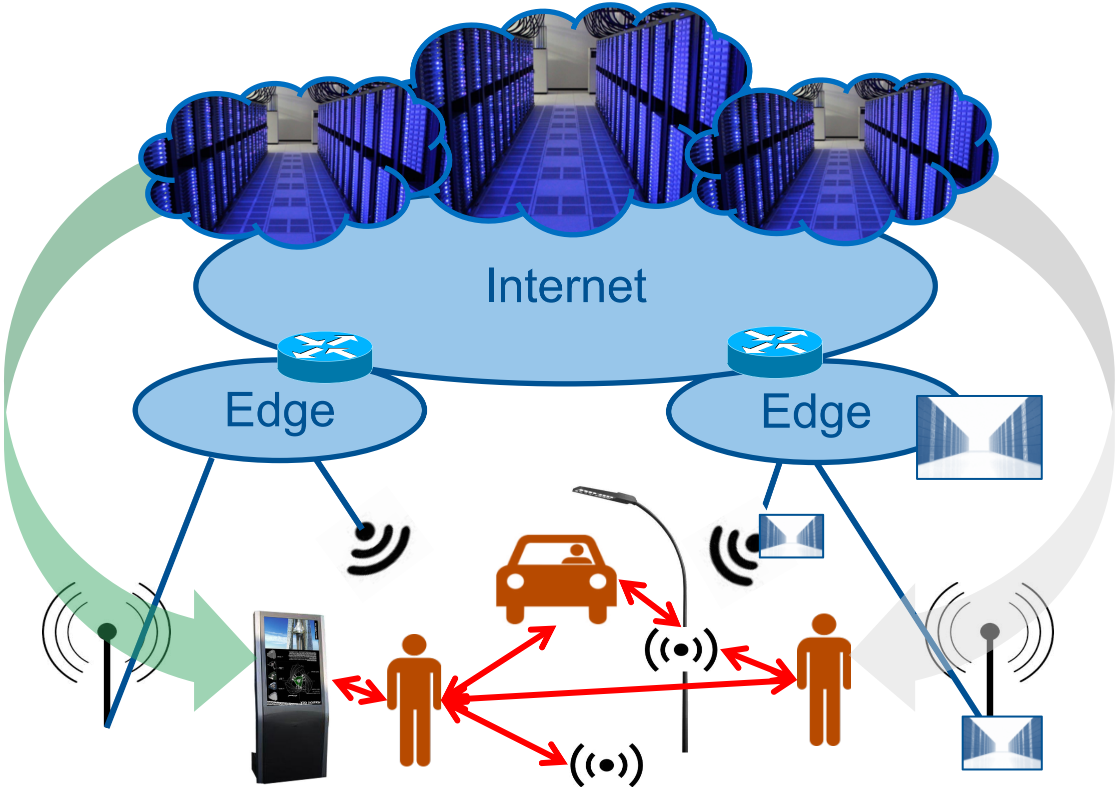
Connecting the (mobile) Dots: Mining User Mobility Patterns from Networks and Social Media

Jörg Ott
Ljubica Kärkkäinen
Leonardo Tonetto

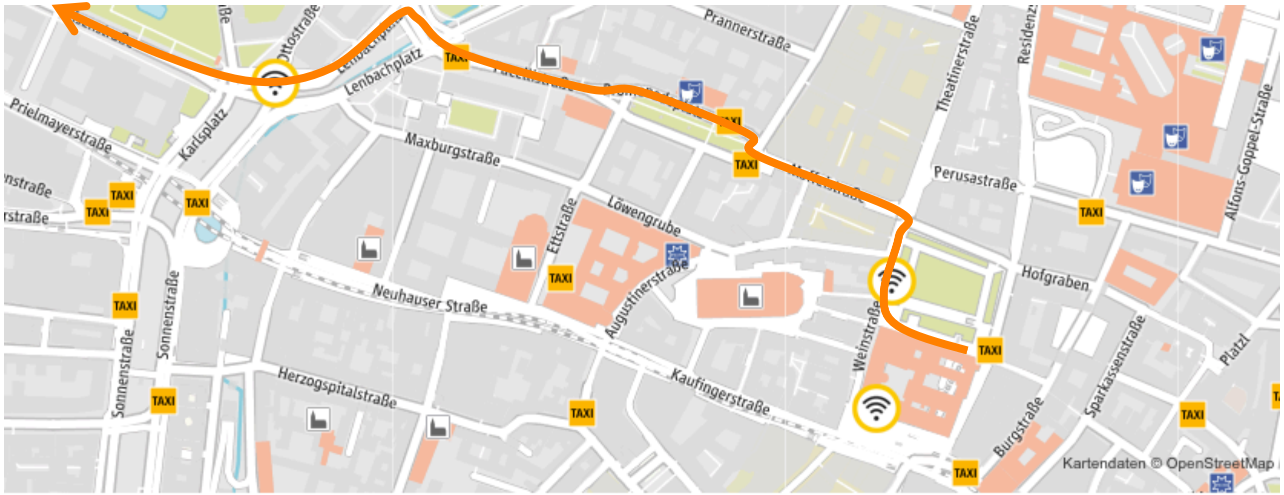
www.cm.in.tum.de

15 December 2017

Why do we care? (1)

