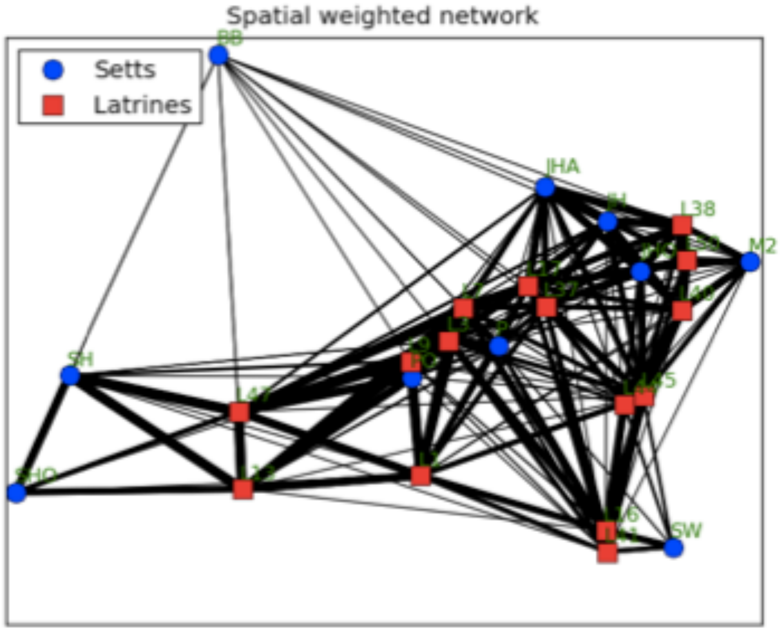
Evolving geo-temporal social networks and their applications

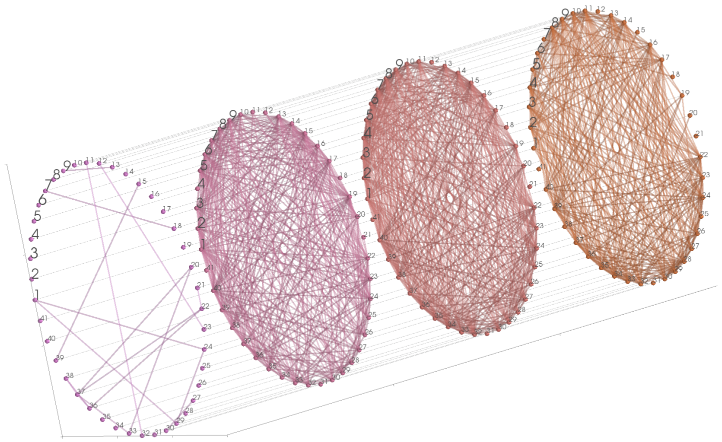
Cecilia Mascolo
Computer Laboratory

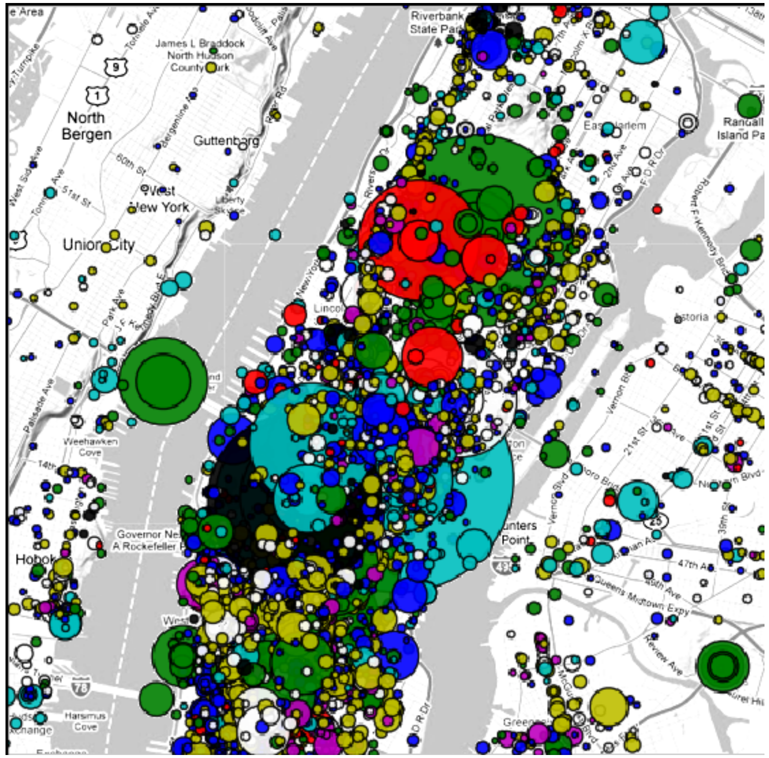
Networks Day
May 2013, Cambridge, UK.

UNIVERSITY OF CAMBRIDGE


SONY
Don't forget all the good times!
Good luck, everyone.

