Human Mobility Science:

Data, Measures, Generative models and Predictive algorithms

https://humanmobility-tutorial.github.io/

SoBigData

EPSRC
Engineering and Physical Sciences
Research Council

Filippo Simini
@ipif

Gianni Barlacchi
@giannibarlacchi

FONDAZIONE BRUNO KESSLER

Luca Pappalardo
@lucpappalard

ECML/PKDD 2018 - Dublin, Ireland

What Human Mobility Science is about

1. Understanding the laws of human motion
   statistical mechanics, pattern mining, network science

   - to what extent are humans predictable?
   - is there a typical traveling distance?
   - can we profile individuals according to their mobility behavior?
   - is an individual's mobility related to other personal characteristics, such as health or income?

What Human Mobility Science is about

2. Develop realistic models of human mobility
   mathematical modelling, inferential statistics, simulation

   - What determines the decision to start a trip?
   - What determines the choice of the destination?
   - What determines the decision to return or to explore?

What Human Mobility Science is about

3. Predicting future whereabouts
   machine learning, predictive analytics

   - Can we predict the future whereabouts of an individual given their mobility history?
   - Which semantic or contextual information can be exploited to improve the prediction?

Relevance and motivation

Mobility has a big impact on society:

- Health
  In the US the typical commute lasts more than 30 minutes.

- Environment
  Transportation is the second source of greenhouse gas emissions to the atmosphere.

- Economy
  In the EU transportation the second largest expenditure category after housing (15-25%).

https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-10

Applications

- Nowcasting
  estimate well-being, unemployment, poverty, recovery from natural disasters, segregation, crime patterns.

- Sustainable transport:
  - ride sharing and vehicle sharing
  - self-driving vehicles: will they change our mobility habits? will there be more or less congestions?

- Epidemic modelling

  simulating the spread of biological or computer viruses

What you will learn in this tutorial

1. The landscape of human mobility data

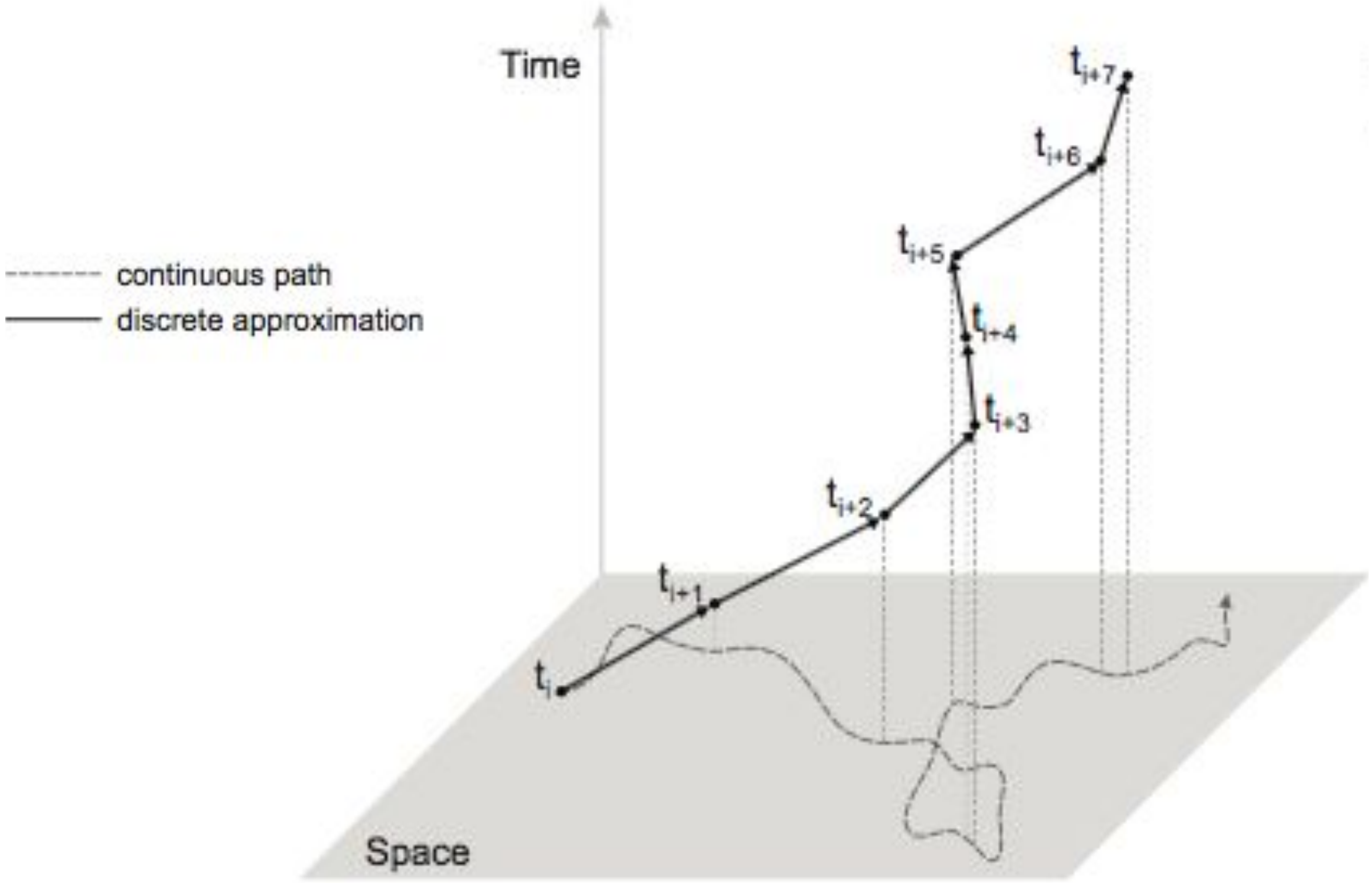
2. Under the microscope:
   measuring mobility patterns

3. Agents on the move:
   simulating mobility patterns

4. Where's next:
   predicting future whereabouts

The Human Mobility Data Landscape

What we want to reconstruct from data



Mobility surveys

- detailed questionnaires on mobility habits
- collected by local authorities or researchers

Examples:

- "My Daily Travel Survey" by the Chicago Metropolitan Agency for Planning
- "California Household Travel Survey" by the National Renewable Energy Laboratory (NREL)

Mobility surveys

- detailed questionnaires on mobility habits
- collected by local authorities or researchers

Household descriptors	Person-level descriptors	Trip details	Transport mode descriptors
household size, income, total trips, ...	age, gender, disability, employment, ...	purpose, mode, travelers, tolls, departure time, arrival time, distance ...	Private vehicles: year, make, model. Public transit trips: boarding locations, type of fare paid, ...

PROs and CONs of mobility surveys

PROs

- **high-resolution**
- information is **rich** and **detailed**

CONs

- **small** sample **size**
- **limited spatial** and **temporal** scale
- they are **static** in time (no “velocity”)
- **expensive** and **time consuming**
- presence of **self-reporting errors**

Official statistics

- periodically collected by governments
- may have information on
  - location of residence and workplace
  - previous address

Examples

- Commuting Flows by the US Census Bureau
- Migration Data by the Internal Revenue Service

PROs and CONs of official statistics

PROs

- **large sample size**
- **large coverage** (entire countries)

CONs

- flows are **aggregated**, usually at the municipality/county level
- **limited information**, only a single trip per respondent (e.g., only commuting)
- **expensive** and **time consuming**

NBC/Instagram

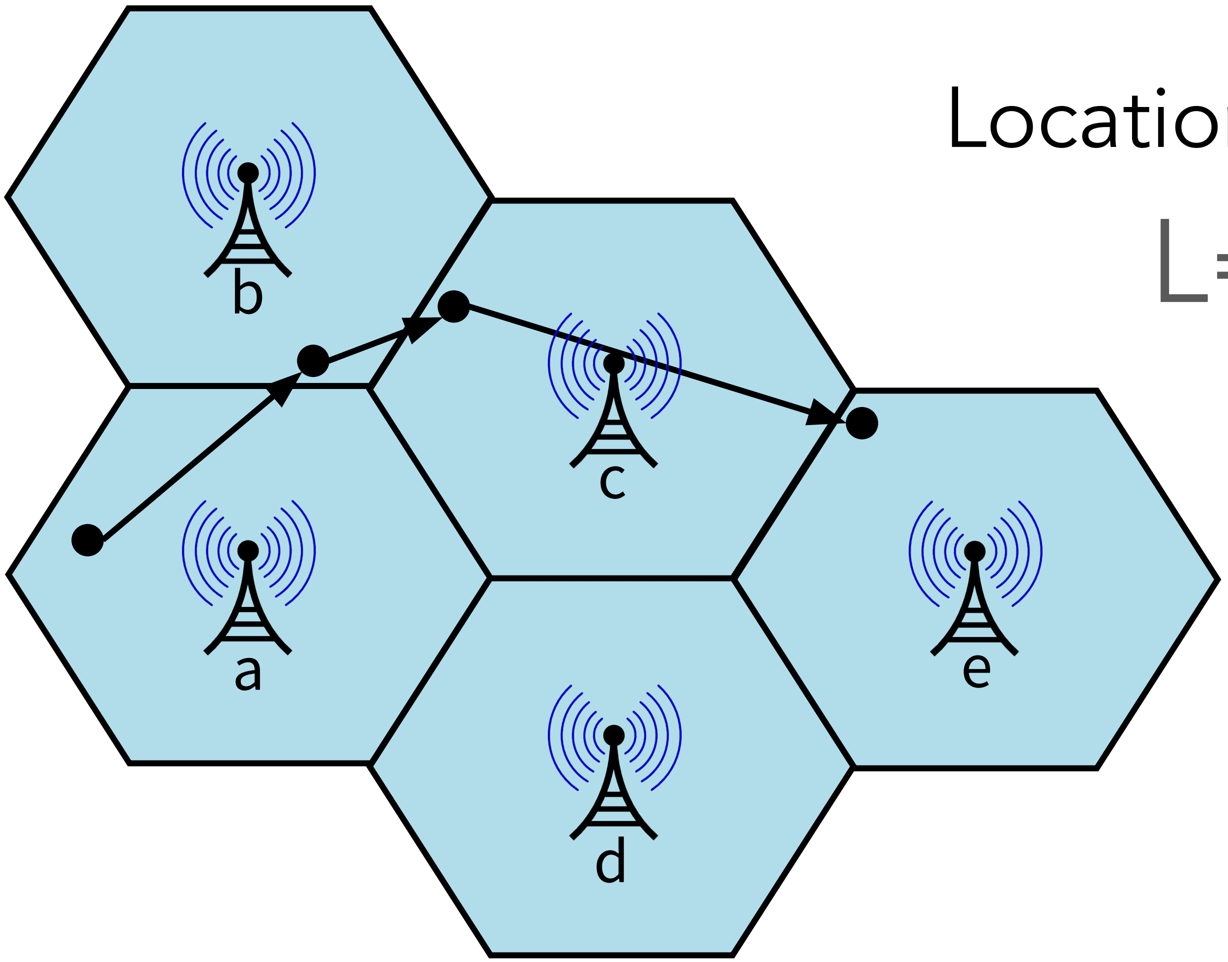
Call Detail Records (CDRs)

- are produced by a phone carrier to store **details of calls** passing through a device

- they contain various attributes of the call:

timestamp	source	destination	tower	duration
10/04/2008 10:12:00	A8563XY	C1123YT	36256	10:32
10/04/2008 10:12:32	B6766CS	H2223IU	43860	03:07
...	...	...	...	...

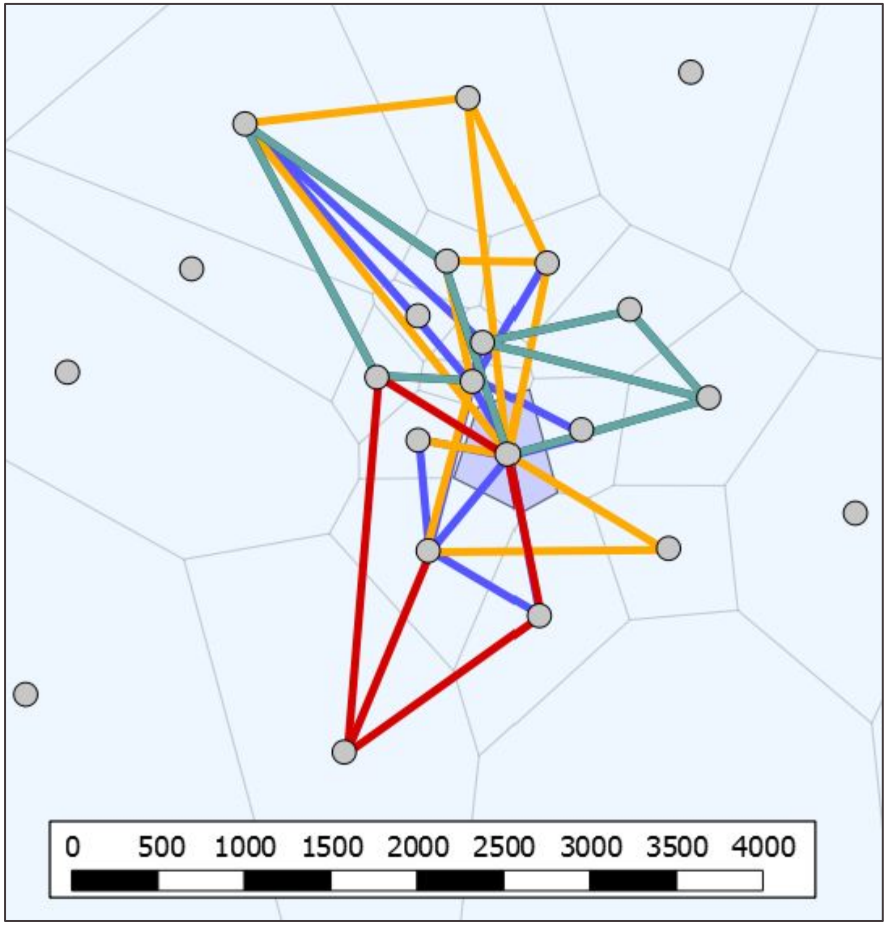
Call Detail Records (CDRs)



Location history

L=abce

Inferring mobility trajectories from CDRs



Pappalardo et al., An analytical framework to nowcast well-being with mobile phone data, JDSA, 2016.

- call → location/tower
- trajectory = sequence of locations
- home = most nighttime calls

PROs and CONs of CDRs

PROs

- mobile phones are **ubiquitous**
- **large** (huge) sample **size**
- **rich** and **multidimensional**
  (social, mobile, time, demographics)

CONs

- **not publicly available** (exception: D4D)
- position is **partially detected**
  (only when calls are made)
- position is known at **tower level** only
  (resolution is variable from m to km)
- ping-pong effect can create **noise**

Filters on CDR data

- non-relevant locations are discarded (e.g., if $n/N < 0.005$)
- users calling from just one locations are discarded (they do not move!)
- users with too low or high calling activity are discarded (e.g. by imposing min and max average calls per day)

References for CDRs

- Understanding individual human mobility patterns (Gonzalez et al., Nature, 2008).

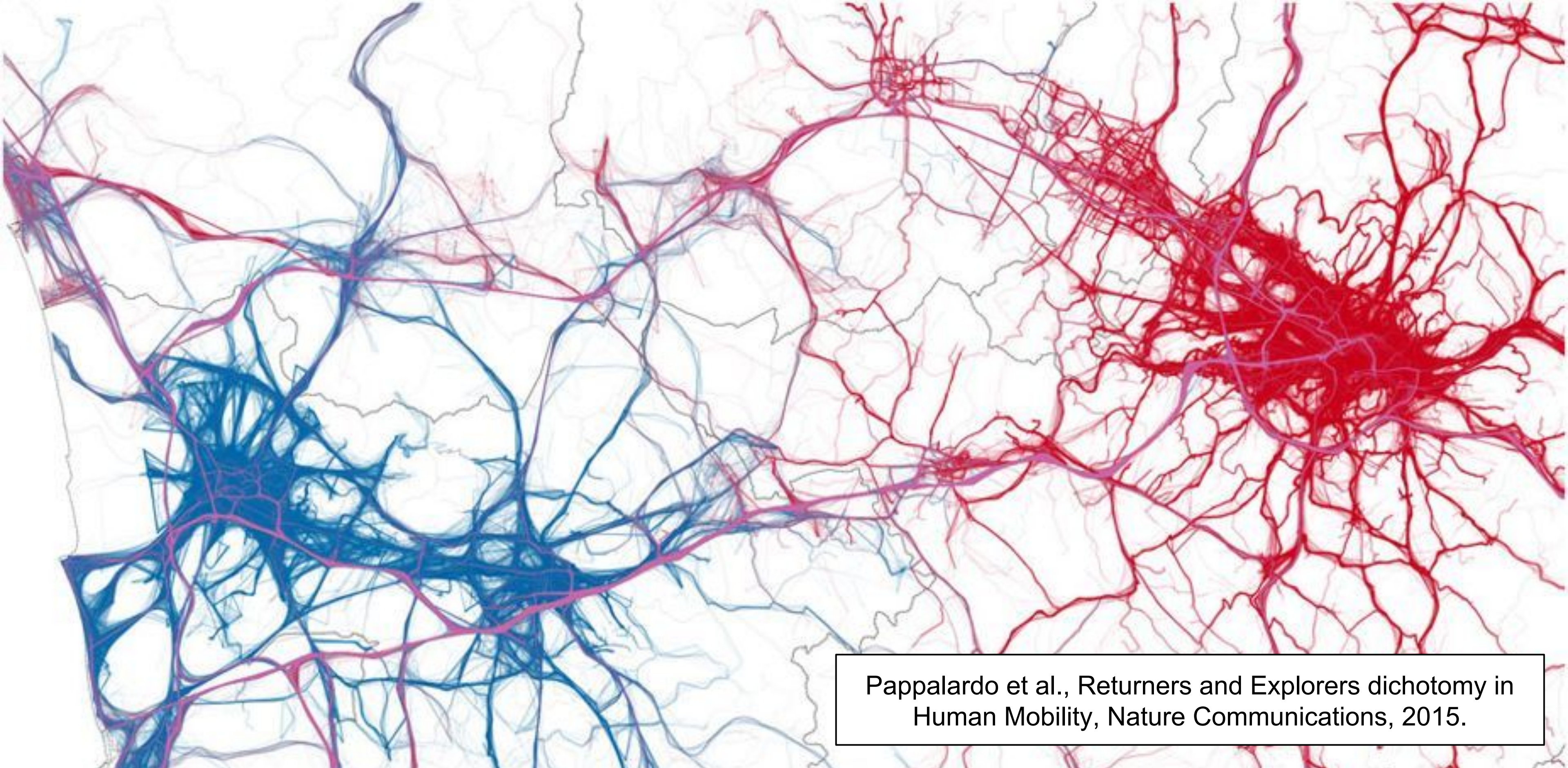
- Human mobility prediction based on individual and collective geographical preferences. (Calabrese et al., IEEE ITSC, 2010).

- Friendship and mobility: user movement in location-based social networks. (Cho et al., ACM SIGKDD, 2011).

- A survey of results on mobile phone datasets analysis (Blondel et al., EPJ DS, 2015).

GPS data

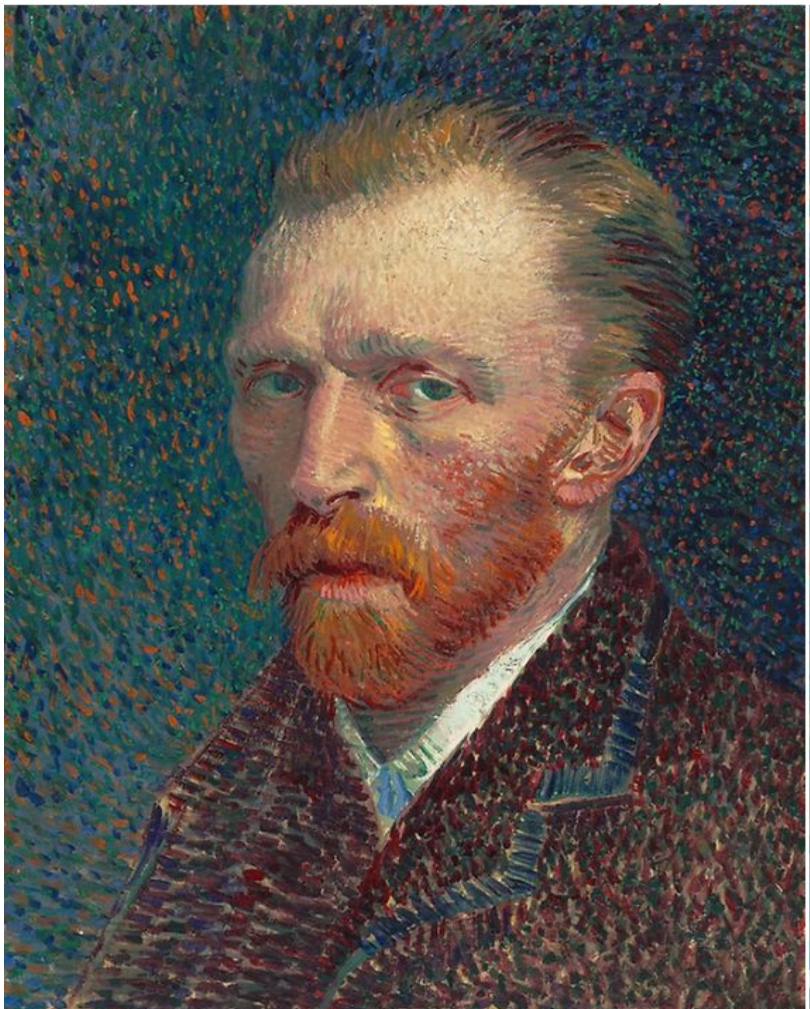
they are produced by **GPS devices** embedded in *smartphones*, *private vehicles*, *boats*, etc.



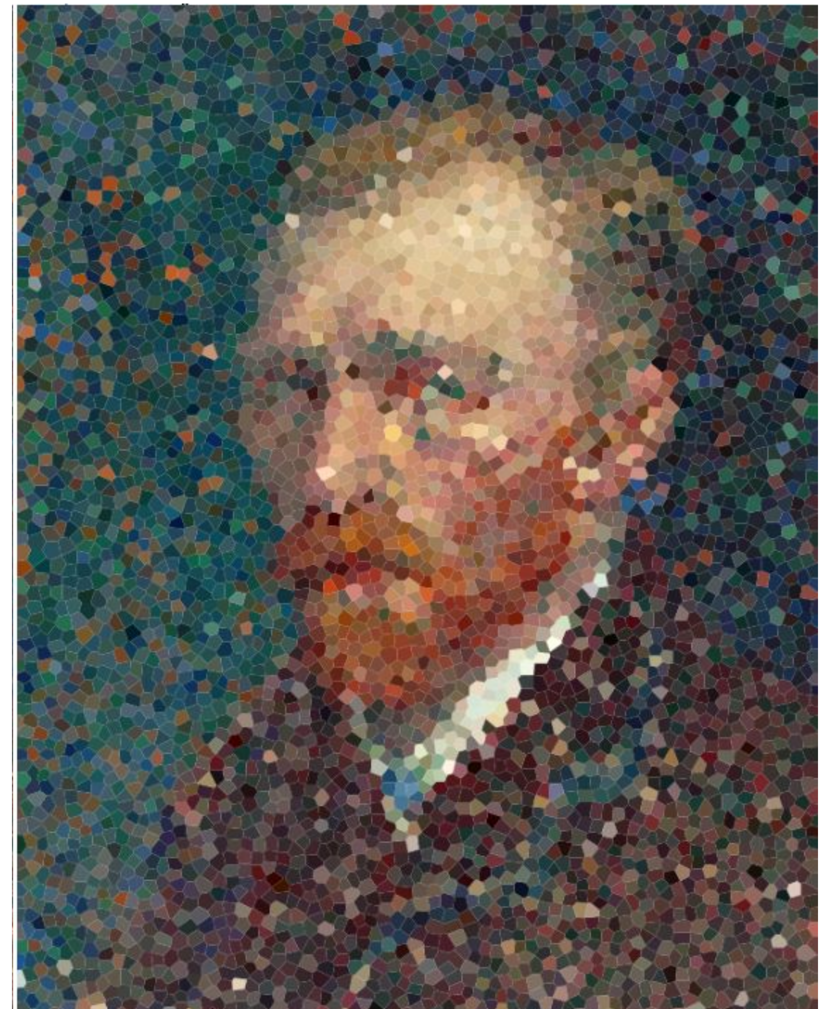
Pappalardo et al., Returners and Explorers dichotomy in Human Mobility, Nature Communications, 2015.

GPS data vs CDRs

GPS



CDRs



PROs and CONs of GPS data

PROs

- trajectory at **high** spatial and temporal **resolution**
- track the **full trajectory** (e.g., all cars movements)

CONs

- **no objective** definition of location (preprocessing is needed)
- **errors** when signal is noisy

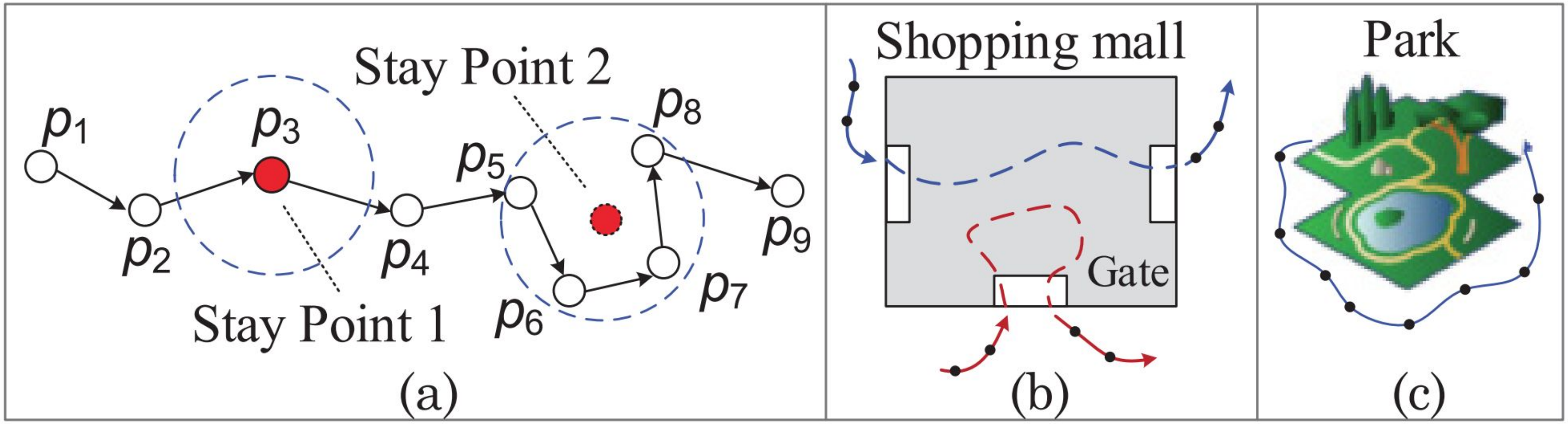
Preprocessing on GPS data

preprocessing steps are need given the nature of GPS data:

- **location detection**
  group points into one meaningful location
- **trajectory segmentation**
  split a trajectory in sub-trajectories

Location detection in GPS data

locations (or stay points) denote places where people have stayed for a while



Zheng, Trajectory Data Mining: an overview, ACM TIST, 2015.

$$P = p_1 \rightarrow p_2 \rightarrow \cdots \rightarrow p_n \implies$$

$$S = s_1 \rightarrow s_2 \rightarrow \cdots \rightarrow s_n$$

Stay Point detection algorithm

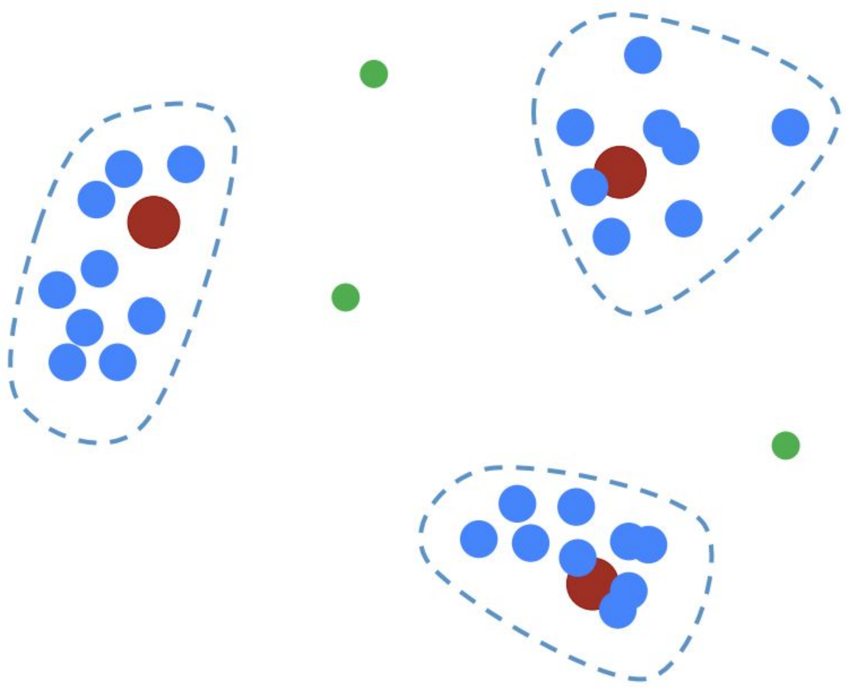
the points are scanned in temporal order,
then if the following condition holds:

$$d(p_{i+1}, p_i) < \delta \wedge t(p_{i+1}, p_i) < \Delta$$

the points belong to the same stay point

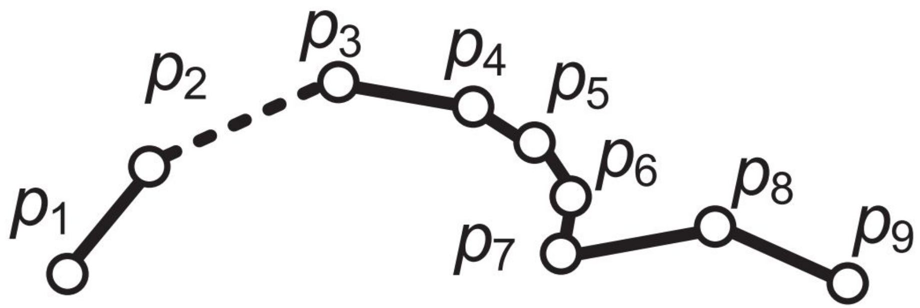
Semantic Location detection

density clustering algorithm, like DBSCAN or OPTICS (time is not considered)

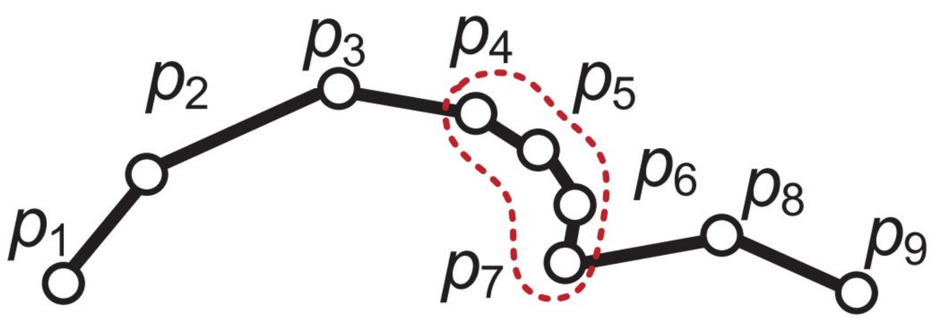


Trajectory segmentation

a trajectory is split into two or more sub-trajectories, with several techniques:



time-interval



stay-point

Zheng, Trajectory Data Mining: an overview, ACM TIST, 2015.

