, 2016

M. Fiore – Mobile Traffic Data Analytics

References (and questions)

M. Fiore – Mobile Traffic Data Analytics

AlgoTel | May 25, 2016 🔀



References

• Surveys

- V.D. Blondel, A. Decuyper, G. Krings, "A survey of results on mobile phone datasets analysis," EPJ Data Science, 4:10, 2015
- D. Naboulsi, M. Fiore, S. Ribot, R. Stanica, "Large-scale Mobile Traffic Analysis: a Survey," IEEE Communications Surveys and Tutorials, 18(1), 2016

• More about the examples in this talk

- D. Naboulsi, R. Stanica, M. Fiore, "Classifying Call Profiles in Large-scale Mobile Traffic Datasets", IEEE INFOCOM, Toronto, Canada, 2014
- A. Furno, R. Stanica, M. Fiore, "Comparative Evaluation of Urban Fabric Detection Techniques Based on Mobile Traffic Data," ACM/IEEE ASONAM, Paris, France, 2015
- M. Gramaglia, M. Fiore, "Hiding Mobile Traffic Fingerprints with GLOVE," ACM CONEXT, Heidelberg, Germany, 2015
- G. Khodabandelou, V. Gauthier, M. El-Yacoubi, M. Fiore, "Population Estimation from Mobile Network Traffic Metadata," IEEE WoWMoM, Coimbra, Portugal, 2016
- A. Furno, D. Naboulsi, R. Stanica, M. Fiore, "Mobile Demand Profiling for Cellular Cognitive Networking," IEEE Transactions on Mobile Computing, to appear
- Advertisement: I am currently looking for postdocs!