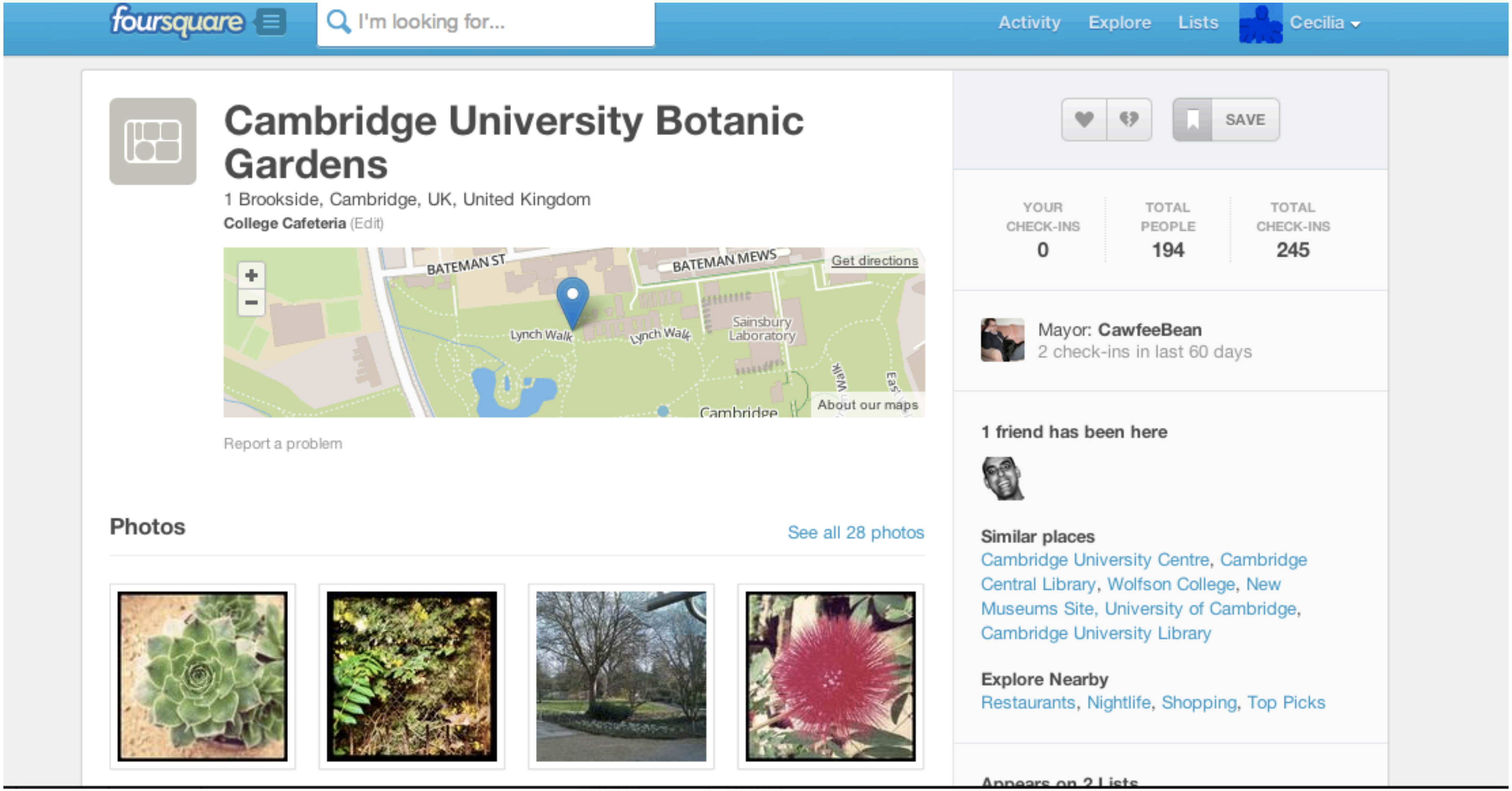


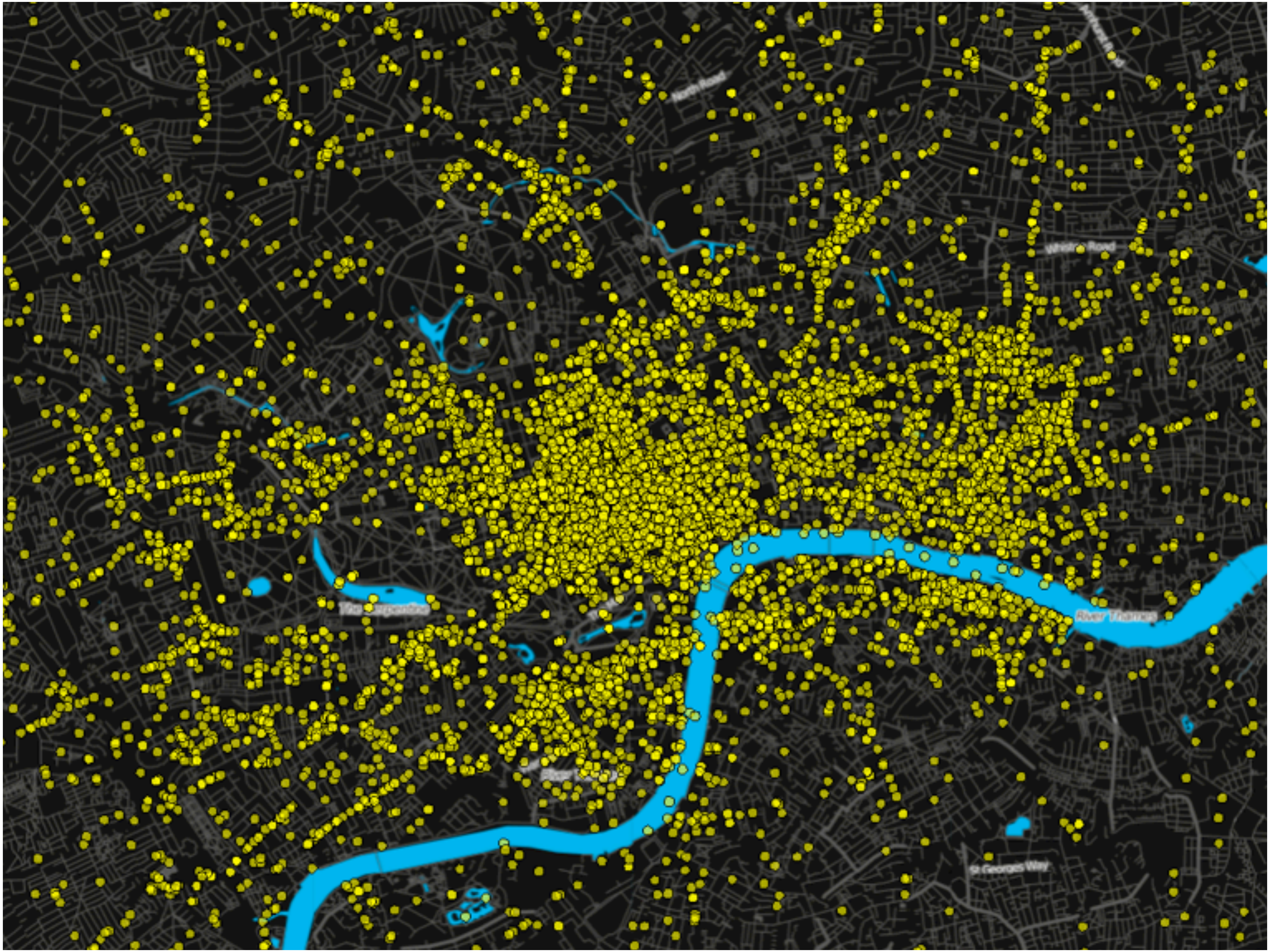


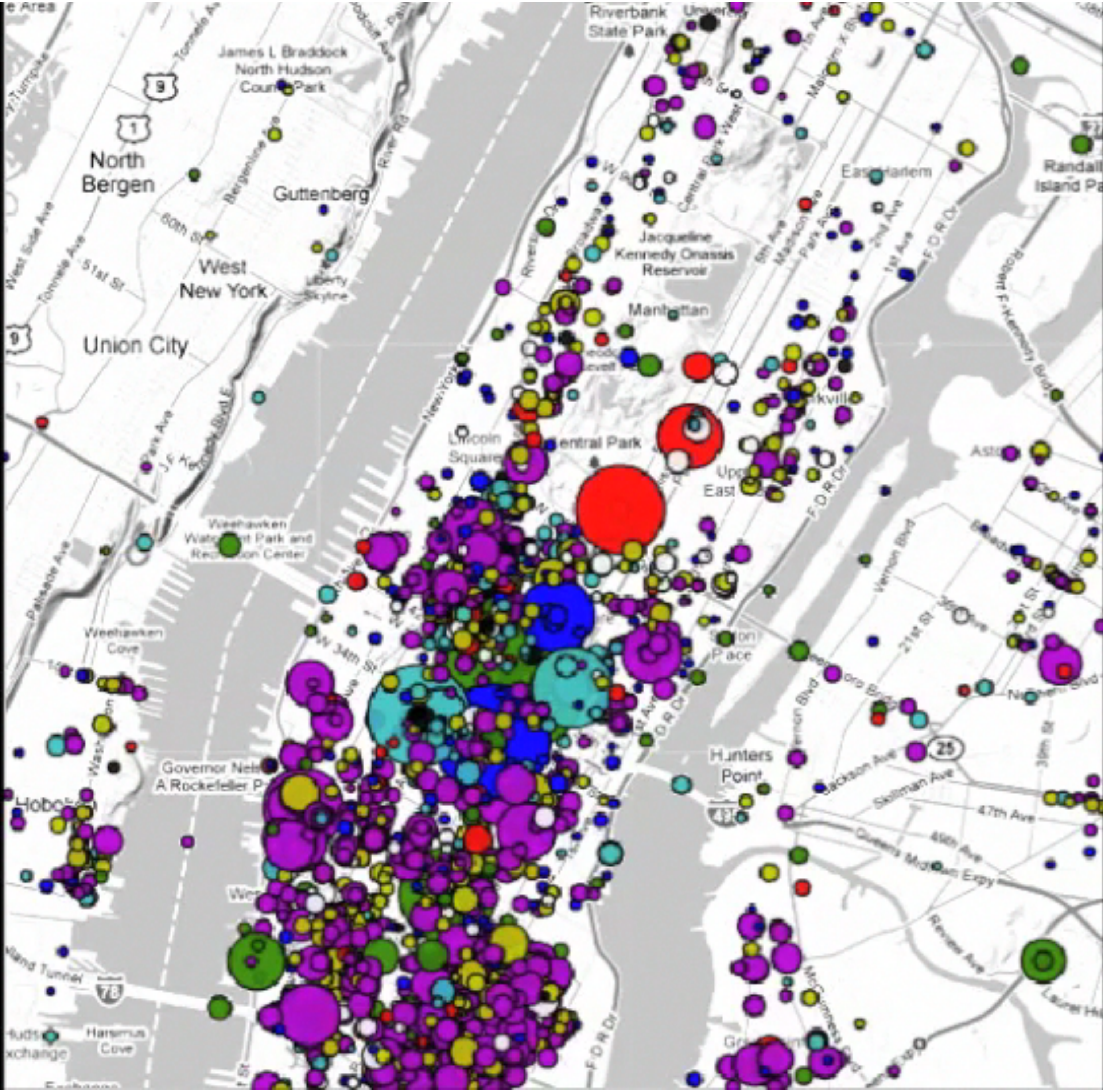


UNIVERSITY OF
CAMBRIDGE

Geo-social network analysis can lead to insights into social behaviour




North Road
River Thames
St Georges Way


Riverbank State Park
James L Braddock North Hudson County Park
North Bergen
Guttenberg
West New York
Union City
East Harlem
Randall Island Pa
Jacqueline Kennedy Onassis Reservoir
Manhattan
Central Park
Lincoln Square
Weehawken Waterfront Park and Recreation Center
Weehawken Cove
Governor Nelson A Rockefeller P
Hunters Point
Hoboken
Harsimus Cove
47th Ave
Queens Midtown Expy
Skillman Ave
Jackson Ave
Vernon Blvd
W 34th St
F D R Dr
Central Park West
Broadway
Madison Ave
5th Ave
1st Ave
2nd Ave
Park Ave
Palisade Ave
Kennedy Blvd E
Park Ave
Bergenline Ave
Tonnelle Ave
West Side Ave
60th St
51st St
Robert F Kennedy Bridge
21st St
36th Ave

What can we study with this type of data?

