Social and Technological Network Analysis

Lecture 11: Spatial and Social Network Analysis

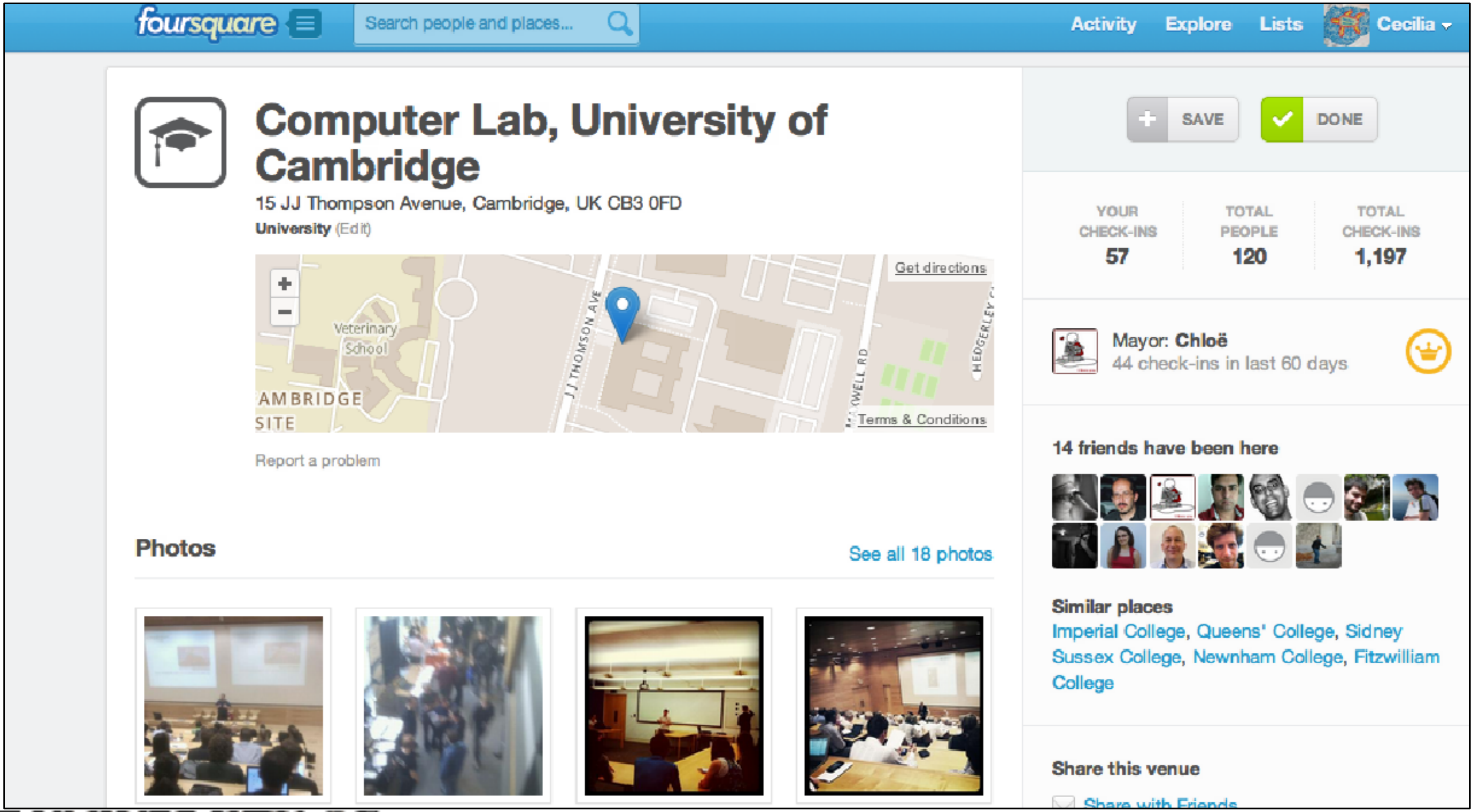
Dr. Salvatore Scellato
Google

UNIVERSITY OF CAMBRIDGE

In This Lecture

- In this lecture we will study spatial networks and geo-social networks through examples from our work.

Places and Friends



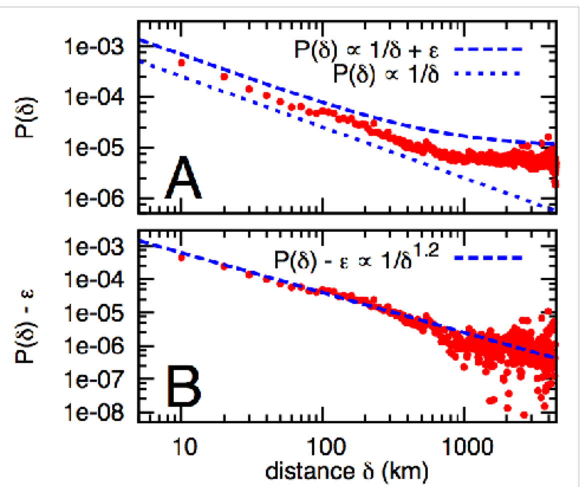
UNIVERSITY OF CAMBRIDGE

Geo-social Network Analysis can lead to insights into social behaviour

UNIVERSITY OF CAMBRIDGE

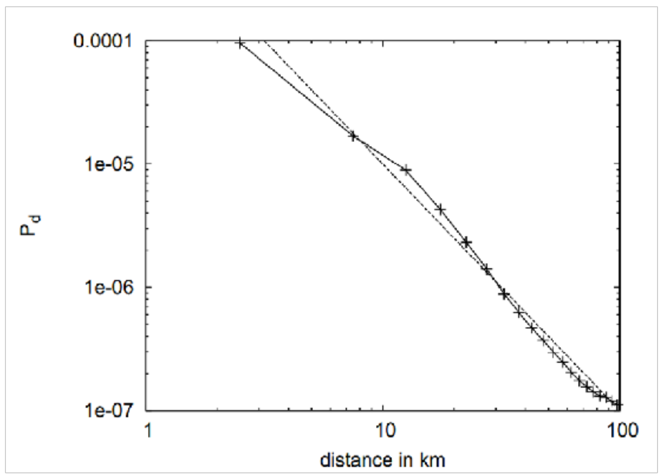
Effect of Distance on Social Connections

LiveJournal (2005)



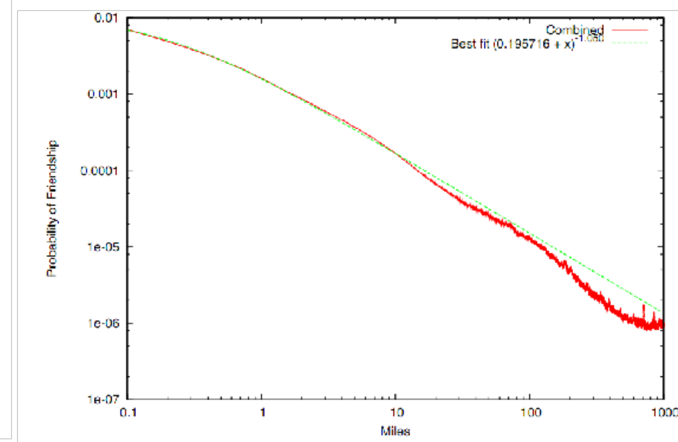
$$P(d) \propto d^{-1} + \epsilon$$

Mobile phones (2008)



$$P(d) \propto d^{-2}$$

Facebook (2010)



$$P(d) \propto d^{-1}$$

UNIVERSITY OF CAMBRIDGE

Effect of geography over social link formation

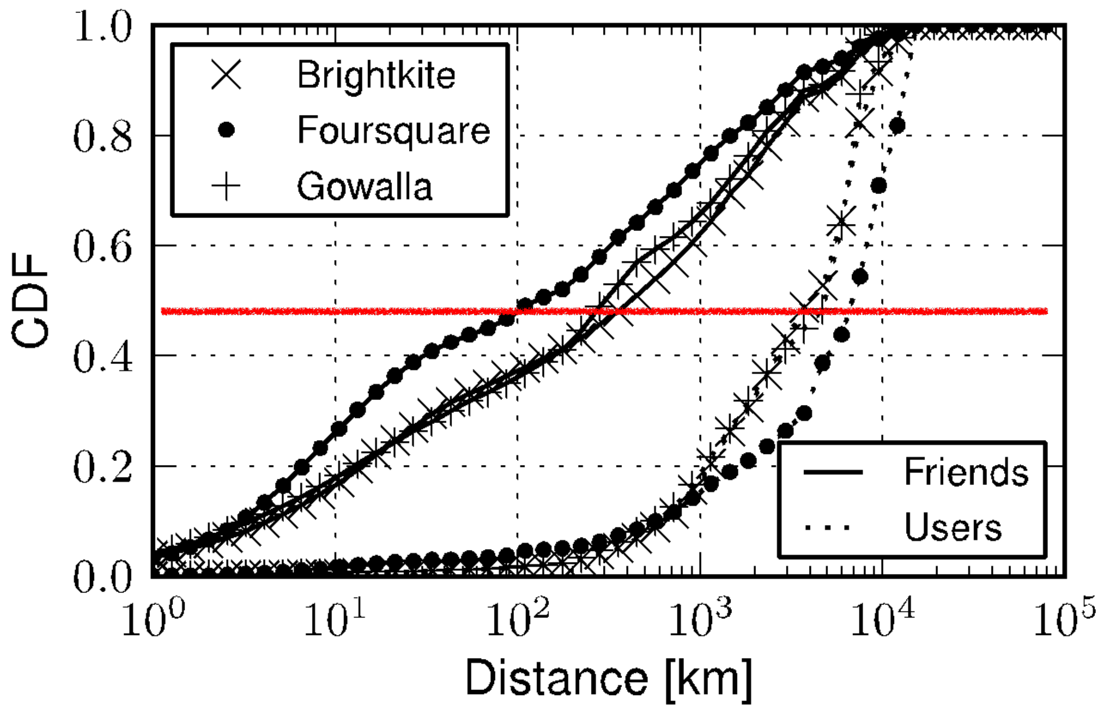
What do we see in Geo-social Networks?

- We have acquired data about the socio-spatial network of **3 real-world location-based services**
- We design **two randomized models** of a socio-spatial network to better understand which factors shape the real networks.
- We study how individual users create their **social links** and their **social triangles** over space.

UNIVERSITY OF CAMBRIDGE

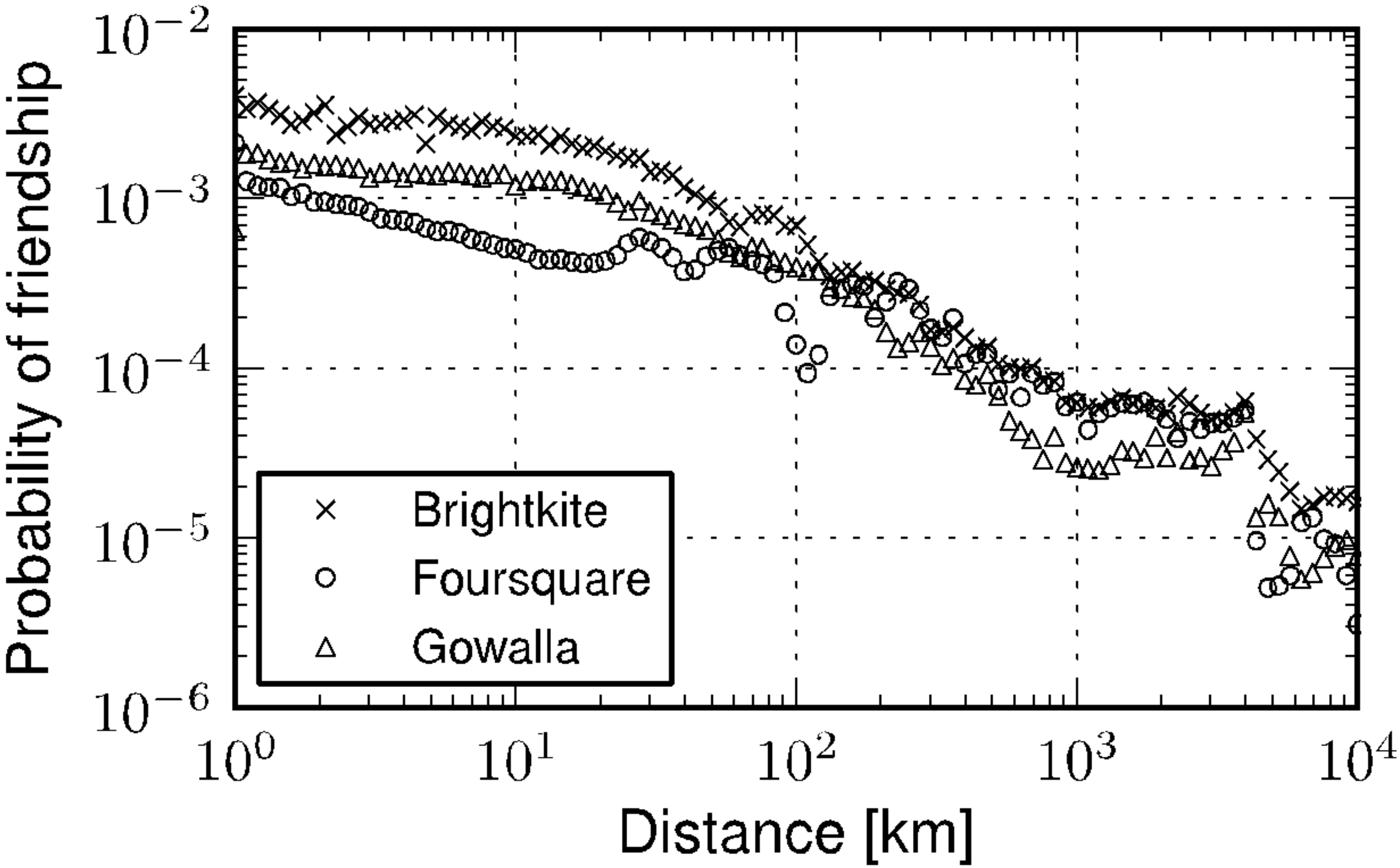
Distance between Friends

Friends tend to be much closer than random users: about 50% of social links span less than 100 km, while about 50% of users are more than 4,000 km apart.