References for GPS data

- **Public Data**:
  - Geolife GPS trajectory dataset
  - Taxi GPS data

- Understanding the patterns of car travel (Pappalardo et al., EPJ ST, 2013).

- Learning travel recommendations from user-generated GPS traces (Zheng et al., ACM TIST, 2011).

- Trajectory Data Mining: an Overview (Zheng, ACM TIST, 2015.)

Location-Based Social Network (LBSN)

they add to an existing social network the **instant location** (***venue***) at a given timestamp (***checkin***)

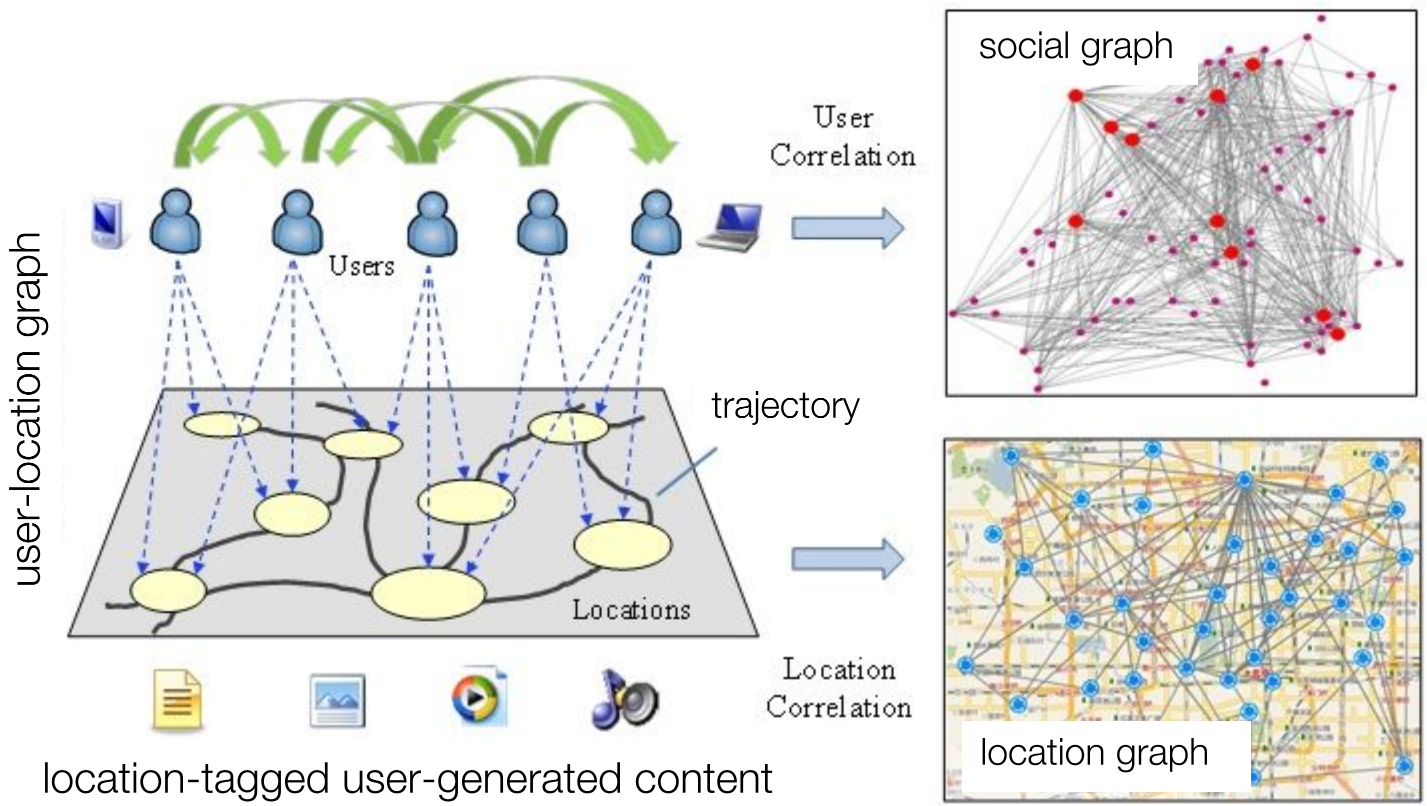
Location-Based Social Network (LBSN)

they add to an existing social network the **instant location** (***venue***) at a given timestamp (***checkin***)

foursquare

Location-Based Social Network (LBSN)

they add to an existing social network the **instant location** (*venue*) at a given timestamp (*checkin*)



PROs and CONs of LBSNs

PROs

- often **publicly available**
- **objective** location and **semantic** information (e.g., restaurant, mall, etc.)

CONs

- movement is **partially detected** (only when checkins are made)
- data **sparsity** (more sparse than CDRs)
- position is known at **location level** only
- **self-selection bias**

References for LBSNs

- Location-Based Social Networks: Users (Y. Zheng, Computing with Spatial Trajectories. Springer.)
- Mobility and geo-social networks (Spinsanti et al., Mobility Data, Cambridge press, 2013).
- A tale of many cities: Universal Patterns in Human Urban Mobility (Noulas et al., PLoS One, 2012).
- Understanding Human Mobility from Twitter (R. Jurdak, PLoS One, 2015).

Public datasets

- Geolife: GPS data https://bit.ly/2Mhy6R5
- YFCC100M: geotagged photos from Flickr http://yfcc100m.appspot.com/
- Foursquare: checkins in different countries https://bit.ly/2pPuOWB
- A multi-source dataset of urban life https://www.nature.com/articles/sdata201555
- Mobility Surveys https://bit.ly/2wPPYbc, https://bit.ly/2QgoPHw
- Taxi traces https://bit.ly/2wTkNMQ, https://bit.ly/2NULerX

SoBigData

- GPS traces of 150k cars
- CDRs of million of users
- LBSN data
- Survey data

http://www.sobigdata.eu/access/transnational

SoBigData TRANSNATIONAL ACCESS

Transnational Access supports **short term scientific mission** (between 2 weeks and 2 months) of researchers or teams at one of the installations of SoBigData that will provide **big data computing platforms**, **big social data resources**, and cutting-edge **computational methods**:

- Interact with the local experts, discuss research questions
  Run experiments on non-public datasets and algorithms
  Present results at workshops/seminars
- **to enable** multi-disciplinary social mining experiments with the Research Infrastructure **assets**: *big data sets, analytical tools, services and skills*.

Site: **http://www.sobigdata.eu/access/transnational**

Contacts: ***Gate*** - Kalina Bontheva <k.bontcheva@sheffield.ac.uk>
SoBigData.it - Claudio Lucchese <claudio.lucchese@isti.cnr.it>
Fraunhofer - Thorsten May <thorsten.may@igd.fraunhofer.de>
UT - Jaan Ubi <jaan.ubi@ut.ee>
L3S - Thomas Risse <risse@L3S.de>
AALTO UNIVERSITY - Michael Mathioudakis <michael.mathioudakis@aalto.fi>
Nervousnet - Iza Moise <izabela.moise@gess.ethz.ch>

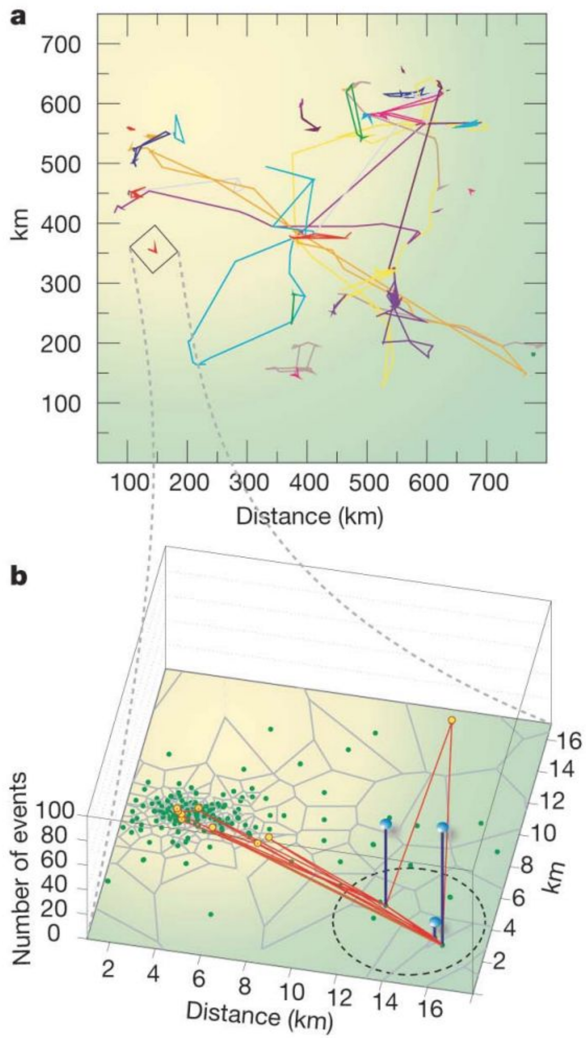
Under the microscope: characterizing human mobility

What… and Why

Analysis of (big) mobility data has revealed distinctive statistical patterns, that can be used for:

- **validating** mathematical models and simulation algorithms
- **training** AI models for predictive tasks (monitoring of well-being, health, location prediction, etc.)

Jump length (or traveling distance)



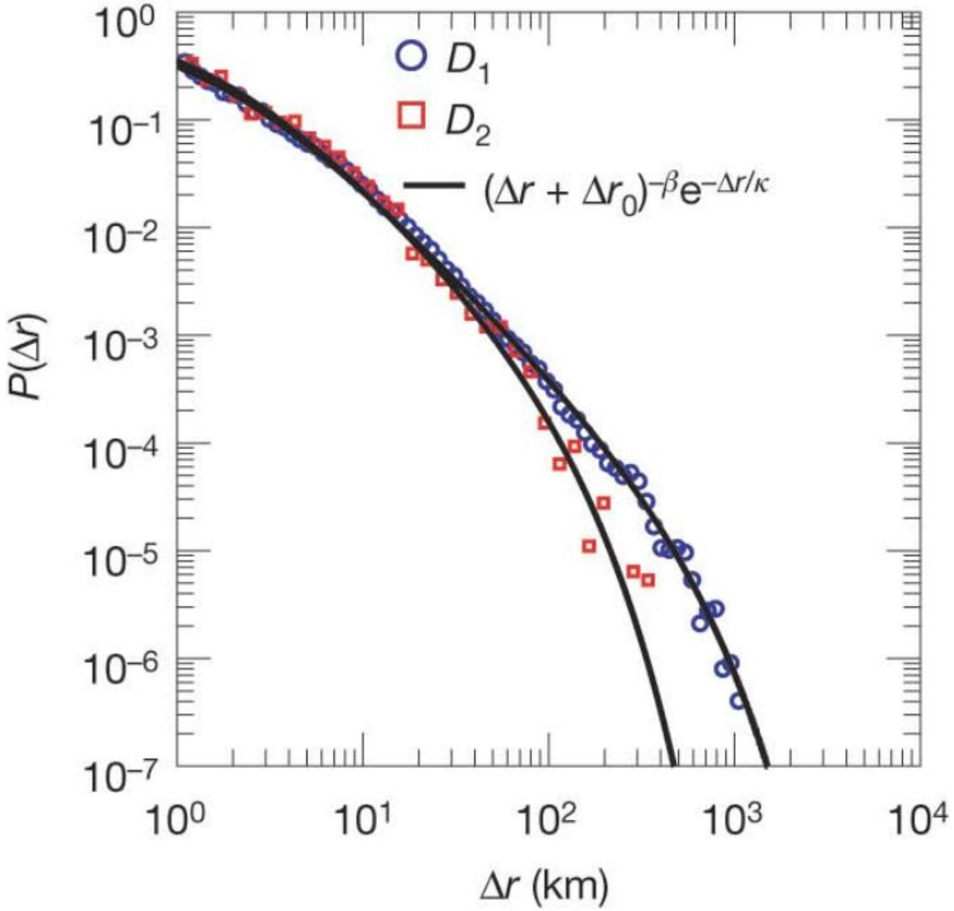
The distance between two consecutive locations visited by an individual

For GPS data, we need to detect locations

Understanding individual human mobility patterns (Gonzalez et al., Nature, 2008).

Jump length (or traveling distance)

Gonzalez et al., Understanding individual human mobility patterns, Nature, 2008.



D1: 100,000 for 6 months
D2: 200, 1 week, every 2h

$$\boxed{\beta = 1.75 \pm 0.15}$$

$$\Delta r_0 = 1.5km$$

$$\kappa = 400km$$

$$P(\Delta r) = (\Delta r + \Delta r_0)^{-\beta} \exp(-\Delta r/\kappa)$$

Jump length (or traveling distance)

- CDRs $\beta \in [1.75, 2.02]$
- GPS $\beta \in [1.35, 1.82]$
- LBSNs $\beta \in [1.5, 1.88]$

http://lauraalessandretti.weebly.com/plosmobilityreview.html

$$P(\Delta r) = (\Delta r + \Delta r_0)^{-\beta} \exp(-\Delta r / \kappa)$$

Radius of gyration

The characteristic distance traveled by an individual

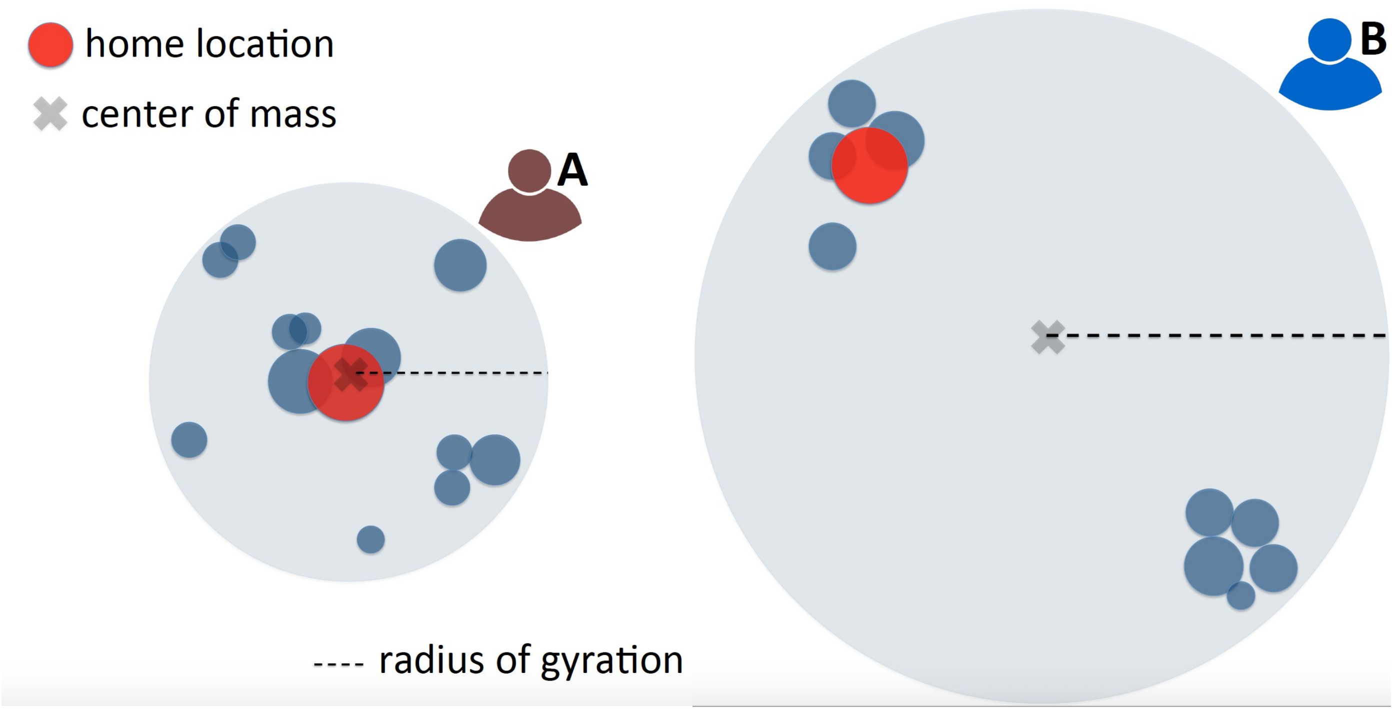
For GPS data, we need to detect locations

$$r_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N} w_i (r_i - r_{cm})^2}$$

$$r_{cm} = \frac{1}{N}\sum_{i=1}^{n} w_i r_i \qquad N = \sum_{i=1}^{n} w_i$$

Radius of gyration

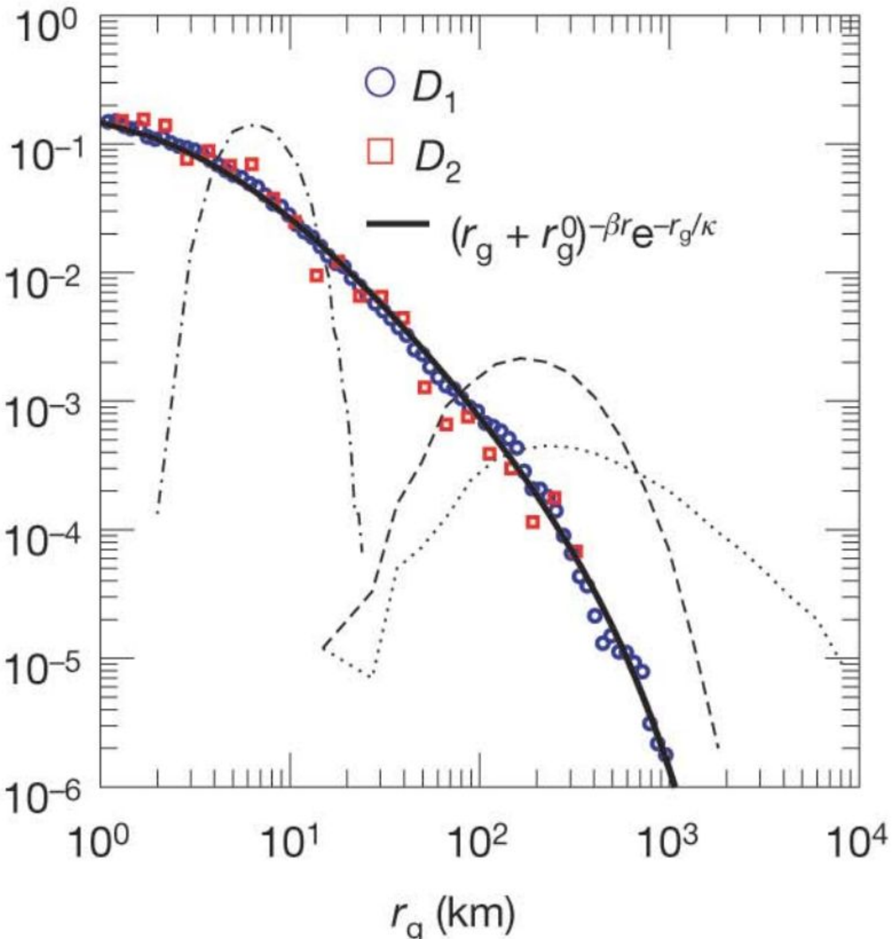
The characteristic distance traveled by an individual



An analytical framework to nowcast well-being with mobile phone data (Pappalardo et al., JDSA, 2016).

Radius of gyration

Gonzalez et al., Understanding individual human mobility patterns, Nature, 2008.



$$r_g^0 = 5.8km$$

$$\boxed{\beta_r = 1.65 \pm 0.15}$$

$$\kappa = 350km$$

$$P(r_g) = (r_g + r_g^0)^{-\beta_r} \exp(-r_g/\kappa)$$

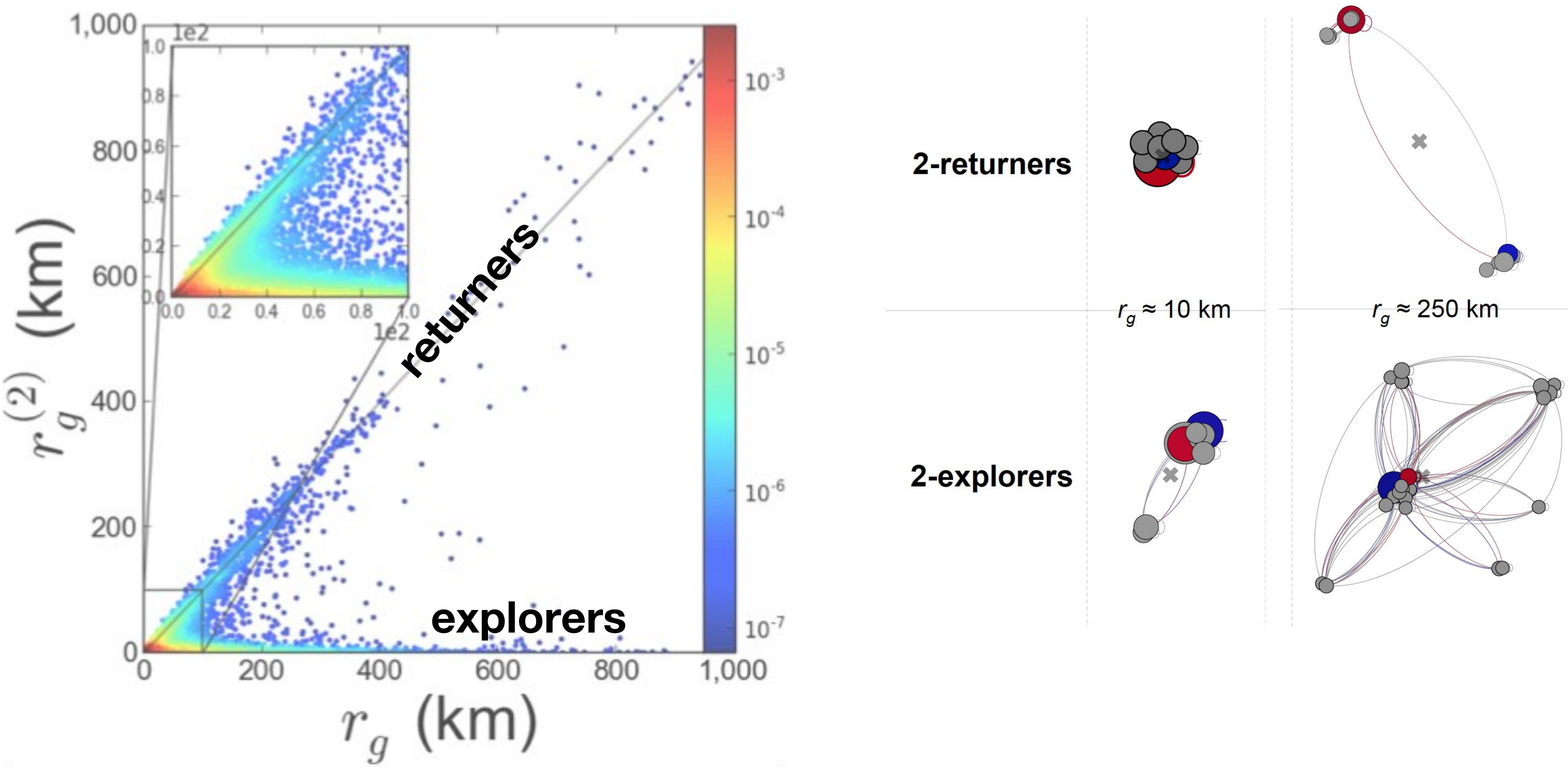
k-radius of gyration

The recurrent characteristic distance traveled by an individual

$$r_g^{(k)} = \sqrt{\frac{1}{N_k}\sum_{i=1}^{k} w_i (r_i - r_{cm}^{(k)})^2}$$

$$r_{cm}^{(k)} = \frac{1}{N_k}\sum_{i=1}^{k} w_i r_i \qquad N_k = \sum_{i=1}^{k} w_i$$

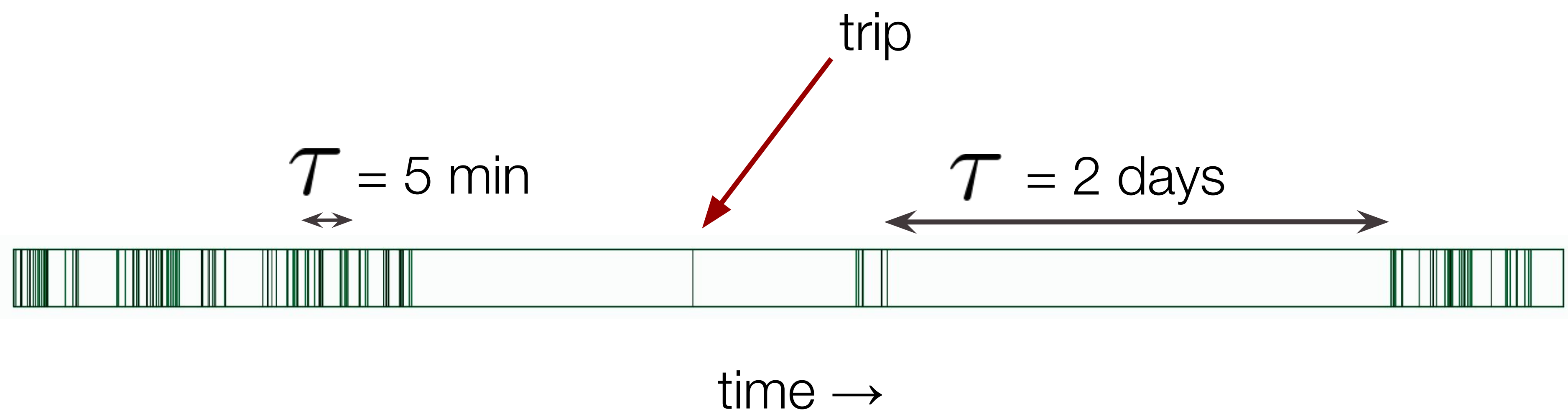
k-radius of gyration



Pappalardo et al., Returners and Explorers dichotomy in human mobility, Nature Communications, 2015.

Waiting times (or inter-event time)

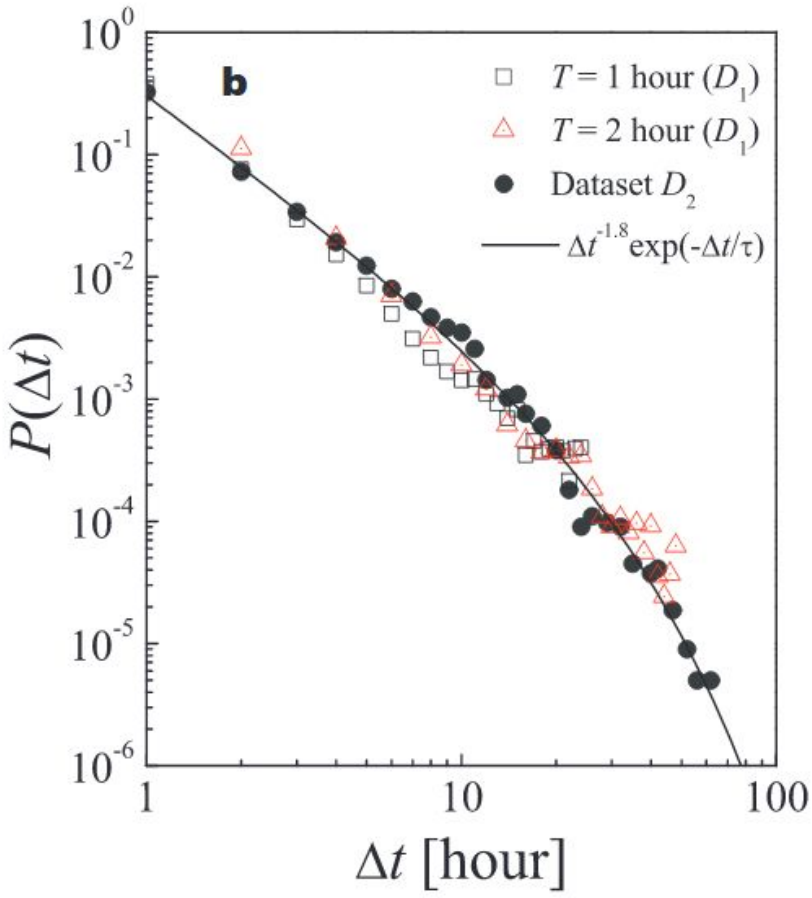
The time between two consecutive trips of an individual



Barabasi., The origin of bursts and heavy tails in human dynamics, Nature, 2005.

Waiting times (or inter-event time)

Modelling the scaling properties of human mobility (Song et al., Nature Physics, 2010).



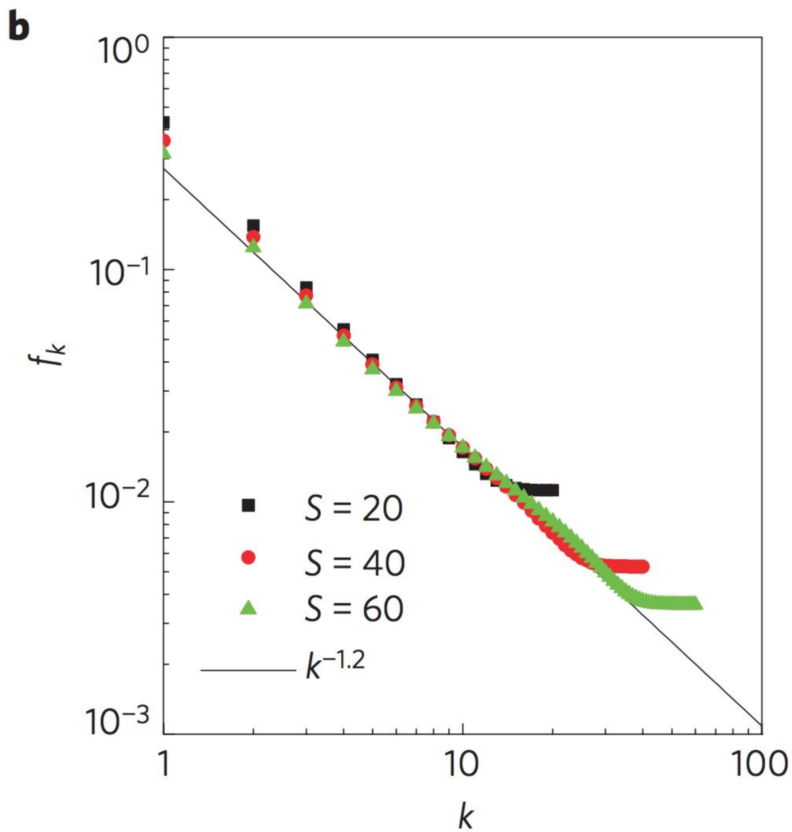
$$\beta_t = 0.8 \pm 0.1$$

$$\tau = 17h$$

$$P(\Delta t) = \Delta t^{-1-\beta_t} \exp(-\Delta t/\kappa)$$

Location frequency

The visitation frequency of the k-th most visited location of an individual



A Zipf's law:

$$f_k = k^{-\xi}$$

$$\xi = 1.2 \pm 0.1$$

Song et al., Modelling the scaling properties of human mobility, Nature Physics, 2010.