- Relationship of friendship and distance
- Relationship of interaction and distance
- **Human mobility and role of places**
- **Models for geo-social network evolution**
- **Evolution of communities in space**

- Applications:
  - Recommendations, advertisement
  - Urban planning
  - Epidemics control

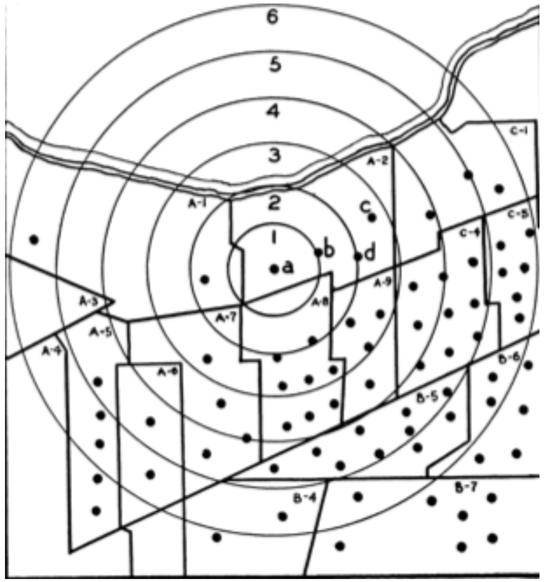
Human Mobility and Role of Places

UNIVERSITY OF CAMBRIDGE

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. doi:10.1371/journal.pone.0037027.

Samuel A. Stouffer

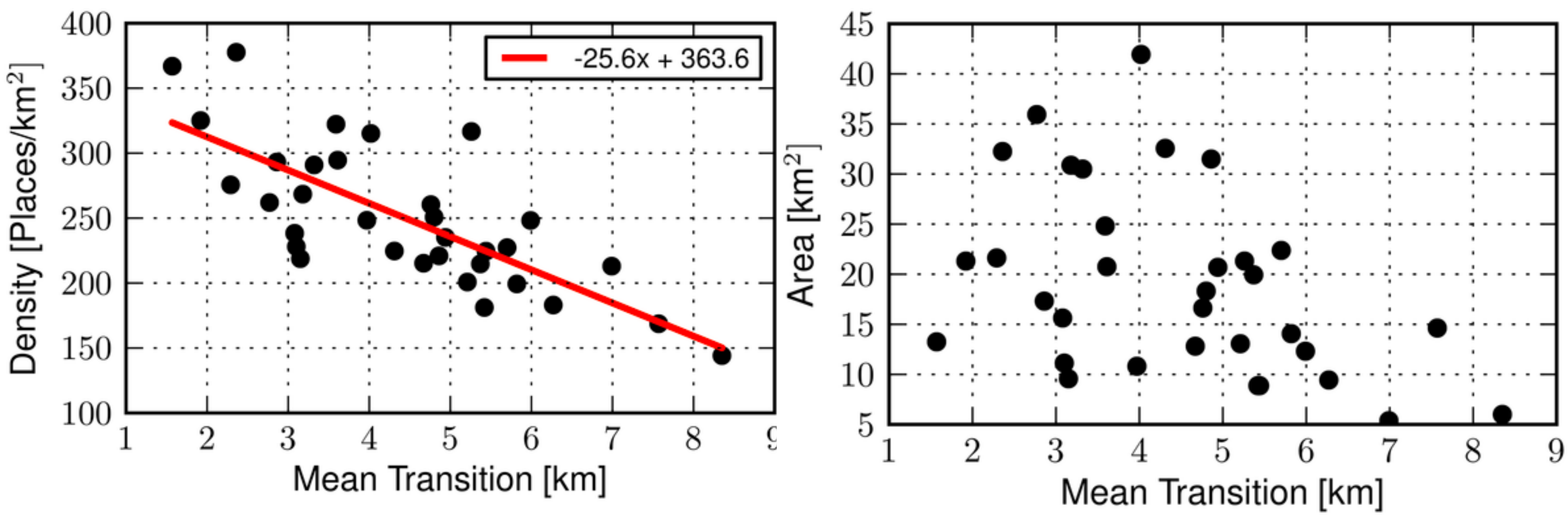
Stouffer's **law of intervening opportunities** states, *"The number of persons going a given distance is directly proportional to the number of opportunities at that distance and inversely proportional to the number of intervening opportunities." **



**Empirically proved over migration data from Cleveland.
Is it true in our data?**

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

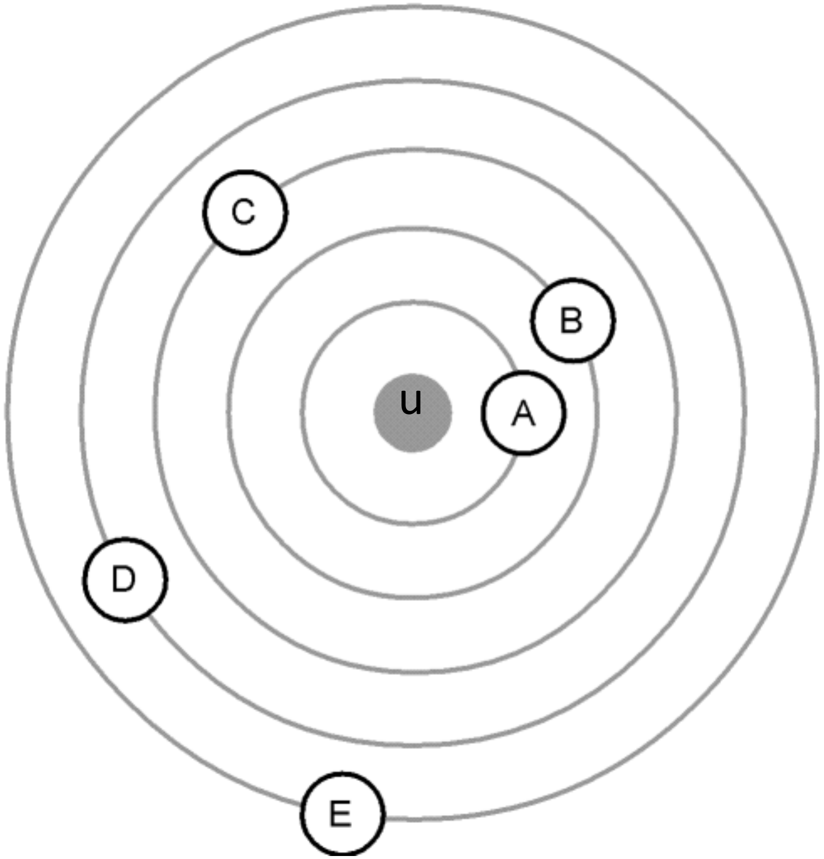
The importance of density



Place density by far more important than city area size with respect to mean length of human movements.