UNIVERSITY OF CAMBRIDGE

Probability of Friendship vs Distance



Network Randomization

	Description	Social properties	Spatial properties
Original data	No modification.	✔	✔
Geo model	Fix node locations and reassign all links according to probability P(d).	✘	✔
Social model	Fix links and shuffle all node locations.	✔	✘

UNIVERSITY OF CAMBRIDGE

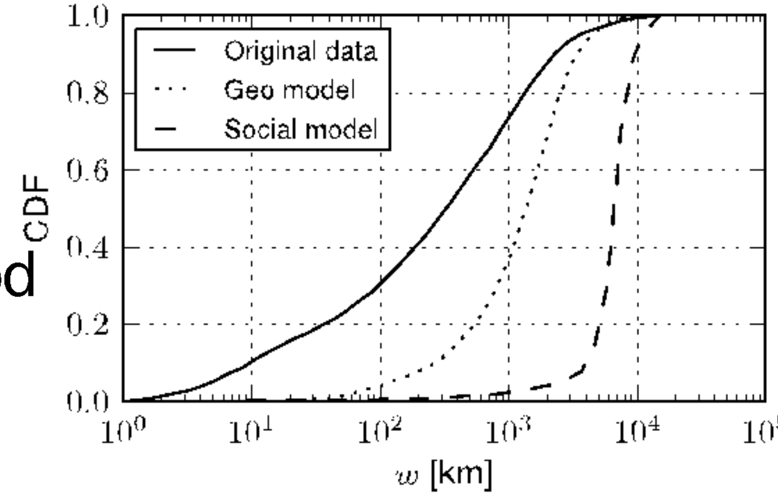
Users have heterogeneous friend distances

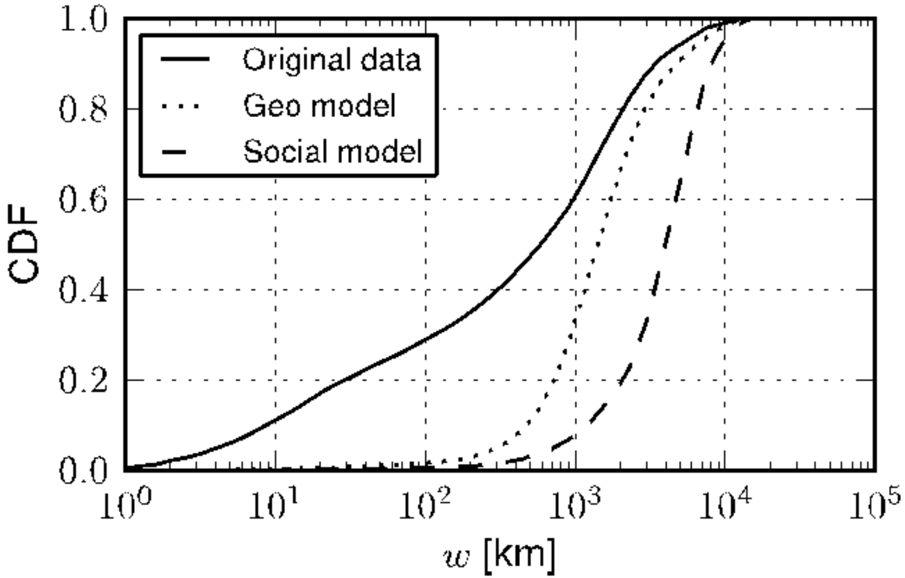
Link length

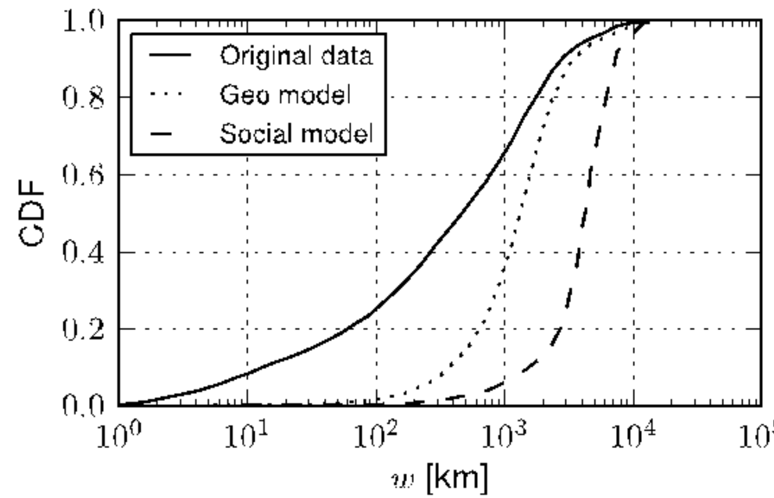
$$w_i = \frac{1}{k_i} \sum_{j \in \Gamma_i} l_{ij}$$