Mobility Entropy

The predictability of an individual's whereabouts

For GPS data, we need to detect locations

w_i/N

uncorrelated entropy |

$$S^{unc} = -\sum_{i=1}^{n} p_i \log_2 p_i$$

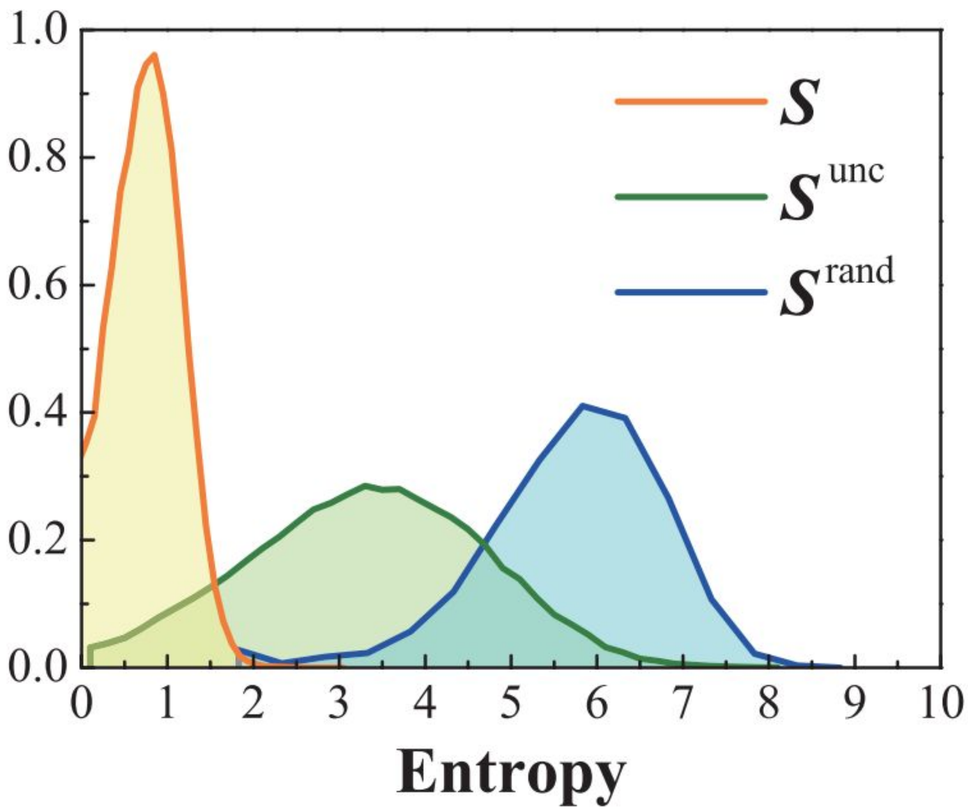
real entropy |

$$S = -\sum_{T_i' \subset T_i} p_{T_i'} \log_2 p_{T_i'}$$

Mobility Entropy

The predictability of an individual's whereabouts

Song et al., Limits of predictability in human mobility, Science, 2010.



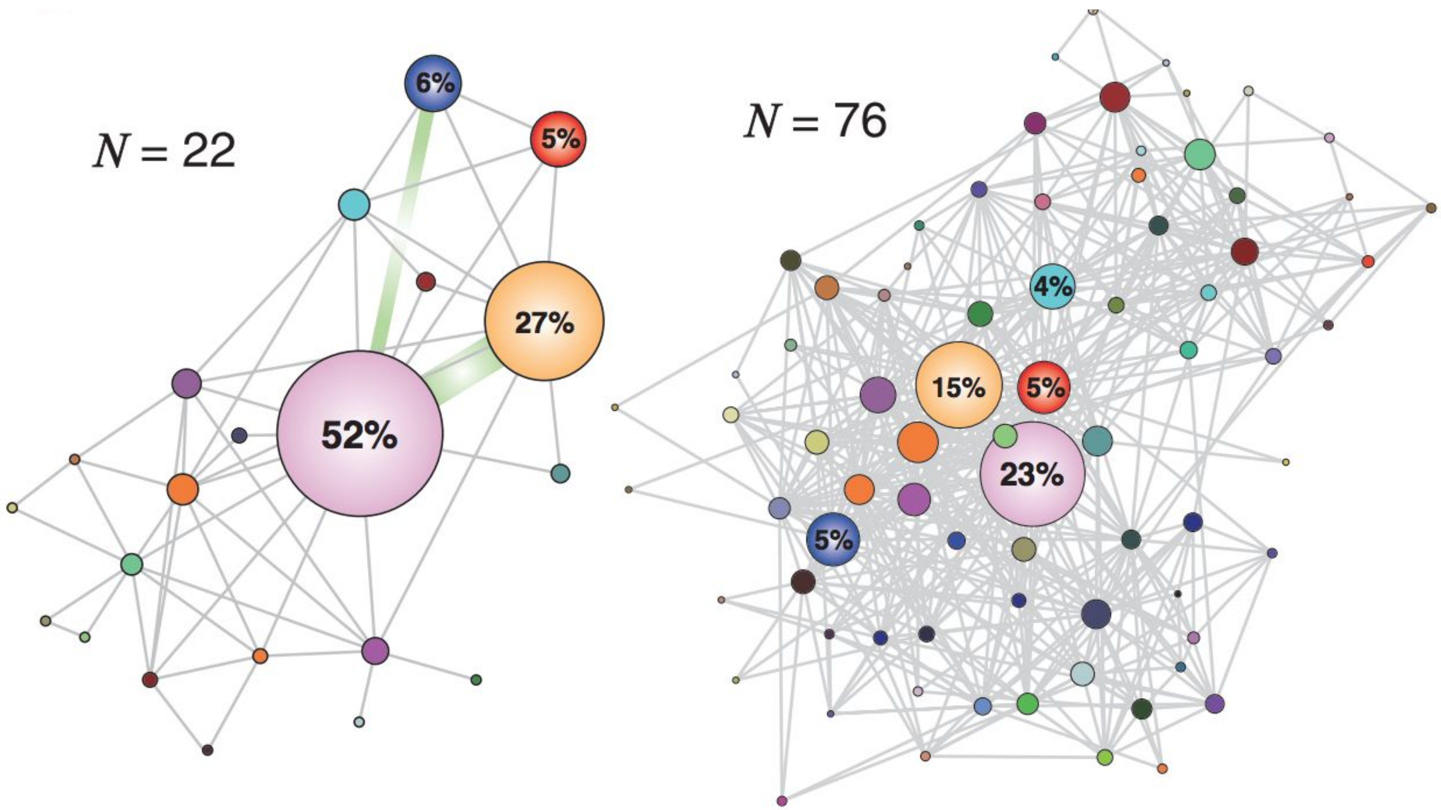
S peaks at 0.8

$$2^{0.8} = 1.74$$

typical uncertainty is less than 2 locations

Individual Mobility Networks (Motifs)

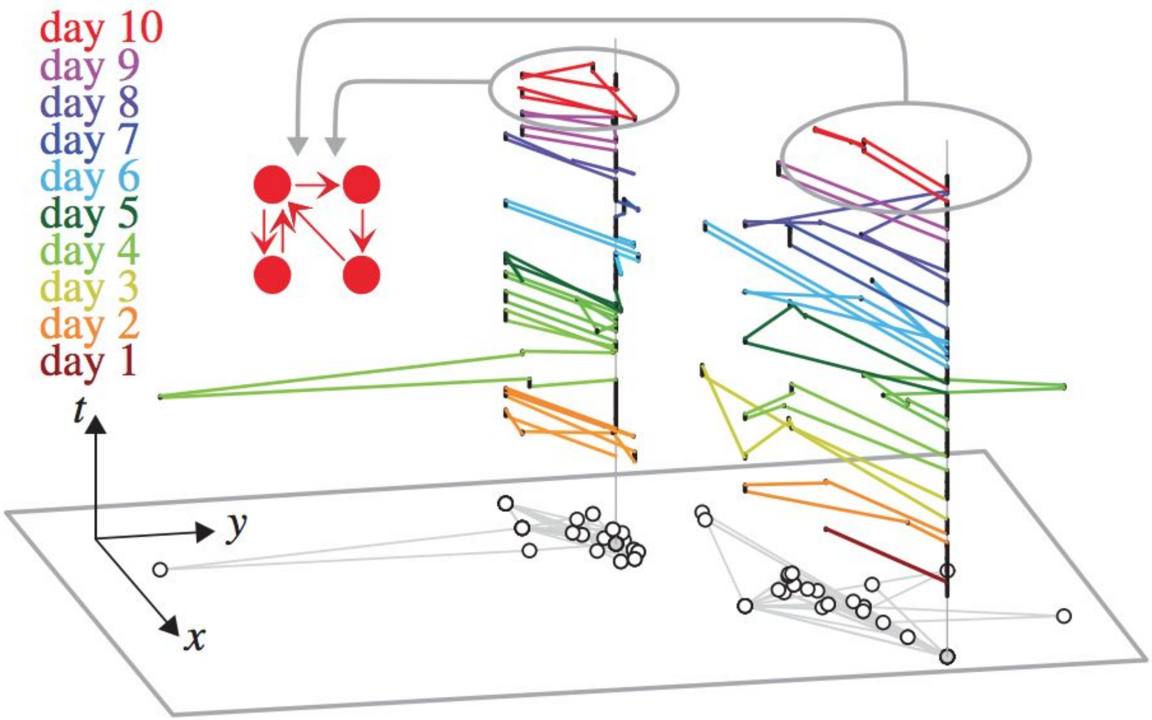
IMN: a network describing the typical movements of an individual



Song et al., Limits of predictability in human mobility, Science, 2010.

Individual Mobility Networks (Motifs)

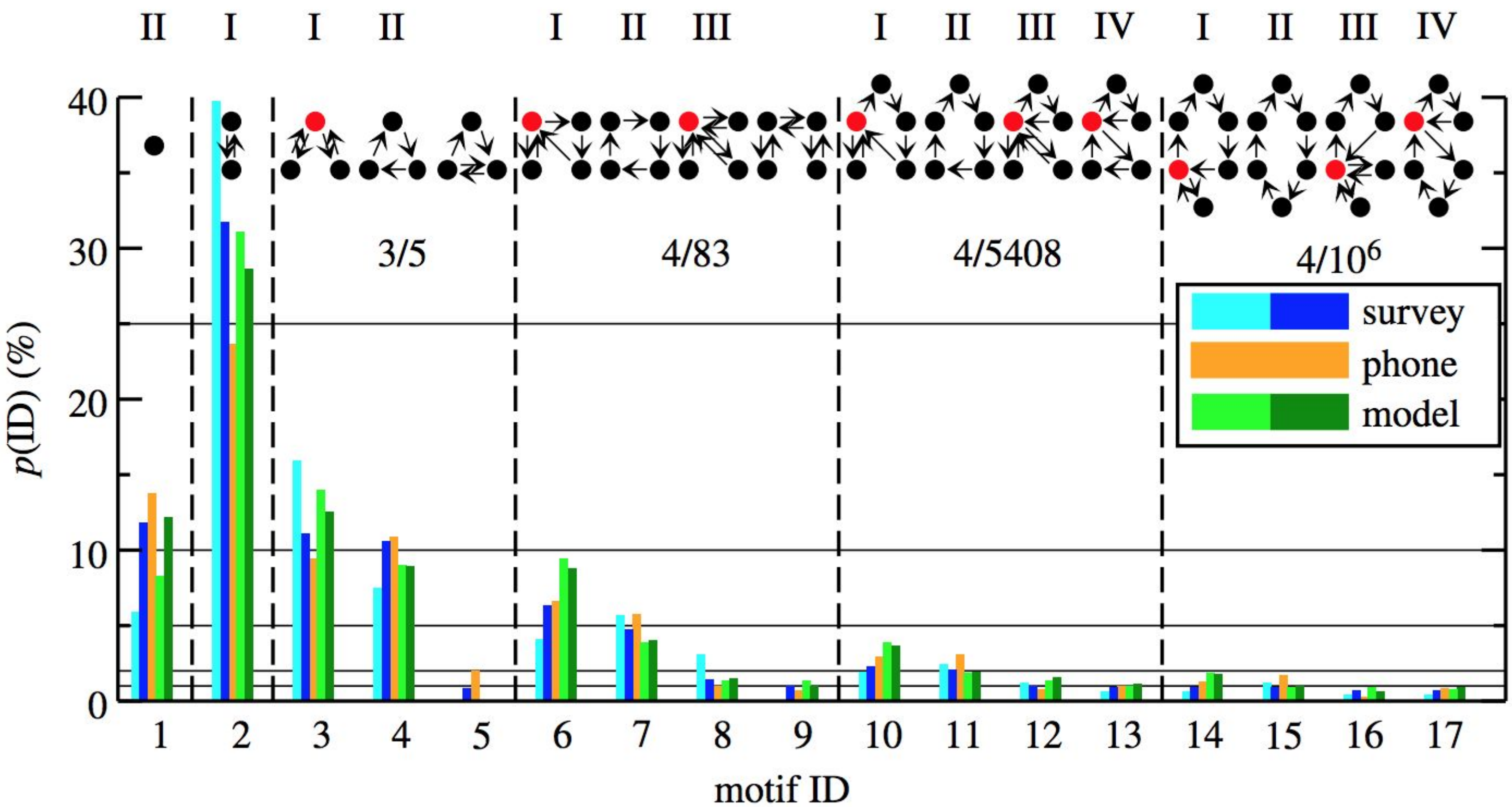
Daily Motif: a frequent network describing the daily movements of an individual



Schneider et al., Unravelling individual daily mobility motifs, Journal of the Royal Society Interface, 2013.

Individual Mobility Networks (Motifs)

17 daily motifs have been found in different mobility data sources



Schneider et al., Unravelling individual daily mobility motifs, Journal of the Royal Society Interface, 2013.

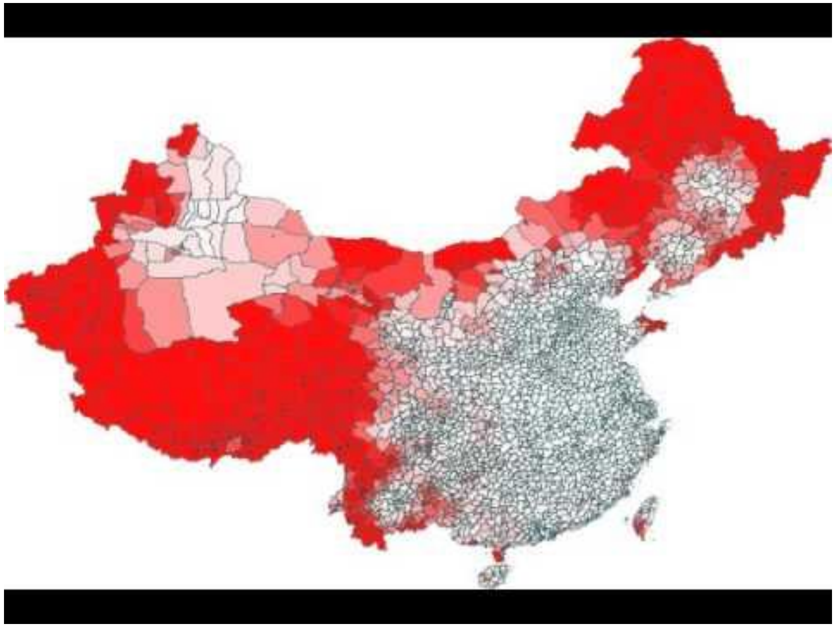
Summarizing

- jump length Δr
- radius of gyration r_g
- k-radius of gyration $r_g^{(k)}$
- waiting time Δt
- location frequency f_k
- entropies S^{unc} S

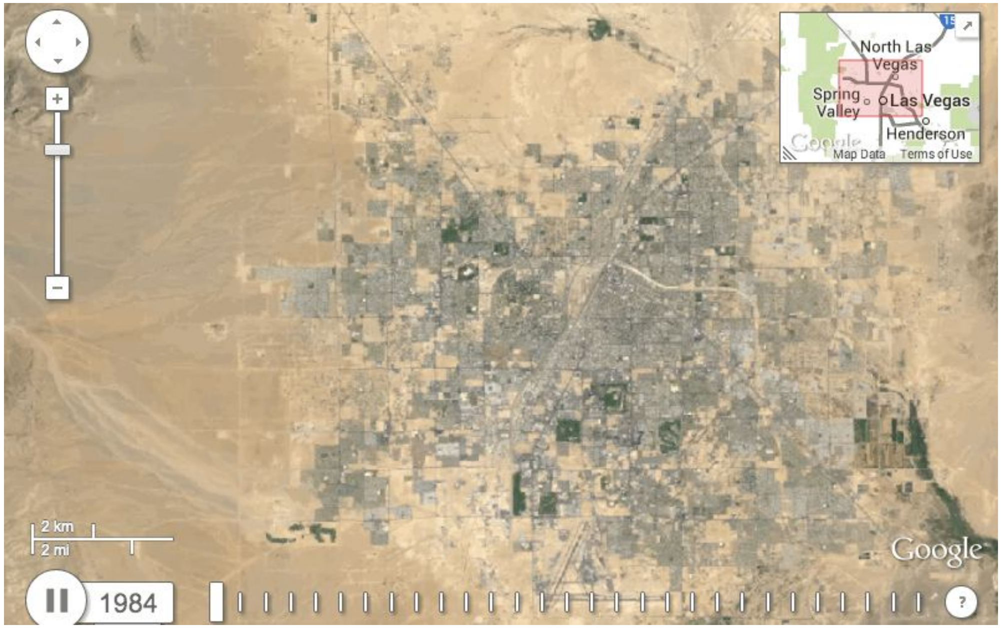
- individual mobility networks
- daily motifs

Agents on the move: simulating mobility patterns

Motivation



Disease outbreak
Transportation
Urban planning



Predictive vs Generative models

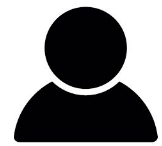
- **Predictive** models
  forecast a real individual's future whereabouts given their past history

- **Generative** models
  generate synthetic individuals with realistic* spatio temporal trajectories

  *realistic = compatible with individual and collective empirical patterns (distribution of time spent in the visited locations, distribution of radii of gyration, …)

Generative models: Individual vs Collective

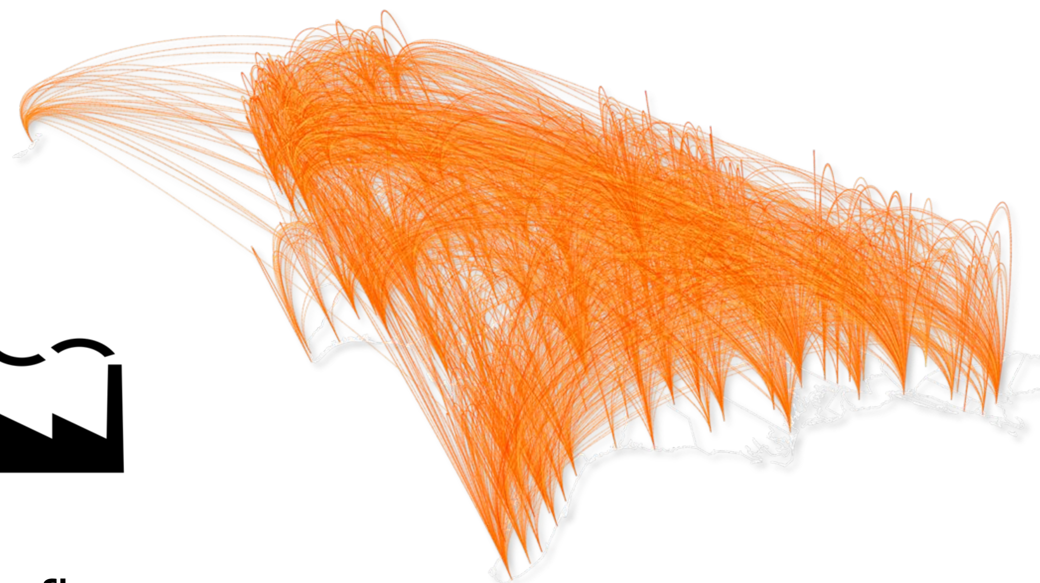
- **Individual** models
  generate the trajectory of a single agent



- **Collective** models
  generate aggregated trips (mobility flows) between locations

Collective models

Goal: generate mobility flows between origins and destinations



Examples:

- Commuting flows

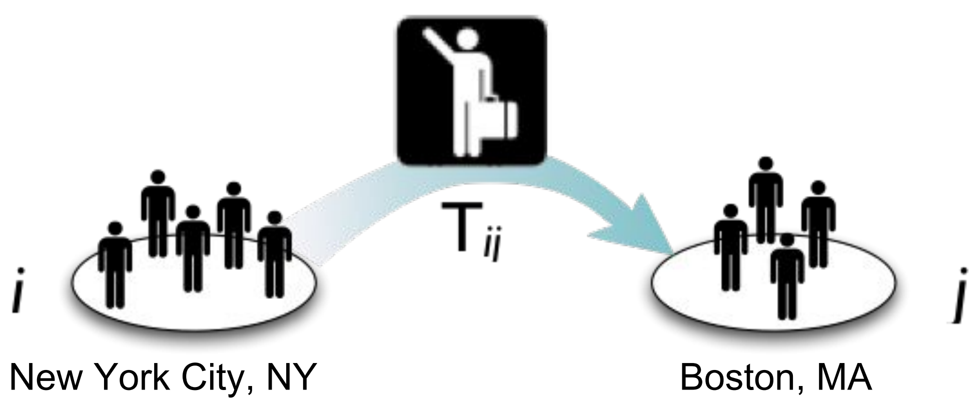
- Migration/relocation flows

https://www.flaticon.com/authors/dave-gandy

Spatial flows and OD matrices

mathematically, spatial flows are represented as a Origin-Destination **(OD) matrix**, **T**:

1. Define locations discretizing space
   (e.g., counties, municipalities)
2. Element T_{ij} is *the number of trips from i to j per unit time*.



OD matrix

		destination					
		a	b	c	d	e	f
origin	a	-	3	27	2	1	0
	b	1	-	4	0	0	5
	c	8	3	-	1	13	6
	d	2	1	5	-	0	2
	e	11	0	6	5	-	1
	f	0	3	2	2	0	-

(self-loops are usually not considered)

total out-flow from i

$$\sum_j T_{ij} = O_i$$

total in-flow to j

$$\sum_i T_{ij} = D_j$$

total flow

$$\sum_{ij} T_{ij} = N$$

Probabilistic models of spatial flows

- assign a **probability** to each possible OD-matrix $\mathbf{T}$

- **fit** model's parameters
  - maximising the likelihood of observed $\mathbf{T}^*$
  - minimising the distance from observed $\mathbf{T}^*$

Probabilistic models of spatial flows

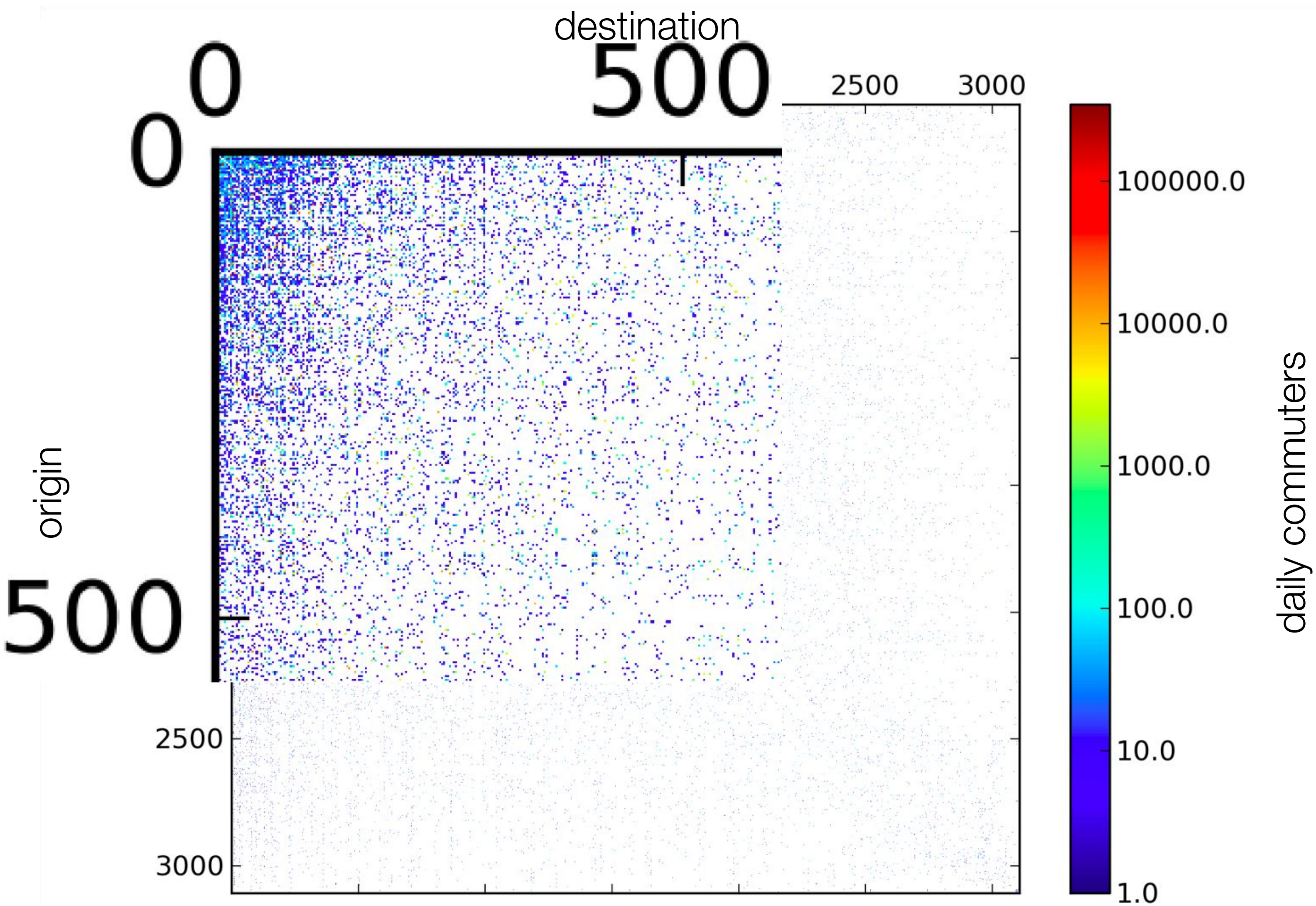
- **Constrained** models
  - *globally* constrained $\sum_{ij} T_{ij} = N$

    (or unconstrained)
  - *origin* constrained $\sum_{j} T_{ij} = O_i \quad \forall i$
  - *destination* constrained $\sum_{i} T_{ij} = D_j \quad \forall j$
  - *doubly* constrained $\sum_{j} T_{ij} = O_i$ and $\sum_{i} T_{ij} = D_j$

singly

A real OD matrix

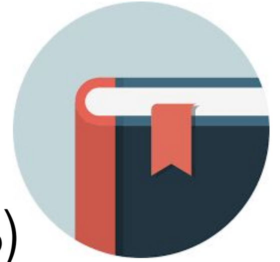
United States' county to county commuting flows



Properties of spatial flows

- Flows **decay** with **distance** (long tail)
- Flows **grow** with **population**

Original references:



- ❖ Henry C. Carey, Principles of Social Science (1858)
- ❖ Ernst Ravenstein, The laws of migration (1885)
- ❖ William J. Reilly, The law of retail gravitation (1931)
- ❖ Samuel A. Stouffer, Intervening opportunities: A theory relating mobility and distance (1940)
- ❖ George K. Zipf, The P1 P2/D Hypothesis: On the Intercity Movement of Persons (1946)

Two main modelling approaches

1. **Gravity** (*G*) models
2. **Intervening opportunities** (*IO*) models

Similarities

They both depend on two types of variables:

- ❖ the **weight** is an attribute of each individual location e.g., population, number of opportunities.
- ❖ the "**distance**" is a quantity relating a pair of locations

Differences

- ❖ **different distance variables** considered: geographical distance (G) vs. # of intervening opportunities (IO).