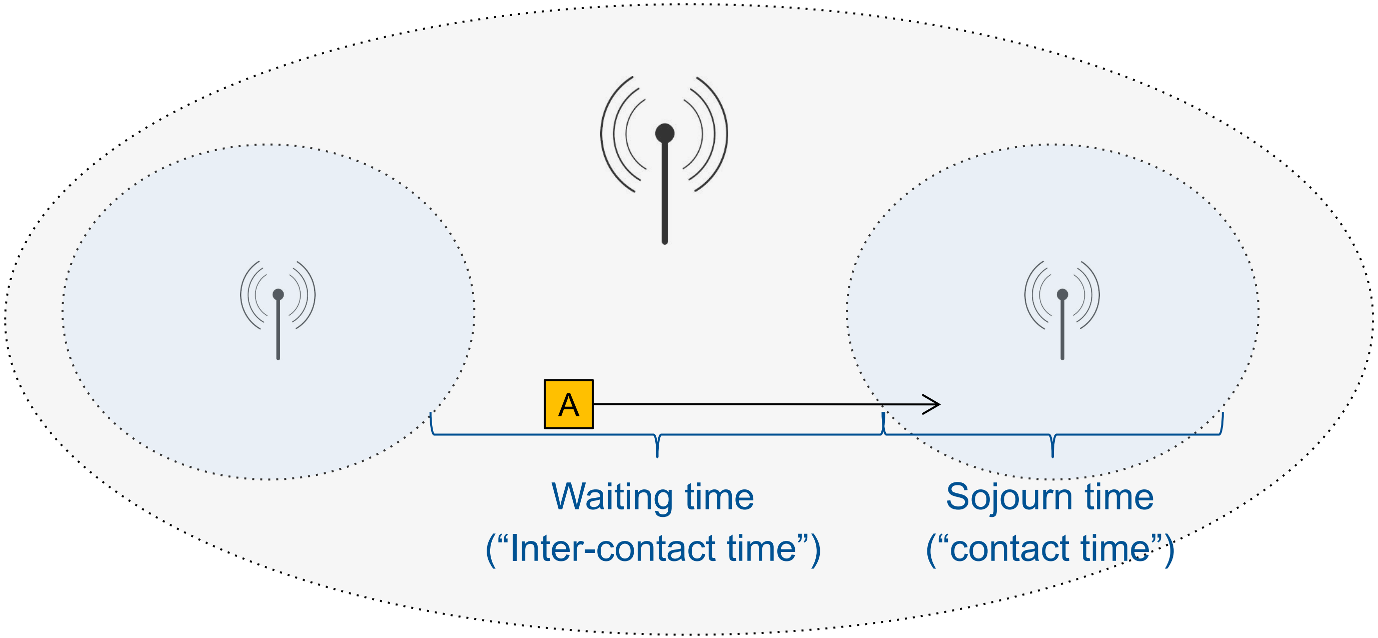


Why do we care? (2)

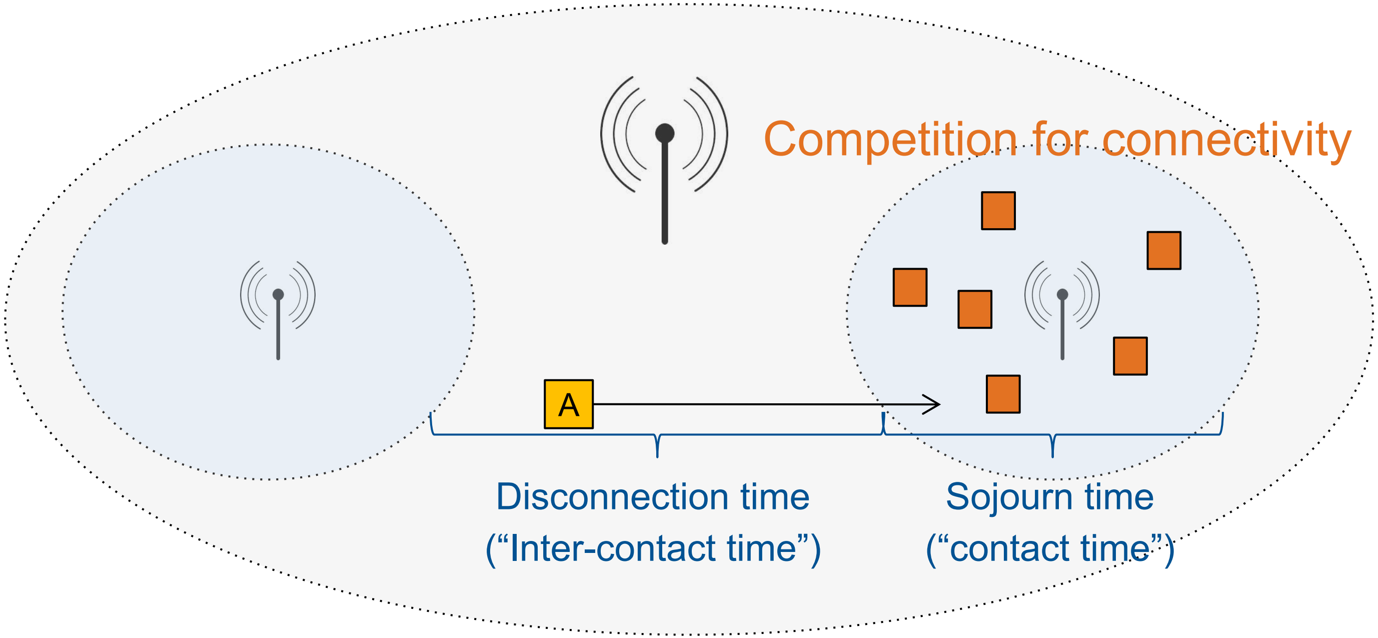


http://www.muenchen.de/leben/wlan-hotspot.html

Why do we care? (3)