Node degree

Node neighborhood

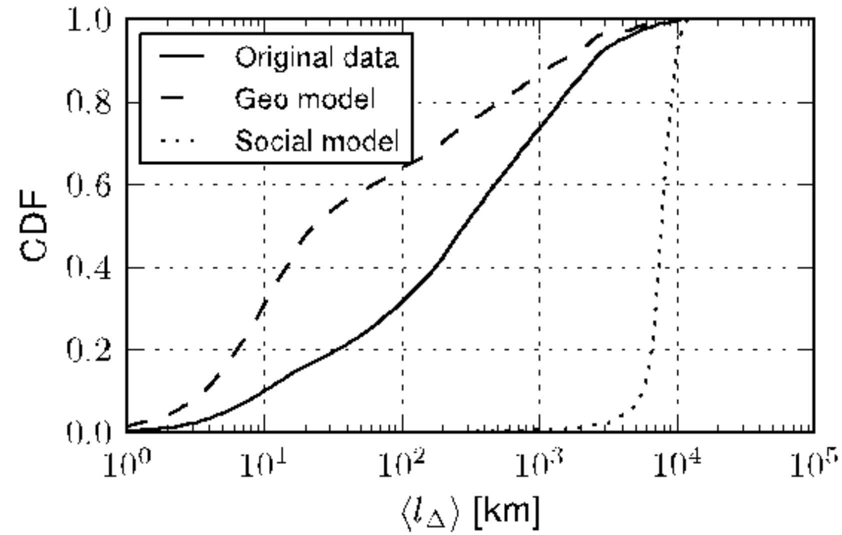


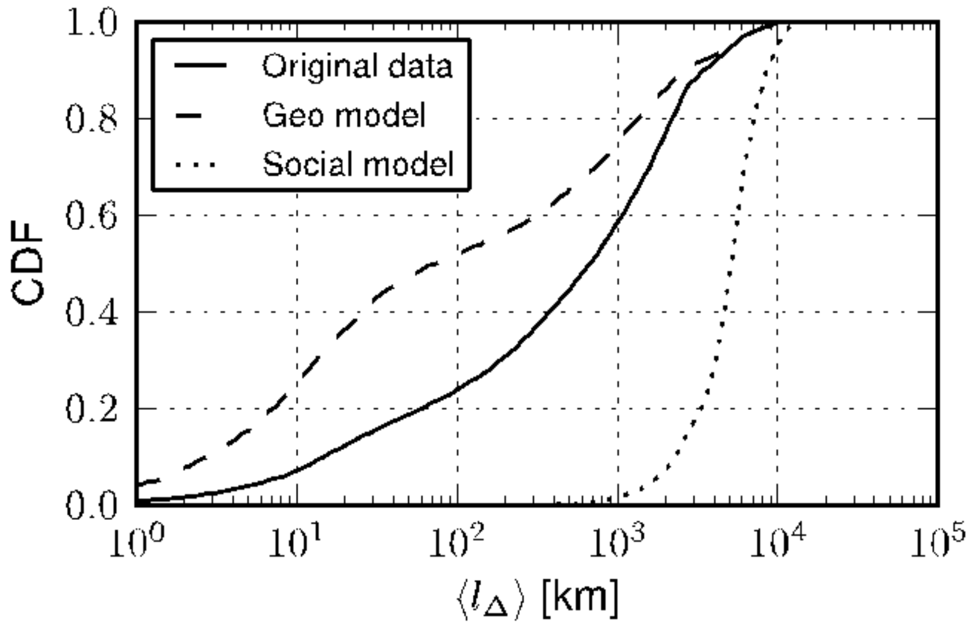


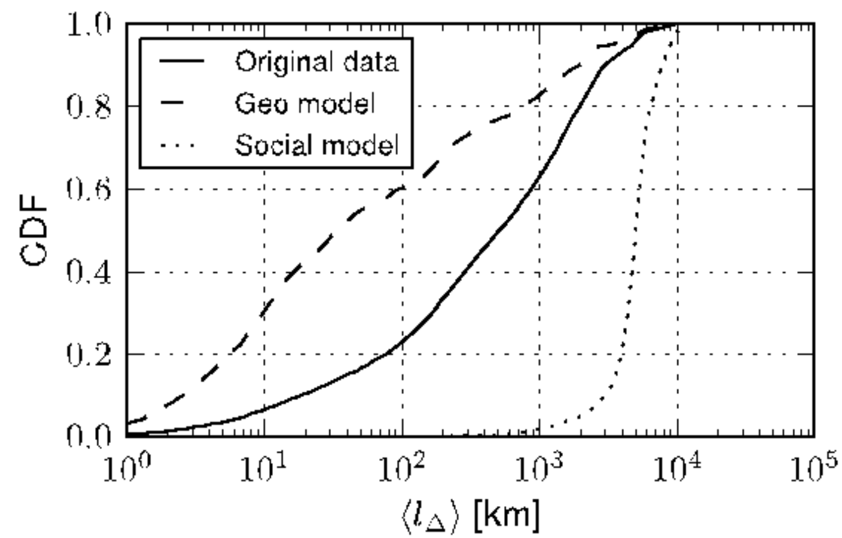


Average triangle geographic length

<lΔ> is the average length of the triangles of a user

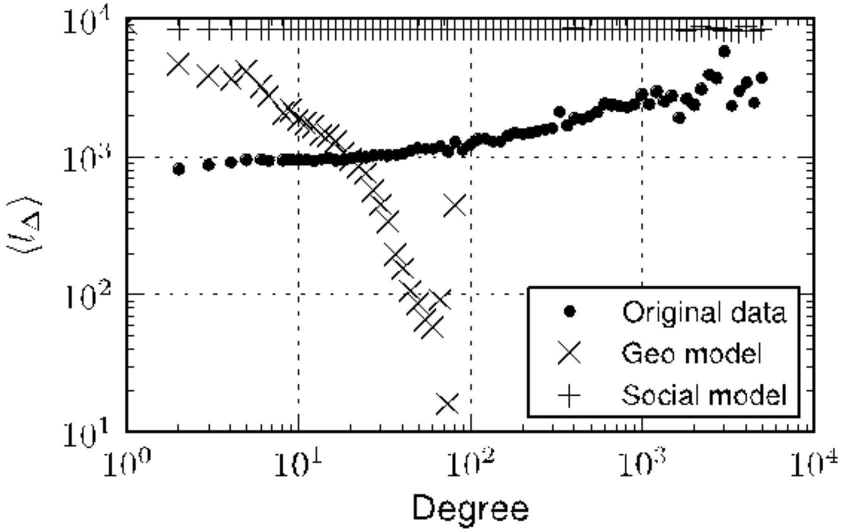


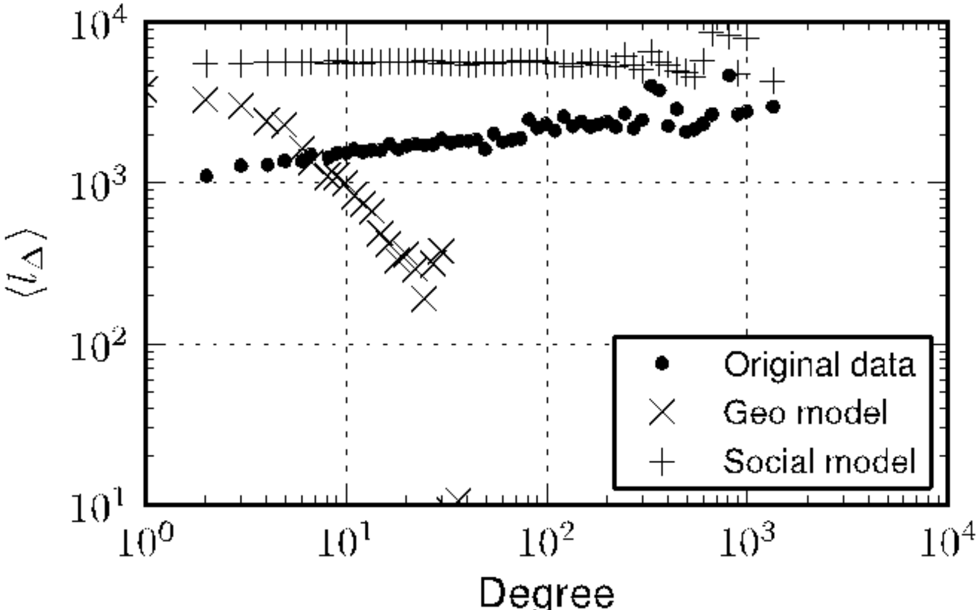


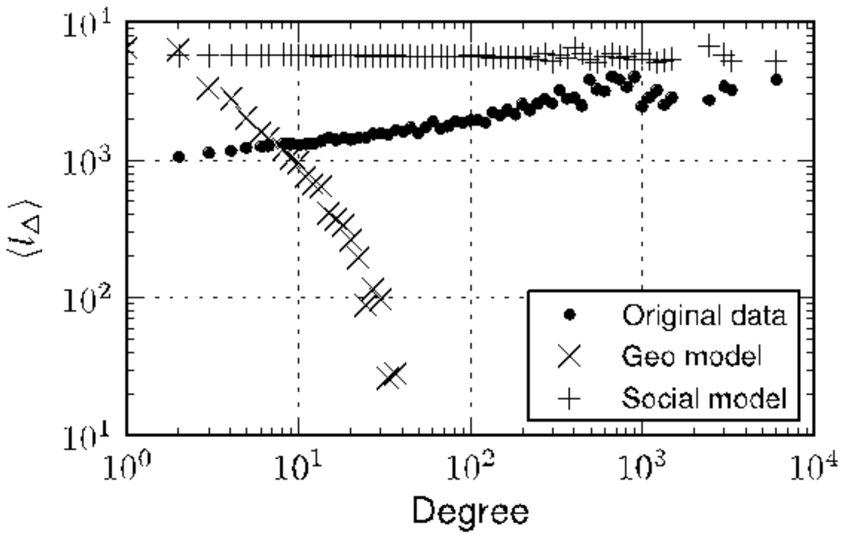


Correlation triangle length/degree

Correlation with degree: users with many friends belong to bigger triangles



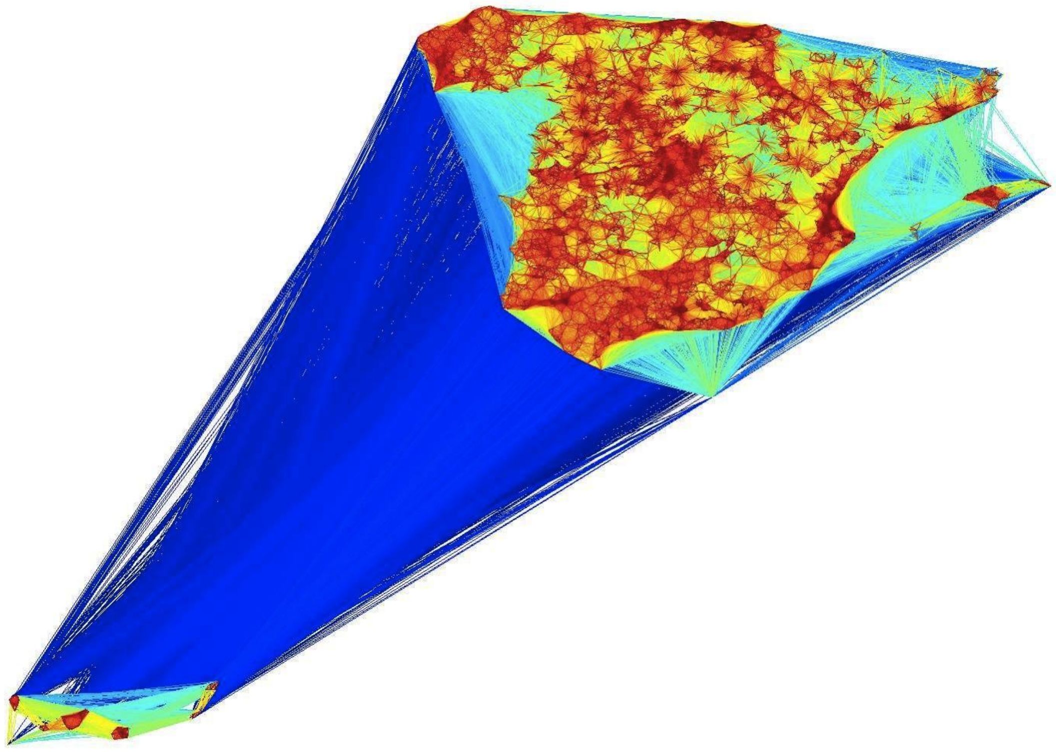




Effect of geography over interaction

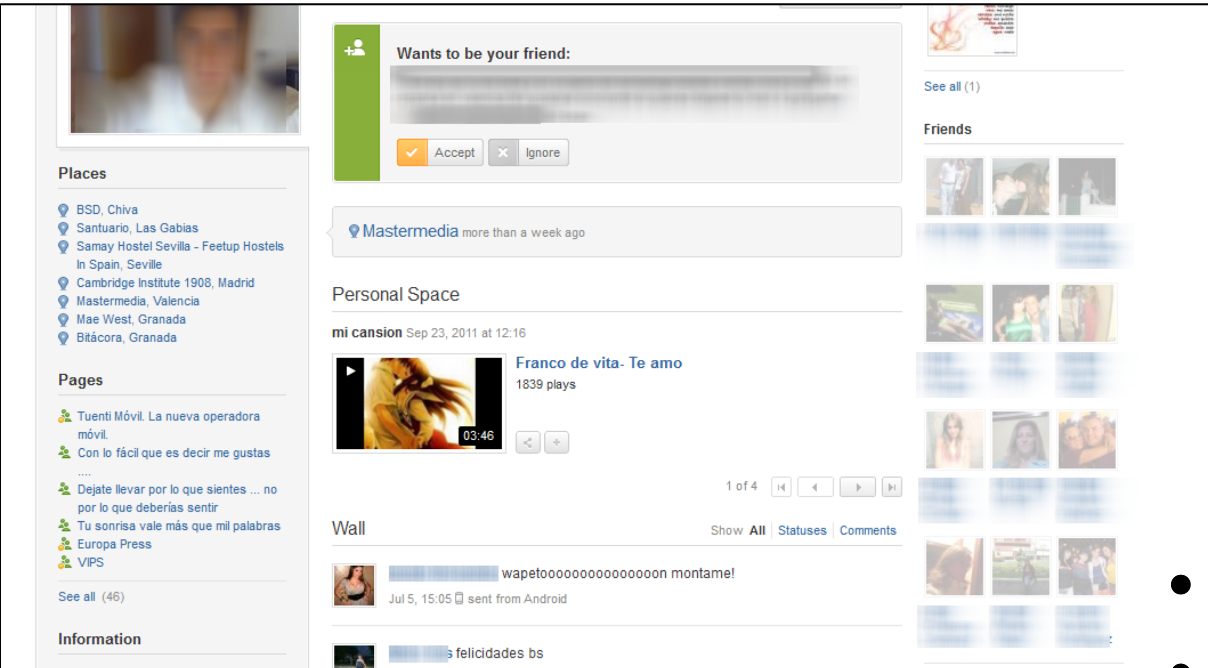
UNIVERSITY OF CAMBRIDGE

How does geography affect interaction?



UNIVERSITY OF CAMBRIDGE

Tuenti



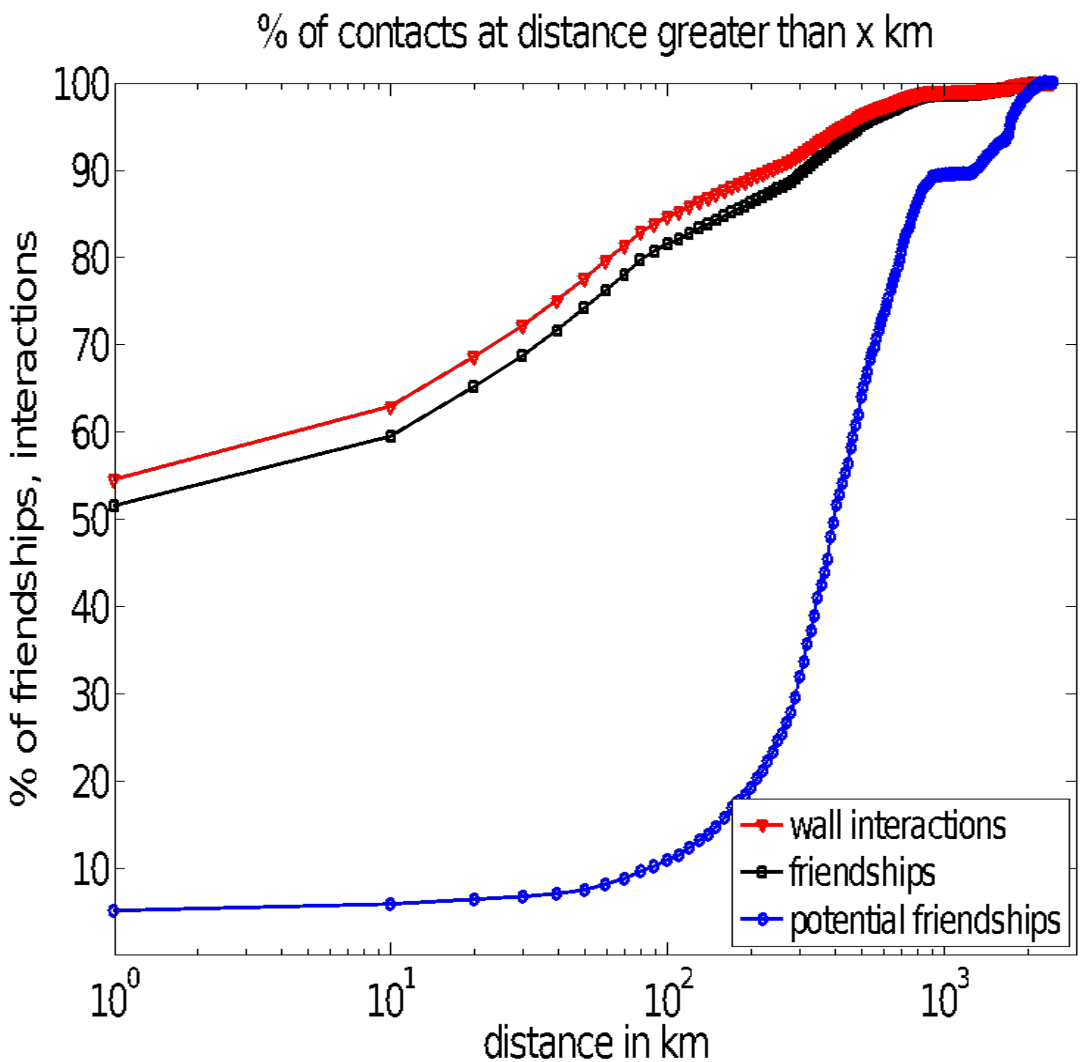
- Tuenti dataset (Nov 11)
- 9.88 million registered users
- ~1 174 million friendship links
- 500 million messages in 3 months

UNIVERSITY OF CAMBRIDGE

Geographic Properties

60% of social connections are at a distance of ≤ 10km.

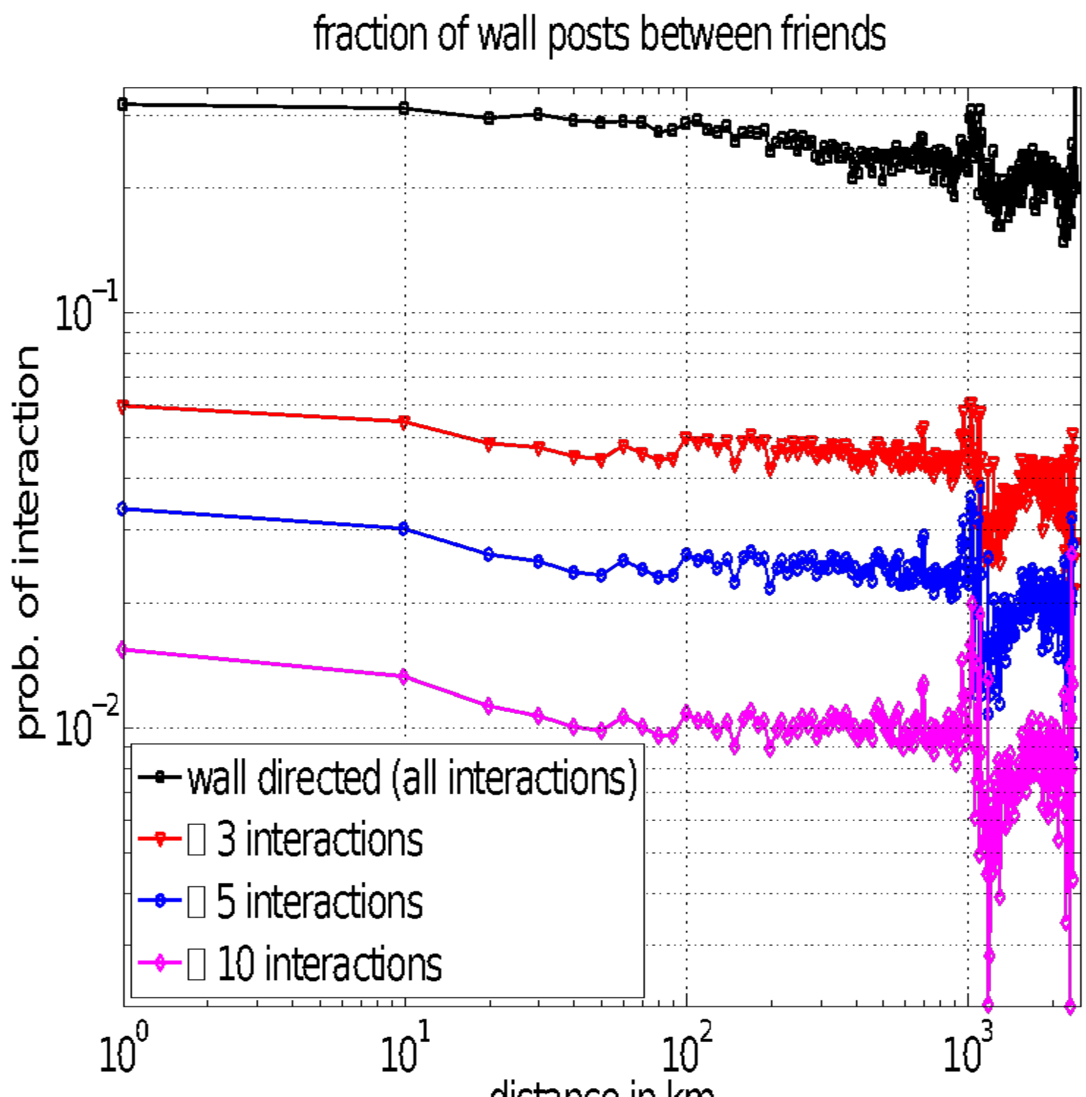
Only 10% of all distances between users are below 100km