UNIVERSITY OF CAMBRIDGE

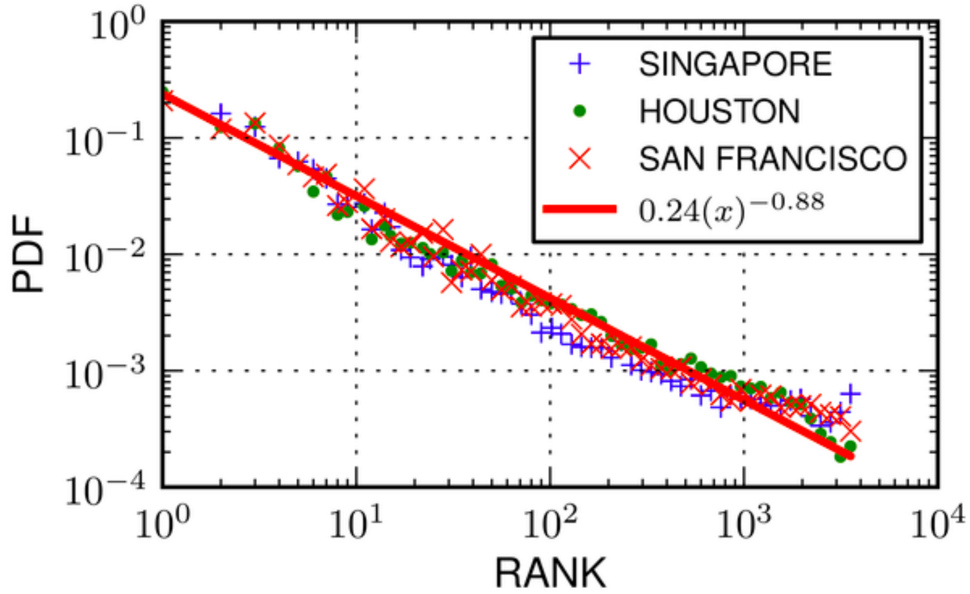
Rank distance

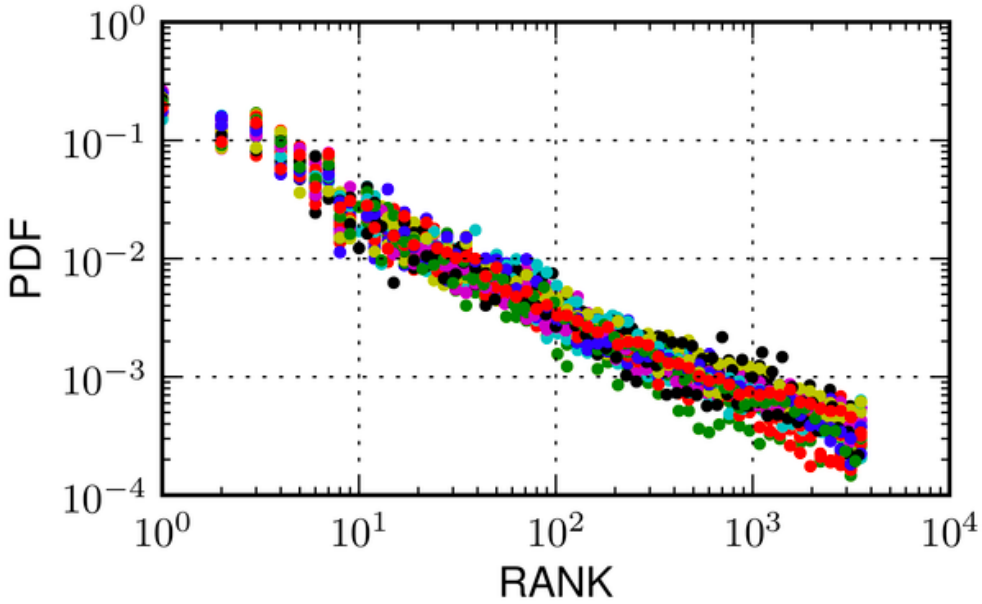


$$rank_u(v) = |\{w : d(u,w) < d(u,v)\}|$$

UNIVERSITY OF CAMBRIDGE

Rank universality

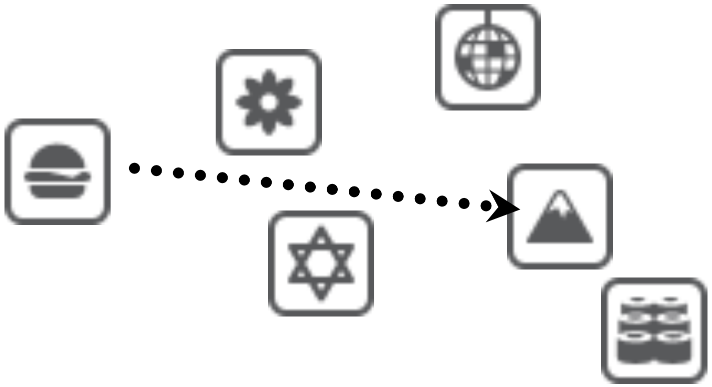




The rank of all cities collapses to a single line.
We have measured a power law exponent α = 0.84 ± 0.07