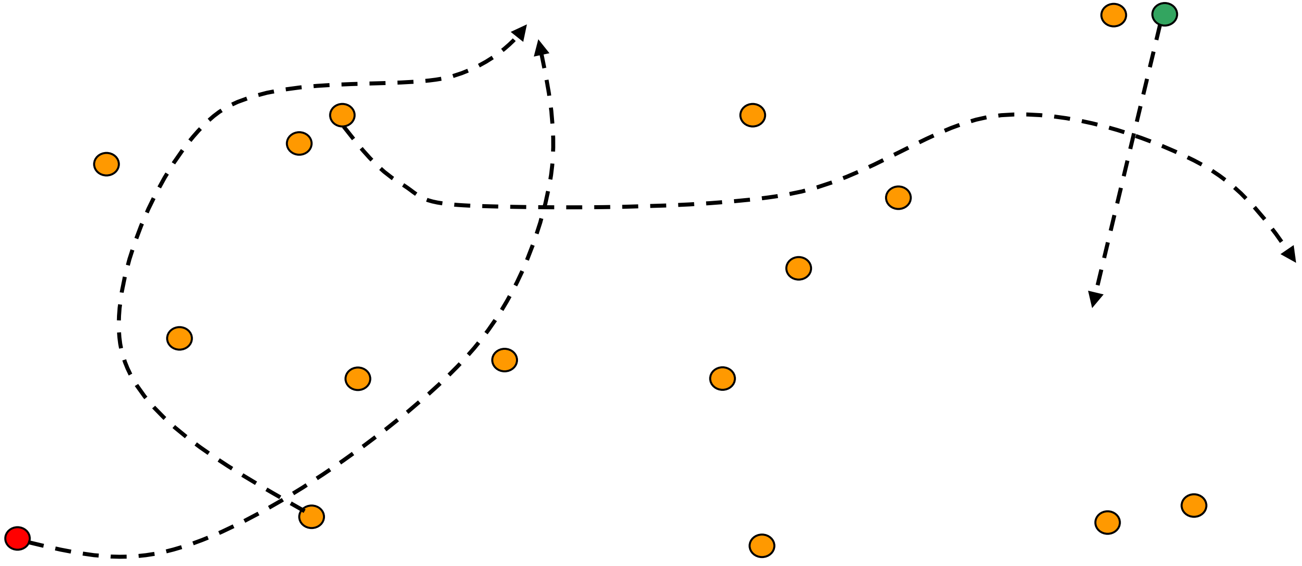
Why do we care? (3)



Network connectivity
(Availability, data rate, disruptions)

Why do we care? (4)

Contact time (duration) Inter-contact time Inter-encounter time Node density



Content spreading performance
(infection rate, coverage, latency)

Content delivery performance
(delivery probability, latency)

Sample application areas

Assessing (cellular) Internet access performance
- Example: complementing cellular networks by Wi-Fi or others

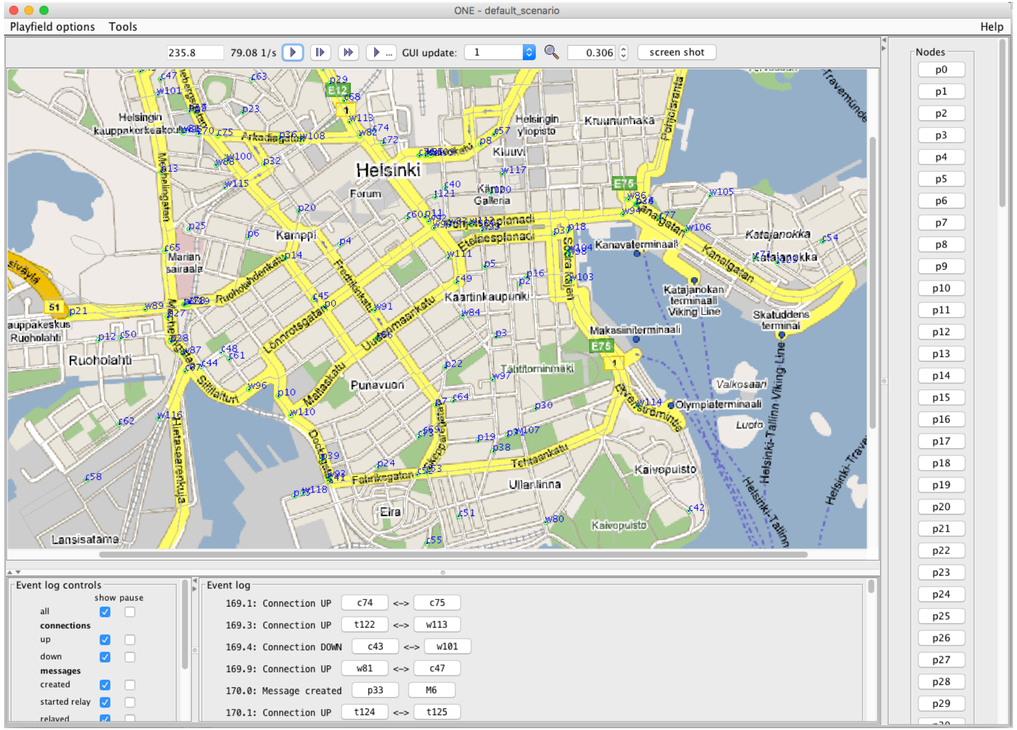
Local content dissemination
- Example: neighborhood networks for content sharing without the cloud

Censorship-resistant information exchange
- Example: Firechat during the Hong Kong protests

Enabling communication in rural or remote areas
- Example: Interconnecting villages through “messengers”

…

The ONE Simulator: mobility & apps



http://akeranen.github.io/the-one/

Working Day Movement Model

Synthetic model for daily routines
- Defines home, work places, favorite locations
  - Map-based movement with different regions
- Considers social groups for evening activities

Multi-modal transportation
- Submodels for walking, cars, trams/buses
- Simple schedules for public transportation
  (can also support real schedules)

Activity models
- Shortest path routing from/to places
- Not much movement at home
- Random movement at work
- Evening activities
- Variable parameters for daily routines (wake time, workday length, etc.)