UNIVERSITY OF CAMBRIDGE

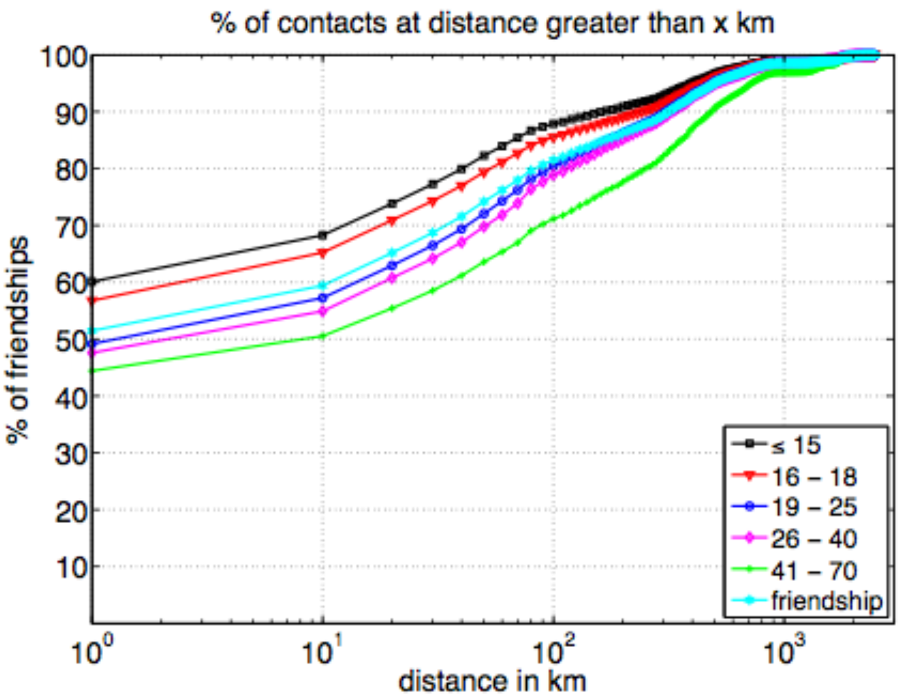
Effects of geography on interactions

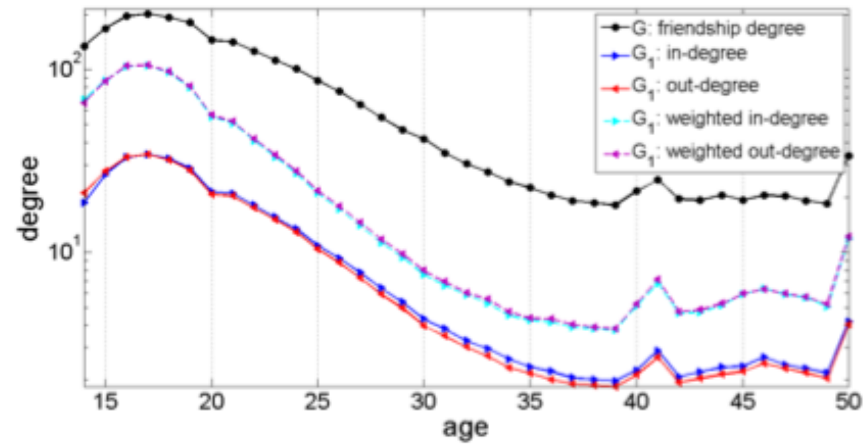
Probability that a message is exchanged over a friendship link seems not to be very affected by distance



UNIVERSITY OF CAMBRIDGE

Effect of Age



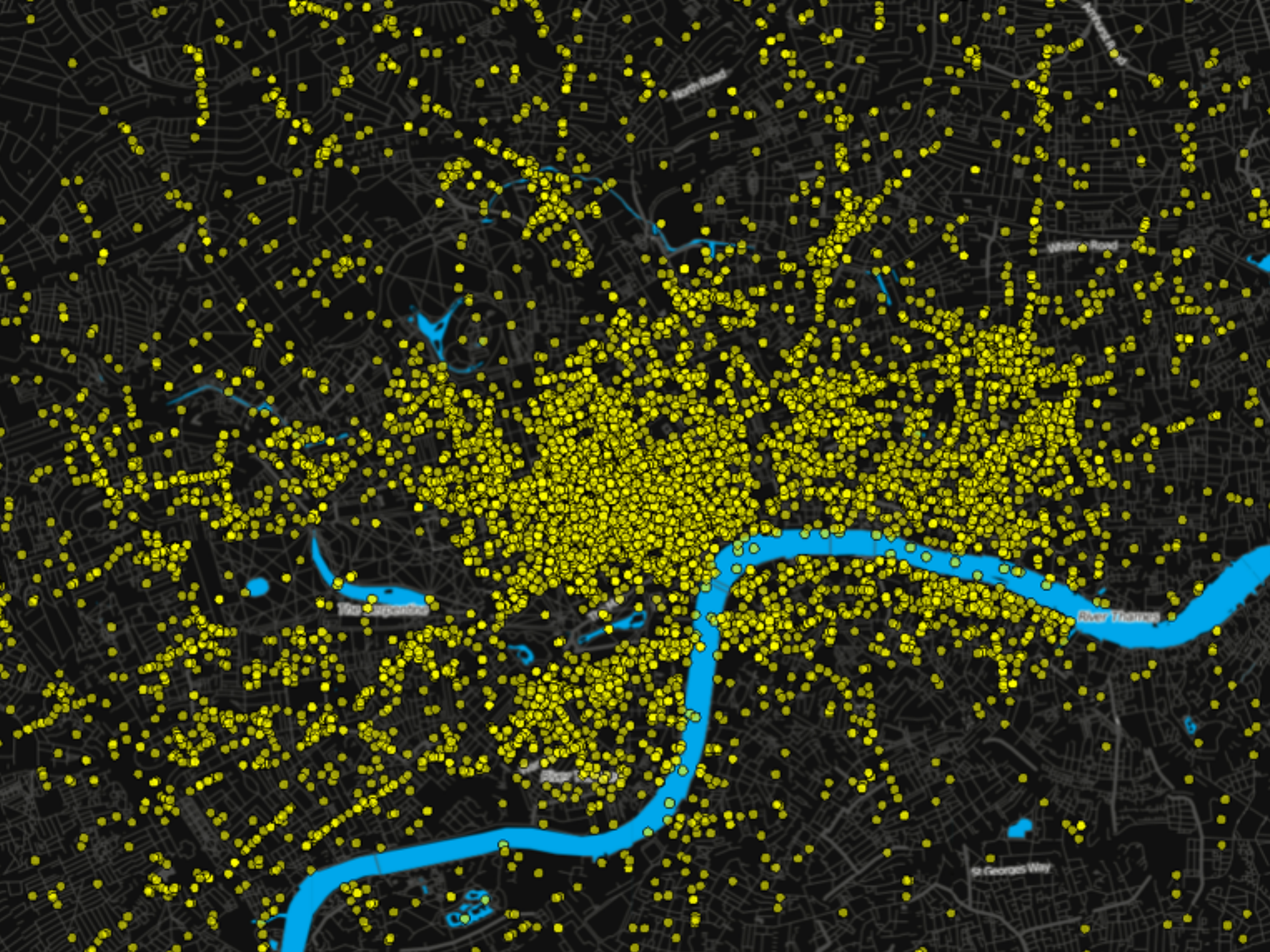


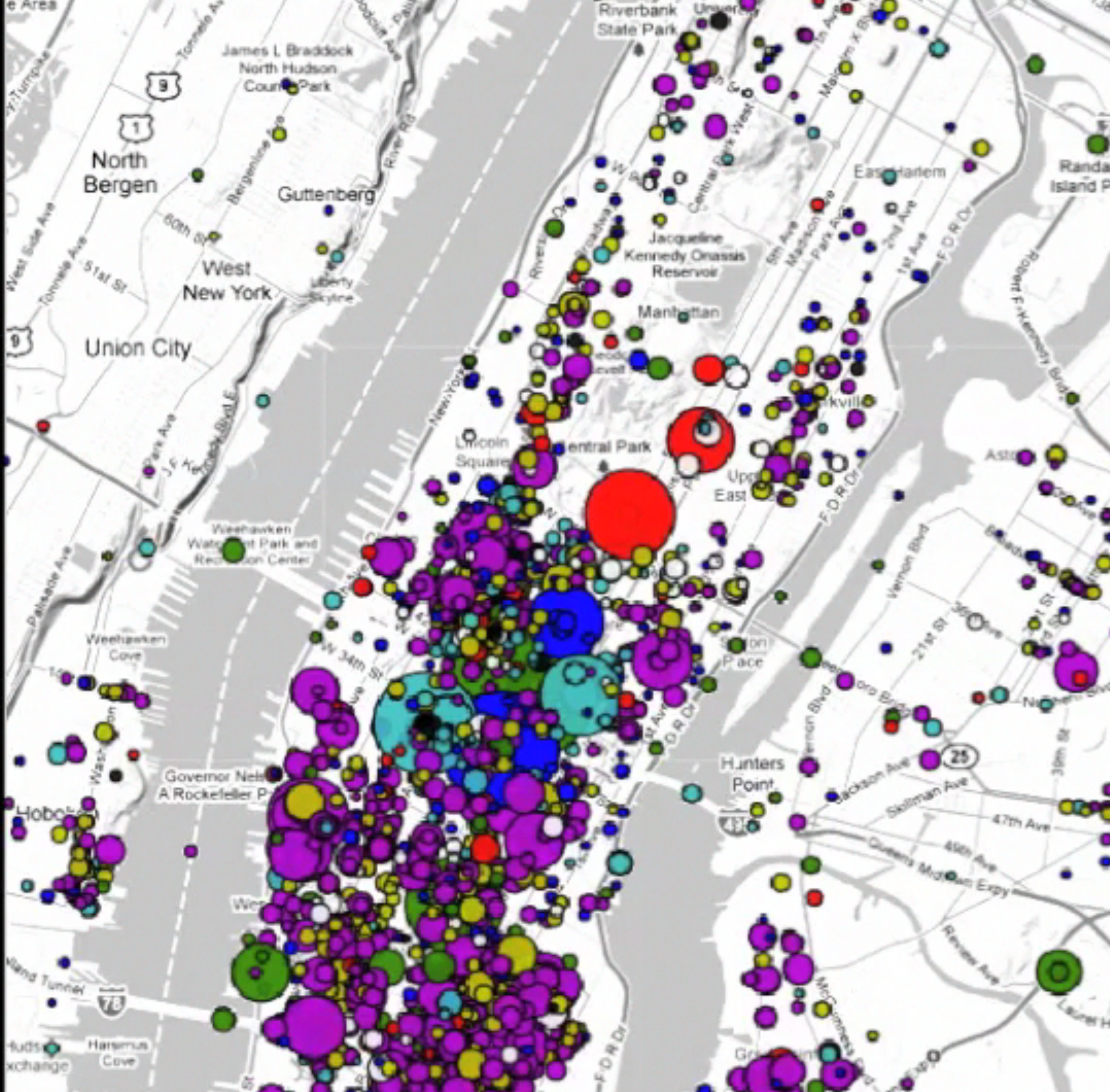
Under 15 have 70% friends within 10km. They also have more friends and interact more.

UNIVERSITY OF CAMBRIDGE

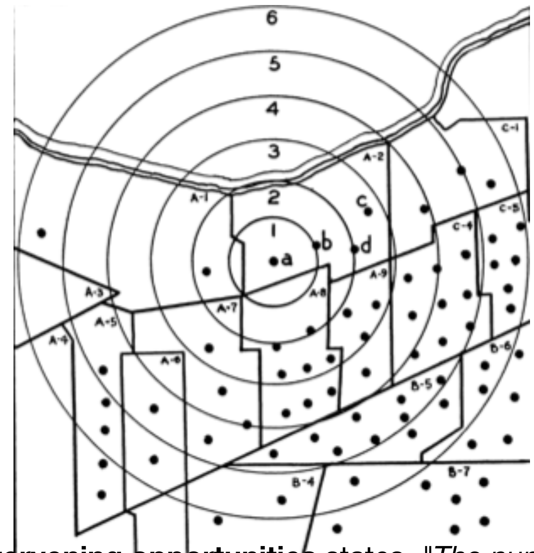
Understanding human mobility

UNIVERSITY OF CAMBRIDGE


North Road
Whiston Road
The Serpentine
River Thames
St Georges Way


Riverbank
State Park
James L Braddock
North Hudson
County Park
North
Bergen
Guttenberg
East Harlem
Randall
Island Pa
Jacqueline
Kennedy Onassis
Reservoir
West
New York
Manhattan
Union City
Central Park
Lincoln
Square
Weehawken
Waterfront Park and
Recreation Center
Weehawken
Cove
W 34th St
Hunters
Point
Governor Nelson
A Rockefeller Park
Hoboken
47th Ave
Harsimus
Cove
F D R Dr

Samuel A. Stouffer



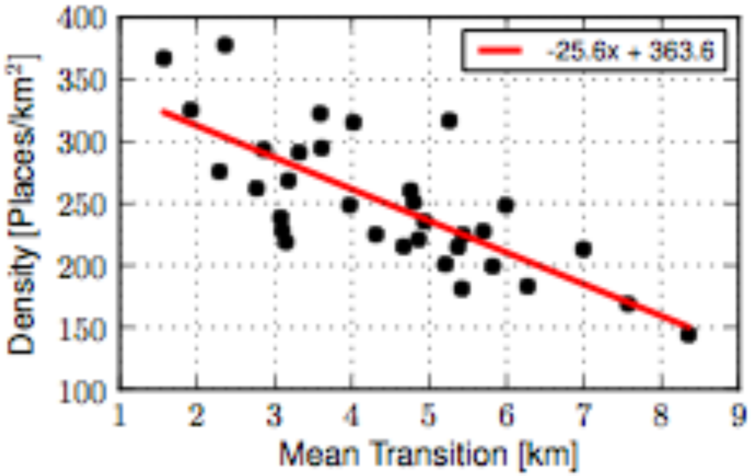
Stouffer's **law of intervening opportunities** states, "*The number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities.*" *

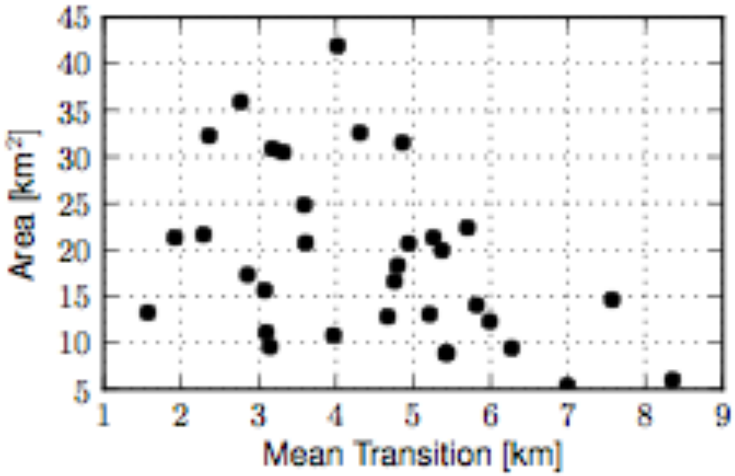
- Empirically proven using data for migrating families in the city of Cleveland.