UNIVERSITY OF CAMBRIDGE

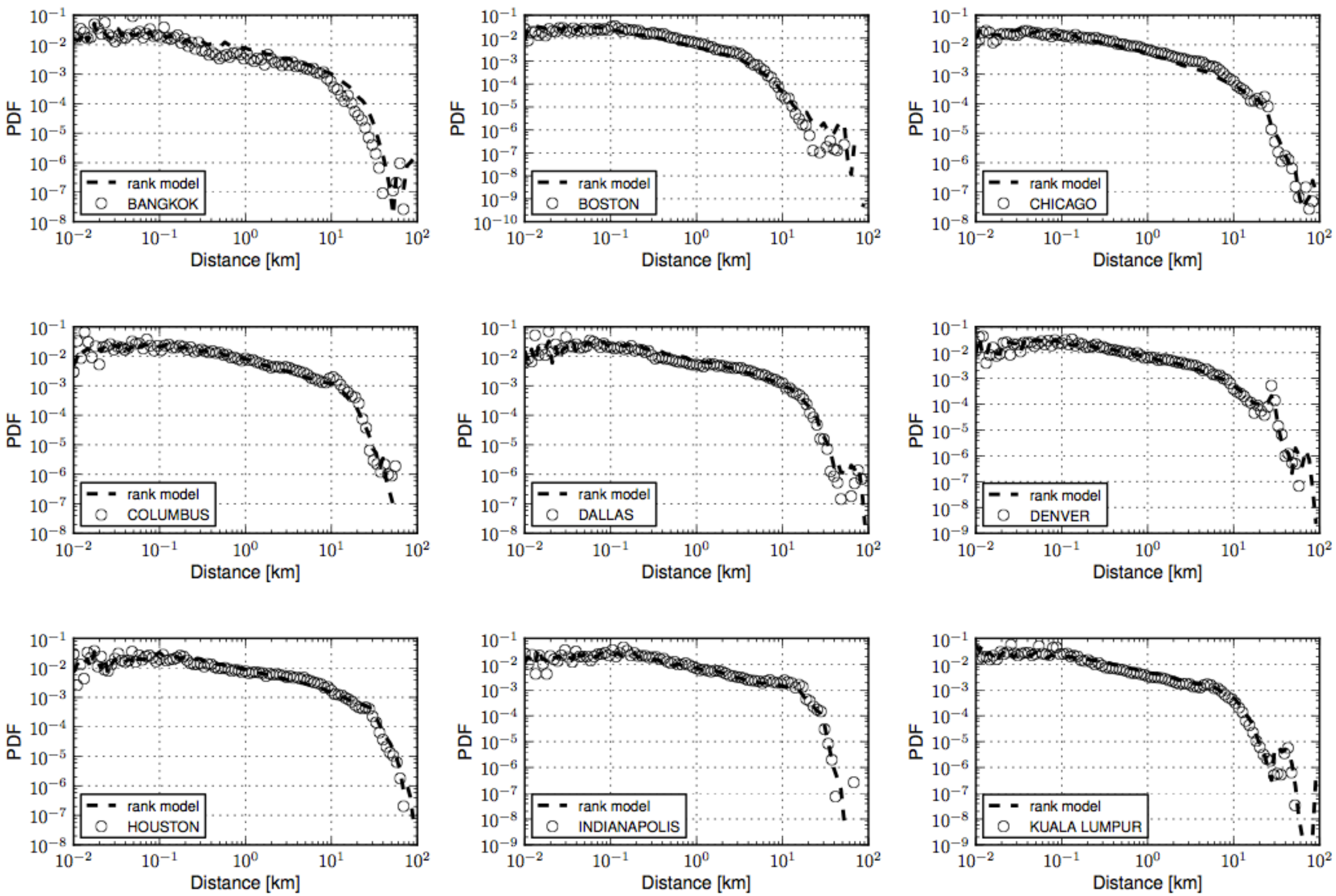
A new model for urban mobility



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

UNIVERSITY OF CAMBRIDGE

Simulation results ...



UNIVERSITY OF CAMBRIDGE

Modelling Geo Social Network Evolution

UNIVERSITY OF CAMBRIDGE

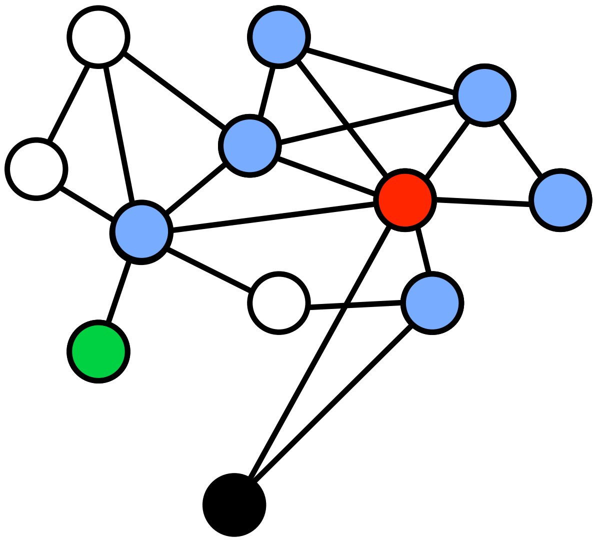
Evolution of a Location-based Online Social Network: Analysis and Models. Militiadis Allamanis, Salvatore Scellato and Cecilia Mascolo. In Proceedings of ACM Internet Measurement Conference (IMC 2012). Boston, MA. November 2012.

Geo-Social network evolution...

- How would Foursquare evolve in 2 months?
- What are the factor that shape geo-social network evolution?
- Why would we be interested in forecasting evolution?
    - Design of distributed storage solutions.
    - Delivery of user generated content.
    - Recommendation.