City Square Model

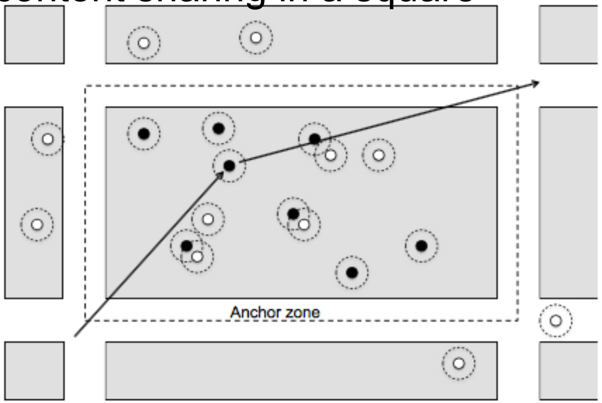
Observation-based model: web cams as sources

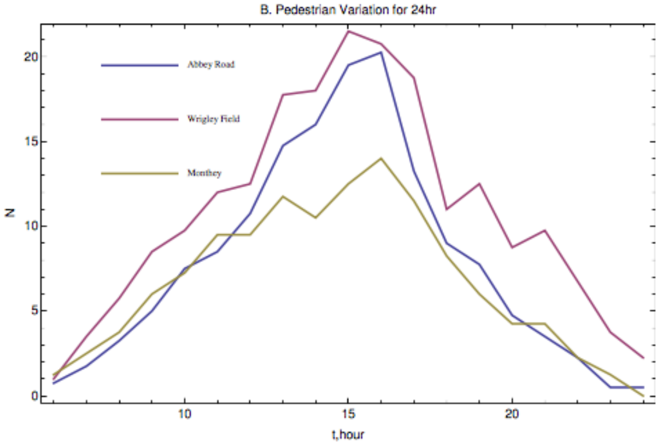
- Manual interpretation (no usable automated algorithms at the time)
- Shown to be automatable for vehicle density)

Computing arrival patterns and sojourn times

- Deriving a simple model with a few parameters

Application: content sharing in a square





More on mobility models

Trace-driven modeling and analysis

- for movement patterns
- for communication and interaction patterns
- how these influence each other

Mobility modeling using traces from the network

Mobility modeling using traces from social media

150 Jahre
culture of
excellence

Connecting the (mobile) Dots: Mining User Mobility Patterns from Networks and Social Media

Jörg Ott

Ljubica Kärkkäinen

Leonardo Tonetto

www.cm.in.tum.de

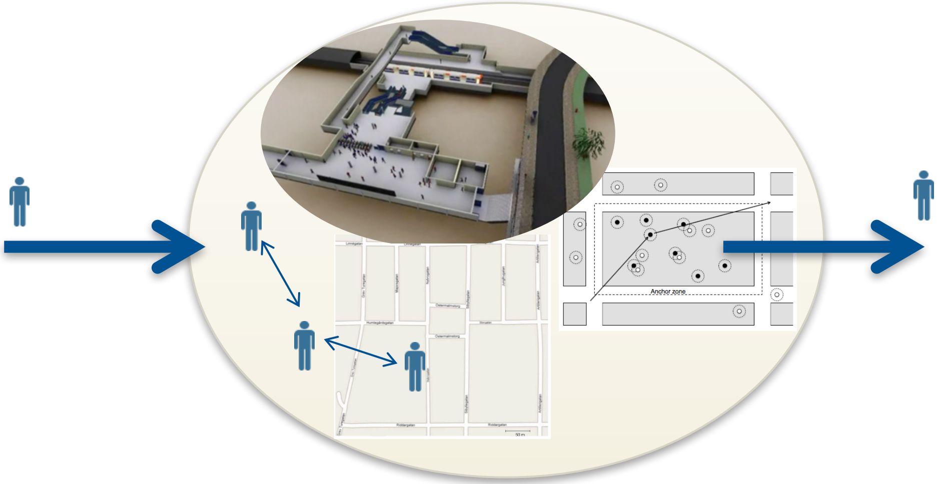
15 December 2017

Outline

- Analytic mobility models
- Mobility modeling from wireless network records
- Modeling user association patterns in a university campus network
- Ongoing work: Mobility prediction

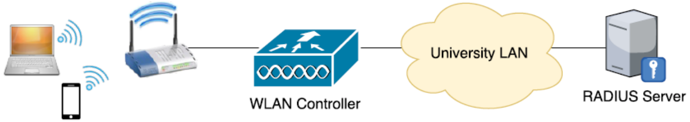
Analytic Mobility Modeling

- Queueing model for mobility of users and their interaction
- Black-box model for (small) urban areas: city square, subway stations, grid of streets, buildings
- Application: infrastructureless content sharing and ephemeral networks