Gravity model

[1] Principles of Social Science (H.C. Carey, Lippincott (1858))
[2] The P1 P2/D hypothesis: on the intercity movement of persons (G.K. Zipf, Am. Sociol. Rev., 11 (1946))

Analogy with Newton's law of gravitation

mass $\longrightarrow$ population

$$F_{ij} \propto \frac{m_i \times m_j}{r_{ij}^2} \quad \longrightarrow \quad T_{ij} \propto \frac{pop_i \times pop_j}{r_{ij}^2}$$

$$T_{ij} = A\, N\, m_i^{\alpha}\, m_j^{\beta}\, f(r_{ij})$$

Gravity model

[1] Principles of Social Science (H.C. Carey, Lippincott (1858))
[2] The P1 P2/D hypothesis: on the intercity movement of persons (G.K. Zipf, Am. Sociol. Rev., 11 (1946))

Analogy with Newton's law of gravitation

mass $\longrightarrow$ population

$$F_{ij} \propto \frac{m_i \times m_j}{r_{ij}^2} \longrightarrow T_{ij} \propto \frac{pop_i \times pop_j}{r_{ij}^2}$$

$$T_{ij} = A\,N\,m_i^{\alpha}\,m_j^{\beta}\,f(r_{ij})$$

deterrence function

$f(r) = r^{\gamma}$ or $f(r) = e^{-r/R}$

power law exponential

Gravity model

Entropy in urban and regional modelling (A. Wilson, Pion, London (1970))

Analogy with Newton's law of gravitation

mass → population

$$F_{ij} \propto \frac{m_i \times m_j}{r_{ij}^2} \longrightarrow T_{ij} \propto \frac{pop_i \times pop_j}{r_{ij}^2}$$

Globally constrained

$$T_{ij} = A\, N\, m_i^{\alpha}\, m_j^{\beta}\, f(r_{ij}) \qquad \sum_{ij} T_{ij} = N$$

Gravity model

Analogy with Newton's law of gravitation

mass → population

$$F_{ij} \propto \frac{m_i \times m_j}{r_{ij}^2} \longrightarrow T_{ij} \propto \frac{pop_i \times pop_j}{r_{ij}^2}$$

Globally constrained

$$T_{ij} = A\,N\,m_i^{\alpha}\,m_j^{\beta}\,f(r_{ij}) \qquad \sum_{ij} T_{ij} = N$$

Singly constrained

$$T_{ij} = A_i\,O_i\,m_j^{\beta}\,f(r_{ij}) \qquad \sum_{j} T_{ij} = O_i \;\forall i$$

Gravity model - parameters fitting

Singly constrained

$$T_{ij} = A_i \, O_i \, m_j^{\beta} \, f(r_{ij})$$

$$\sum_j T_{ij} = O_i \quad \forall i$$

$$T_{ij} = O_i \, p_{ij} = O_i \, \frac{m_j^{\beta} \, f(r_{ij})}{\sum_k m_k^{\beta} \, f(r_{ik})}$$

$$\sum_j p_{ij} = 1 \; \forall i$$

$$P(\{\mathbf{T}_i\} = P(\{T_{i0}, T_{i1}, \ldots, T_{iN}\}) = Multinomial(O_i, p_{ij})$$

Probability to observe a given set of flows from *i*

Gravity model - parameters fitting

Singly constrained

$$T_{ij} = A_i \, O_i \, m_j^{\beta} \, f(r_{ij})$$

$$\sum_j T_{ij} = O_i \quad \forall i$$

$$T_{ij} = O_i \, p_{ij} = O_i \, \frac{m_j^{\beta} \, f(r_{ij})}{\sum_k m_k^{\beta} \, f(r_{ik})}$$

$$\sum_j p_{ij} = 1 \; \forall i$$

$$P(\{\mathbf{T}_i\} = P(\{T_{i0}, T_{i1}, \ldots, T_{iN}\}) = Multinomial(O_i, p_{ij})$$

Probability to observe a given set of flows from *i*

$$LogL(\mathbf{T}) = \ln \left(\prod_i P(\{\mathbf{T}_i\}) \right) \propto \sum_{ij} T_{ij} \, \ln \left(\frac{e^{(\beta \ln m_j + \ln f(r_{ij}))}}{\sum_k e^{(\beta \ln m_k + \ln f(r_{ik}))}} \right)$$

LogLikelihood of the OD matrix **T**

(assuming that flows from different locations are independent variables)

Gravity model - parameters fitting

If the deterrence function, ***f***, is an **exponential** or a **power law**,

the gravity model is a **Generalised Linear Model**

Efficient parameters fitting using a Poisson regression

[see: *Categorical data analysis (*Agresti, 2003)]

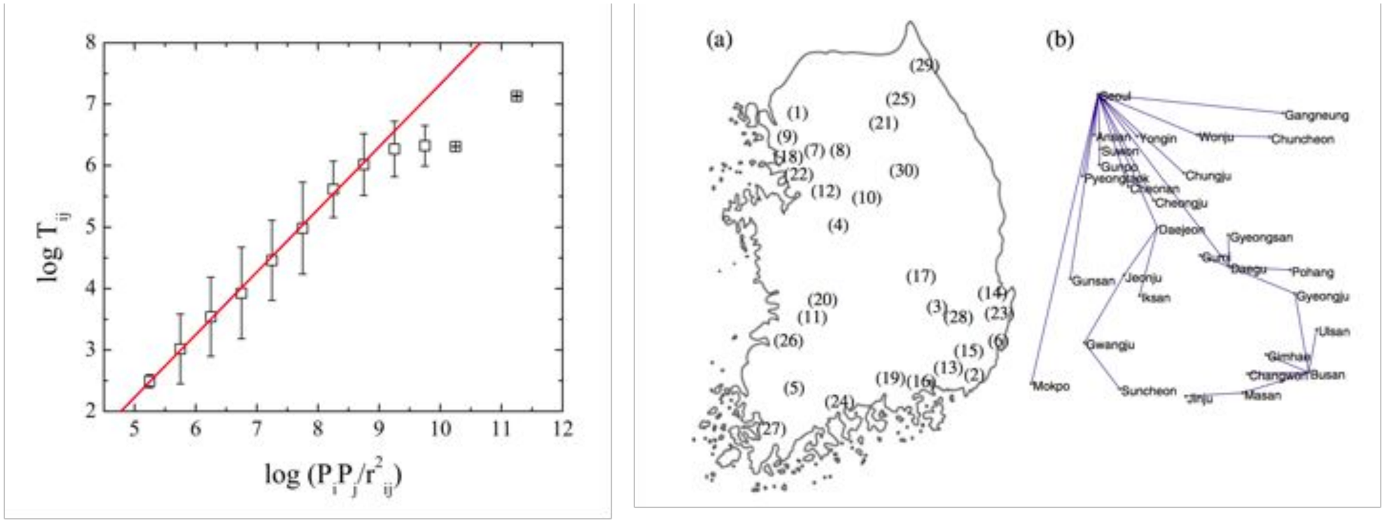
Note:

this is the loglikelihood of a multinomial logistic regression.

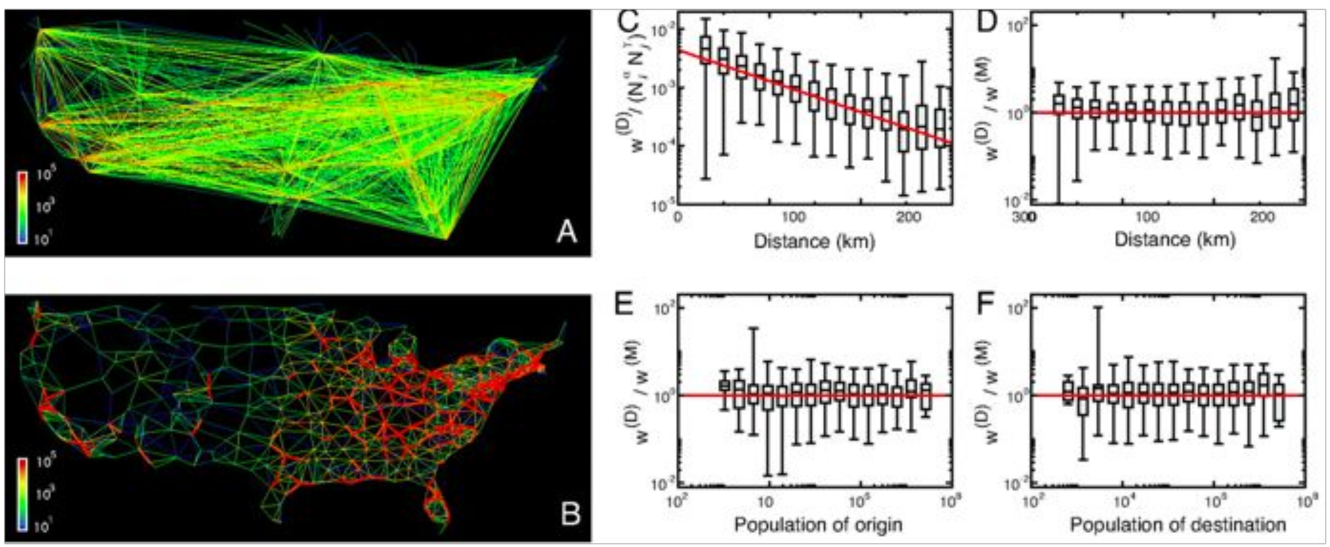
$$LogL(\mathbf{T}) = \ln\left(\prod_i P(\{\mathbf{T}_i\})\right) \propto \sum_{ij} T_{ij} \ln\left(\frac{e^{(\beta \ln m_j + \ln f(r_{ij}))}}{\sum_k e^{(\beta \ln m_k + \ln f(r_{ik}))}}\right)$$

LogLikelihood of the OD matrix **T**

Gravity model - applications



Jung, W. S., Wang, F., & Stanley, H. E. (2008). Gravity model in the Korean highway. *EPL (Europhysics Letters)*, *81* (4), 48005.



Balcan, D., et al. "Multiscale mobility networks and the spatial spreading of infectious diseases." *PNAS* 106.51 (2009): 21484-21489.

PROs and CONs of Gravity models

PROs

- parameters are **easy to fit**
  (for globally and singly constrained)
- state-of-the-art **performance**
- **versatility** and wide applicability

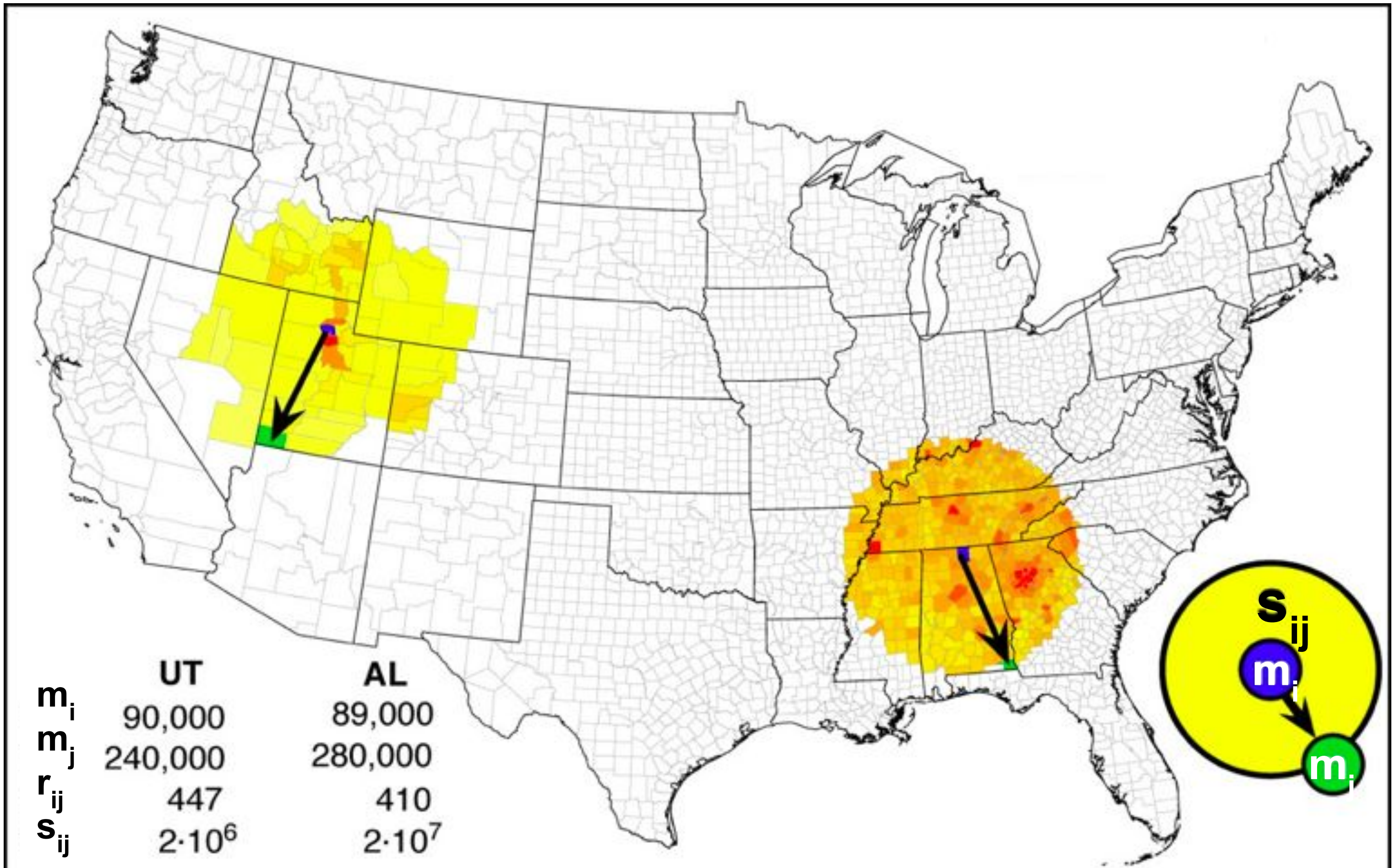
CONs

- **underfitting** and **overdispersion**
- **low generalisation** power

Intervening Opportunities model

Intervening opportunities: a theory relating mobility and distance (Stouffer, S. A. (1940)).

Flows do not decrease with distance, but with the number of **intervening opportunities**, **S**, between origin and destination



		UT	AL
origin	m_i	90,000	89,000
destin.	m_j	240,000	280,000
distance	r_{ij}	447	410
# IO	s_{ij}	$2 \cdot 10^6$	$2 \cdot 10^7$

Original IO model

Gravity models and trip distribution theory (Schneider, M. (1959), Papers in Regional Science, 5(1), 51-56).

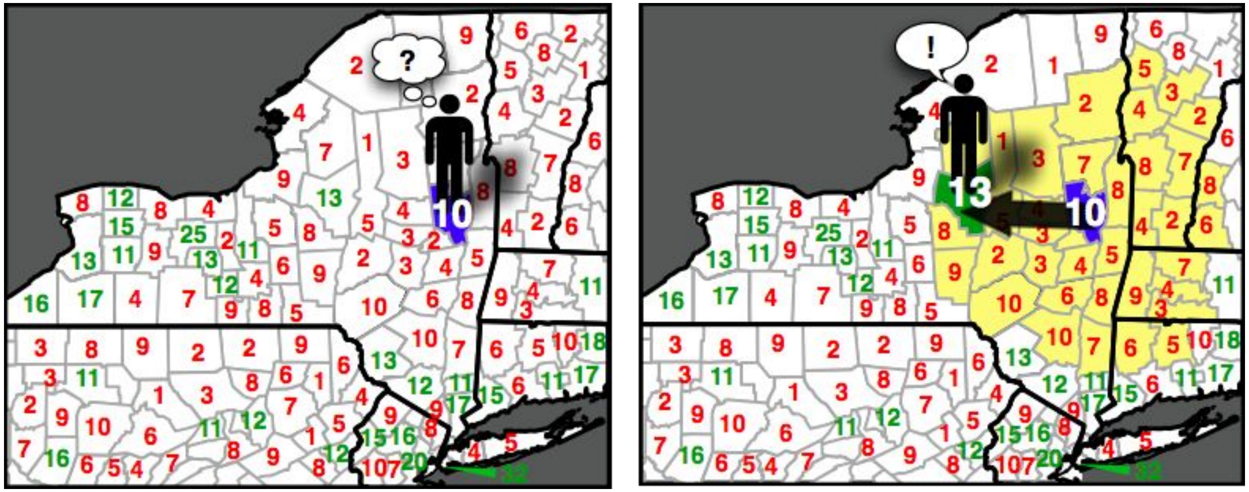
- The probability to choose any opportunity is λ
- Consider the closest opportunity and choose it with probability λ
- If the opportunity is not chosen, consider the next closest opportunity
- ...

$$p_{ij} \propto (1-\lambda)^{(m_i+s_{ij})} - (1-\lambda)^{(m_i+s_{ij}+n_j)}$$

(opportunities are proportional to the population)

Radiation model

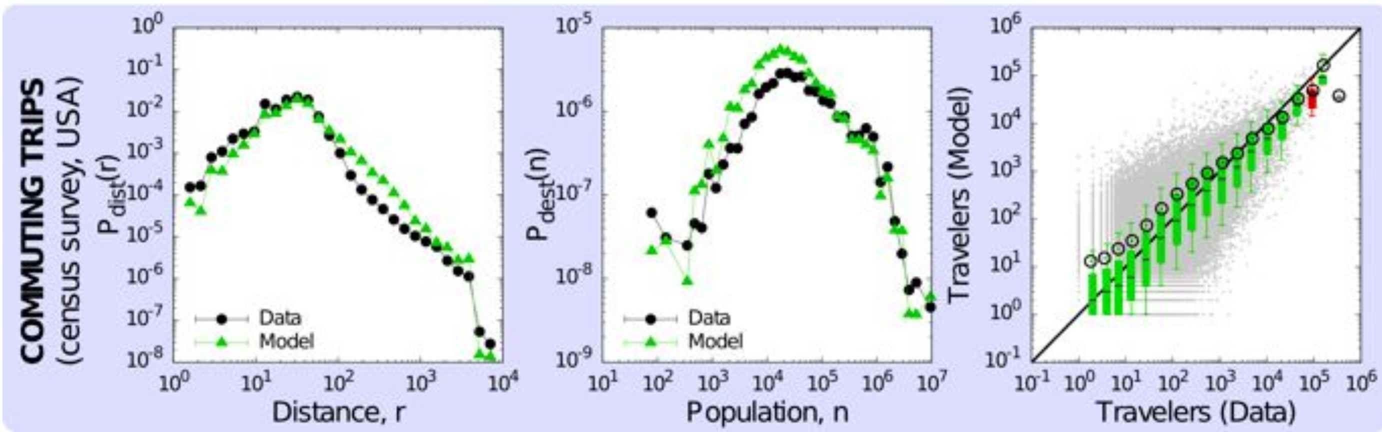
A universal model for mobility and migration patterns. (Simini, F, et al., Nature 484.7392 (2012): 96).



- Each opportunity has a “value”, extracted from some distribution.
- Each individual has expectations, extracted from the same distribution.
- *Principle of least effort*: each individual choose the **closest** opportunity that meets their expectations.

Radiation model

A universal model for mobility and migration patterns. (Simini, F, et al., Nature 484.7392 (2012): 96).



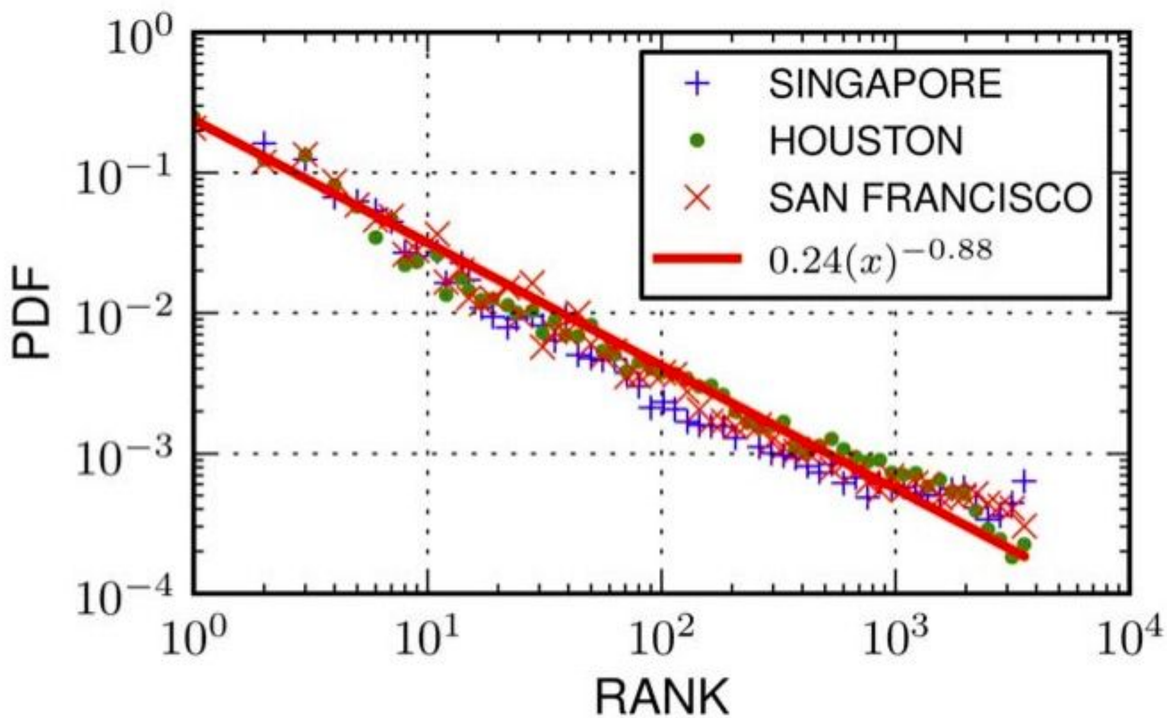
$$p_{ij} = \frac{m_i m_j}{(m_i + s_{ij})(m_i + s_{ij} + n_j)}$$

<u>Parameter-free</u>: the model depends only on the populations

Rank-distance model

A tale of many cities: universal patterns in human urban mobility (Noulas, A, et al., PloS one 7.5 (2012): e37027).

- Data: Foursquare check-ins
  - 925,030 users over 6 months
  - 5 million places, 34 cities, 4 continents, 11 countries



$$p_{ij} \propto \frac{1}{(m_i + s_{ij})^\alpha}$$

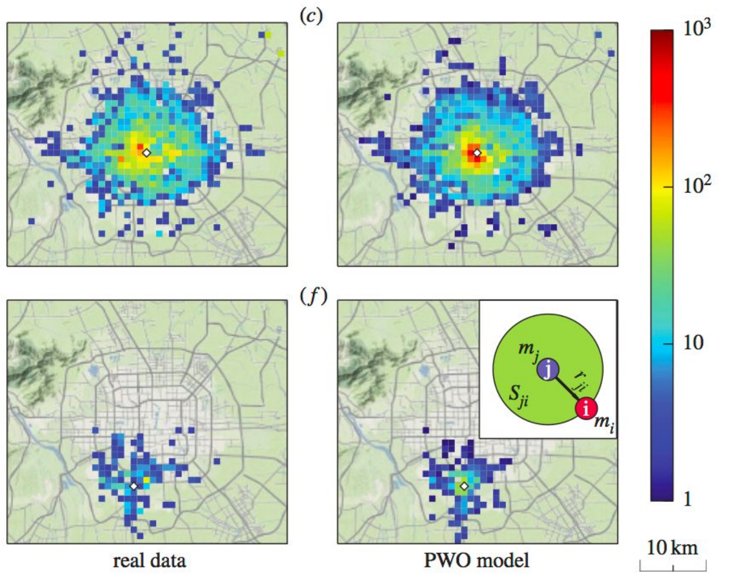
Population Weighted Opportunities model

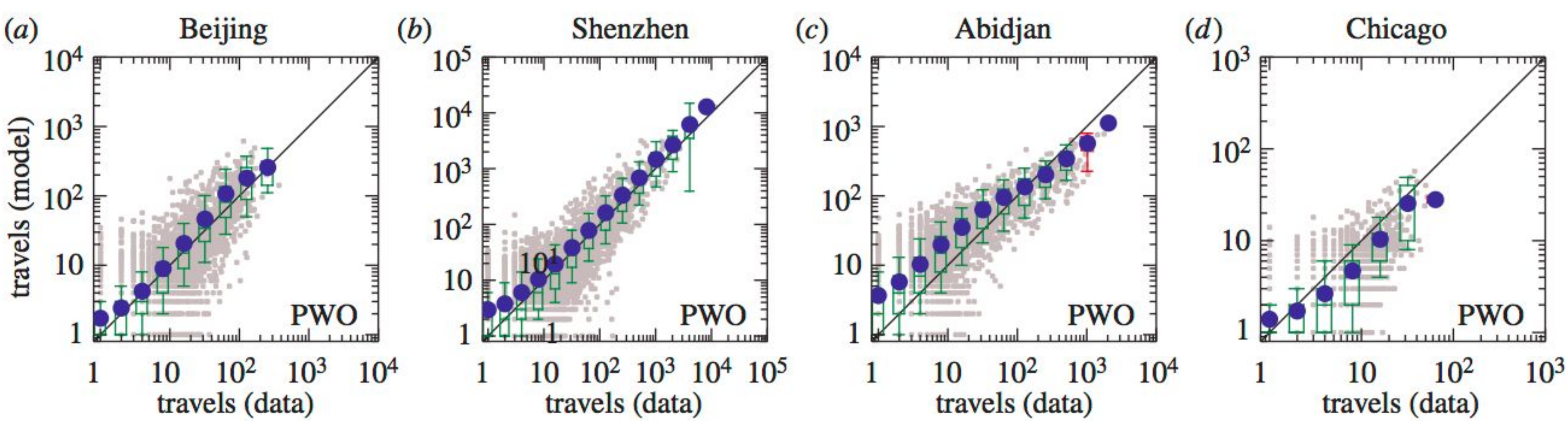
Universal predictability of mobility patterns in cities. (Yan, X. Y., et al., 2014, Journal of The Royal Society Interface, 11(100)).

The PWO model considers the intervening opportunities **centered at the destination**:

$$p_{ij} \propto m_j \left(\frac{1}{m_i + m_j + s_{ji}} - \frac{1}{M} \right)$$

<u>Parameter-free</u>





PROs and CONs of IO models

PROs

- **parameter-free**
  (Radiation and PWO)
- **performance**
  (comparable to Gravity models)

CONs

- **underfitting**
- **overdispersion**

Validation of collective models

Comprehensive survey on distance/similarity measures between probability density functions. (Cha, S. H., 2007, City, 1(2))

Common metrics to compare OD matrices

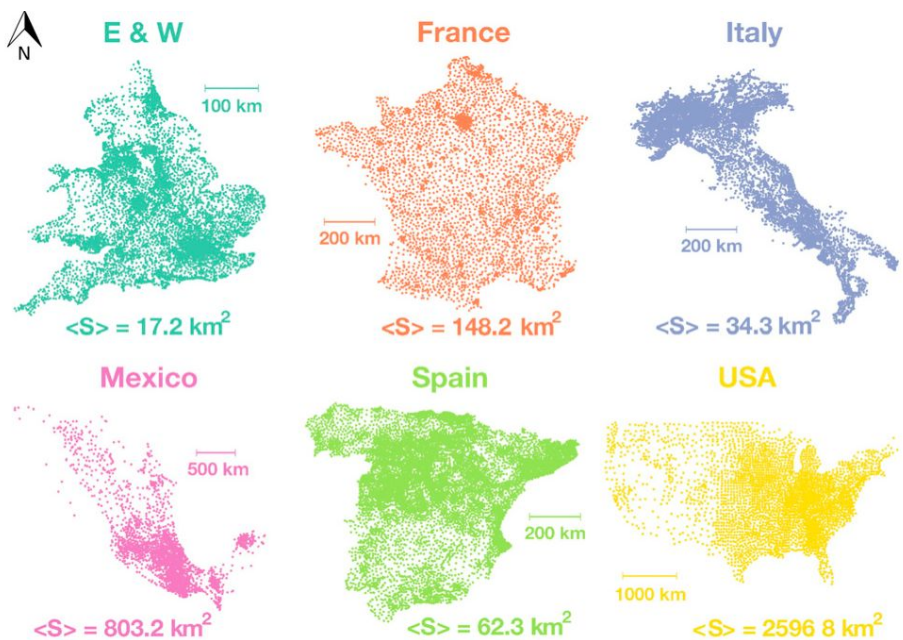
- Sorensen-Dice similarity (Common part of commuters) $$\frac{\sum_{ij} \min(T^e_{ij}, T^m_{ij})}{\sum_{ij} T^e_{ij}}$$

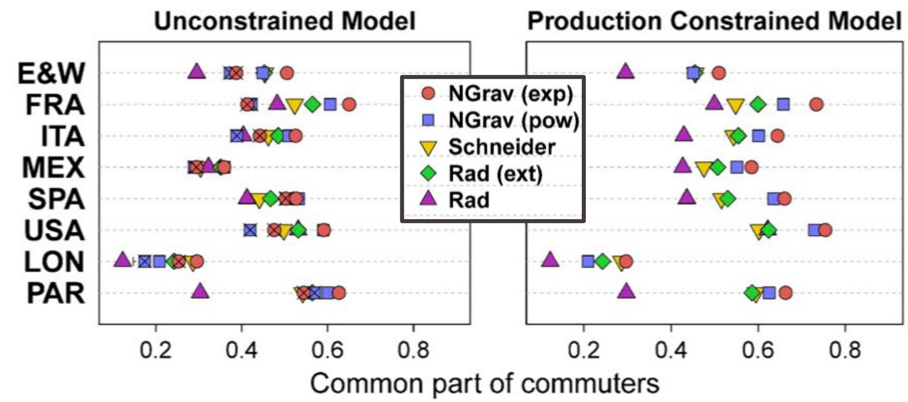
- Root Mean Squared Error $$\sqrt{\frac{\sum_{ij}(T^e_{ij} - T^m_{ij})^2}{n^2}}$$

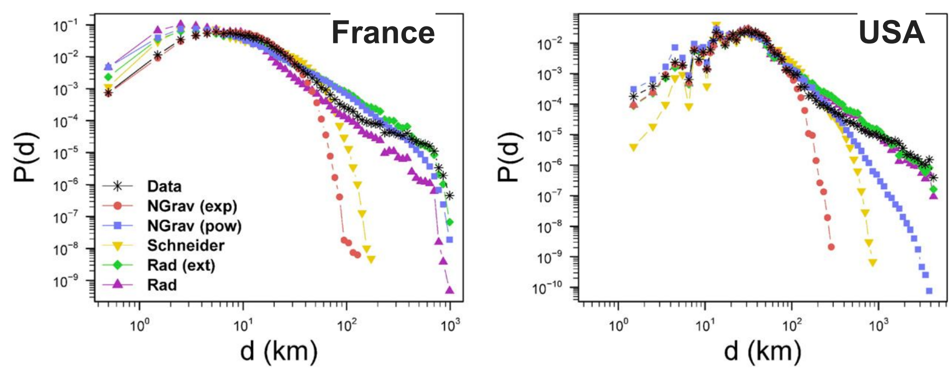
- More (cosine similarity, correlation, …)

Validation of collective models

Systematic comparison of trip distribution laws and models. (Lenormand, M., et al., 2006, Journal of Transport Geography)







Open issues

- Results and performance depend on the size of locations.
  What is the “correct” spatial scale?
- Underfitting and overdispersion:
  - missing variables
  - models are too simplistic