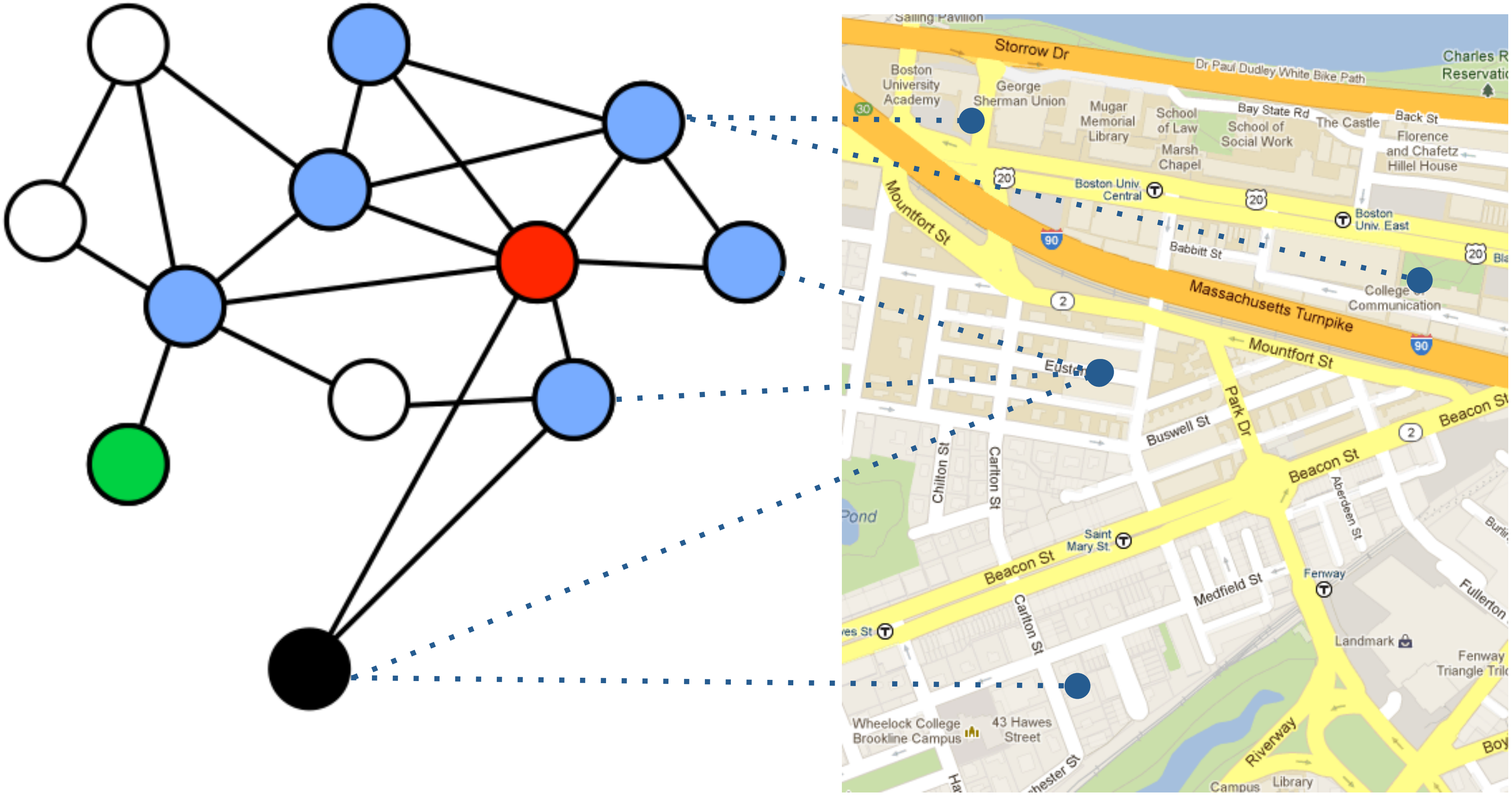
UNIVERSITY OF CAMBRIDGE

Modelling the growth of (online) social networks



- A common feature of social (and other) networks is a skewed degree distribution, and **preferential attachment** can reproduce it: popular users attract more links.
- This fails to generates triangles, which are instead present: **triadic closure** needs to be introduced.
- Several variations of these ideas have been explored to model social networks.

Attachment in geo-social networks



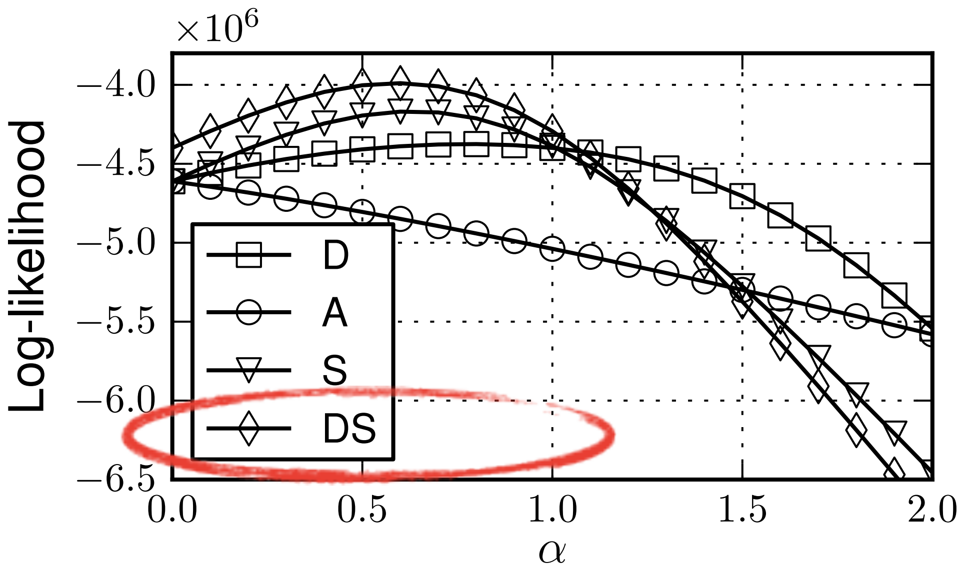
Temporal evolution of a spatial social network

- Daily snapshots of Gowalla data May to Aug. 2010. Information about user profiles, friends and check-ins.
- We study temporal network growth:
  - social edges creation and speed
  - social triangles creation
  - mobility and space impact

Properties at the end of measurement period	
Nodes	122,414
Social links	580,446
Average degree	9.48
Average clustering coefficient	0.254
Average distance between friends [km]	1,792
Average distance between users [km]	5,663

Global attachment models

For each **new edge** created from user A to user B we compute the probability of being created according to different models (for different parameters).