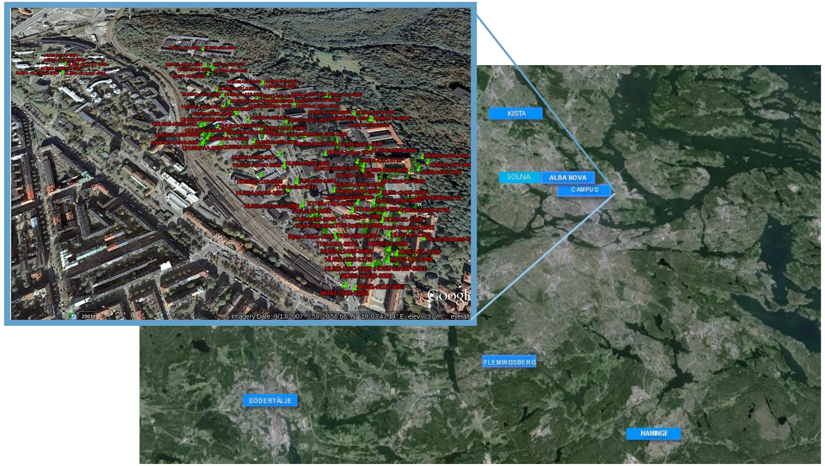
Trace-driven Mobility Modeling

- Objective: extract mobility traces from wireless network traces
- Wireless network records: Eduroam association events
- Trace obtained from the authentication server



- Description:
  - Trace duration: 16 months (January 2014—April 2015)
  - 250-300K associations from 13-15K users per day
  - ~1000 access points, located in 54 buildings on 5 sites
  - Not really big data, but…

has some challenges!

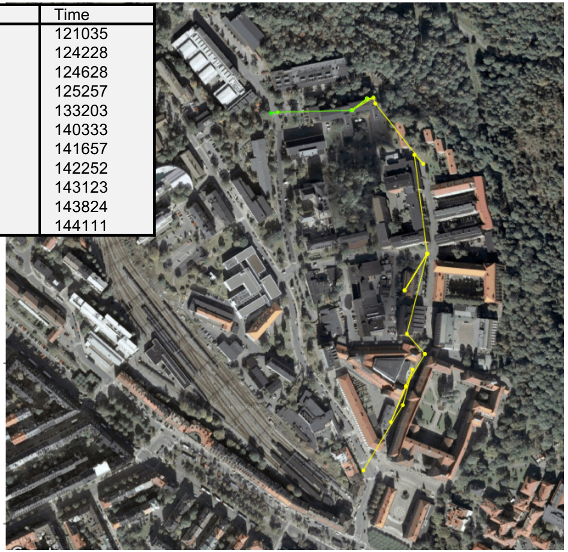


Trace-driven Mobility Modeling

- Objective: extract mobility traces from wireless network traces (location, duration of visiting time)
- Challenges: no accounting information, user devices are anonymized, access points labeling inconsistent, ping-pong effects, filtering short associations…

- Approach
  - Data cleaning
  - Capturing ground truth (*warwalking*, syslog)
  - Deriving heuristics: utilizing wireless coverage map, authentication state machine, infrastructure/device timers

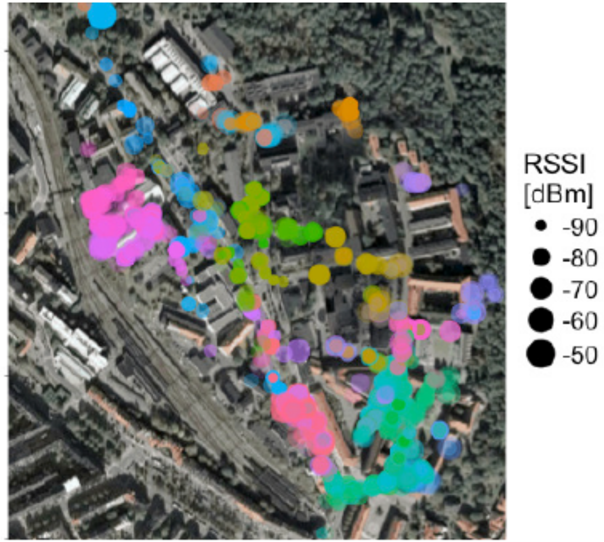
Access Point	Time
KTHB-r1201-0915	121035
KTHB-r1209-0917	124228
KTHB-r1209-0917	124628
KTHB-r1209-0917	125257
KTHB-r1209-0917	133203
KTHB-r1201-0607	140333
HUS20-r4055-0307	141657
QHUS15-kApl4-0285	142252
KTHB-r4212-0894	143123
KTHB-r4212-0894	143824
KTHB-r4212-0894	144111



Trace-driven Mobility Modeling

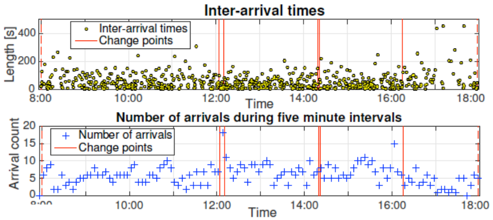
- Objective: extract mobility traces from wireless network traces (location, duration of visiting time)
- Challenges: no accounting information, user devices are anonymized, access points labeling inconsistent, ping-pong effects, filtering short associations…