- Is this true in our data?

* S. Stouffer (1940) Intervening opportunities: A theory relating mobility and distance, American Sociological Review 5, 845-867

The importance of density

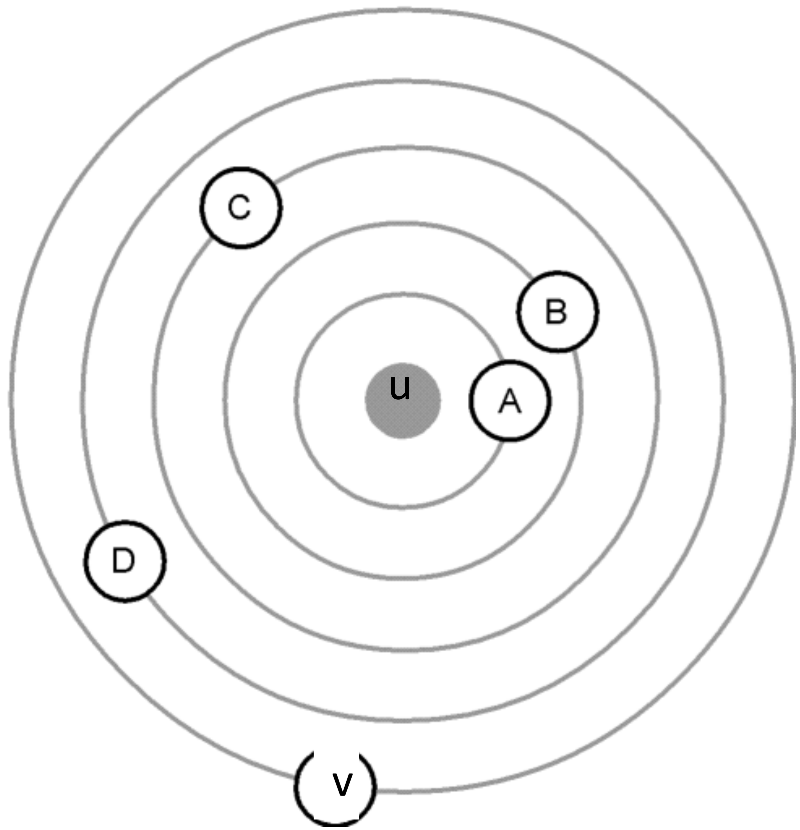




Place density by far more important than city area size with respect to mean length of human movements (R^2 = 0.59 and 0.19 respectively).

UNIVERSITY OF CAMBRIDGE

Rank Distance

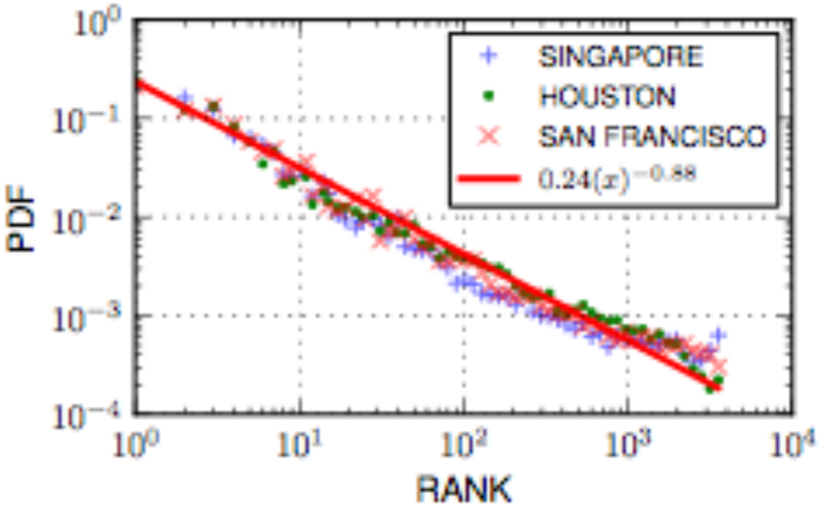


$$rank_u(v) = |\{w : d(u, w) < d(u, v)\}|$$

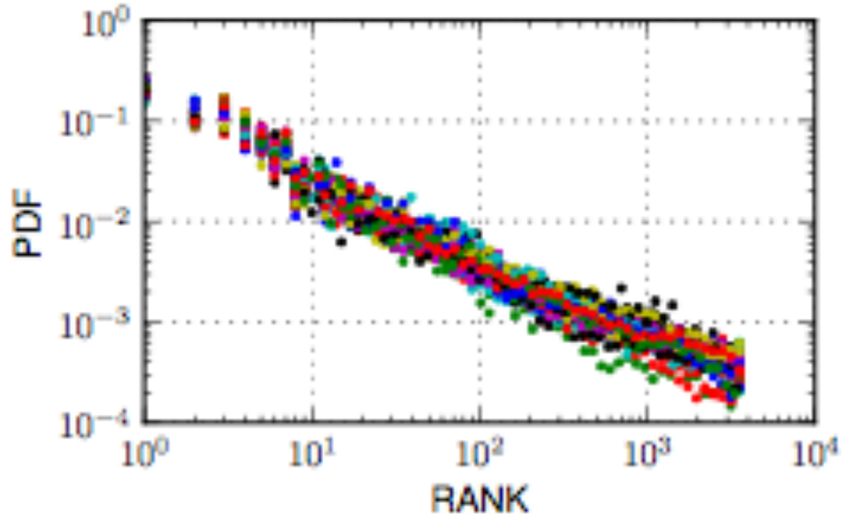
UNIVERSITY OF CAMBRIDGE

Rank universality

three cities



all cities

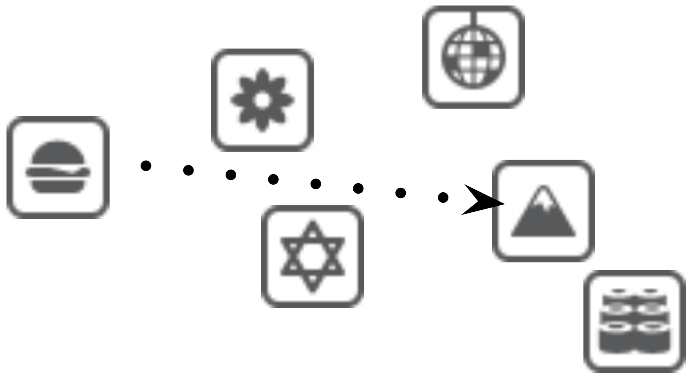


The rank of all cities collapses to a single line.

We have measured a power law exponent $\alpha = 0.84 \pm 0.07$

UNIVERSITY OF CAMBRIDGE

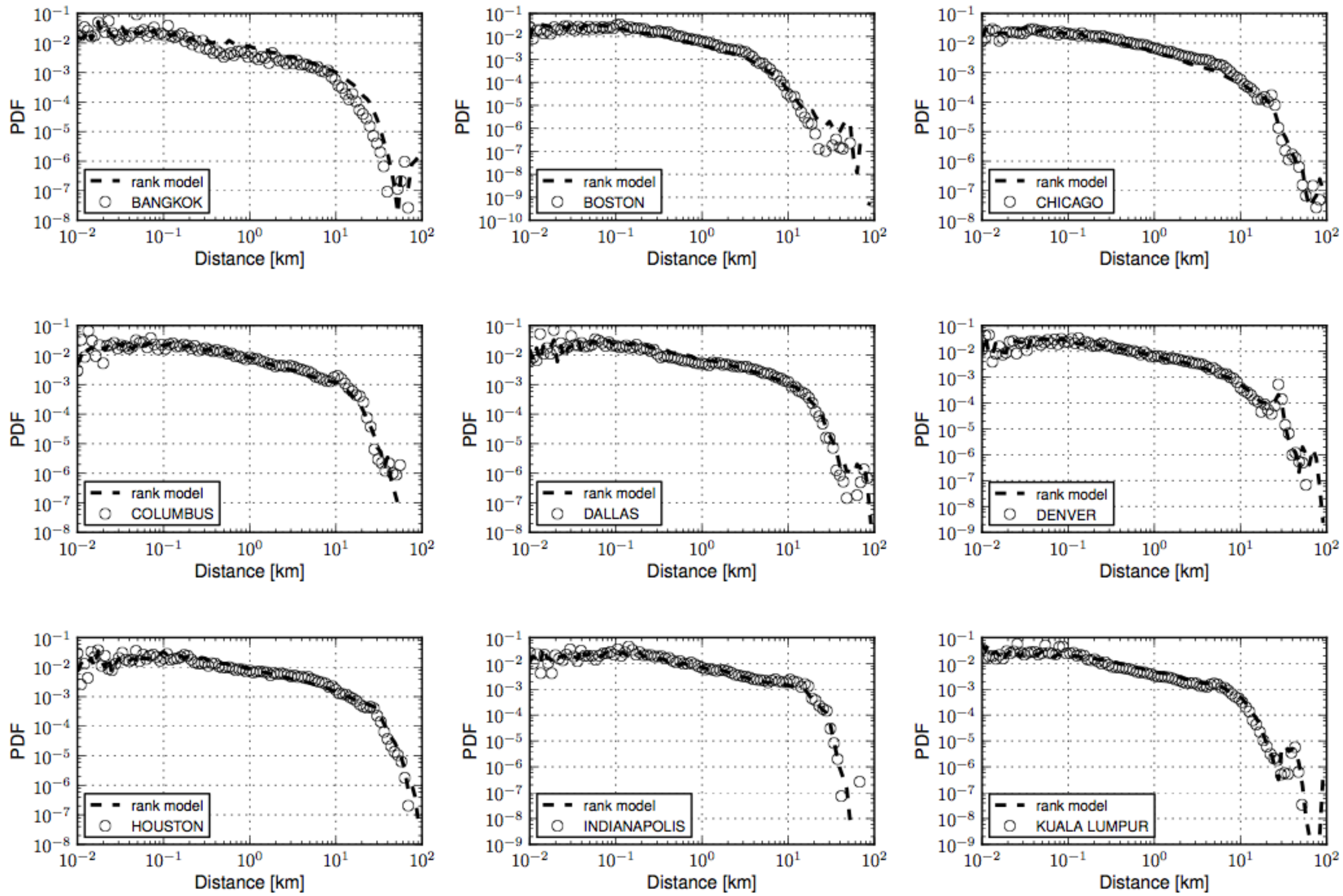
A new model for urban mobility



$$Pr[u \to v] \propto \frac{1}{rank_u(v)^a}$$

UNIVERSITY OF CAMBRIDGE

Simulation Results …



UNIVERSITY OF CAMBRIDGE

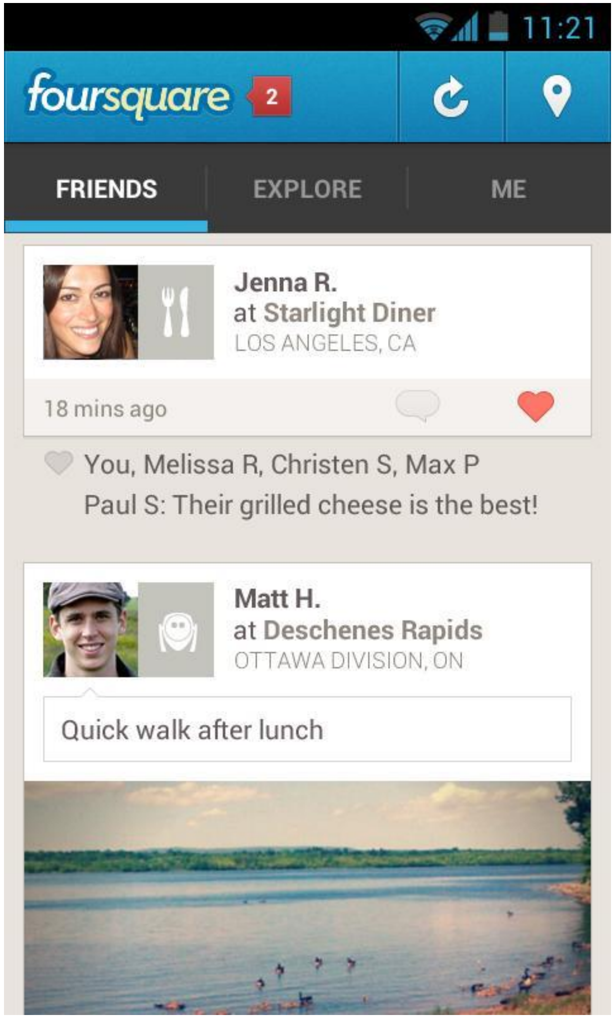
Understanding communities and role of places

UNIVERSITY OF CAMBRIDGE

City-scale social networks

- We look at **intra-city** social networks
- People who have checked in at a place in a given city, and their friends **who have also checked in** at those places

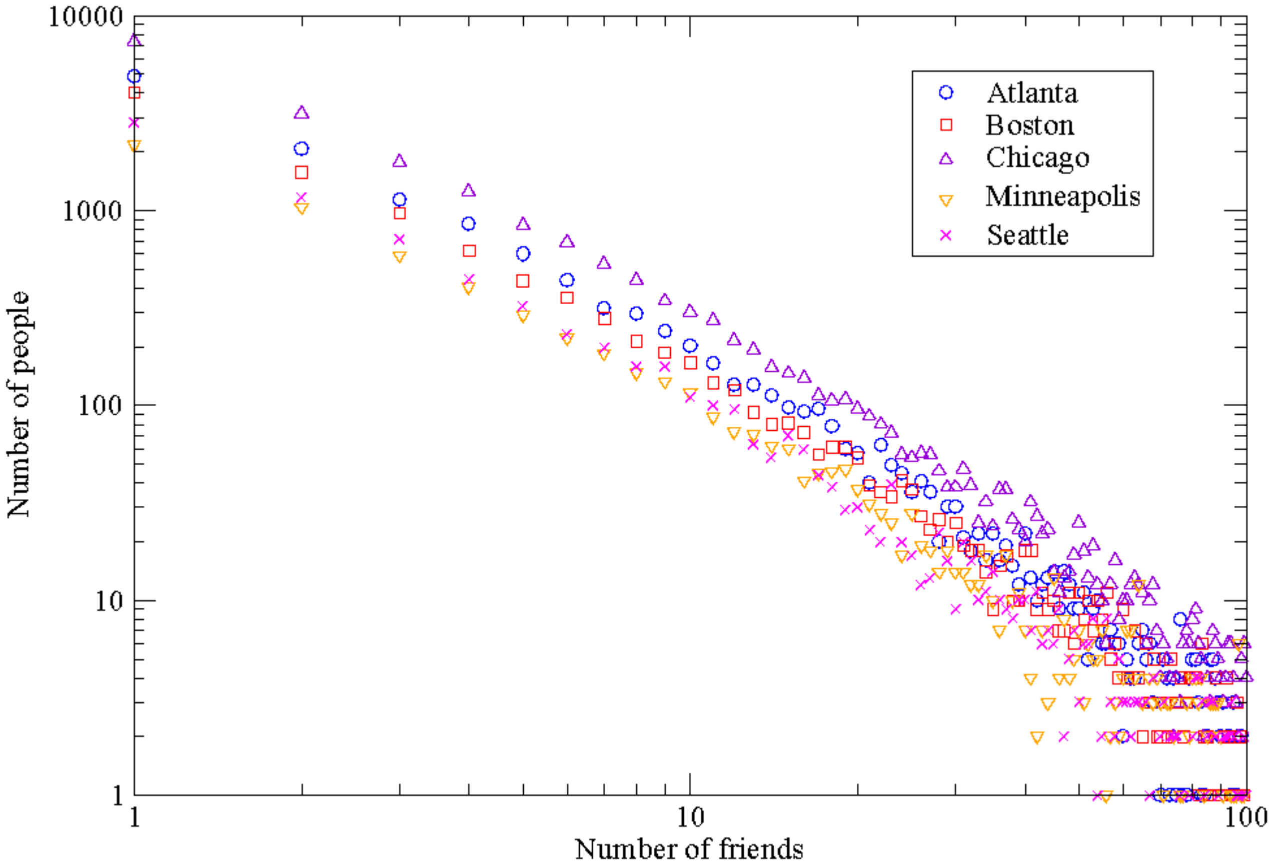
What do these place-based social networks look like?