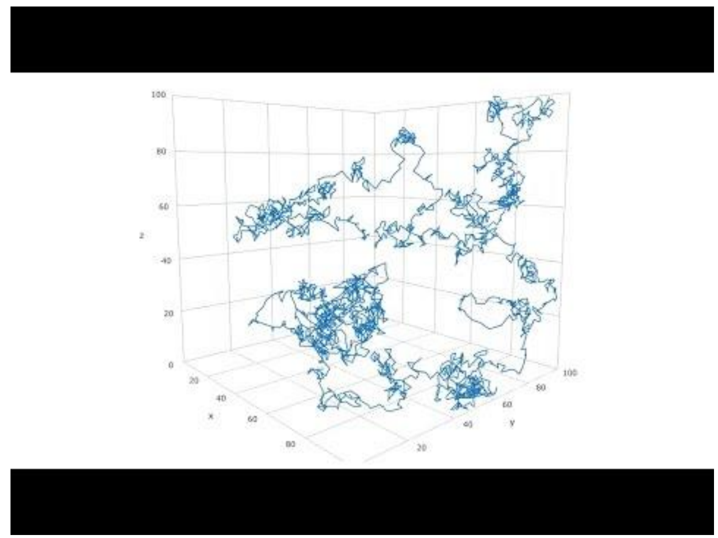
Individual models

Individual generative models

Goal: generate spatio temporal **trajectories**

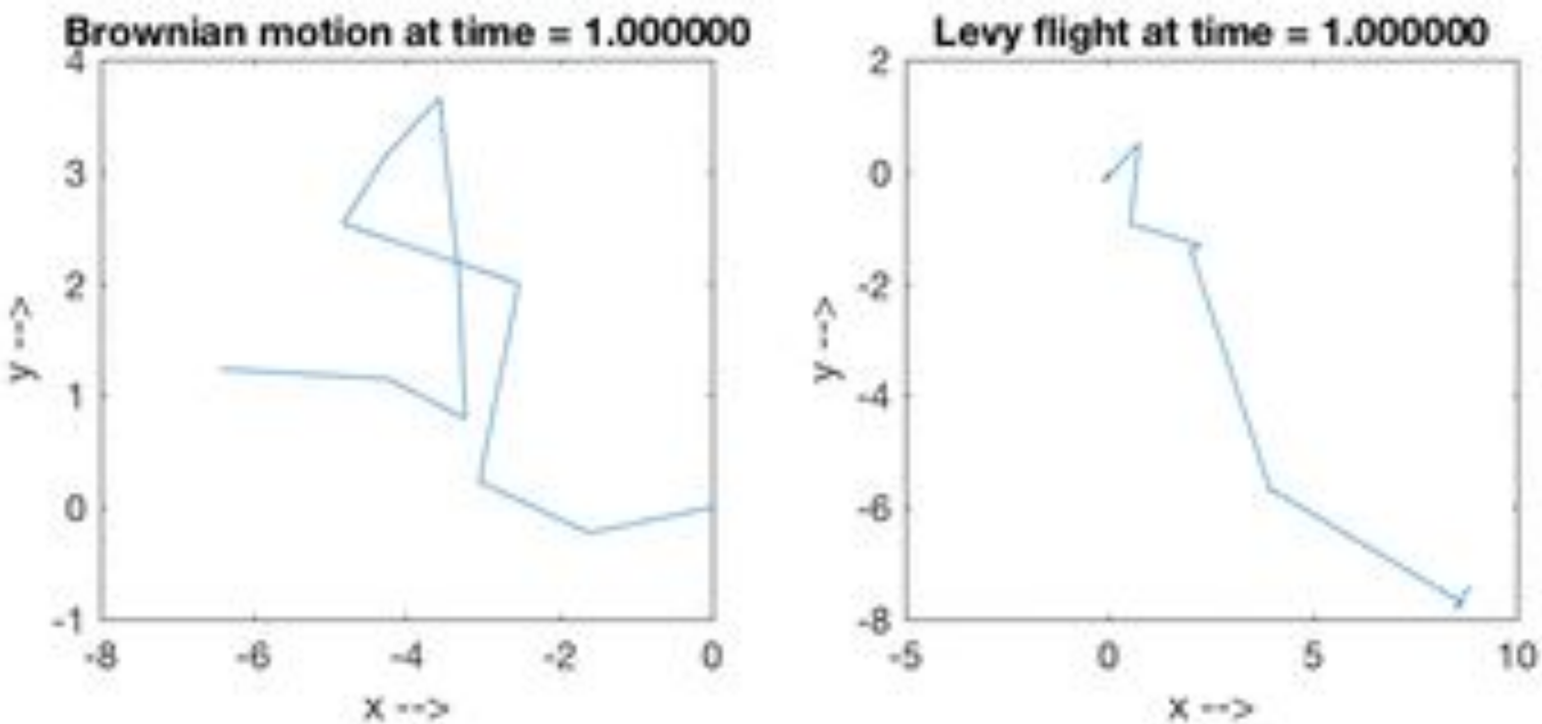
- Two components
  - **spatial** component: *"Where do I go next?"*
  - **temporal** component: *"How long do I stay?"*

- Simplest model: **Random walk**
  - "Where next?"
    1 unit step in a random direction
  - "How long?"
    1 unit time (constant)



Continuous-time random walk (CTRW)

- Jump length distribution $P(\Delta r) \propto \Delta r^{-(1+\alpha)}$
- Wait time distribution $P(\Delta t) \propto \Delta t^{-(1+\beta)}$



Ultraslow diffusion

Mean Squared Displacement

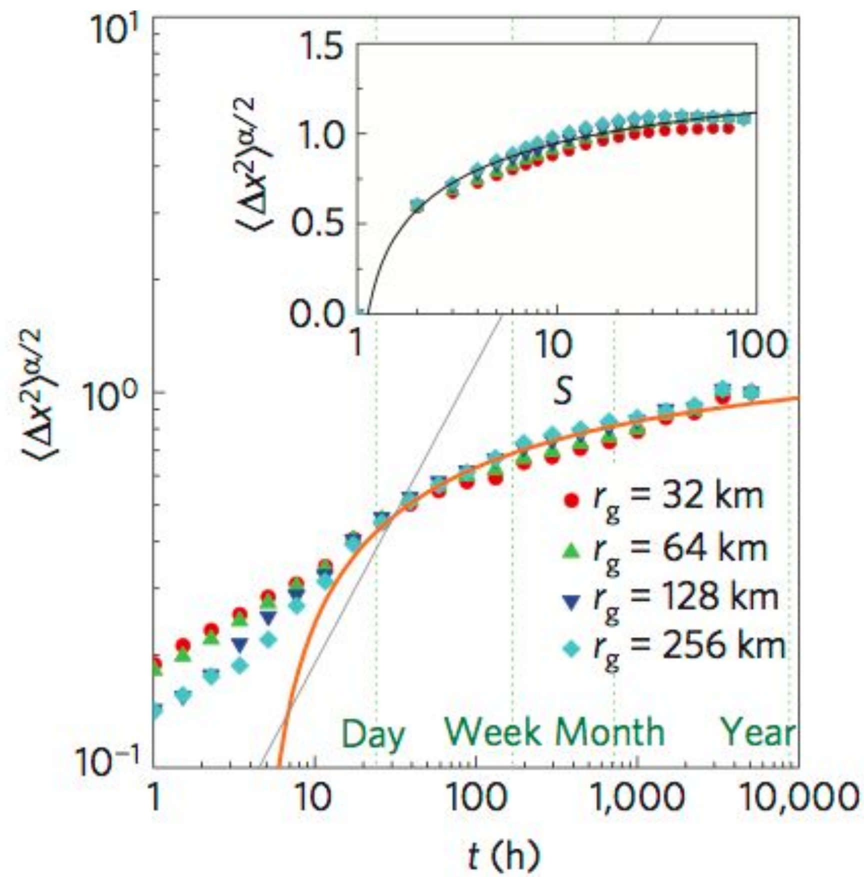
$$\langle \Delta x(t)^2 \rangle \equiv \langle (x(t) - x(0))^2 \rangle$$

CTRW

$$\langle \Delta x(t)^2 \rangle \sim t^{2\beta/\alpha} \sim t^3$$

Data

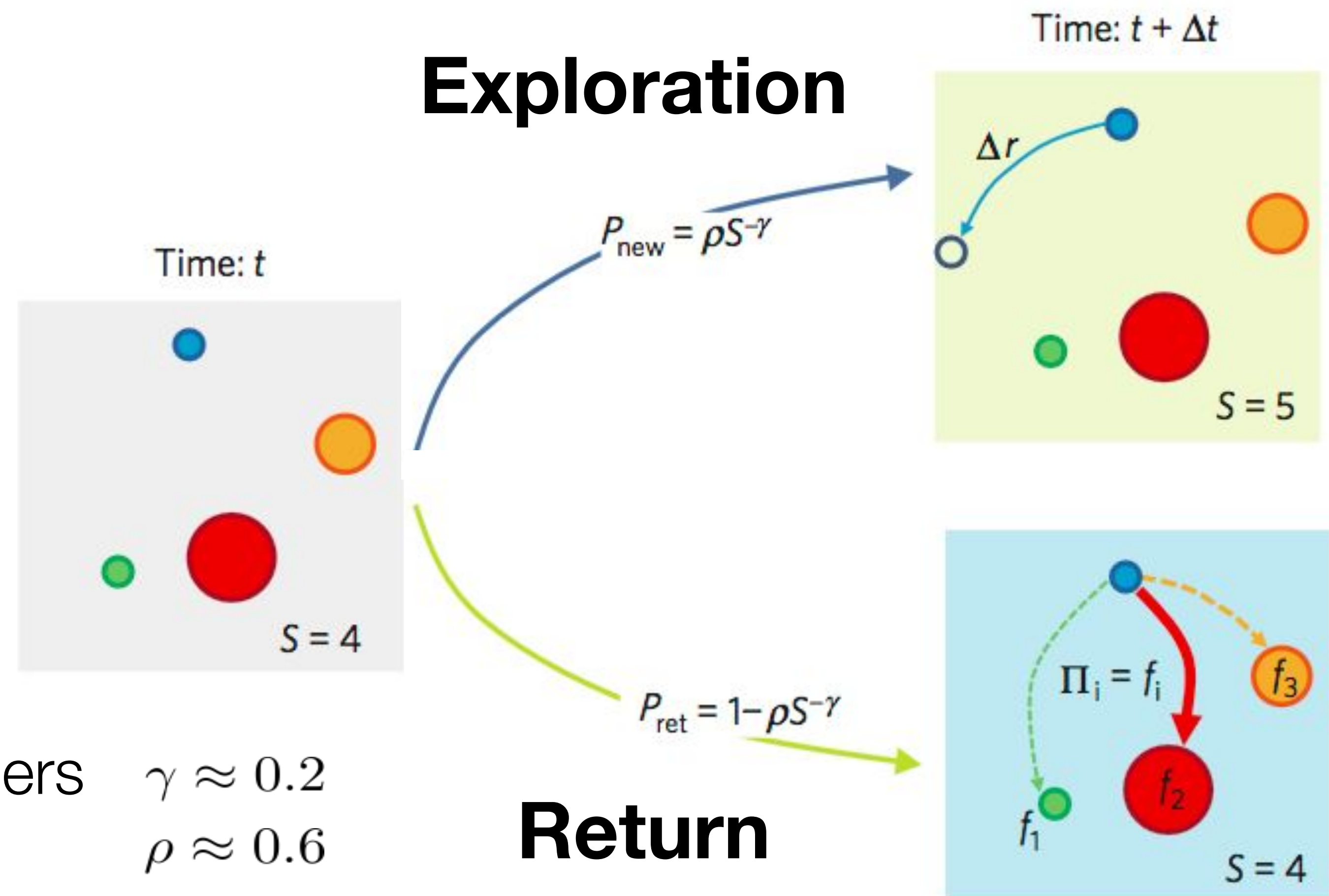
$$\langle \Delta x(t)^2 \rangle \sim (\ln \ln t)^{2/\alpha}$$



Modelling the scaling properties of human mobility. (Song, C., et al., 2010, Nature Physics, 6(10), 818).

Exploration and Preferential Return (EPR)

Modelling the scaling properties of human mobility. (Song, C., et al., 2010, Nature Physics, 6(10), 818).



Parameters

$\gamma \approx 0.2$

$\rho \approx 0.6$

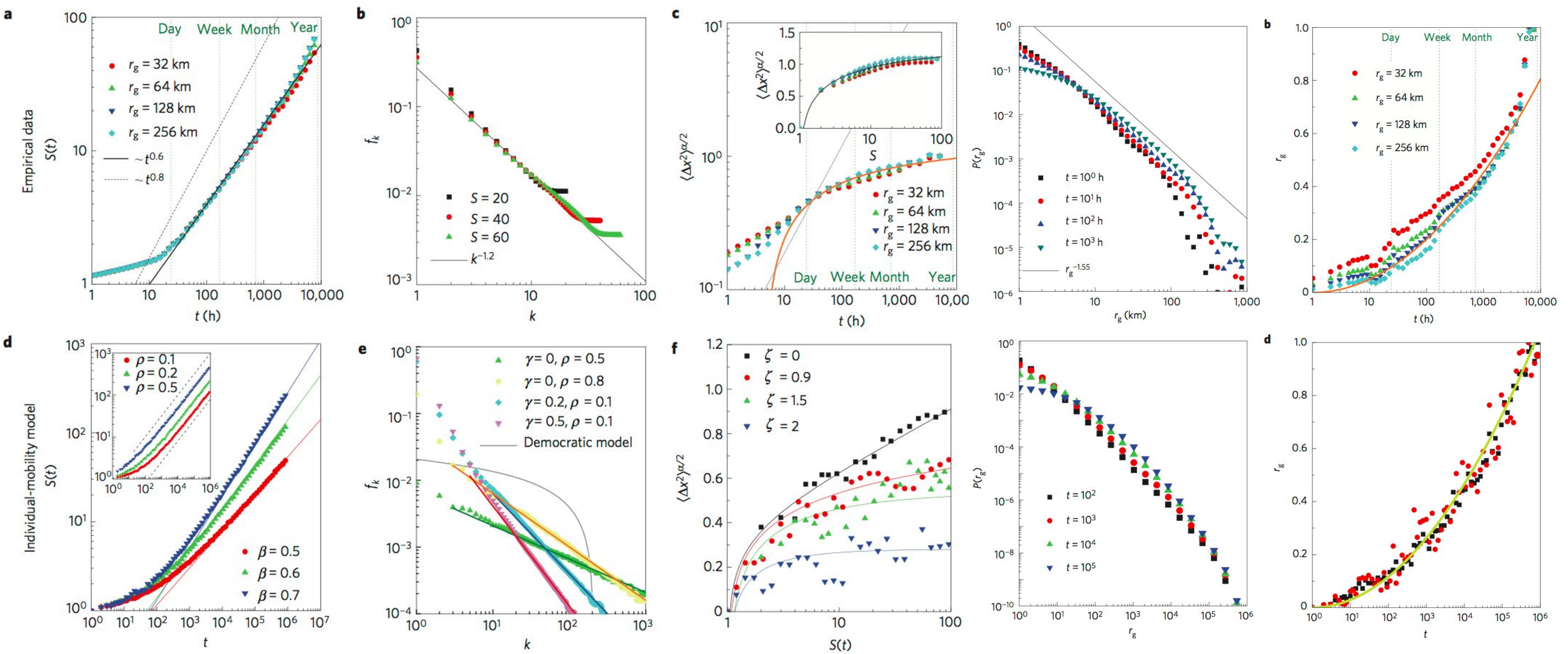
$\alpha \approx 0.55$

$\beta \approx 0.8$

EPR - results

Modelling the scaling properties of human mobility. (Song, C., et al., 2010, Nature Physics, 6(10), 818).

Number of visited locations vs time	**Visitation frequency vs rank**	**Ultraslow diffusion**	**Distribution of radii of gyration**	**Radius of gyration vs time**



$$S \sim t^{\frac{\beta}{(1+\gamma)}}$$

$$f_k \sim k^{-(1+\gamma)}$$

$$\langle \Delta x(t)^2 \rangle \sim (\ln \ln t)^{2/\alpha}$$

$$P(r_g) \sim r_g^{-(1+\alpha)}$$

$$r_g(t) \sim (\ln t)^2$$

PROs and CONs of EPR model

PROs

- **simple** and **elegant**
- turned into a **modeling framework**
- **accurate** in reproducing basic statistics

CONs (spatial component)

- does not consider **population** density in the exploration phase
- cannot reproduce the **returners/explorers** dichotomy

PROs and CONs of EPR model

PROs

- **simple** and **elegant**
- turned into a **modeling framework**
- **accurate** in reproducing basic statistics

CONs (temporal component)

- does not capture **periodicity**
- does not capture **recency**
- cannot decribe **changes** in preferences
- does not capture **limits** in exploration

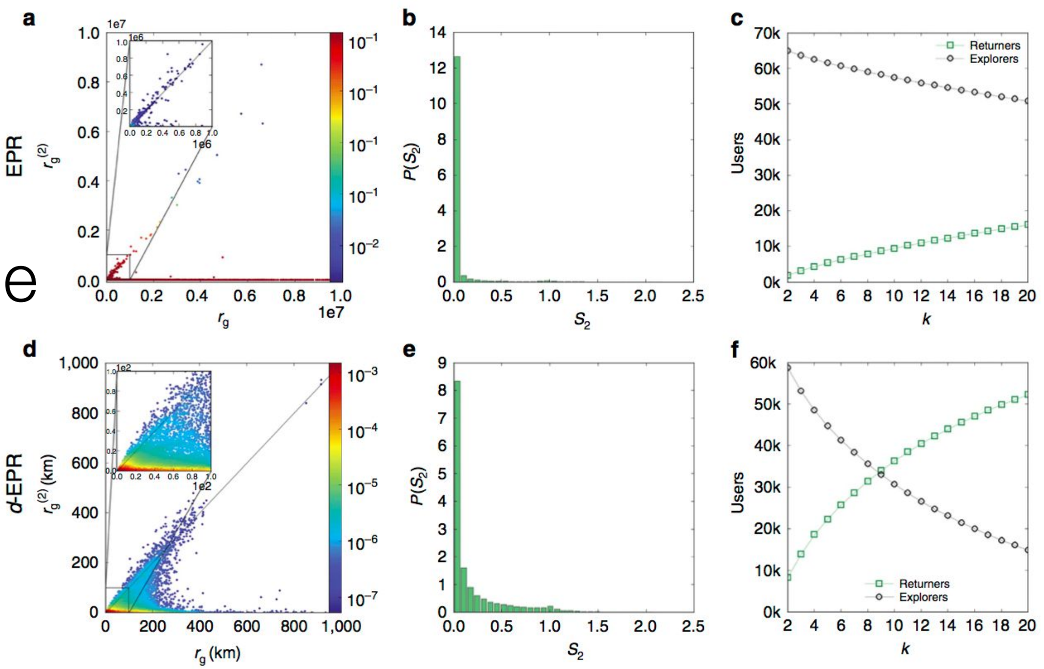
Density EPR (d-EPR)

[1] Returners and explorers dichotomy in human mobility. (Pappalardo, L., et al., 2015,Nature communications, 6, 8166).
[2] Human Mobility Modelling: exploration and preferential return meet the gravity model (Pappalardo, L. et al. (2016))

Modification of the **Spatial** component:

in the exploration phase, use a **gravity model**.

$$p_{ij} \propto O_i \, m_j \, f(r_{ij})$$



<u>Outcomes</u>

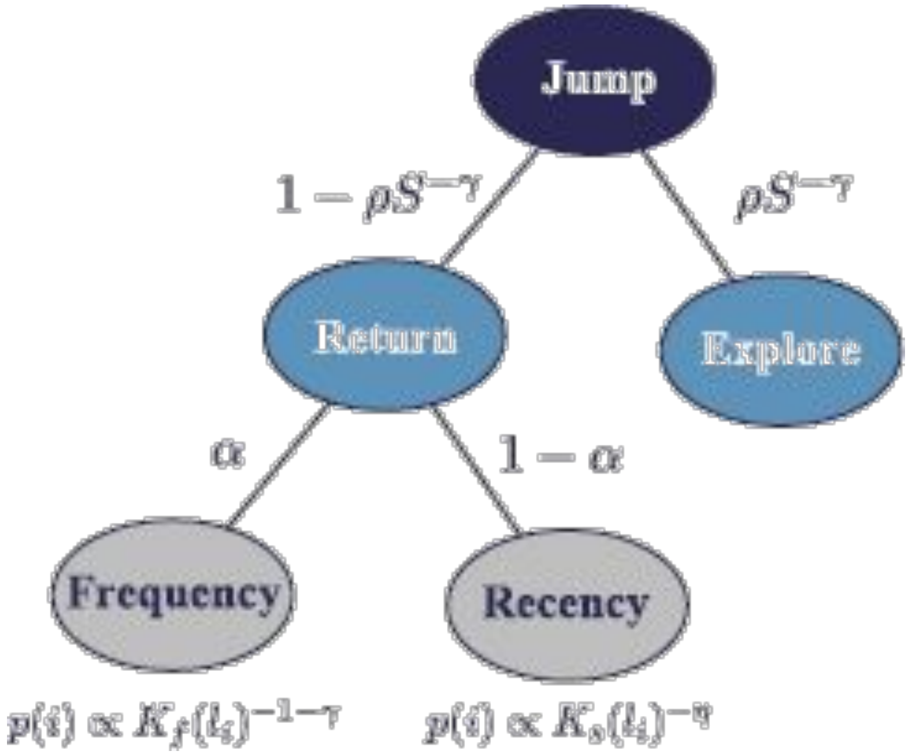
- Takes into account the geography
- Reproduces the statistics of returner and explorers

Recency EPR (r-EPR)

The effect of recency to human mobility. (Barbosa, H., et al., 2015, EPJ Data Science, 4(1), 21).

Modification of the **Temporal** component:

the probability of returning to a location is based on both visitation *frequency* and *recency* ranks



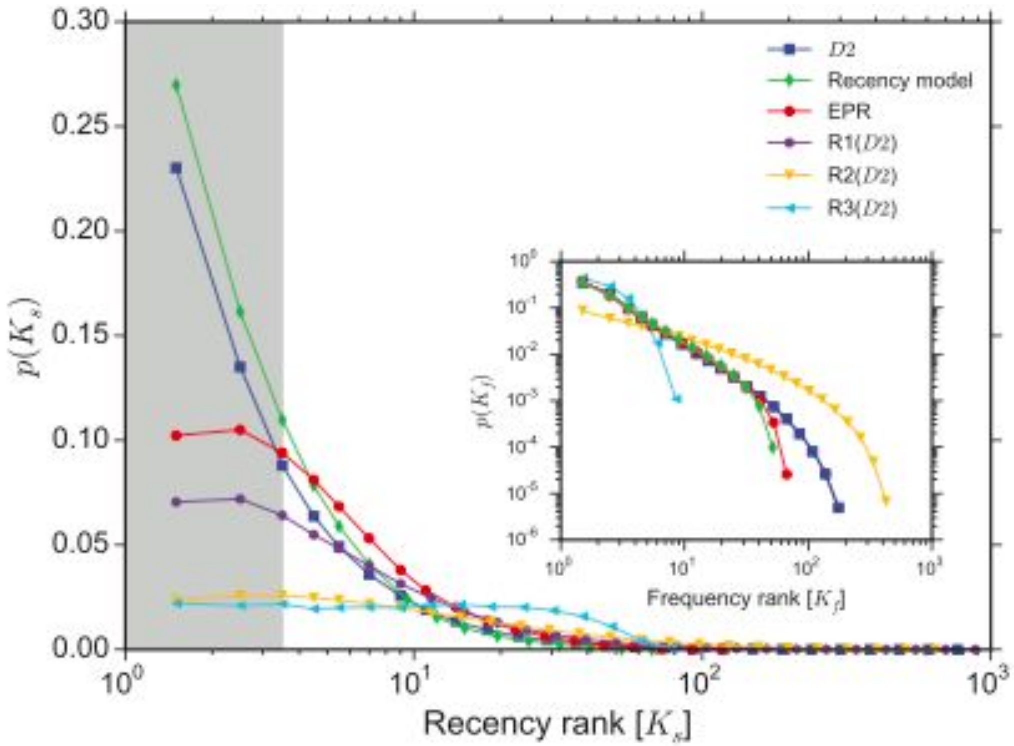
$\alpha = 0.1$
$\nu = 1.6$

Recency EPR (r-EPR)

The effect of recency to human mobility. (Barbosa, H., et al., 2015, EPJ Data Science, 4(1), 21).

Modification of the **Temporal** component:

the probability of returning to a location is based on both visitation *frequency* and <u>*recency*</u> ranks



<u>Outcomes:</u> captures the higher probability to revisit recently visited locations

$$\alpha = 0.1$$
$$\nu = 1.6$$

Memory EPR (m-EPR)

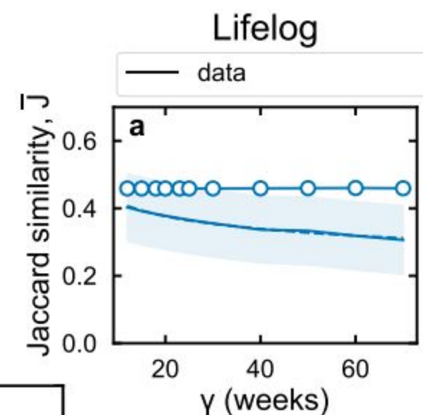
Evidence for a conserved quantity in human mobility. (Alessandretti, L., et al., 2018, Nature Human Behaviour).

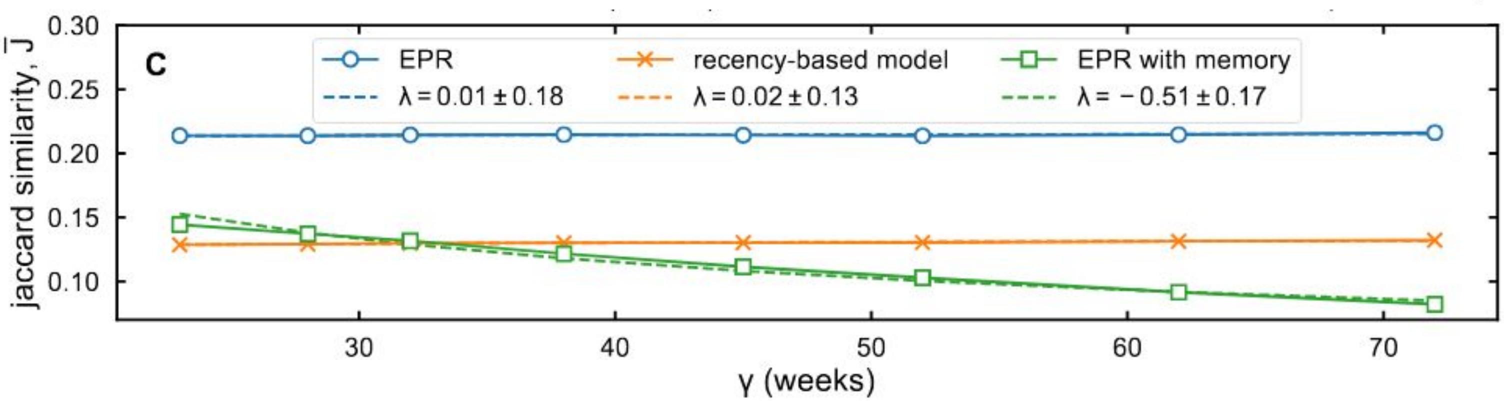
Modification of the **Temporal** component:

the probability of returning to a location is based on the number of visits occurred in the last *M* days

<u>Outcomes</u>

- Captures the change over time of the set of frequently visited locations





TimeGeo

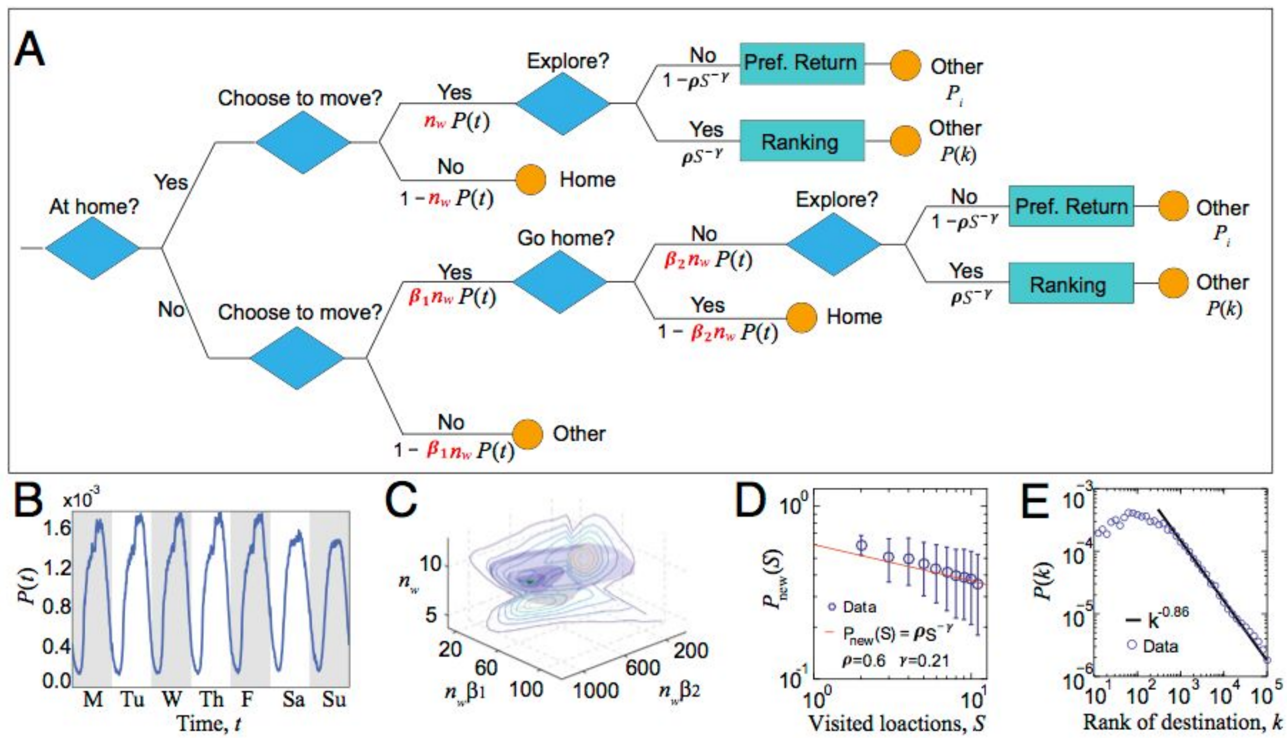
The TimeGeo modeling framework for urban mobility without travel surveys. (Jiang, S., et al., 2016, PNAS, 113(37))

Modification of the **Temporal** component:

the probability of travelling depends on the Travel Circadian Rhythm; home is a special location.

Modification of the **Spatial** component:

in the exploration phase, use a **rank IO model**.