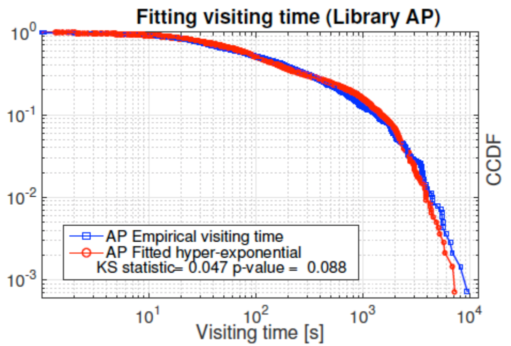
- Approach
  - Data cleaning
  - Capturing ground truth (*warwalking*, syslog)
  - Deriving heuristics: utilizing wireless coverage map, authentication state machine, infrastructure/device timers

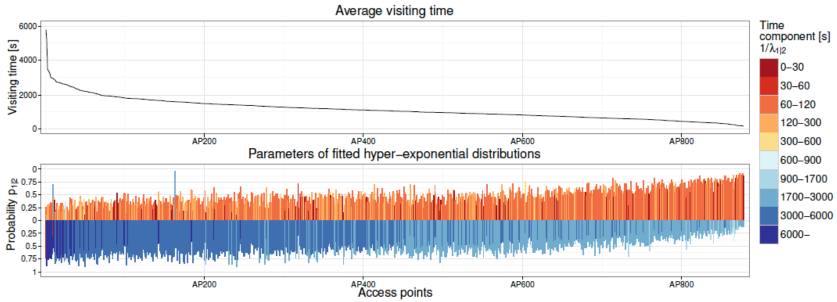


Modeling User Association Patterns

- Analysis and modeling of users' arrival patterns and visiting time at network access points
- Application: resource allocation, wireless protocol design, network dimensioning, abnormality detection
- Findings: Nonhomogeneous Poisson arrivals, two-stage hyper-exponential visiting time
- Tractable (and simple!) models, but time-varying and location-specific

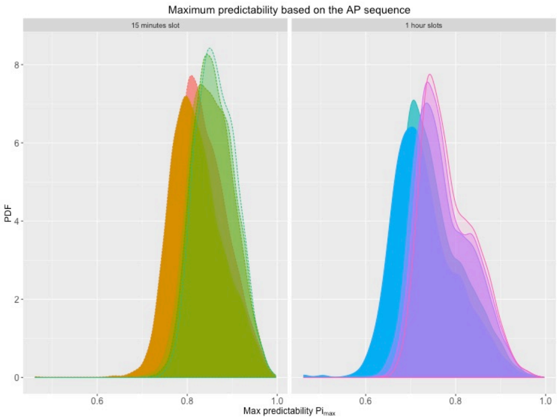


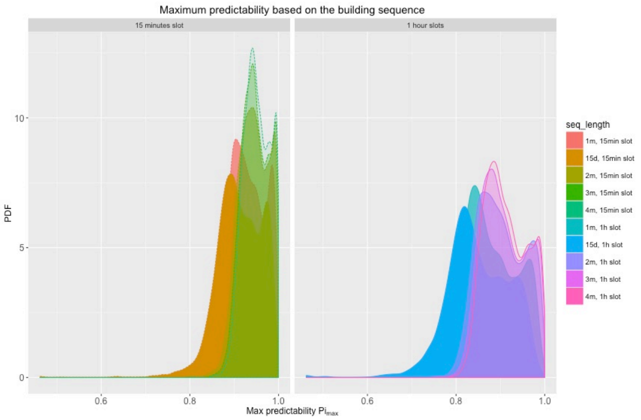


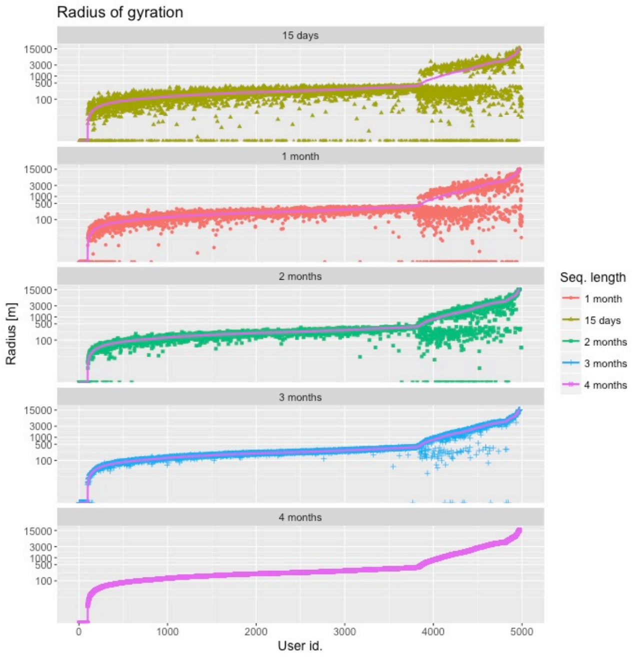


Ongoing Work: Mobility Prediction

- Estimating achievable predictability of the user's location based on entropy







Connecting the (mobile) Dots: Mining User Mobility Patterns from Networks and Social Media