UNIVERSITY OF CAMBRIDGE

City-scale social networks

- **Degree**: power-law distribution



UNIVERSITY OF CAMBRIDGE

City-scale social networks

- **Degree**: power-law distribution
- **Clustering coefficient**: high (between 0.1 and 0.2, in random graph of the same size <0.001)
- **Average shortest path length**: small (about 4 hops), comparable to random graph (Clustering coeff. + average path length = “small world”)
- **Community structure** (modularity > 0.4)

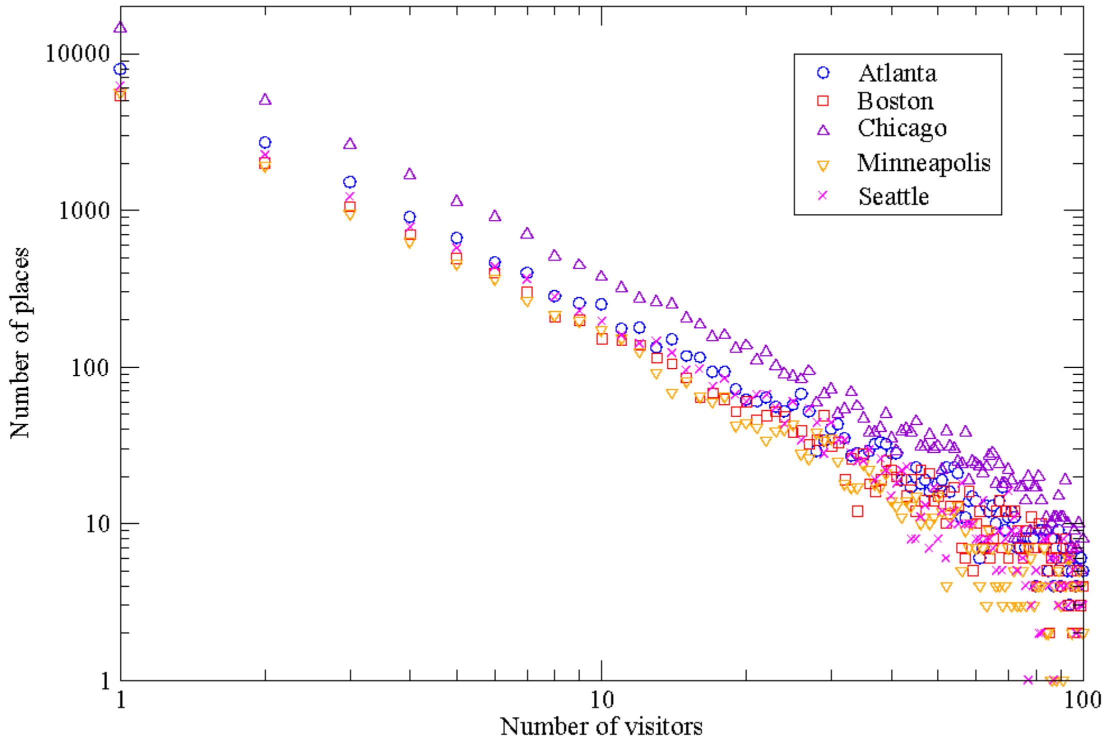
Our city-level graphs have well-known structural properties of social networks.

UNIVERSITY OF CAMBRIDGE

The role of places

- Power-law distribution of **place popularity**

…like the degree distribution in the social network



UNIVERSITY OF CAMBRIDGE

Places vital for tie formation

- Power-law distribution of **place popularity**
- **Analyze triangles**: >70% of triangles have one place shared between all three people
  → **clustering around certain places**

These places could act as foci for tie formation…

UNIVERSITY OF CAMBRIDGE

Role of categories…

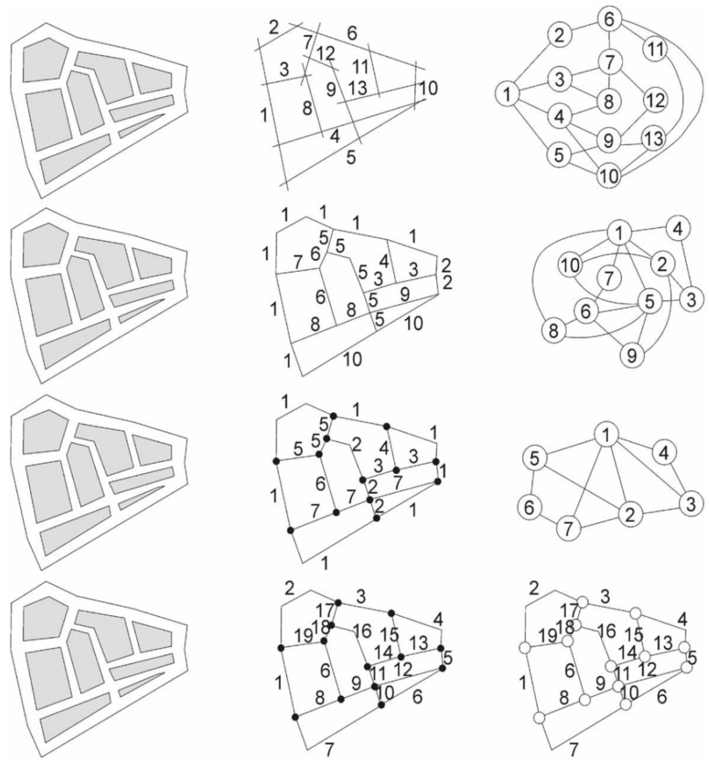
- What is the role of place categories?
  - Probability of friendship between **colocated people** at places in each Foursquare **category**
  - Some kinds of places are **much more likely to reinforce friendship** than others

UNIVERSITY OF CAMBRIDGE

Urban analysis with spatial networks

UNIVERSITY OF CAMBRIDGE

Model cities with networks



DUAL

- Named Street Approach
- Axial Line, Space Syntax
- Good Continuation Approach

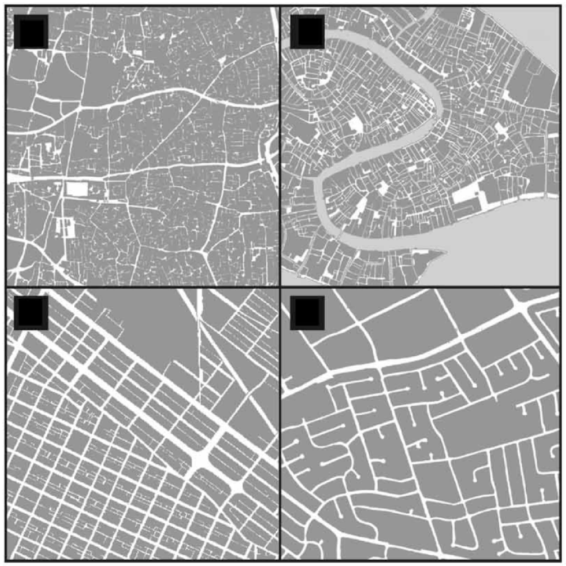
PRIMAL

Pure Primal

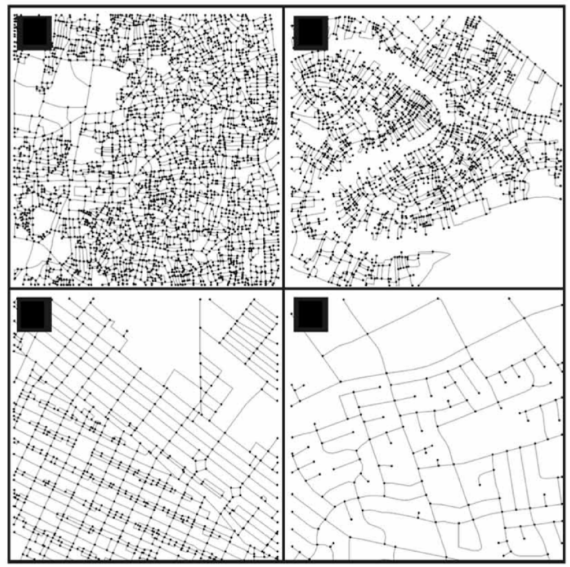
Porta S, Crucitti P, Latora V. (2006), The network analysis of urban streets: a primal approach, «Environment and Planning B: planning and design», 33

UNIVERSITY OF CAMBRIDGE

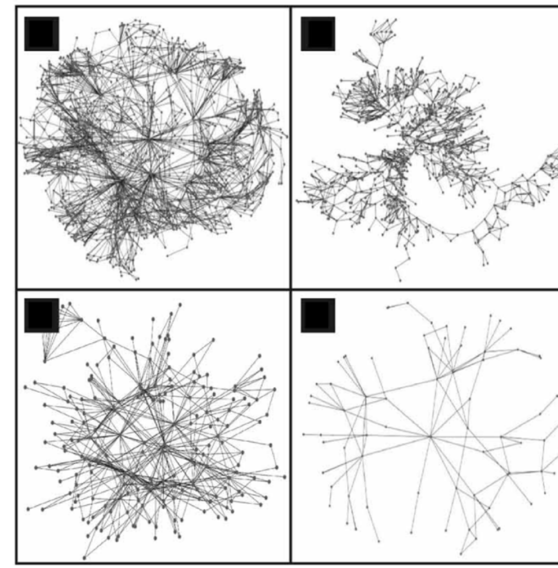
Street networks models



Urban Map



Primal Representation

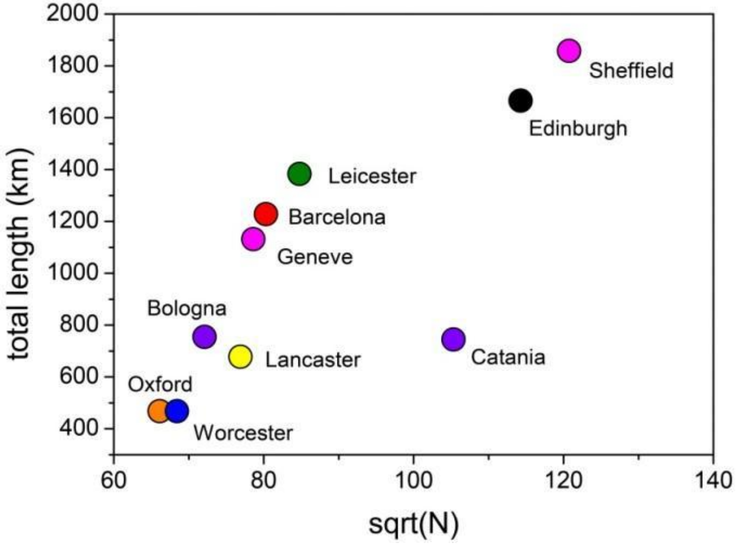


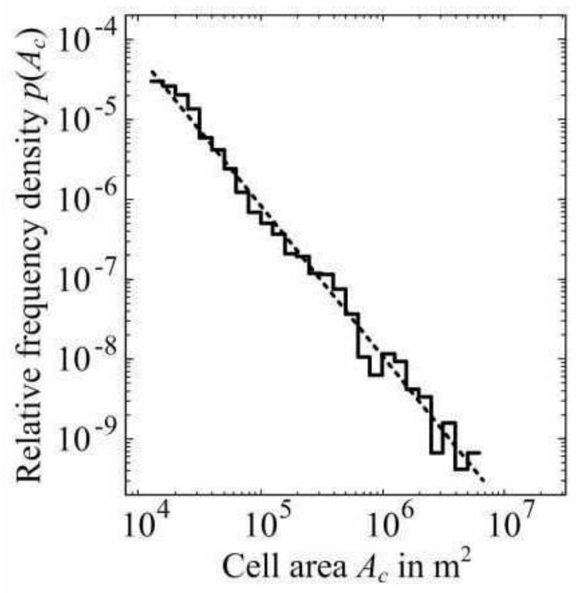
Dual Representation

Porta S, Crucitti P, Latora V. (2006), The network analysis of urban streets: a primal approach, «Environment and Planning B: planning and design», 33

UNIVERSITY OF CAMBRIDGE

Street networks analysis





- Barthelemy, M. Spatial Networks, Physic Report, 2011
- Strano, E. Viana, M. Cardillo, A. Porta, S. Da Costa, L Latora, V. "Urban street networks, a comparative analysis of ten European cities." Environment and Planning B (in print).