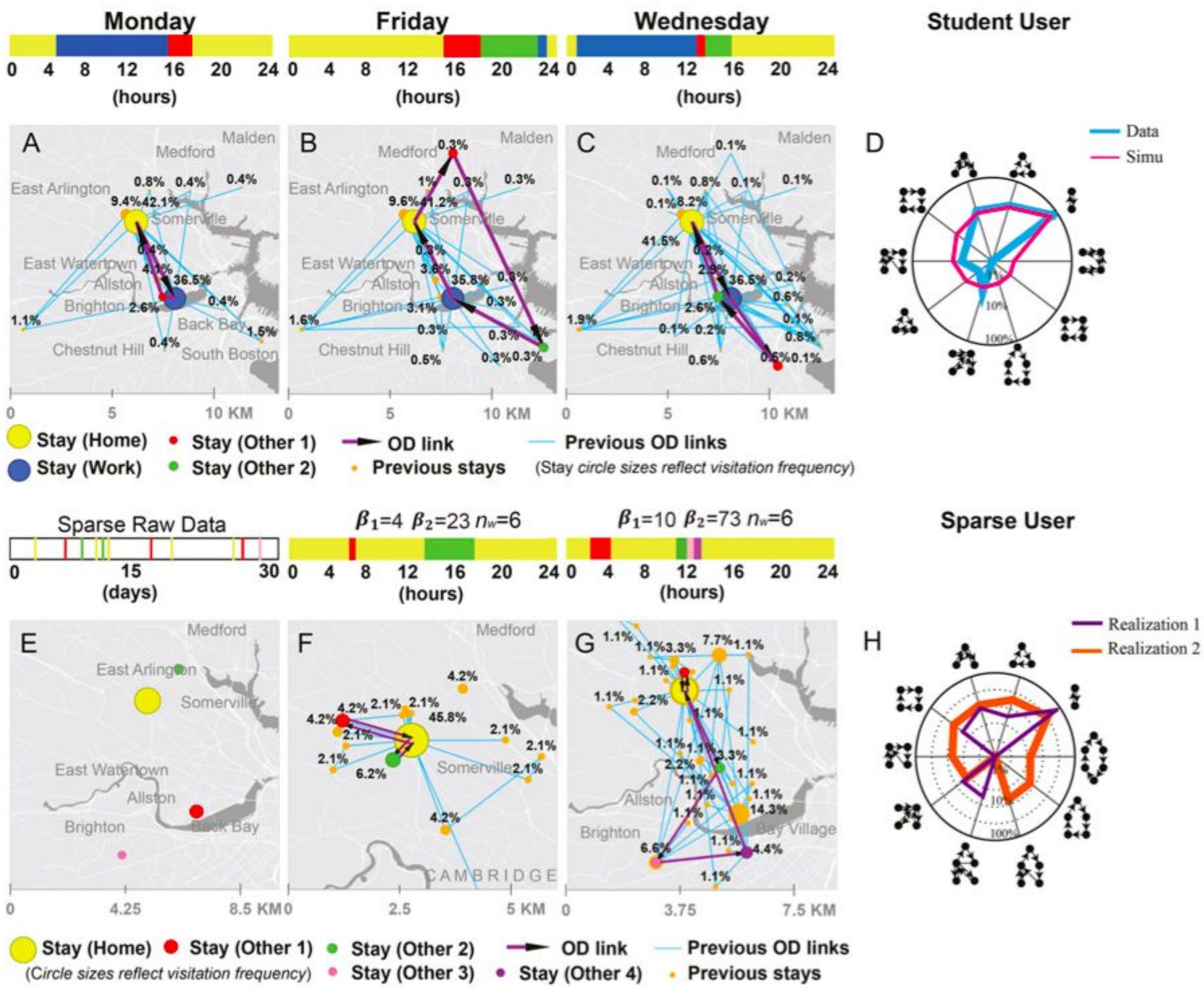


TimeGeo

The TimeGeo modeling framework for urban mobility without travel surveys. (Jiang, S., et al., 2016, PNAS, 113(37))

Outcomes

- Periodic temporal patterns
- Reproduces the statistics of mobility motifs



DITRAS

Data-driven generation of spatio-temporal routines in human mobility. (Pappalardo, L., et al., 2018, Data Mining and Knowledge Discovery).

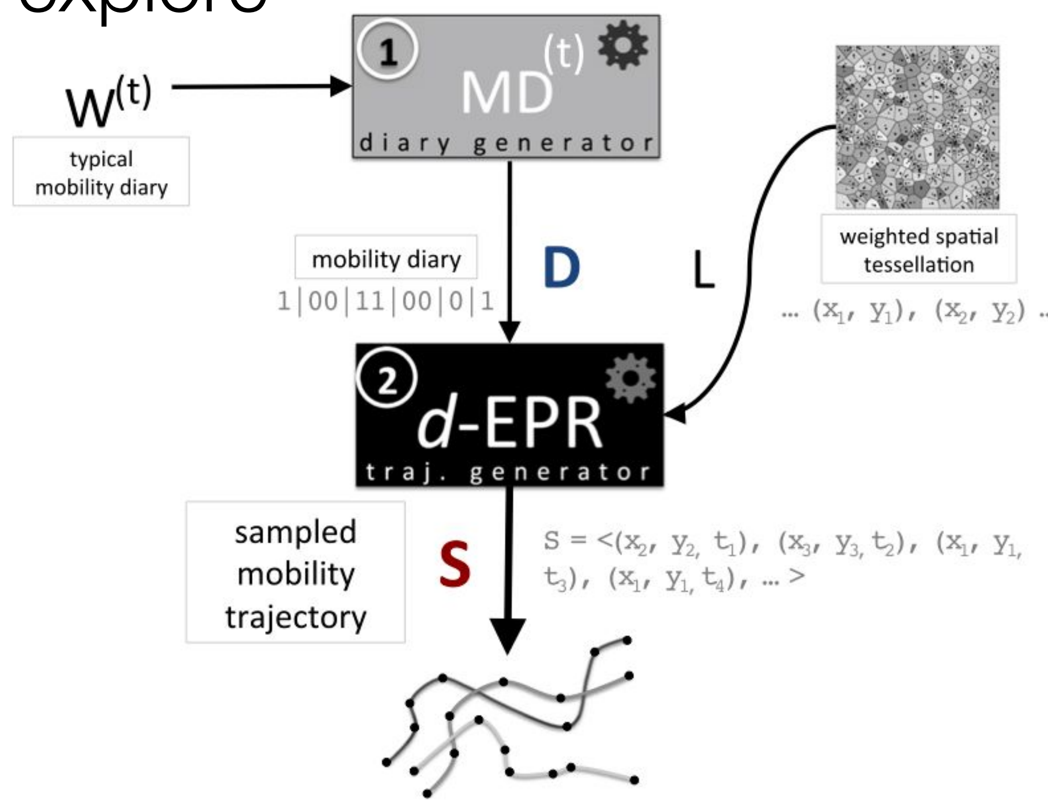
Modification of the **Temporal** component:

Markov chain determines the probability to follow the typical mobility or explore

Spatial component:
d-EPR

<u>Outcomes</u>

- Realistic distributions of locations per user, trips per hour and day, stay duration

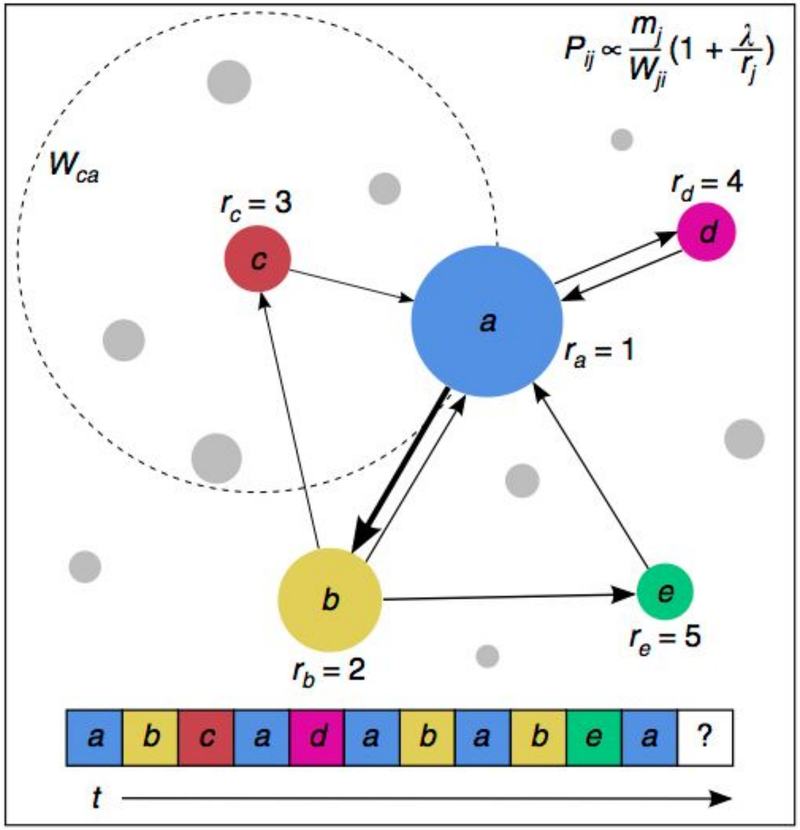


Individual PWO

Universal model of individual and population mobility on diverse spatial scales. (Yan, X. Y., et al., 2017, Nature Communications, 8(1), 1639.)

combines the **P**opulation **W**eighted **O**pportunity model and a **memory effect** based on rank

$$p_{ij} \propto \frac{m_j}{m_i + s_{ji} + m_j}\left(1 + \frac{\lambda}{r_j}\right)$$



<u>Outcomes</u>

- Return time distribution
- Visitation frequency
- Number of visited locations vs time
- Motifs' frequencies
- No temporal mechanism

Open issues

- long-term mobility
  (holiday, illness, relocations)

- <u>jointly</u> model the spatial and temporal components

- capture out-of-routine behavior
  (like DITRAS)

- incorporate social interactions

Where's next: predicting human movements

Human Mobility and Machine Learning

Human mobility patterns, both at individual and collective level, can be used for several **predictive tasks**:

Movements Prediction

- Next Location prediction

  DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Human Behaviour Prediction

- Predicting future health condition

  Are you getting sick? Predicting influenza-like symptoms using human mobility behaviors *(Barlacchi et al.)*

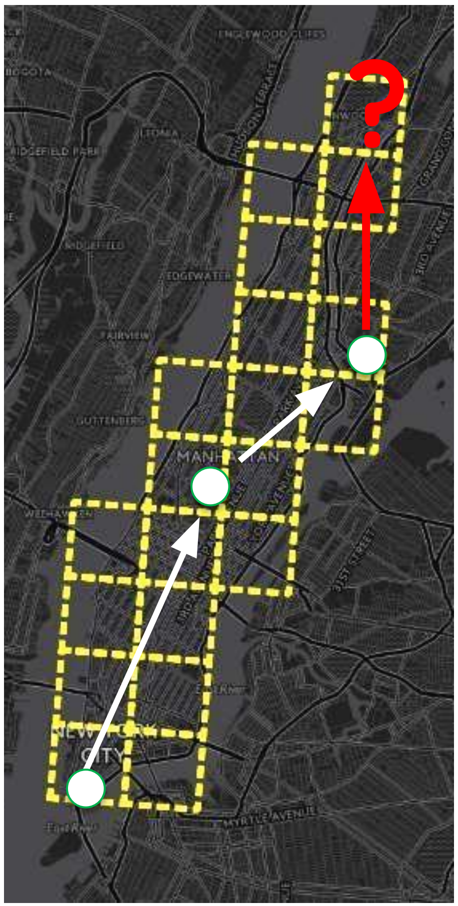
Urban Planning

- Discover functional regions

  Discovering regions of different functions in a city using human mobility and POIs (Yuan et al.)

Next Location Prediction

How to predict human movement from trajectories?



Mobility trace of the user — **Input**

↓

Predict the next location the user will visit — **Task**

An important problem
in recommendation systems

Problem Definition

For each user *u*, their mobility history is:

$$Q^u = \{q^u_{t_1}, q^u_{t_2}, \dots\}$$

where $q^u_{t_i}$ denotes where user *u* is at time t_i.

The history of all users is denoted as

$$Q^U = \{Q^{u_1}, Q^{u_2}, \dots\}$$

Given historical records of a user, the task is to predict where they will go next.

Classification VS Regression

Classification → location

Classify the next location the user will visit

- Human Mobility Prediction based on Individual and Collective Geographical Preferences. (Calabrese et al.)
- Next Place Prediction using Mobility Markov Chains (Gambs et al.)
- WhereNext: a location predictor on trajectory pattern mining (Monreale et al.)
- DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Regression → lat, lon

Predict the latitude and longitude values of the next location

Artificial Neural Networks Applied to Taxi Destination Prediction
(de Brébisson et al.)

Evaluation metrics

In the context of recommendation systems (**classification**) we are most likely interested in recommending top-*k* items to the user.

Accuracy@k: the number of correct location predictions, where a location prediction is correct if the real location is within the top-k recommended locations.

Recall@k: the number of correct location predictions over the number of all possible locations

$$recall = \frac{number\ of\ real\ prediction\ locations}{total\ number\ of\ ground\ truth\ locations},$$

Evaluation metrics

Predictability
theoretical measure representing the degree to which the mobility of an individual is predictable based on his Markov Chain Model

probability of being in a particular state

$$Pred = \sum_{k=1}^{l} (\pi(k) \times P_{max\_out}(k, *))$$

Max outgoing probability from the k^{th} state

Individual-Collective model

Human Mobility Prediction based on Individual and Collective Geographical Preferences. (Calabrese et al.)

Idea: Predict the location of a user over time based on **individual** and **collective** behaviors.

$\alpha(k)$ parameter

$$P(x_{k+1}(u) = j | x_k(u) = i) =$$

$$(1 - \alpha(k))P_I(x_{k+1} = j | x_k = i) +$$

individual component

$$\alpha(k)P_C(x_{k+1} = j | x_k = i)$$

collective component

Collective behavior

helps in predicting the likelihood the user changes location and the type of place they will visit

1. Distances being traveled (**D**)

2. Land Use (**LU**)
   describe the primary use of an area (e.g., commercial)

3. Point Of Interest (**POI**)
   collection of information regarding a venue (e.g., school, restaurant, police department)

Collective component

$P_C(x_{k+1} = j | x_k = i)$ is composed by three terms taking into account geographic data

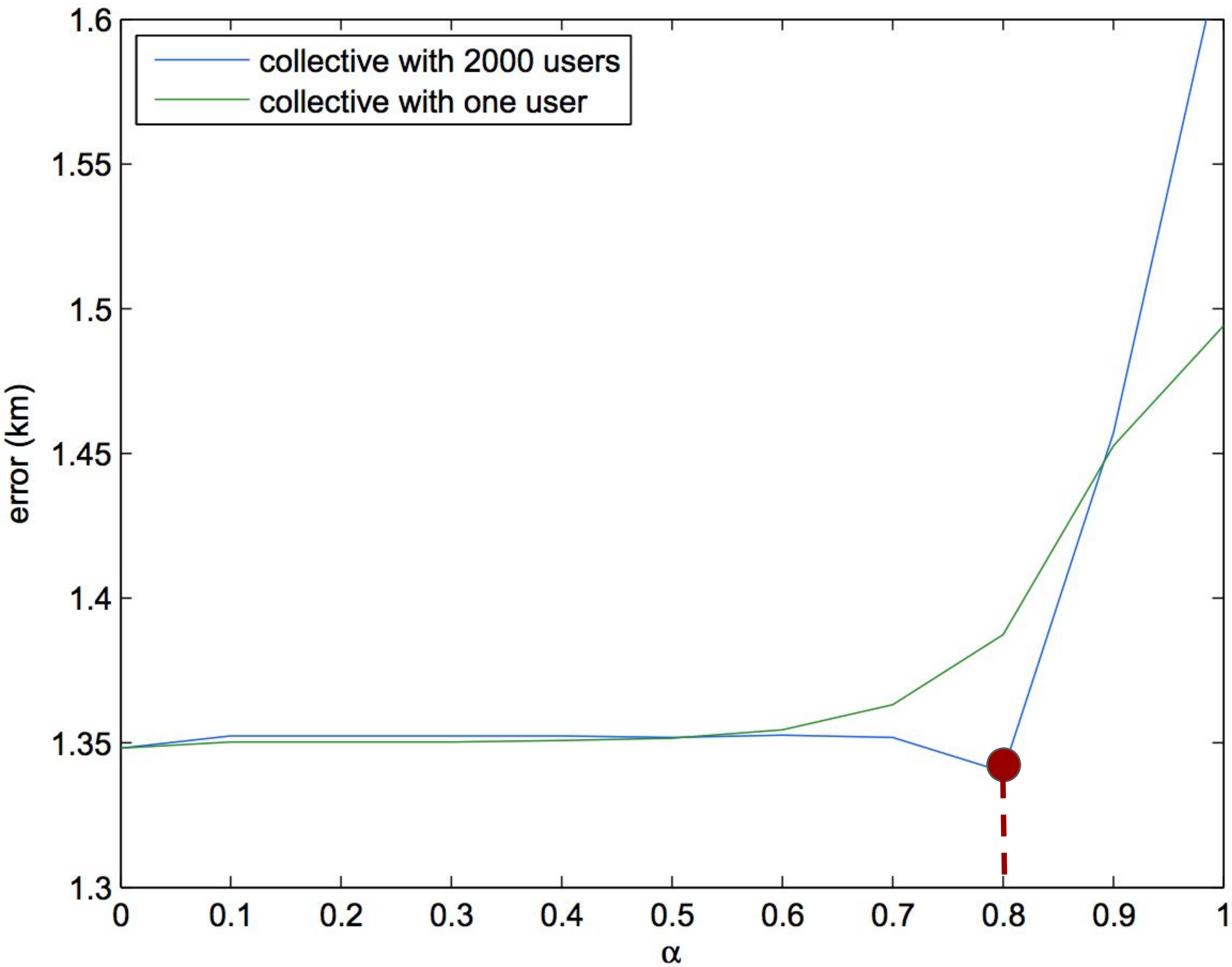
$P_{POI}(x_{k+1} = j)$ probability to find a user in location j at time *k* given the POIs in that location

$P_{LU}(x_{k+1} = j)$ probability to find a person in location j at time k given the percentage of land use in location j

$P_D(x_{k+1} = j | x_k = i)$ probability of moving between locations at distance d(i, j)

Human Mobility Prediction based on Individual and Collective Geographical Preferences. (Calabrese et al.)

Experimental Results

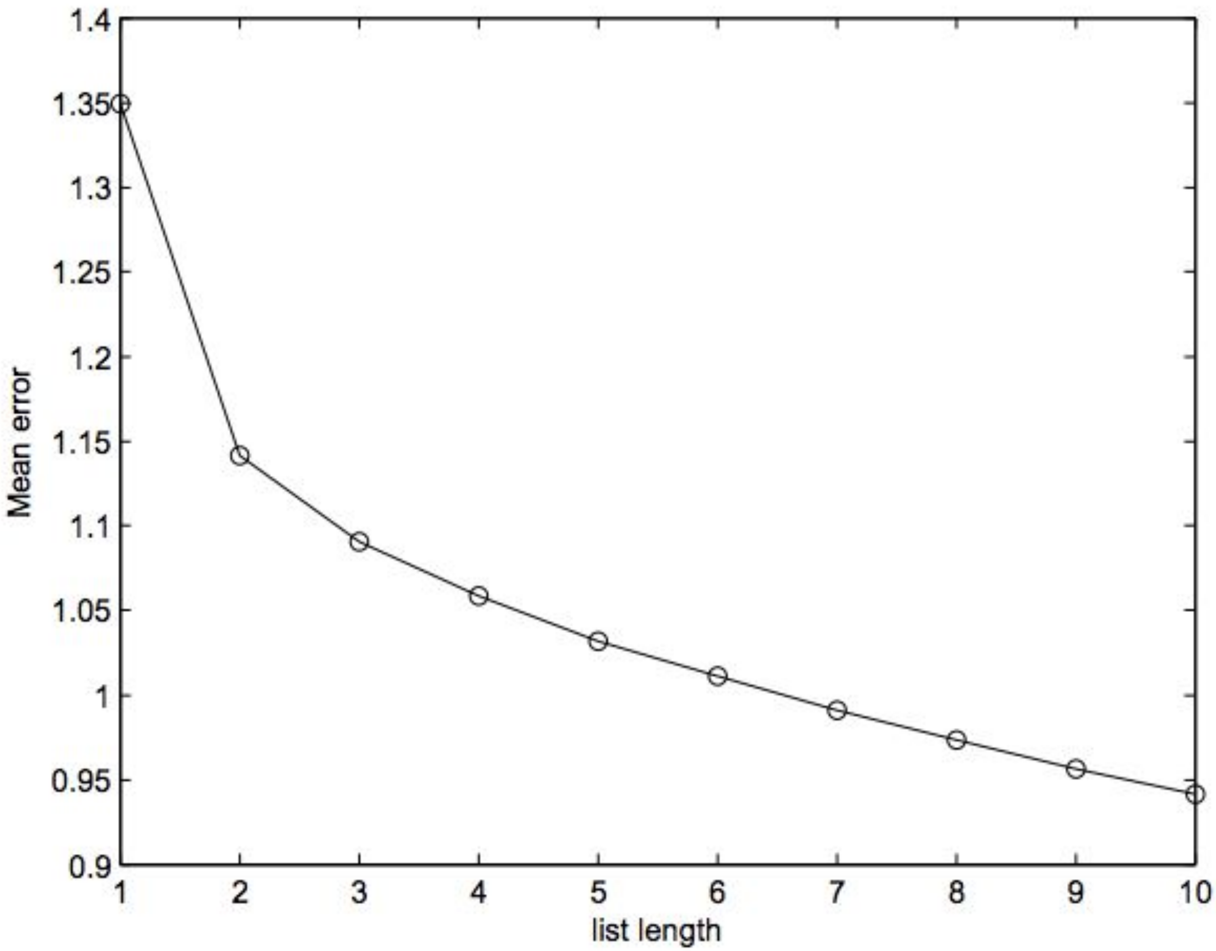


the collective component is prominent

$$\alpha = 0.8$$

is optimal

Experimental Results



- the higher the k the lower the error
- real location is usually within the first 3 locations

Mean error = distance between predicted location and real location

Next Location Prediction with Mobility Markov Chain: MMC and n-MMC

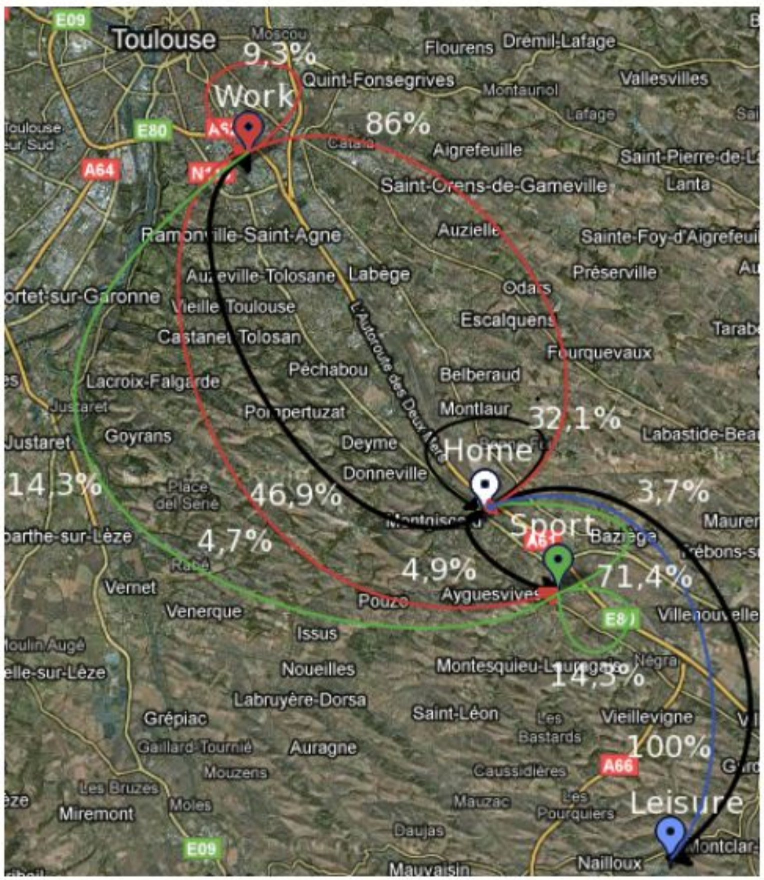
Idea:

- Modeling human mobility with a Markov Chain [1].
- Extend the mobility model called Mobility Markov Chain [2].
- MMC incorporating the n previous visited locations (n-MMC).

[1] Show me how you move and I will tell you who you are. (S. Gambs)
[2] Next Place Prediction using Mobility Markov Chains (Gambs et al.)

Mobility Markov Chain (MMC)

Models the mobility behavior of a user as a discrete stochastic process



MMC is composed by:

- Set of states
  P = {p1, . . . pk},
  each state is a POIs
- Set of transitions $t_{i,j}$,
  probabilities of moving from i to j

Mobility Markov Chain (MMC)

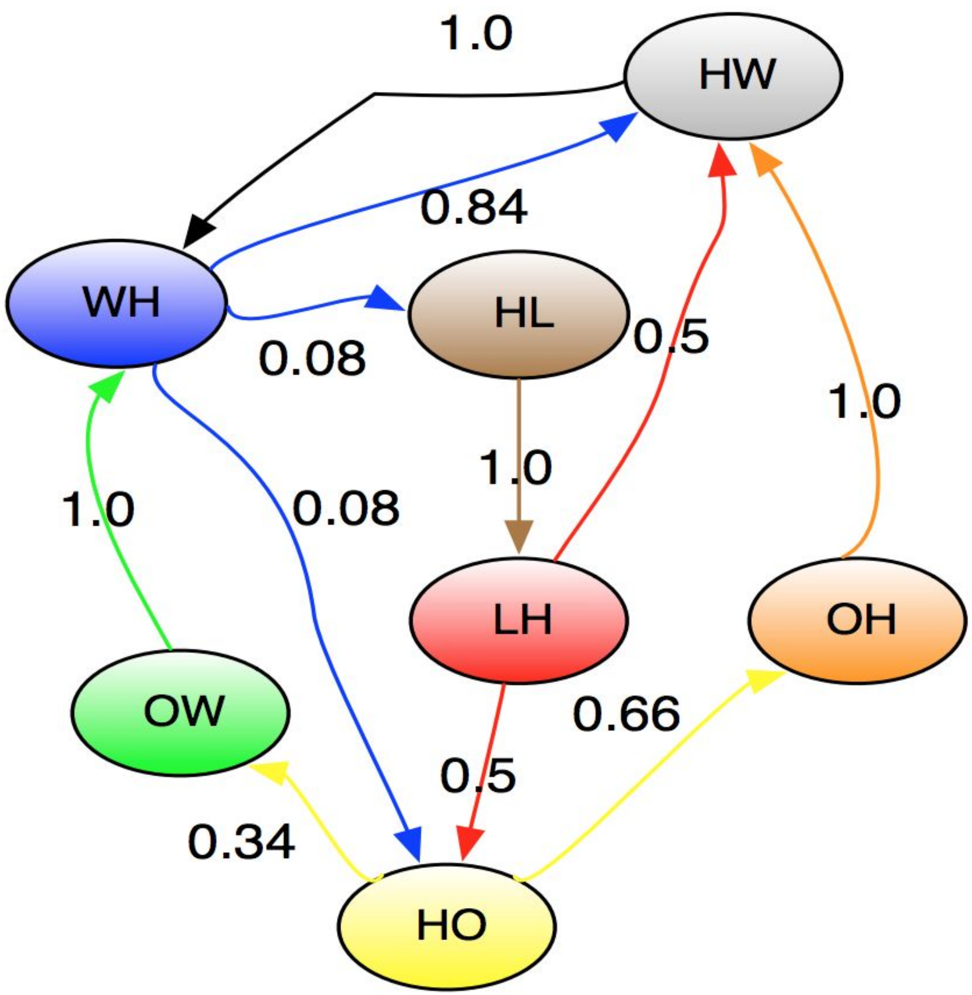
PROs

- **Simple** modeling
- Computationally **efficient**

CONs

- **Memoryless**:
  only the past state is considered
- **Low accuracy**

n-MMC: adding memory to MMC



- the past n states are remembered
- rows in the transitions matrix are all possible combinations of the *n* previous locations

Predicting next location with n-MMC

Example with n=2

Return the POI corresponding to the column with P_{max}

Source/Dest.	H	W	L	O
H W	1,00	0,00	0,00	0,00
H L	1,00	0,00	0,00	0,00
H O	0,64	0,34	0,00	0,00
W H	0,00	0,84	0,08	0,08
L H	0,00	0,50	0,00	0,50
O H	0,00	1,00	0,00	0,00
O W	1,00	0,00	0,00	0,00

The recommended location is H

Search the row r corresponding to the n previous visited locations

Find the column corresponding to the maximum probability of transition

Datasets for the experiments

Phonetic dataset

Mobility traces from 6 researchers sampled at a rate of 1 to 5 minutes from October 2009 to January 2011.

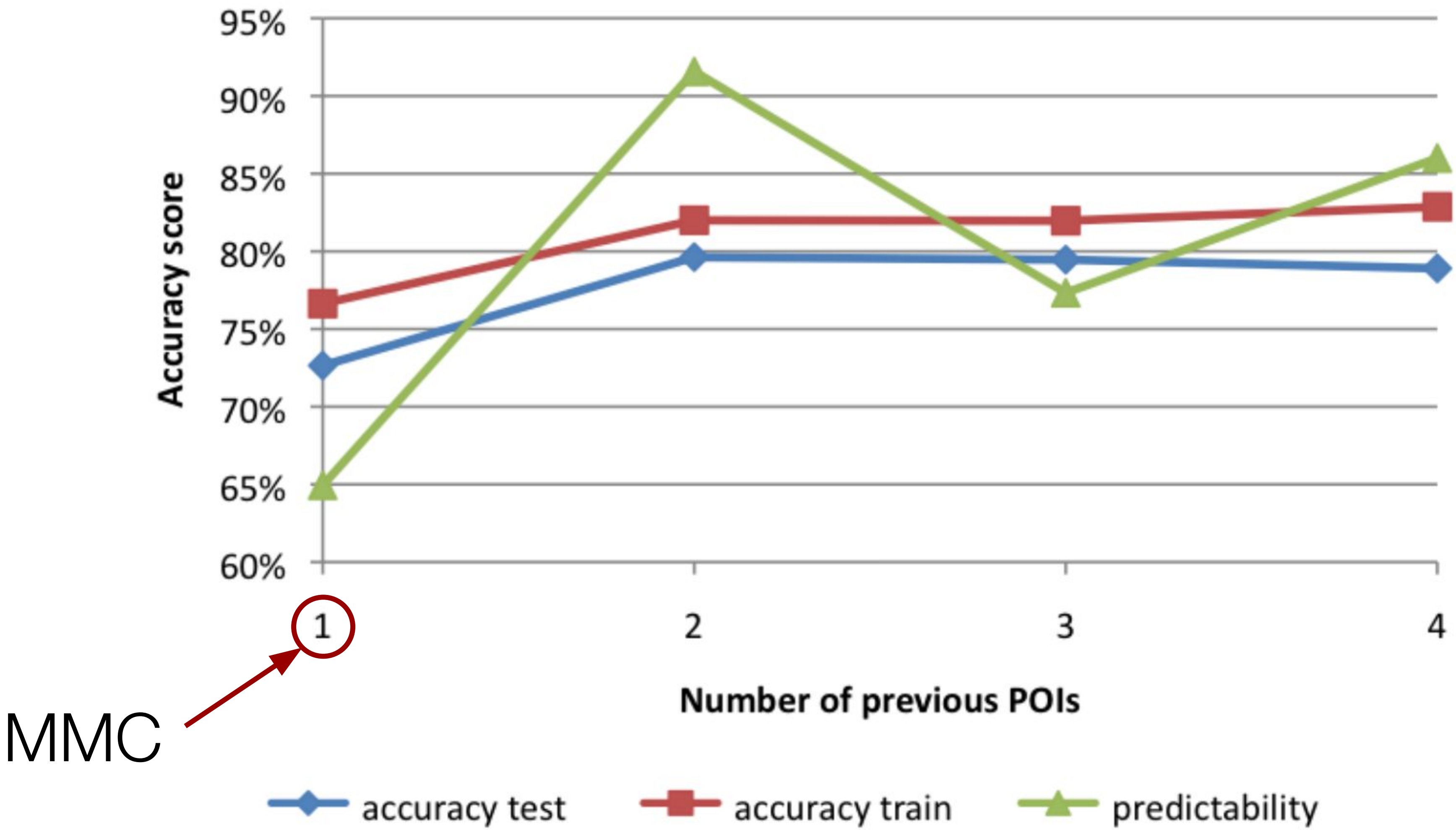
Geolife dataset

Gathered by researchers from Microsoft Asia and it consists of mobility traces collected from April 2007 to October 2011 using GPS-enabled devices, mostly in the area of Shanghai.