Jörg Ott
Ljubica Kärkkäinen
Leonardo Tonetto

www.cm.in.tum.de

15 December 2017

150 Jahre
culture of
excellence

Pervasive Games and Human Mobility

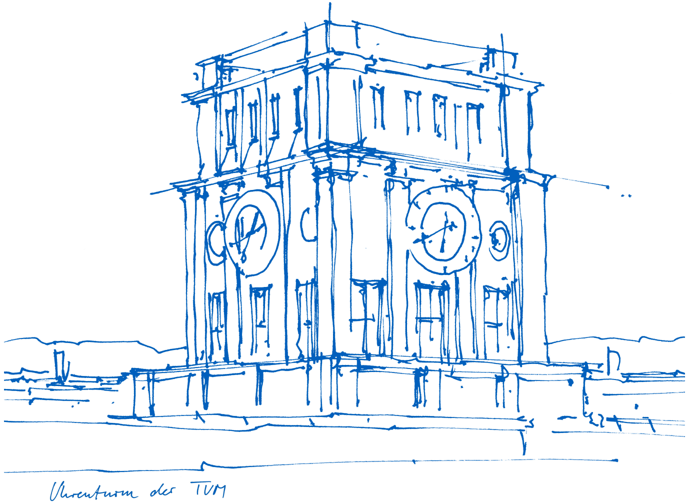
Leonardo Tonetto

Technical University of Munich

Department of Computer Science

Chair of Connected Mobility

Munich, 15. December 2017



What we found

- Increase of up to **2 km** in daily displacements, persistent after the game
- Gamers visit new **locations**, close to their past trajectories
- Gamers play for up **20 days** longer when playing often on cellular network

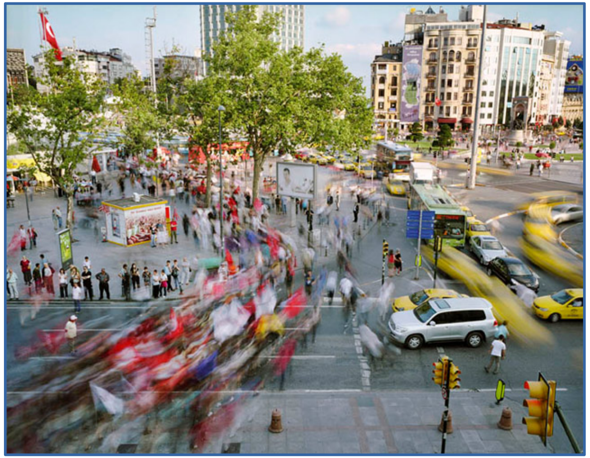
Background

- **Pervasive Games**

- Augmented Reality and Location Aware
- *Virtual*-world & *Real*-world
- Examples:

- **Human Mobility**

- Data driven modeling
- Understanding and prediction of users behavior
- Wireless network deployment, urban planning, ...



Our Datasets

- **Twitter**

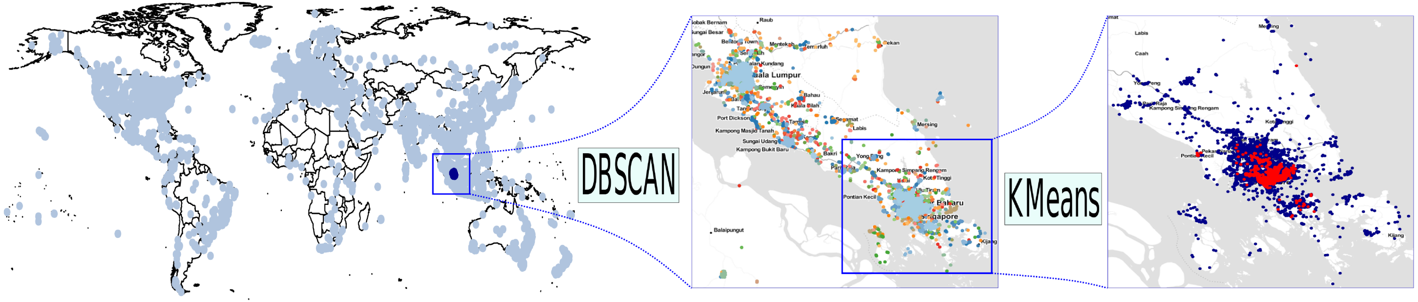
- 8.7M tweets from 21500 users
- 15 countries (18 cities)
- 8900 gamers with "#pokemongo"
- Bot detection with Botometer* (~3.1%)
- Spatial granularity: Fine
- Time granularity: Coarse

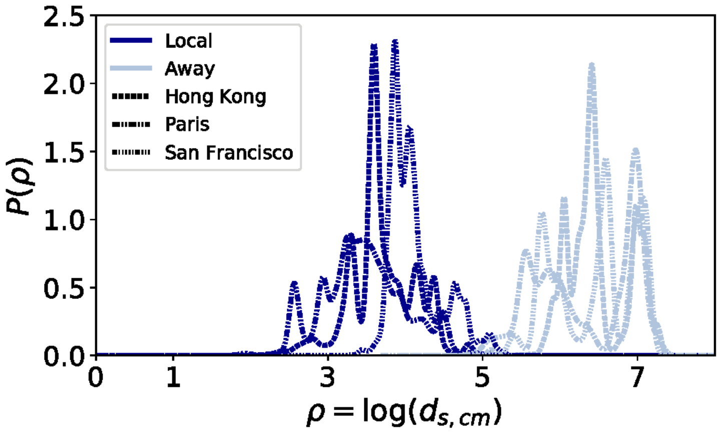
- **Carat**

- 62.8M records 58000 users
- +100 countries
- 3392 gamers
- Info about phone status/behavior
- Spatial granularity: 1D
- Time granularity: Fine (1% battery change)