- S. Lämmer, B. Gehlsen, and D. Helbing (2006) Scaling laws in the spatial structure of urban road networks. *Physica A* **363**(1) 89-95.

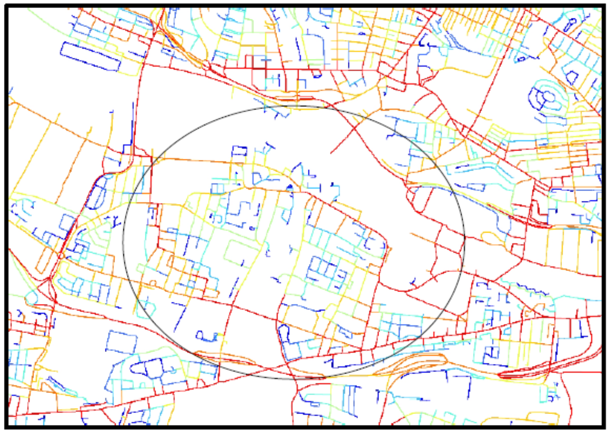
UNIVERSITY OF CAMBRIDGE

Betweenness Centrality

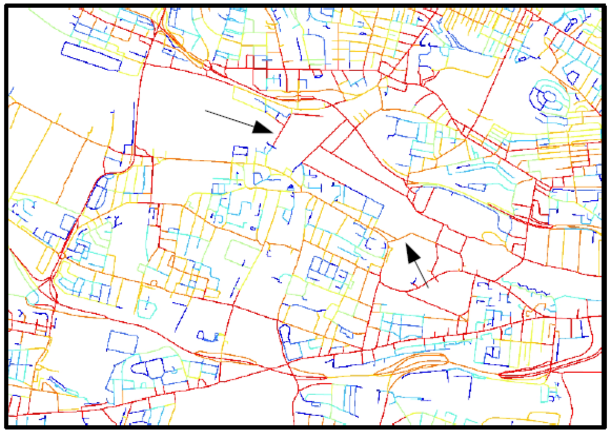
Betweenness

BetC is based on the idea that a node is more central when it is traversed by a larger number of shortest paths connecting all couples of nodes in the network. njk is the number of shortest paths between nodes j and k, and njk(i) is the number of these shortest paths that contain node i.

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j=1;k=1;j\neq k\neq i}^{N} \frac{n_{jk}(i)}{n_{jk}}$$



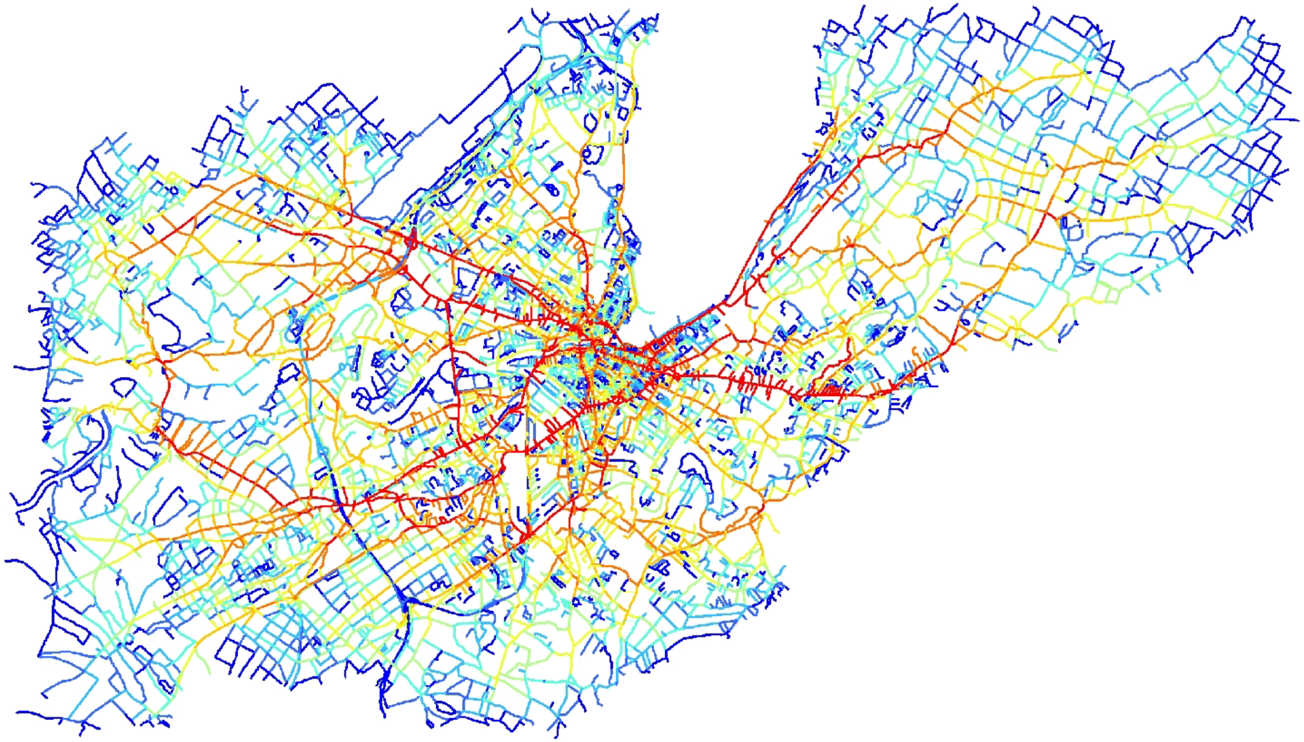
Current situation



Simulation

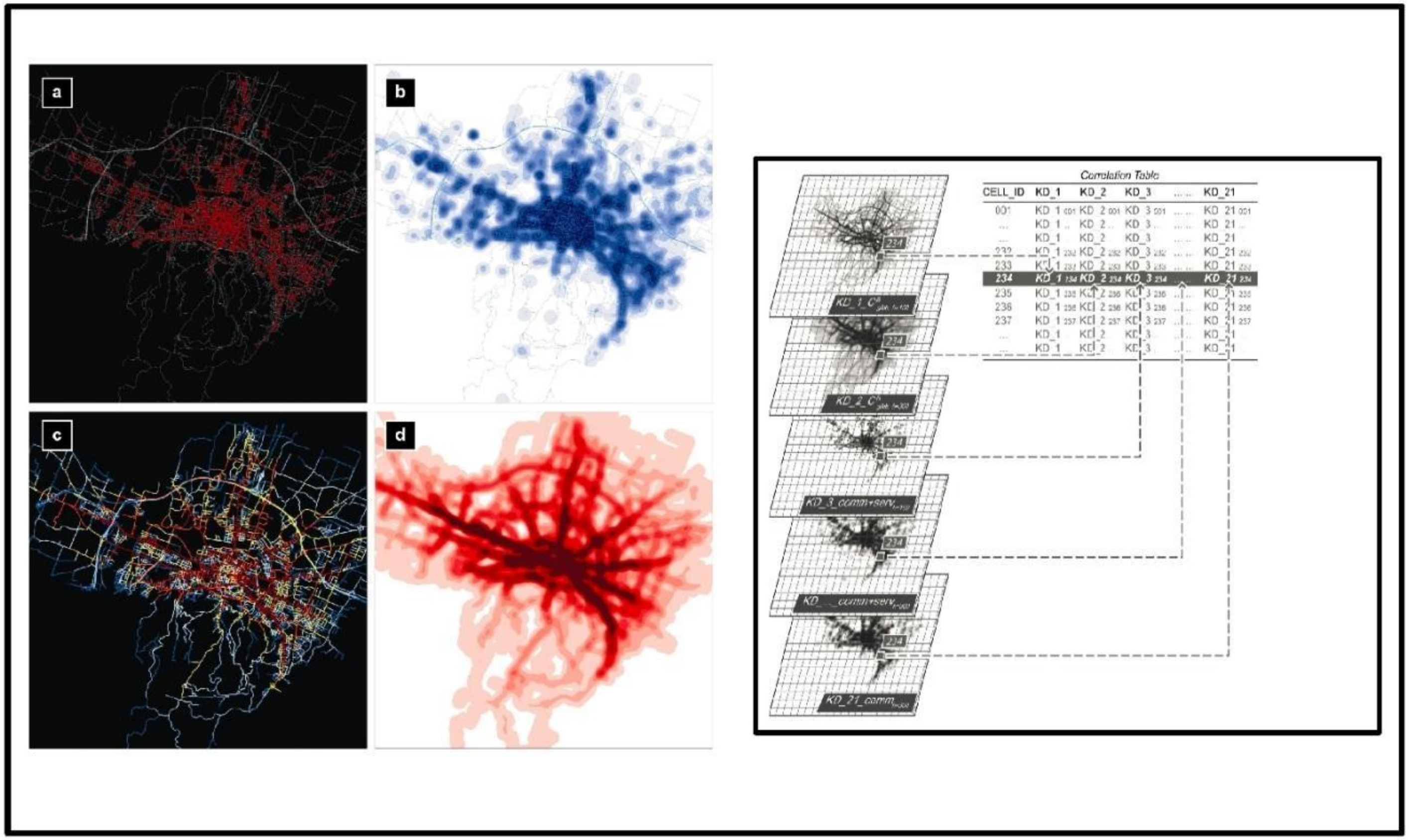
UNIVERSITY OF CAMBRIDGE

Betweenness Centrality



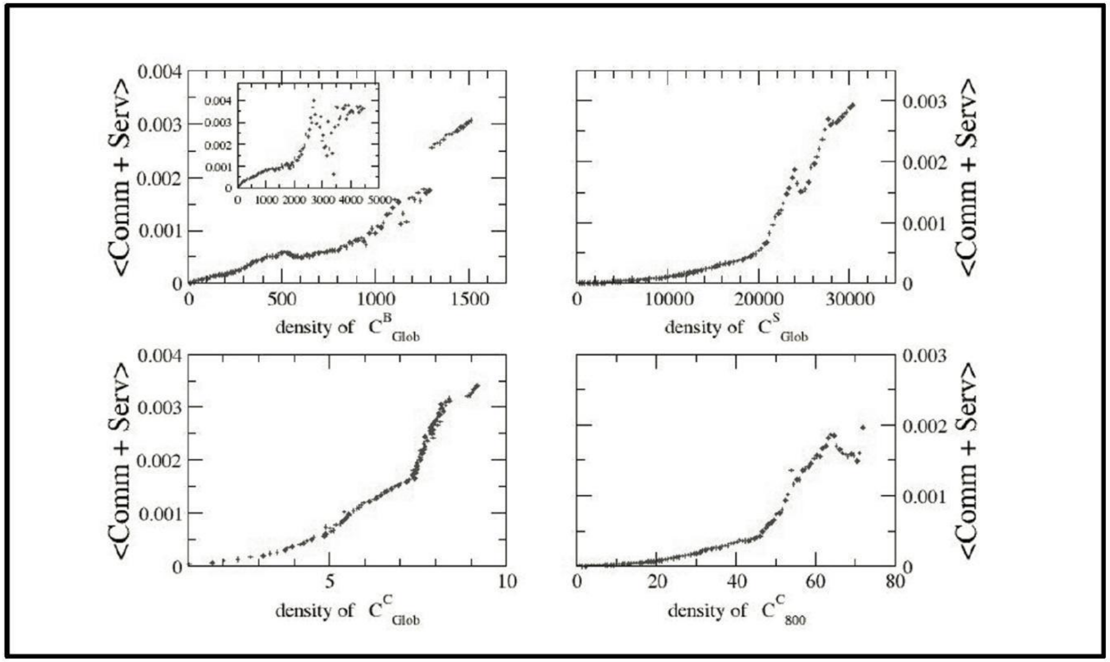
UNIVERSITY OF CAMBRIDGE

Example 1: Betweenness and Economic Activities



UNIVERSITY OF CAMBRIDGE

Example 1: Betweenness and Economic Activities

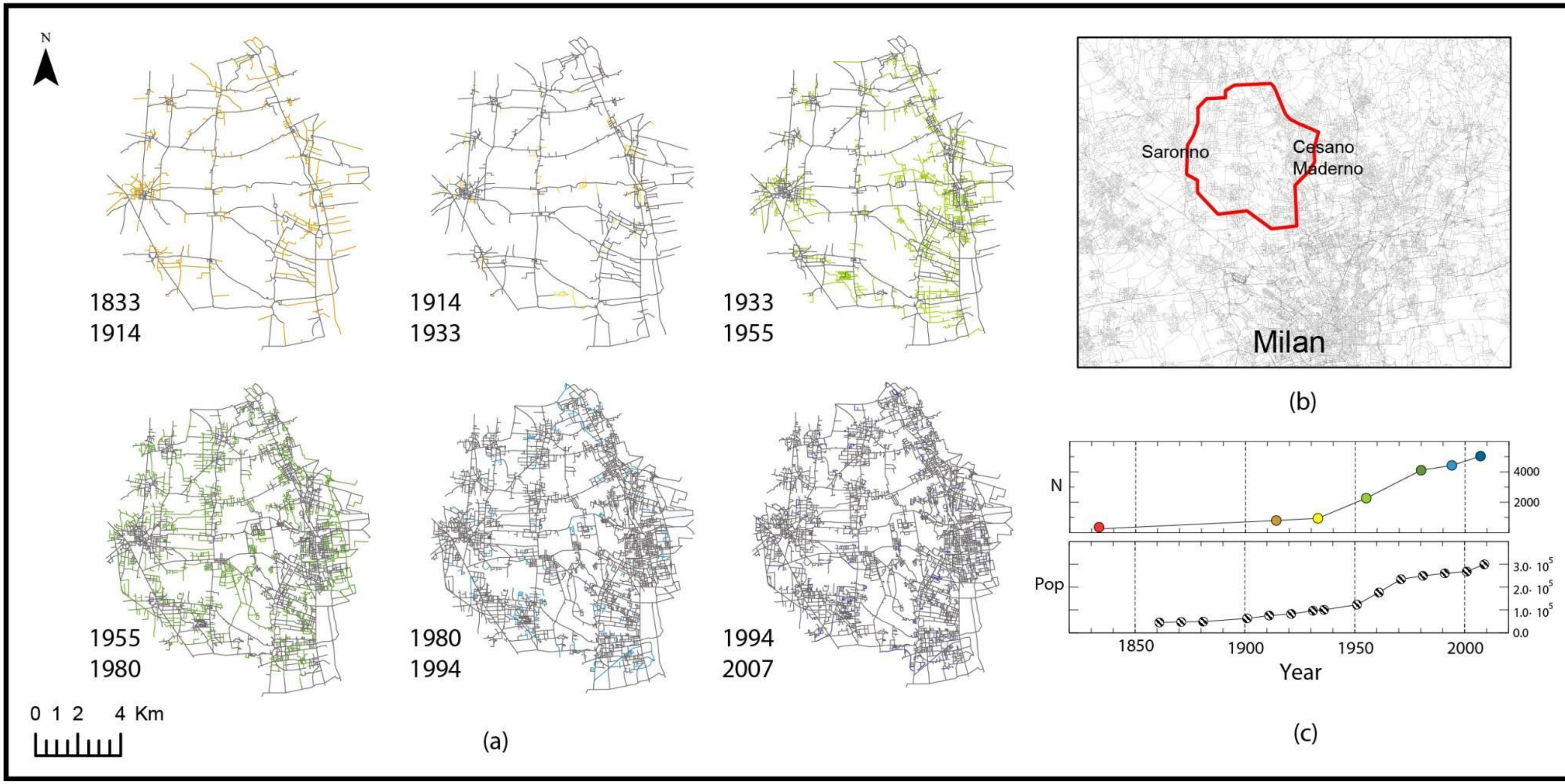


It seems that centralities are correlated to the location of micro economic activities

Porta, S. Latora, V, Strano,E, et al "Street centrality and densities of retail and services in Bologna, Italy", Environment and Planning B: Planning and design, 36,3,450-465,2009.