Synthetic dataset

Generated out of the first user of the Geolife dataset and we use it mainly as a sanity check to verify the behavior of the prediction algorithm. For instance, the user of this dataset corresponds to the first user of Geolife obtained by duplicating the mobility traces of the first user and then applying a translation in the time domain.

Experimental Results of n-MMC



Other probabilistic approaches

- **Pattern mining approaches**
  Use movement patterns previously extracted by using the *Trajectory Pattern* algorithm.
  Monreale et al., WhereNext: a location predictor on trajectory pattern mining, 2009

- **Social-based approaches**
  embed the social dimension to improve the location prediction (or vice versa)
  Wang et al., Human Mobility, social ties and link prediction, ACM SIGKDD 2012.

A Pattern Mining based approach

Monreale et al., WhereNext: a location predictor on trajectory pattern mining

Idea

Use movement patterns previously extracted by using the *Trajectory Pattern* algorithm.

- Concise representation of behaviors of moving objects as sequences of regions frequently visited with a typical travel time.

The, T-patterns, are combined in a prefix tree called T-pattern Tree, which is then used to predict the next location.

The algorithm

1. **Data Selection**
Selection of a spatial area and a time period, in order to take only the portion of trajectories crossing that area in that time period.

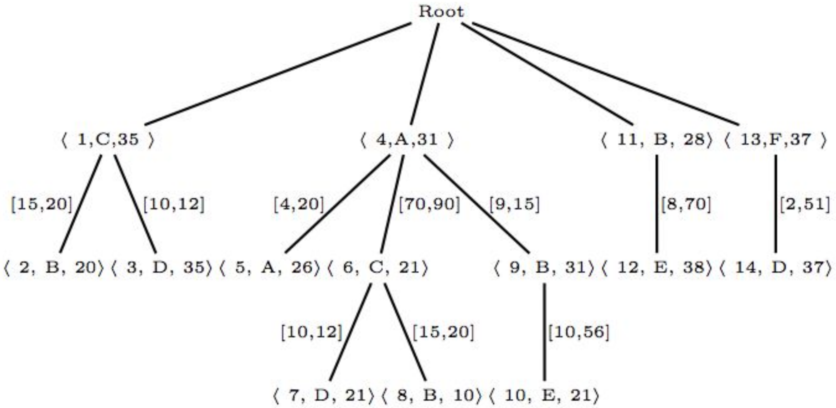
2. **Local Models Extraction**
Extraction of frequent movement patterns (Trajectory Pattern) from the selected trajectories.

Monreale et al., WhereNext: a location predictor on trajectory pattern mining

The algorithm

3. T-Pattern Tree Building
T-patterns, are combined in a prefix tree called Tpattern Tree.

Each common prefix of T-patterns becomes common path on the tree. T



Monreale et al., WhereNext: a location predictor on trajectory pattern mining

The algorithm

4. Prediction

The T-pattern Tree is used to predict the future location of a moving object.

- Find the best path on the tree, namely the best T-pattern that matches the given trajectory.

The algorithm uses a defined distance to find the best matching pattern and to predict the next movement location.

Monreale et al., WhereNext: a location predictor on trajectory pattern mining

PROs and CONs of probabilistic methods

PROs

- High **interpretability**
- **No need big** dataset for training
- **Simple** models

CONs

- **Memory** is **limited**: history does not take into account all movements (e.g. Markov Model)
- **Not state-of-the-art** in next location prediction
- Require a lot of **feature engineering**

Neural Networks for next location prediction

- Neural Networks can be trained to predict the latitude and longitude of the next location.

Artificial Neural Networks Applied to Taxi Destination Prediction (de Brébisson et al.)

Neural Networks for next location prediction

- Neural Networks can be trained to predict the latitude and longitude of the next location.

- Winning model for the

  ECML/PKDD 15 Discovery Challenge:
  Taxi Trajectory Prediction (I)

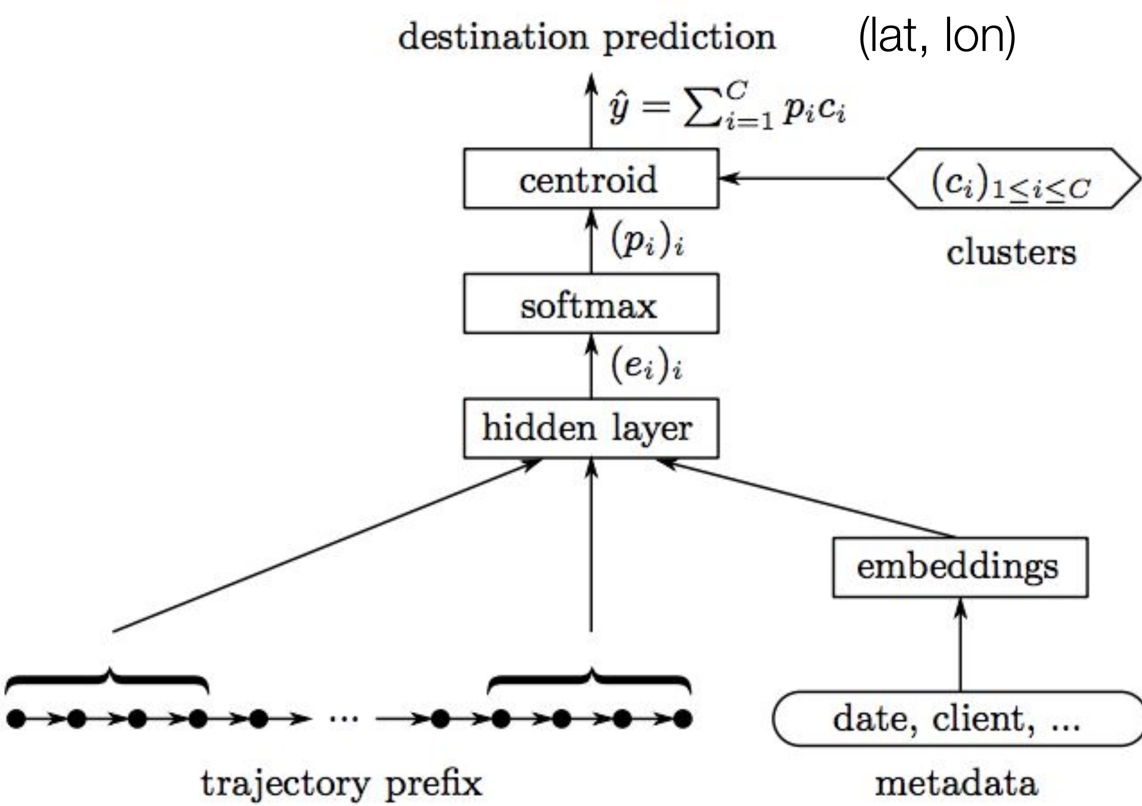
  In this challenge, we ask you to build a predictive framework that is able to infer the final destination of taxi rides in Porto, Portugal based on their (initial) partial trajectories. The output of such a framework must be the final trip's destination (WGS84 coordinates).

  Data available at
  https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i

Artificial Neural Networks Applied to Taxi Destination Prediction (de Brébisson et al.)

Model Architecture

https://github.com/adbrebs/taxi



Cost: distance of the prediction to the true destination

~3,000 **destination clusters** (lat, lon)

Metadata embeddings

Input data:

- Trajectory prefix
- Trip metadata

Metadata	Number of possible values	Embedding size
Client ID	57106	10
Taxi ID	448	10
Stand ID	64	10
Quarter hour of the day	96	10
Day of the week	7	10
Week of the year	52	10

Artificial Neural Networks Applied to Taxi Destination Prediction (de Brébisson et al.)

Results on Taxi Destination Prediction

Cost: distance of the prediction to the true destination

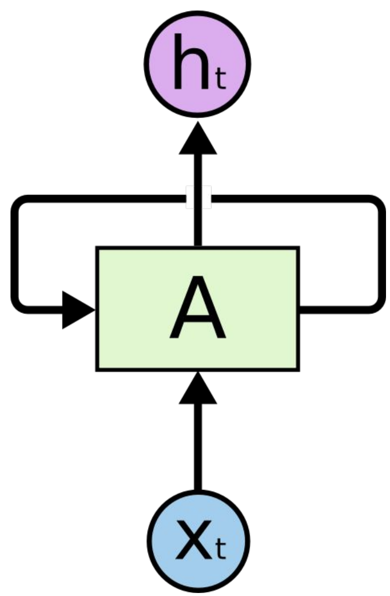
Model	Custom Test	Kaggle Public	Kaggle Private
1 MLP, clustering (winning model)	2.81	**2.39**	**1.87**[8]
2 MLP, direct output	2.97	3.44	3.88
3 MLP, clustering, no embeddings	2.93	2.64	2.17
4 MLP, clustering, embeddings only	4.29	3.52	3.76
5 RNN	3.14	2.49	2.39
6 Bidirectional RNN	3.01	2.65	2.33
7 Bidirectional RNN with window	**2.60**	3.15	2.06
8 Memory network	2.87	2.77	2.20
Second-place team	-	2.36	2.09
Third-place team	-	2.45	2.11
Average competition scores[9]	-	3.02	3.11

- Embeddings and clusters significantly improve the model.
- Testing on a larger dataset suggests RNNs perform better.

Artificial Neural Networks Applied to Taxi Destination Prediction (de Brébisson et al.)

Recurrent Neural Network

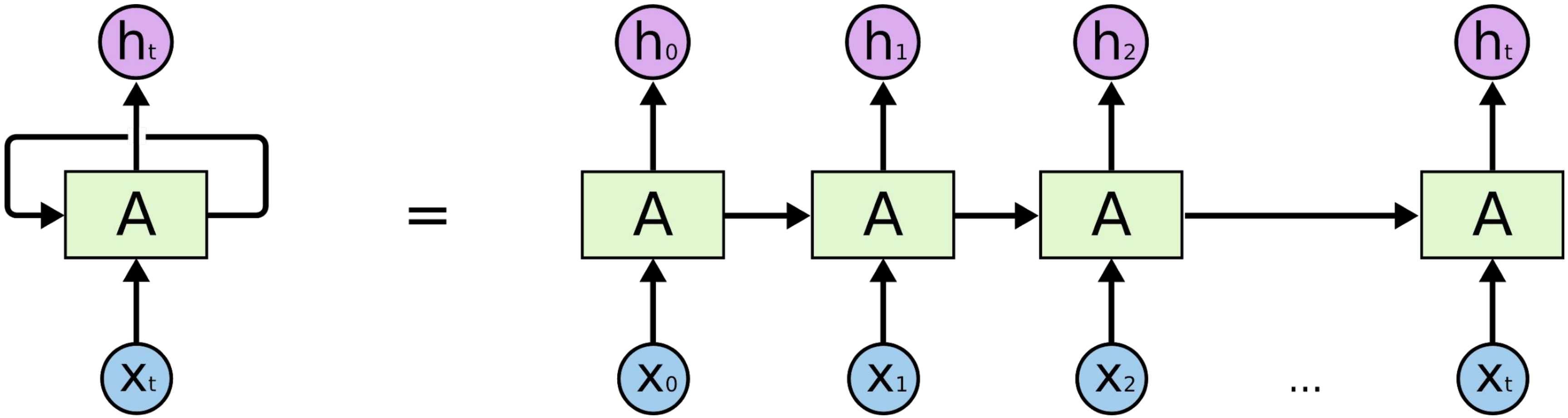
- Neural network with connections between nodes in hidden layers



http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Recurrent Neural Network

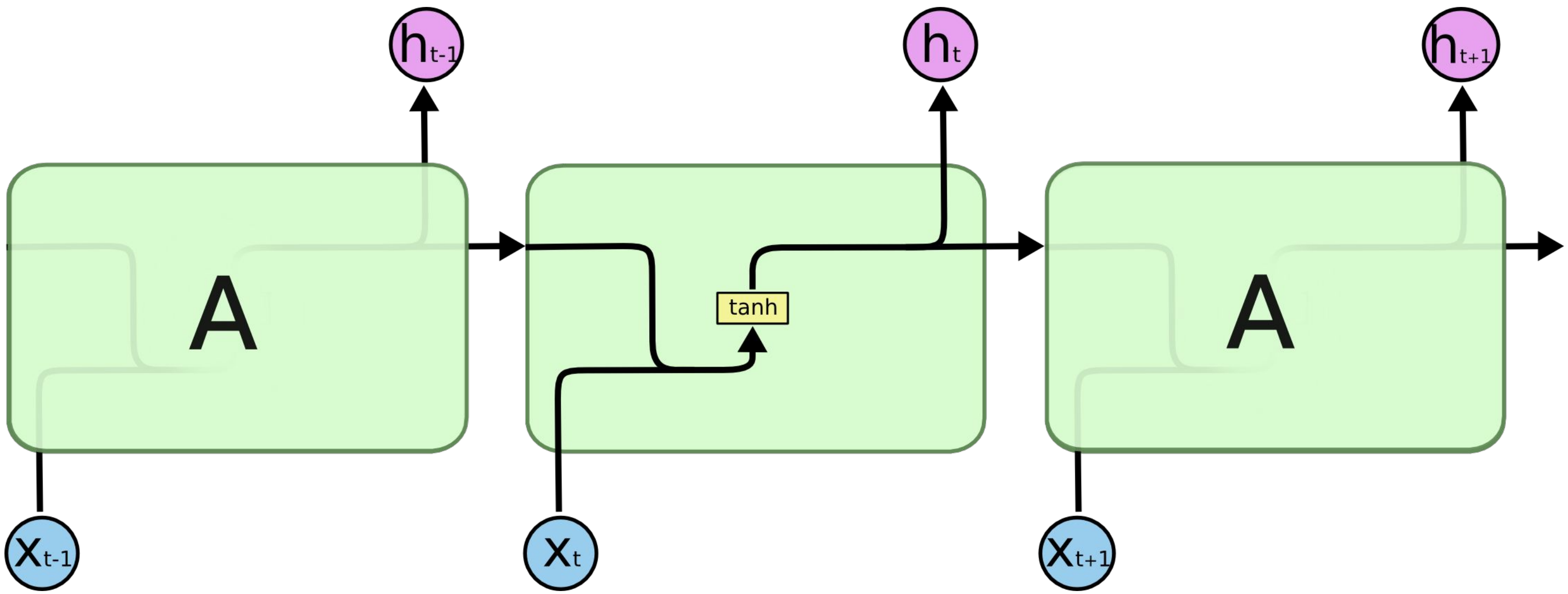
- Neural network with connections between nodes in hidden layers
- Used to learn sequences of data (time series, text, …), hence naturally appropriate to model trajectories



http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Recurrent Neural Network

Simple RNNs are not able to learn long-term time dependencies because of the vanishing gradient problem

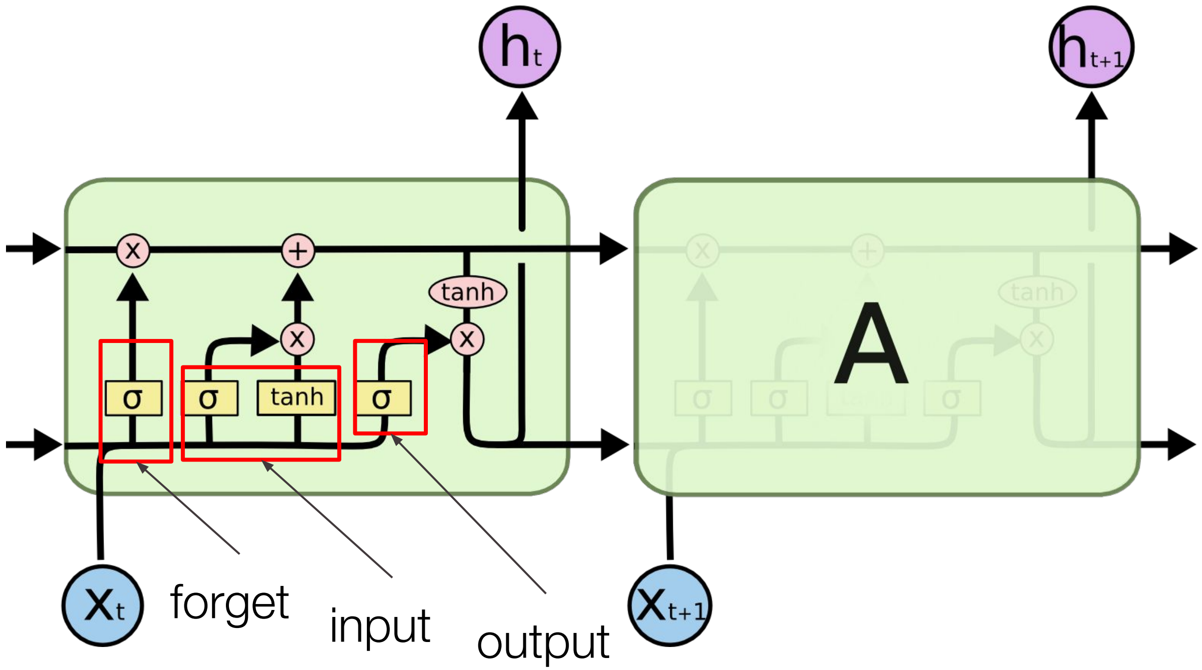


http://colah.github.io/posts/2015-08-Understanding-LSTMs/

http://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM and GRU units

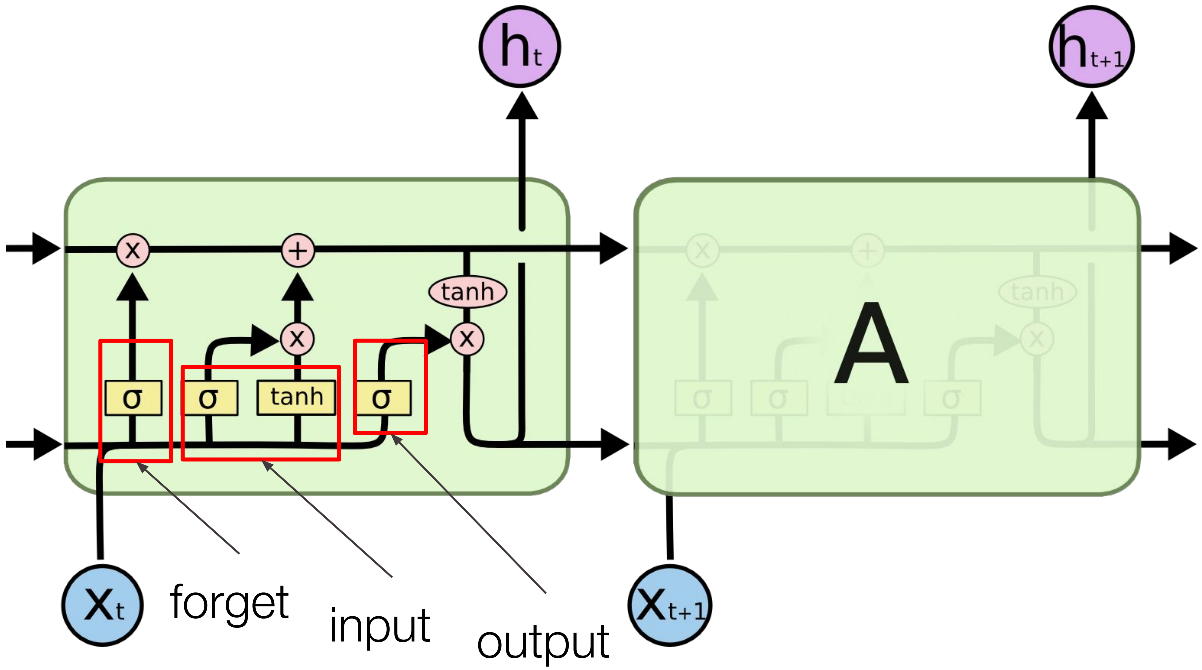
LSTM
Long
Short-Term
Memory



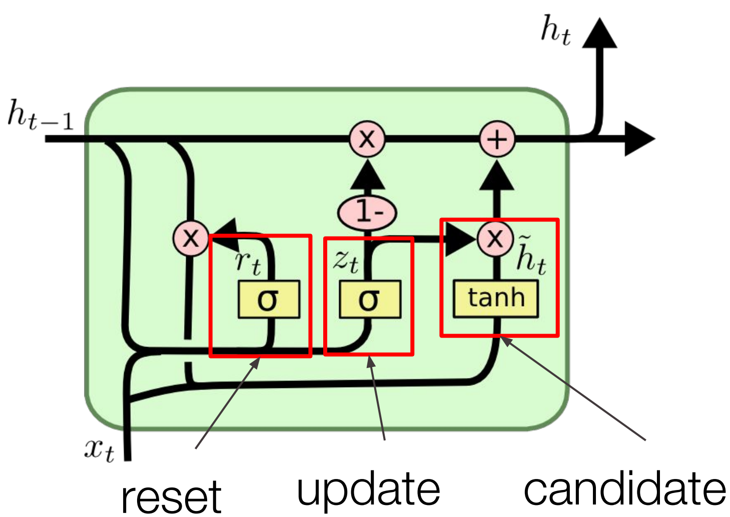
http://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM and GRU units

LSTM
Long
Short-Term
Memory



GRU
Gated
Recurrent
Unit



$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Human Mobility with Attentional Recurrent Networks

https://github.com/vonfeng/DeepMove

Goal

Given the past trajectory sampled every 30 minutes, predict the location in the next 30 minutes.

Challenges

1. complex sequential transition regularities
   (e.g. the probability of moving from home to office is high in workday mornings but low in weekend mornings)
2. multi-level periodicity
   (daily routines, weekend leisure, yearly festivals, and even other personal periodic activities)

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Human Mobility with Attentional Recurrent Networks

https://github.com/vonfeng/DeepMove

Goal

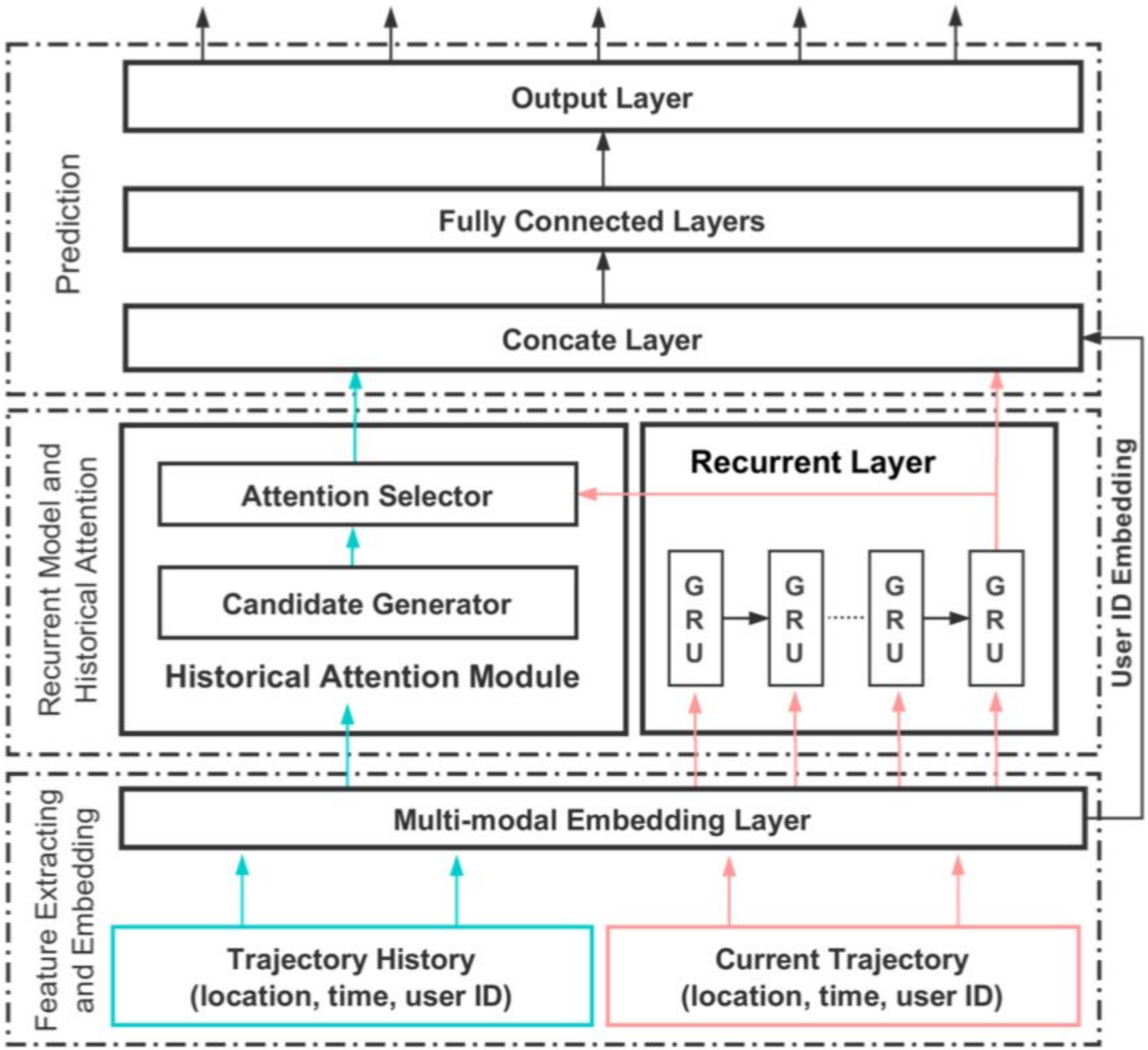
Given the past trajectory sampled every 30 minutes, predict the location in the next 30 minutes.

Challenges

1. complex sequential transition regularities → **RNN**
   (e.g. the probability of moving from home to office is high in workday mornings but low in weekend mornings)
2. multi-level periodicity → **Attention**
   (daily routines, weekend leisure, yearly festivals, and even other personal periodic activities)

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

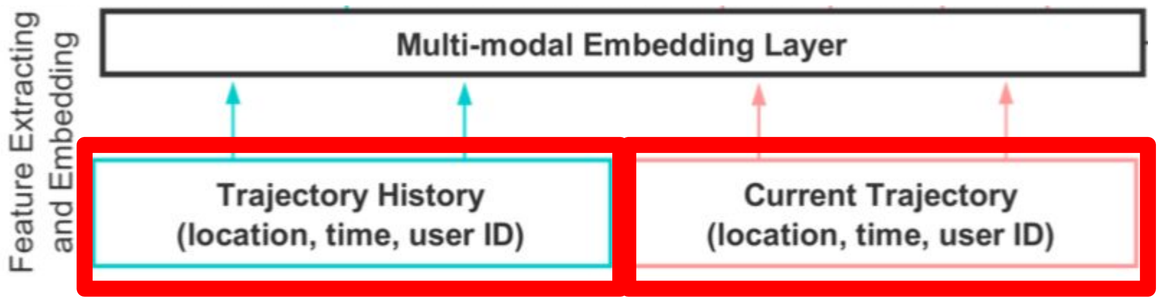
Attention



RNN

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Multi-modal Embedding Layer

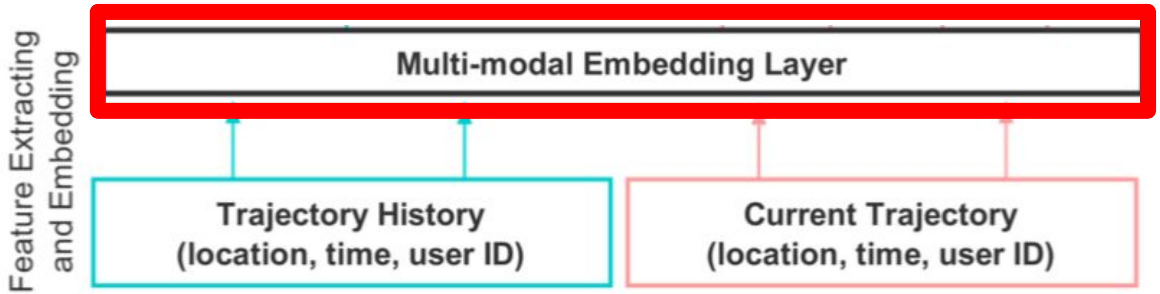


Trajectories are partitioned into two parts:

- current trajectory (e.g. mobility in the last 24 hours)
- historical trajectory (previous mobility)

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

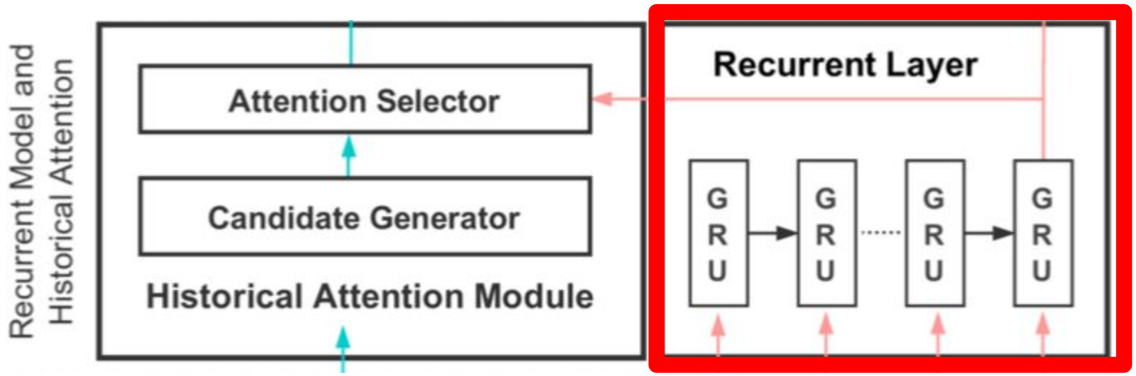
Multi-modal Embedding Layer



- The numbered features of the trajectory are translated into one-hot vectors and inputted to the multi-modal embedding module.
- The Multi-modal embedding layer jointly embed the vector into dense representations.