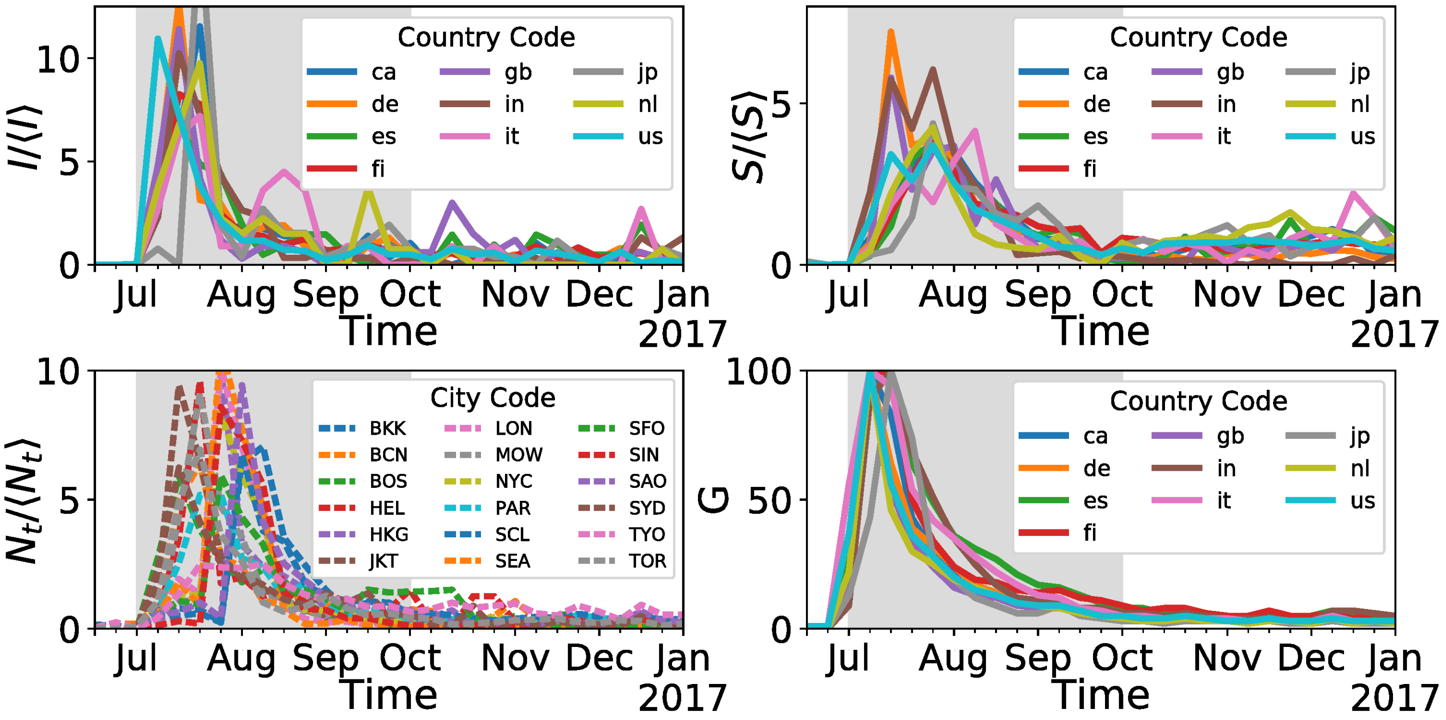
* https://botometer.iuni.iu.edu/

Spatial Clustering: Local vs. Away





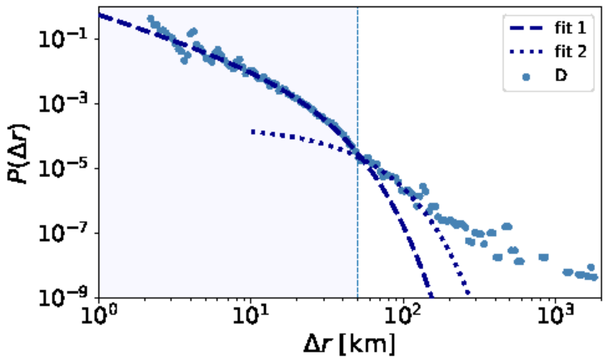
When was it trendy?



How each dataset was studied?

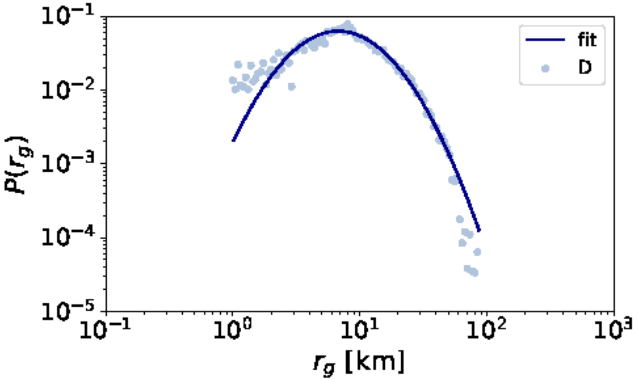
- **Carat**

- Displacements of consecutive records (Δr)
- Gaming sessions & installations
  - Average 2.3 km per session



- **Twitter**

- Radius of Gyration (r_g)
- Isotropy ratio
- *<u>Location</u>* visitation
- Displacements of consecutive records (Δr)



Combined Analysis (Carat & Twitter)

- Consistent scale between active days
  - Twitter: 59.2 days
  - Carat: 83.8 days
- Increase in daily mobility observed on Twitter
  - Supporting the observation in Carat

Conclusion

- Flow of people → Flow of information
- Synthetic Models vs. Real-World Data
- Mobility might be affected by exterior factors (Mobile Games)