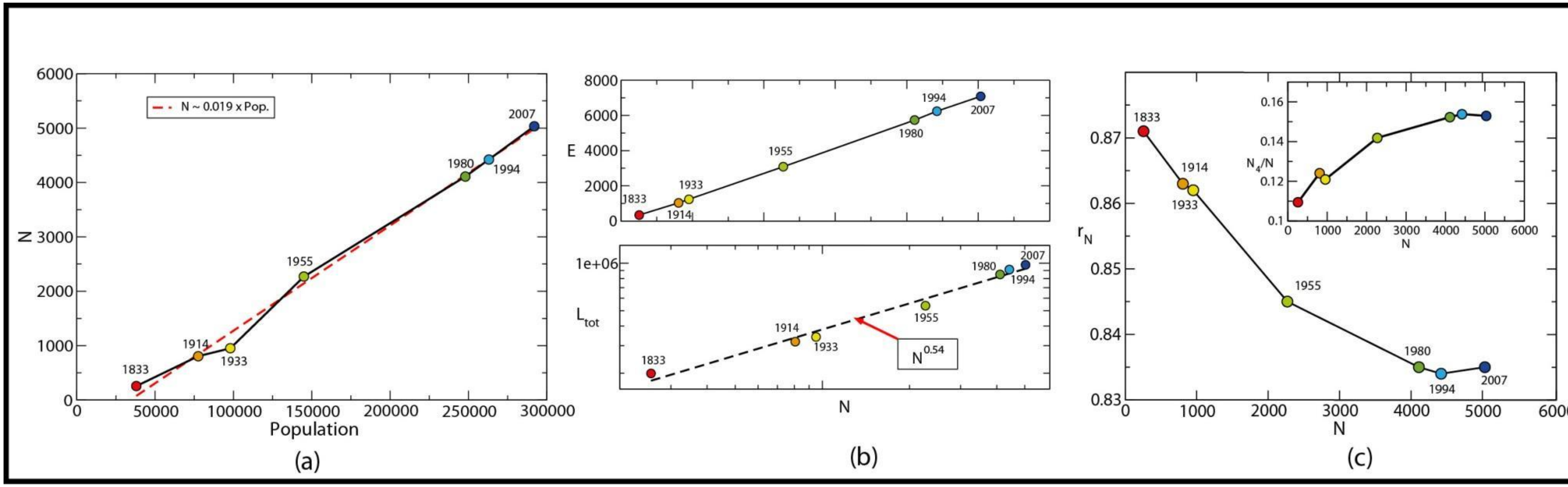
UNIVERSITY OF CAMBRIDGE

Example 2: Network Evolution and urbanization



UNIVERSITY OF CAMBRIDGE

Example 2: Network Evolution and urbanization



Pop Vs N, it seems that we all own a fraction of a street intersection.

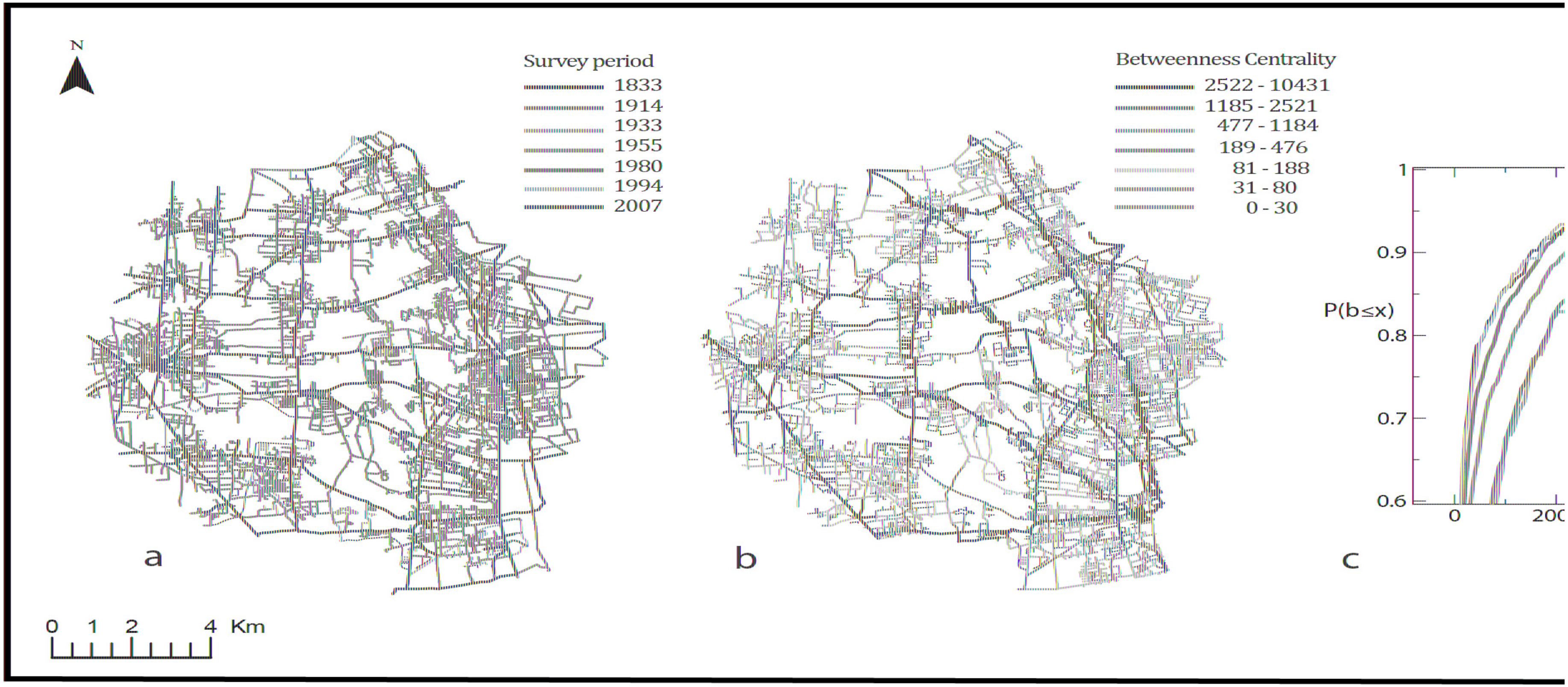
$$L_{tot} \sim N^{\gamma}$$

$$\gamma = 0.54$$

$$r(N) = \frac{N_1 + N_2}{\sum_{k \neq 2} N_k}$$

UNIVERSITY OF CAMBRIDGE

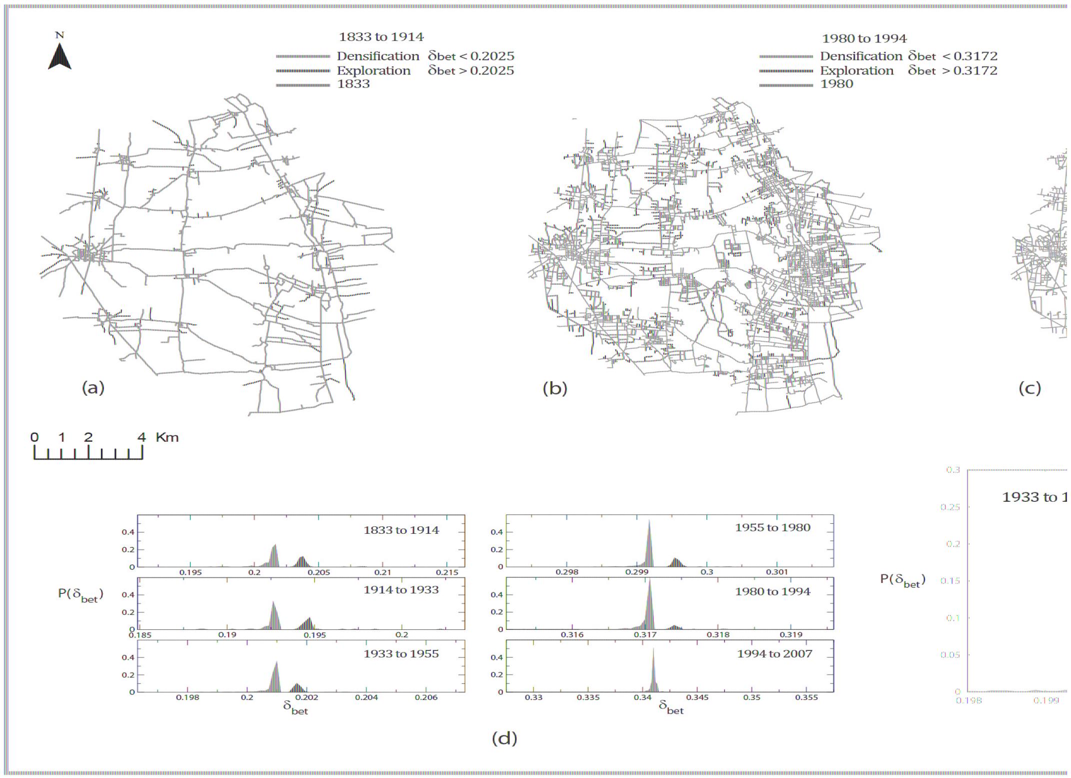
Example 2: Network Evolution and urbanization



Betweenness Centrality is strongly correlated to the age of each street!

UNIVERSITY OF CAMBRIDGE

Example 2: Network Evolution and urbanization



Betweenness contribution of each new street can be assigned to two main processes: exploration and densification

Strano et al: Elementary Processes Governing the evolution of Street Network. Nature Scientific Report (2012)

UNIVERSITY OF CAMBRIDGE

What can all this be used for?

- **Friendship Recommendation:**
  - **Exploiting Place Features in Link Prediction on Location-based Social Networks**, Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo. In Proceedings of 17th ACM International Conference on Knowledge Discovery and Data Mining (KDD 2011). San Diego, USA. August 2011.
- **Place Recommendation:**
  - **Mining User Mobility Features for Next Place Prediction in Location-based Services.** Anastasios Noulas, Salvatore Scellato, Neal Lathia and Cecilia Mascolo. In Proceedings of IEEE International Conference on Data Mining (ICDM 2012). Short Paper. Brussels, Belgium. December 2012.
  - **A Random Walk Around the City: New Venue Recommendation in Location-Based Social Networks**. Anastasios Noulas, Salvatore Scellato, Neal Lathia and Cecilia Mascolo. In Proceedings of ASE/IEEE International Conference on Social Computing (SocialCom). Amsterdam, The Netherlands. September 2012.

UNIVERSITY OF CAMBRIDGE

What can all this be used for?

- **More Modelling:**
  - **Talking Places: Modelling and Analysing Linguistic Content in Foursquare**. Sandro Bauer, Anastasios Noulas, Diarmuid Ó Séaghdha, Stephen Clark and Cecilia Mascolo. In Proceedings of ASE/IEEE International Conference on Social Computing (SocialCom). Amsterdam, The Netherlands. September 2012.
  - **Exploiting Foursquare and Cellular Data to Infer User Activity in Urban Environments**. Anastasios Noulas, Cecilia Mascolo and Enrique Frias-Martinez. In Proceedings of 14th International Conference on Mobile Data Management (MDM 2013). Milan, Italy. June 2013.
  - **Evolution of a Location-based Online Social Network: Analysis and Models.** Militiadis Allamanis, Salvatore Scellato and Cecilia Mascolo. In Proceedings of ACM Internet Measurement Conference (IMC 2012). Boston, MA. November 2012.

UNIVERSITY OF CAMBRIDGE

References

- **The length of bridge ties: structural and geographic properties of online social interactions.** Yana Volkovich, Salvatore Scellato, David Laniado, Cecilia Mascolo and Andreas Kaltenbrunner. In Proceedings of Sixth International AAAI Conference on Weblogs and Social Media (ICWSM 2012). Ireland. June 2012.
- **Far from the eyes, close on the Web: impact of geographic distance on online social interactions.** Andreas Kaltenbrunner, Salvatore Scellato, Yana Volkovich, David Laniado, Dave Currie, Erik J. Jutemar, Cecilia Mascolo. In ACM SIGCOMM Workshop on Online Social Networks (WOSN 2012). Finland. 2012.
- **A tale of many cities: universal patterns in human urban mobility**. Anastasios Noulas, Salvatore Scellato, Renaud Lambiotte, Massimiliano Pontil, Cecilia Mascolo. In PLoS ONE. PLoS ONE 7(5): e37027.
- **The Importance of Being Placefriends: Discovering Location-focused Online Communities**. Chloë Brown, Vincenzo Nicosia, Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo. In ACM SIGCOMM Workshop on Online Social Networks (WOSN 2012). August 2012.

UNIVERSITY OF CAMBRIDGE