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

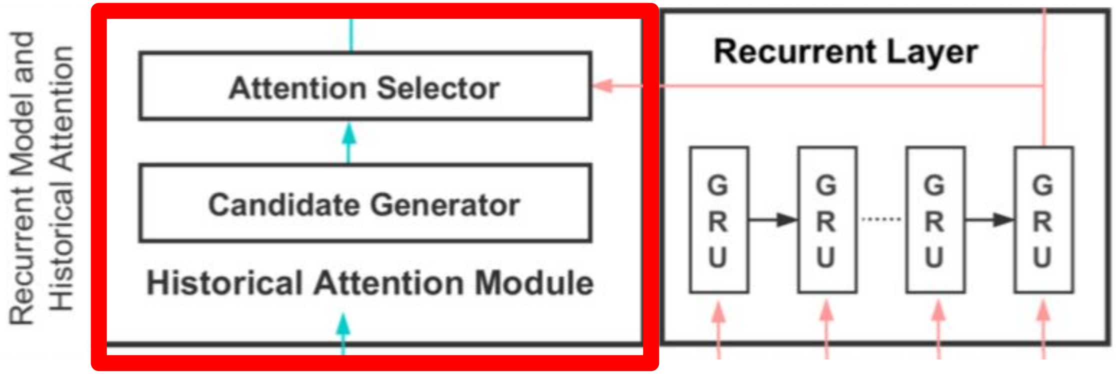
Recurrent Module



The recurrent module aims to capture the complicated sequential information or long-range dependencies contained in the current trajectory
(e.g. *office - lunch - office* or *office - gym - home*)

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Historical Attention Module



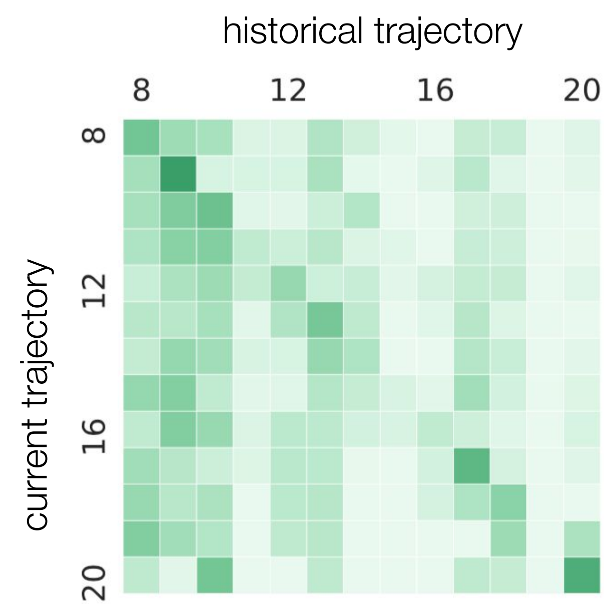
The attention module selects from the historical trajectory the most related historical records to the current mobility status.

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Historical Attention Module

$$c_t = \sum \alpha_i s_i,$$
$$\alpha_i = \sigma(f(h_t, s_i)),$$
$$f(h_t, s) = tanh(h_t W s),$$

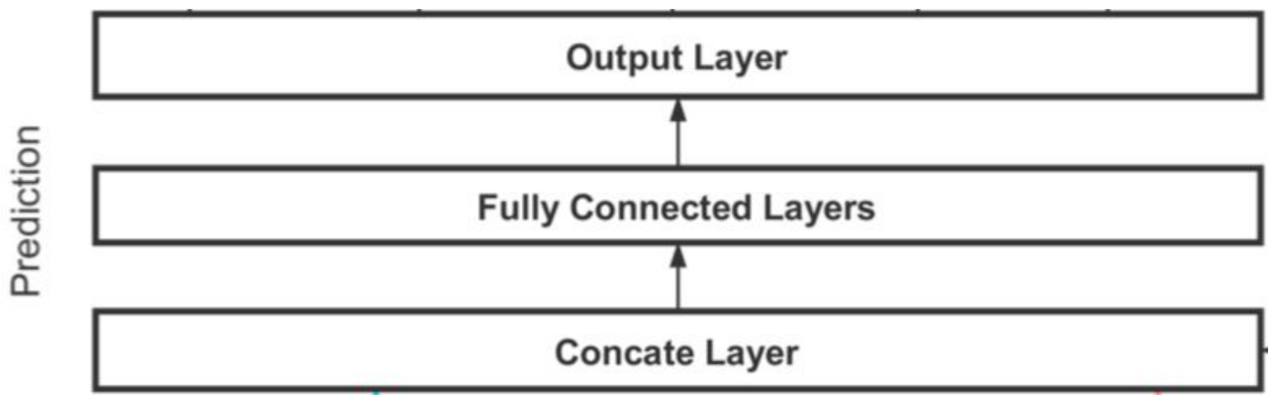
- s represents the historical features
- W is the learnable parameters
- h is the query vector
- f represents the score function
- σ is the soft-max function and
- c is the context output representing the periodicity related to the current mobility status.



Normalized attention weights, α_i

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Prediction Module



The Prediction Layer combines the context from different modules to complete the prediction task.

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Datasets for the experiments

Foursquare Check-in Data

Collected from Foursquare API from Feb. 2010 to Jan. 2011 in New York. Every record in the data consists of user ID, timestamp, GPS location and POI ID

Mobile Application Data

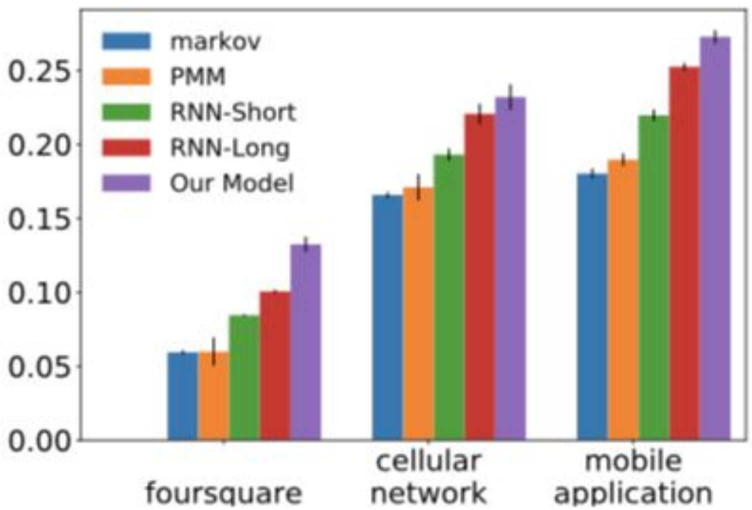
Location of users whenever they request the localization service in the applications. The data is collected from 17 Nov. 2016 to 31 Oct. 2016

Call Detail Records Data

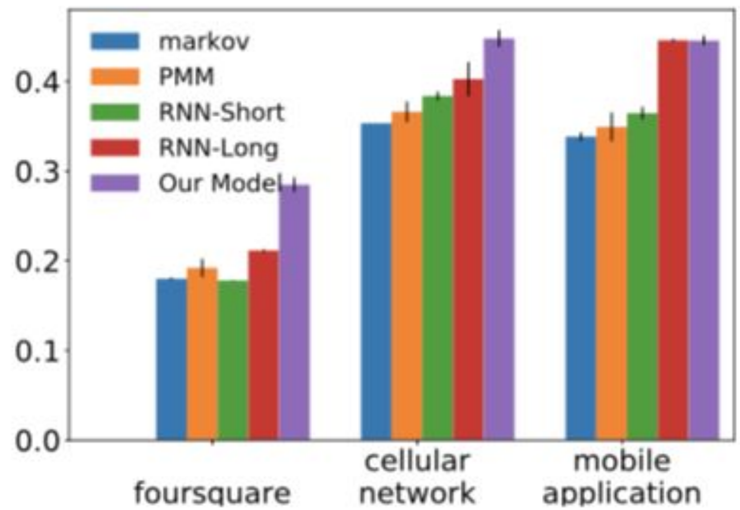
It records the spatiotemporal information of users when they access the cellular networks The data is collected from 1 Jan. 2016 to 31 Jan. 2016.

DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Results with Historical Attention Module



(a) top-1 prediction accuracy

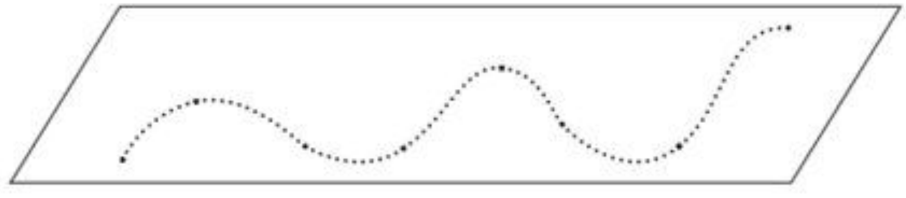


(b) top-5 prediction accuracy

- with attention, accuracy increases by ~30% for Foursquare and by ~6% for mobile phone data

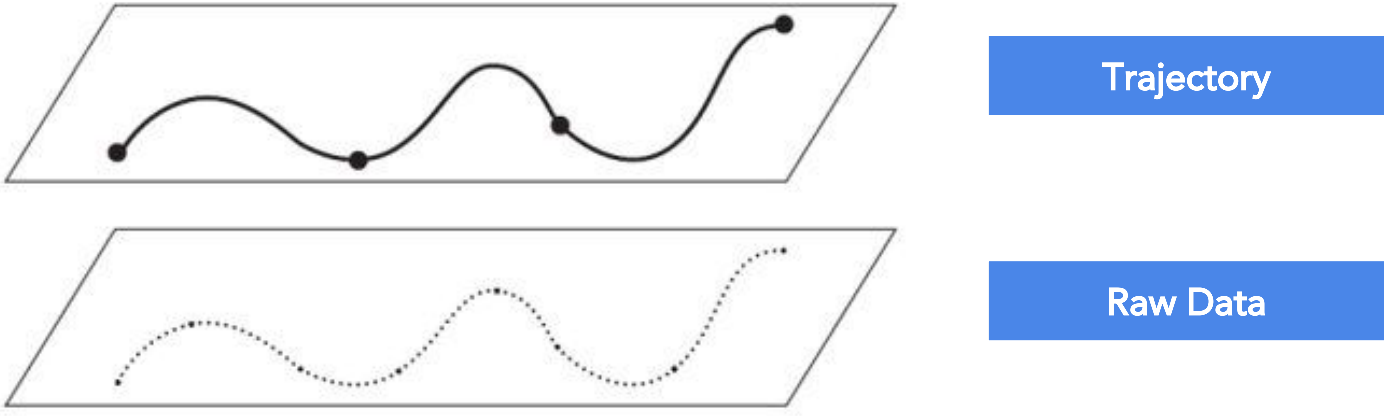
DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Next Location Prediction with Semantic Trajectories

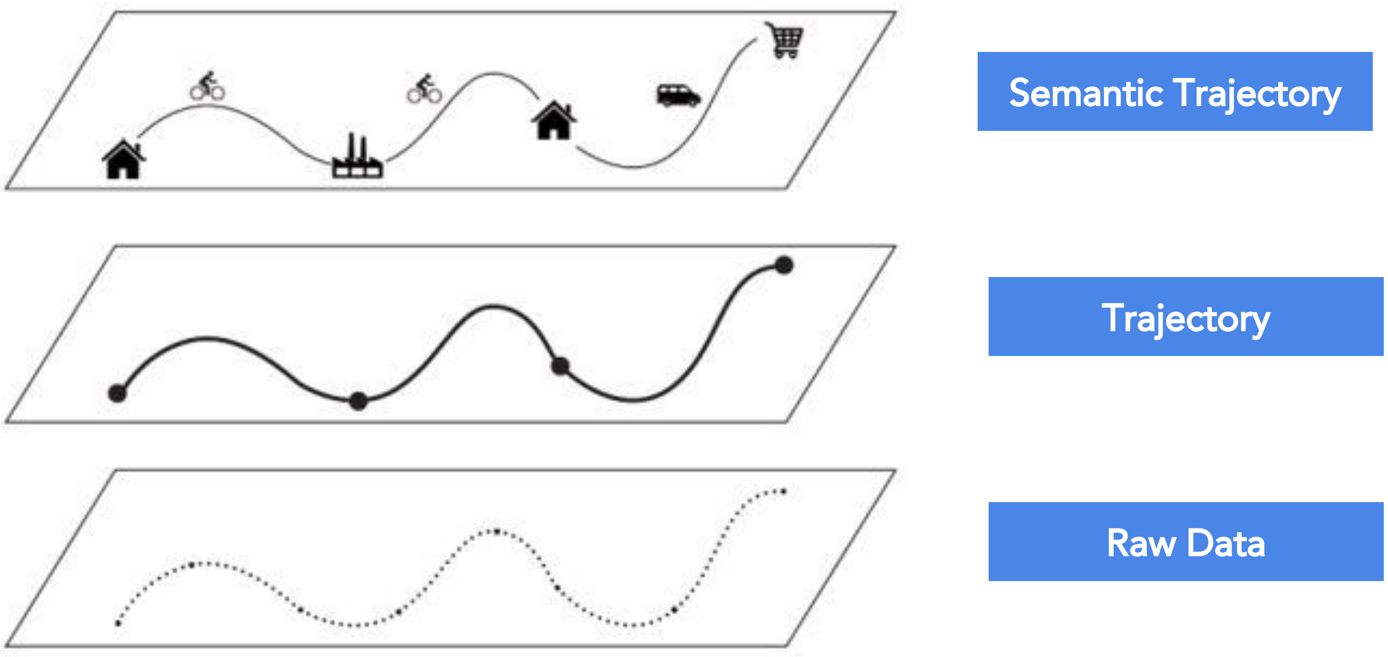


Raw Data

Next Location Prediction with Semantic Trajectories



Next Location Prediction with Semantic Trajectories



A semantic trajectory includes timestamp, location, and a **short text** describing the user's activity

Build a semantic Trajectory

Explicit location in the text

Build a semantic Trajectory

Implicit Location in the text

Location Sequence in Semantic Trajectories

The **location sequence** for each user $u_i \in \mathscr{U}$ is

$$S(u_i) = \{r_1(u_i), r_2(u_i), \ldots, r_S(u_i)\}$$

Where $r_k(u_i) \in S(u_i)$ is a tuple $(\boxed{t_k}, \boxed{l_k}, \boxed{c_k})$

text message describing the activity of the user

timestamp

location

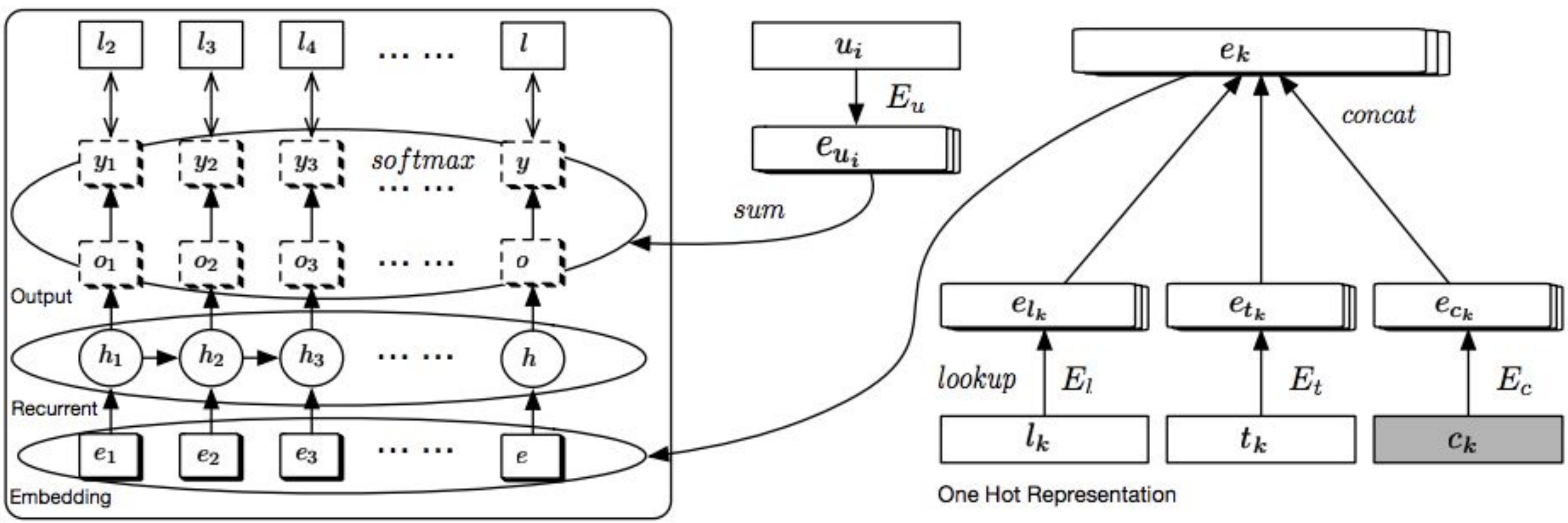
Goal

- Predict the ground-truth location l_K based on the preceding sequence $\{r_1(u_i), r_2(u_i), \ldots, r_{K-1}(u_i)\}$

SERM: A Recurrent Model for Next Location Prediction in Semantic Trajectories (Yao et al.)

SERM: A Recurrent Model for Next Location Prediction in Semantic Trajectories

https://github.com/yaodi833/serm



SERM: A Recurrent Model for Next Location Prediction in Semantic Trajectories (Yao et al.)

SERM Results on Twitter and Foursquare

accuracy distance

Data	Method	HR@1	HR@5	HR@10	HR@20	δ_d/m
NY	NL	0.1630	0.2455	0.2998	0.4386	2903
	MF	0.1690	0.4326	0.5013	0.5358	1963
	HMM	0.1763	0.4298	0.5251	0.5518	1952
	ST-RNN	0.1942	0.4421	0.5381	0.6053	1602
	SERM*	0.2181	0.4398	0.5401	0.6107	1563
	SERM	**0.2535**	**0.4507**	**0.5433**	**0.6237**	**1457**
LA	NL	0.3745	0.4516	0.4704	0.4911	6061
	MF	0.3646	0.5810	0.6354	0.6877	2647
	HMM	0.3921	0.5935	0.6331	0.6732	2521
	ST-RNN	0.4311	0.6013	0.6521	0.6980	2384
	SERM*	0.4452	0.6147	0.6590	0.6973	2377
	SERM	**0.4625**	**0.6265**	**0.6670**	**0.7026**	**2177**

0.3 million Foursquare **check-ins** from 2011-01 to 2012-01

1.4 million **tweets** from 2014-08 to 2014-11

SERM* is a version of SERM without textual information

SERM: A Recurrent Model for Next Location Prediction in Semantic Trajectories (Yao et al.)

PROs and CONs

PROs

- High **accuracy**
- **Flexibility** of the model
- **No feature engineering**

CONs

- Low **interpretability**
- Need **big** training **dataset**
- **Transfer** learning is not possible
- **Complex** model

Human Mobility and Machine Learning

Statistical mobility properties of group of people can be used for:

Movements Prediction

- Next Location prediction
  [3] DeepMove: Predicting Human Mobility with Attentional Recurrent Networks (Feng et al.)

Human Behaviour Prediction

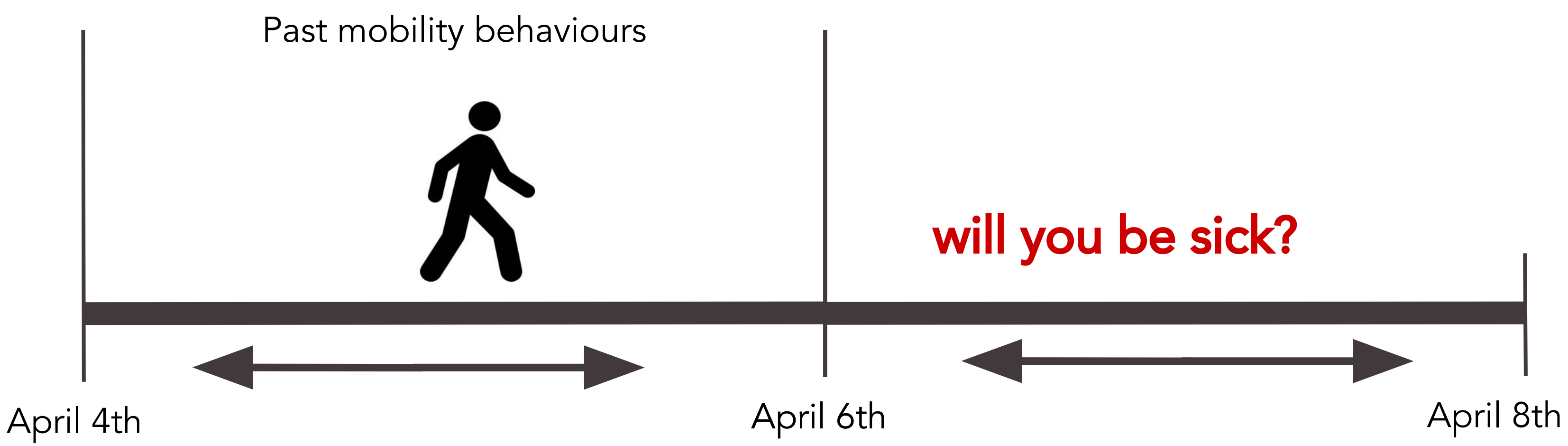
- Predicting future health condition
  [2] Are you getting sick? Predicting influenza-like symptoms using human mobility behaviors *(Barlacchi et al.)*

Urban Planning

- Discover functional regions
  [1] Discovering regions of different functions in a city using human mobility and POIs (Yuan et al.)

Human mobility to predict flu-like symptoms

Individual mobility behaviours is used to predict the future presence of flu-like symptoms (e.g. cold, fever and cough).



Are you getting sick? Predicting influenza-like symptoms using human mobility behaviors (Barlacchi et al.)

Human mobility to predict flu-like symptoms

We characterize the mobility of an individual by computing a set of features based on its movements:

Distance traveled related features

- Total traveled distance, total displacements

Movements related features

- Radius of gyration, number of geo-locations points

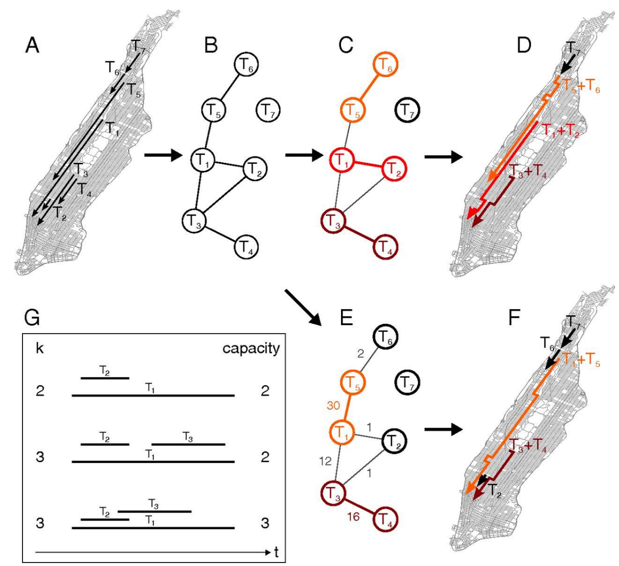
Visited places related features

- Diversity of visited places, number of unique visited places

Are you getting sick? Predicting influenza-like symptoms using human mobility behaviors (Barlacchi et al.)

Other applications

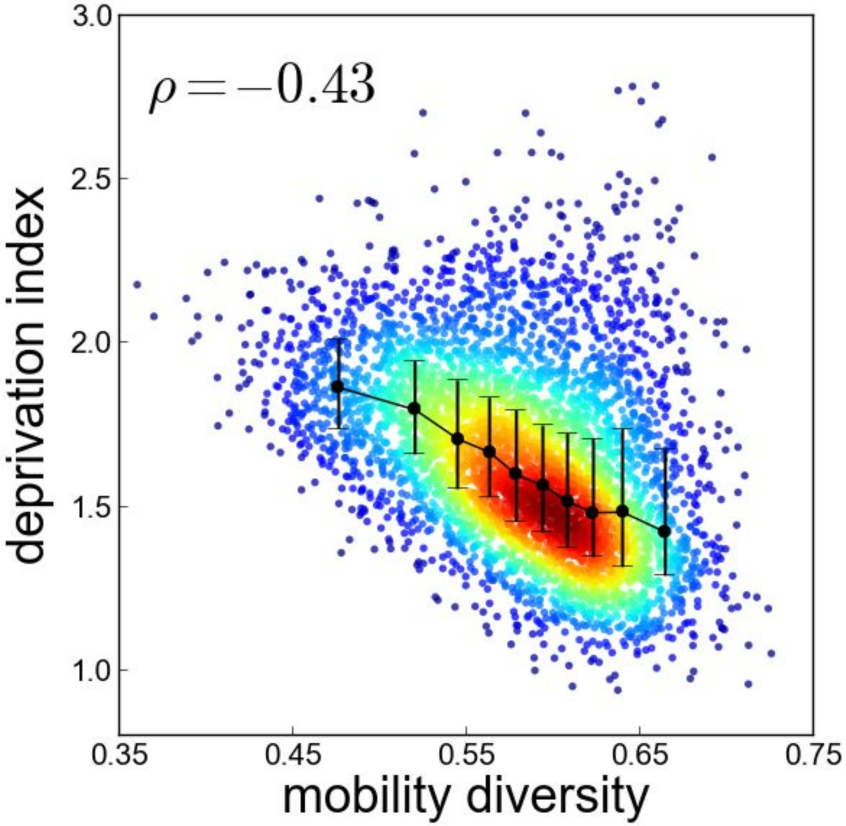
- **Urban Planning**
  Inferring the functionality of area in a city by using mobility patterns extracted from taxi movements.
  [1] Discovering Regions of Different Functions in a City Using Human Mobility and POIs (Yaun et al.)



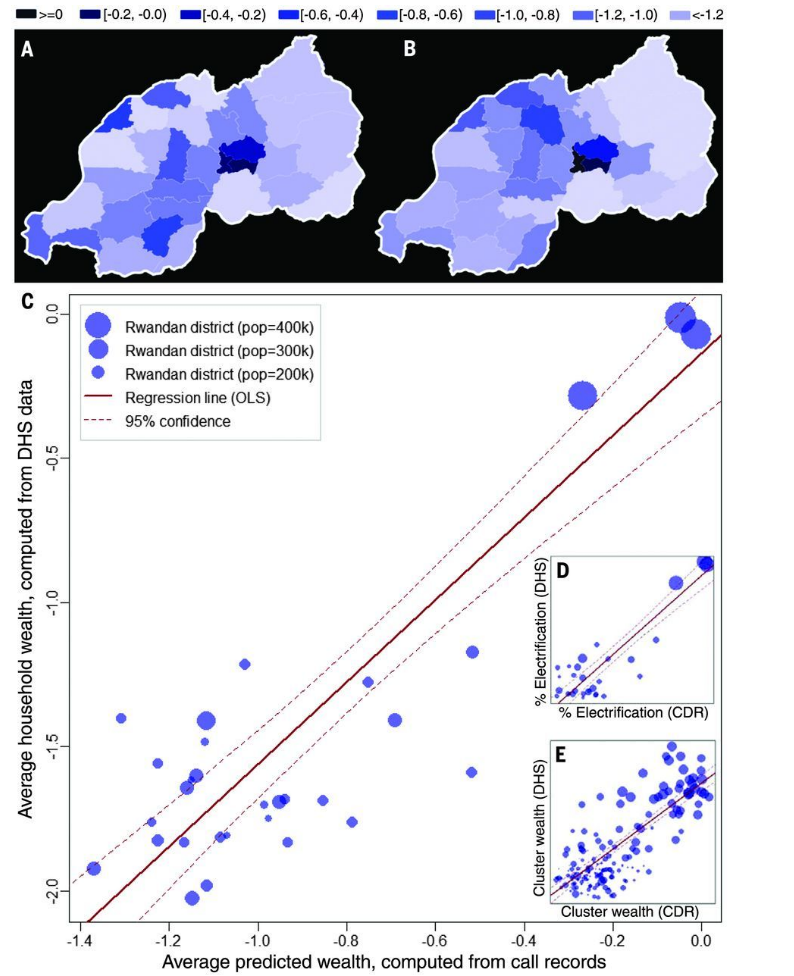
- **Vehicle sharing**
  Given a collection of trips (origin, destination and start time), determine the minimum number of vehicles needed to serve all the trips without incurring any delay to the passengers.
  [2] Addressing the minimum fleet problem in on-demand urban mobility (Vazifeh et al.)
  [3] Quantifying the benefits of vehicle pooling with shareability networks (Santi et al.)

Other Applications

Well-being prediction



[3] An analytical framework to nowcast well-being with mobile phone data (Pappalardo et al.)



[4] Predicting poverty and wealth from mobile phone metadata (Blumenstock et al.)

Other applications

- **Credit scoring**
  Mobility features (e.g., radius of gyration) are combined with other behavioural features to estimate the user's financial risk from mobile phone usage data.

  [5] MobiScore: Towards Universal Credit Scoring from Mobile Phone Data (San Predro et al)

- **Crime Prediction**

Special Issue on Individual and Collective Human Mobility: Description, Modelling, Prediction

EPJ Data Science

https://www.springeropen.com/collections/HumanMobility

Lead Guest Editor

Filippo Simini
University of Bristol, UK
f.simini@bristol.ac.uk

Co-Guest Editors

Gourab Ghoshal
University of Rochester, USA
gghoshal@pas.rochester.edu

Luca Pappalardo
University of Pisa, Italy
lpappalardo@di.unipi.it

Michael Szell
Hungarian Acad. of Sci., Hungary
michael.szell@gmail.com

Philipp Hövel
Institute of Theoretical Physics of TU Berlin
phoevel@physik.tu-berlin.de

- Kadar et al., *Mining large-scale human mobility data for long-term crime prediction*
- Conrad et al., *Human mobility and innovation spreading in ancient times: a stochastic agent-based simulation approach*
- Najem et al., *Debye–Hückel theory for refugees' migration*
- Steentoft et al., *The canary in the city: indicator groups as predictors of local rent increases*
- Scherrer et al., *Travelers or locals? Identifying meaningful sub-populations from human movement data in the absence of ground truth*
- Cuttone et al., *Understanding predictability and exploration in human mobility*
- D'Silva et al., *Predicting the temporal activity patterns of new venues*

Further reading

- Karamshuk et al., *Human mobility models for opportunistic networks*, IEEE Communications Magazine, 2011.
- Blondel et al., *A survey of results on mobile phone datasets analysis*, EPJ DS, 2015.
- Zheng, *Trajectory Data Mining: an overview*, ACM TIST, 2015.
- Lenormand et al., *Systematic comparison of trip distribution laws and models.* Journal of Transport Geography, 2016.
- Alessandretti et al., *Multi-scale spatio-temporal analysis of human mobility*, PLoS One, 2017.
- Barbosa et al., *Human mobility: Models and applications*. Physics Reports, 2018.
- Wu et al., *Location prediction on trajectory data: A review*, Big Data Mining and Analytics, 2018.

Human Mobility Science:

Data, Measures, Generative models and Predictive algorithms

https://humanmobility-tutorial.github.io/

Thanks!

Filippo Simini
@ipif

Gianni Barlacchi
@giannibarlacchi

Luca Pappalardo
@lucpappalard