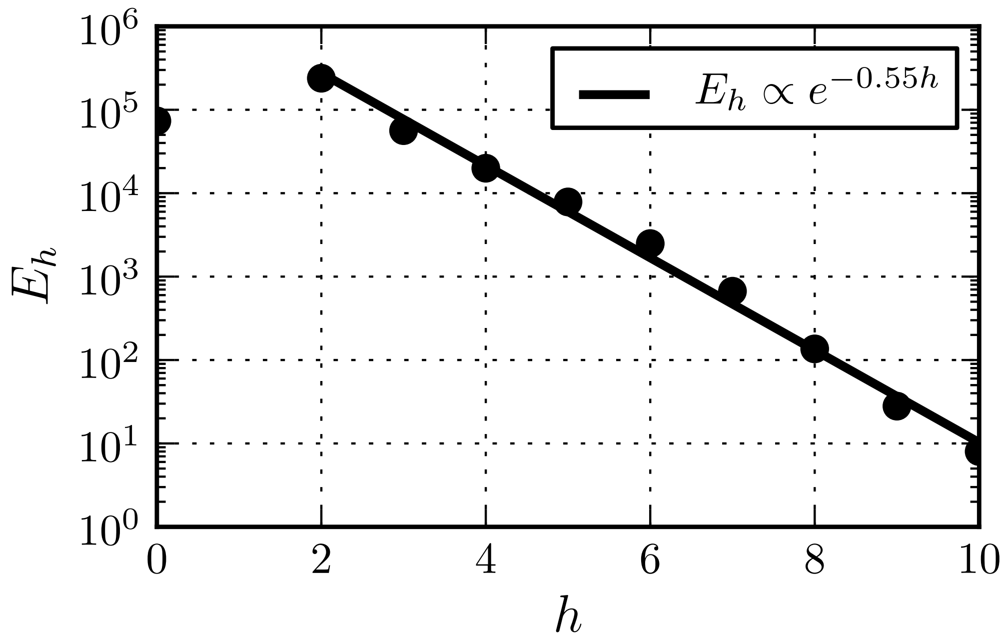


- **D**: proportional to a power α of the **degree** of user B
- **A**: prop. to a power α of the **age** of user B
- **S**: inversely prop. to a power α of the geographic **distance** between A and B
- **DS**: prop. to the **degree** of user B and inversely prop. to a power α of the geographic **distance** between A and B.

UNIVERSITY OF CAMBRIDGE

Predominance of triangle-closing links

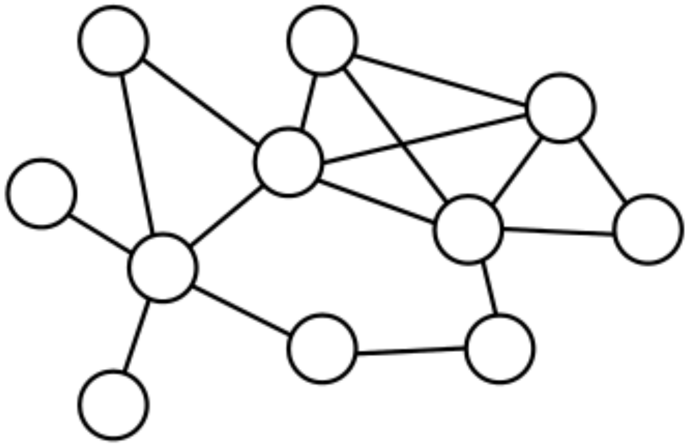
Number of new links



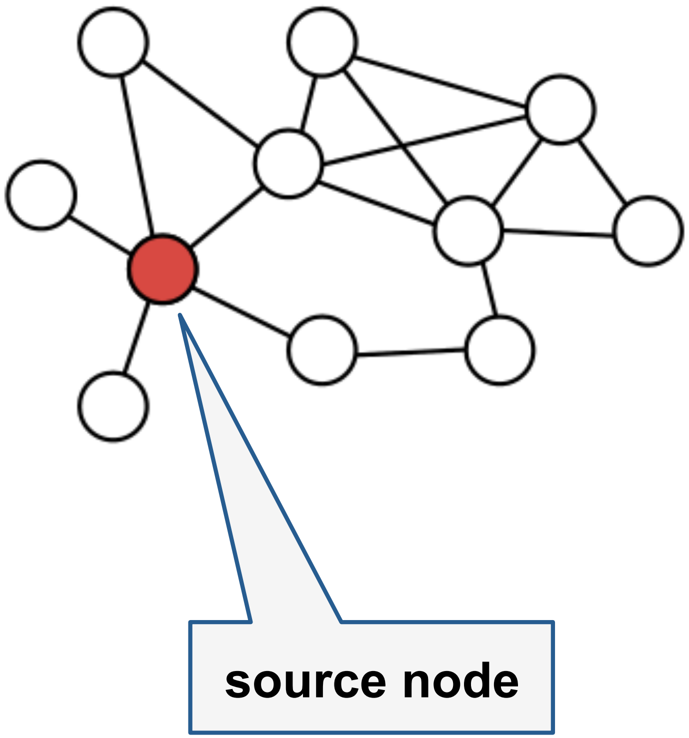
New edges are exponentially more likely to connect people sharing at least one friend, creating social triangles.

UNIVERSITY OF CAMBRIDGE

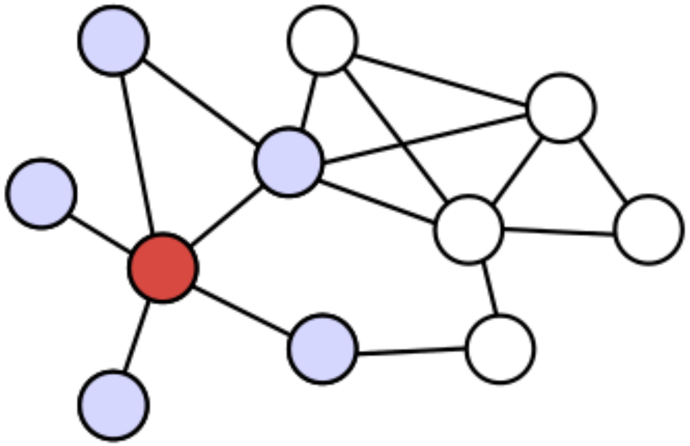
Social attachment



Social attachment

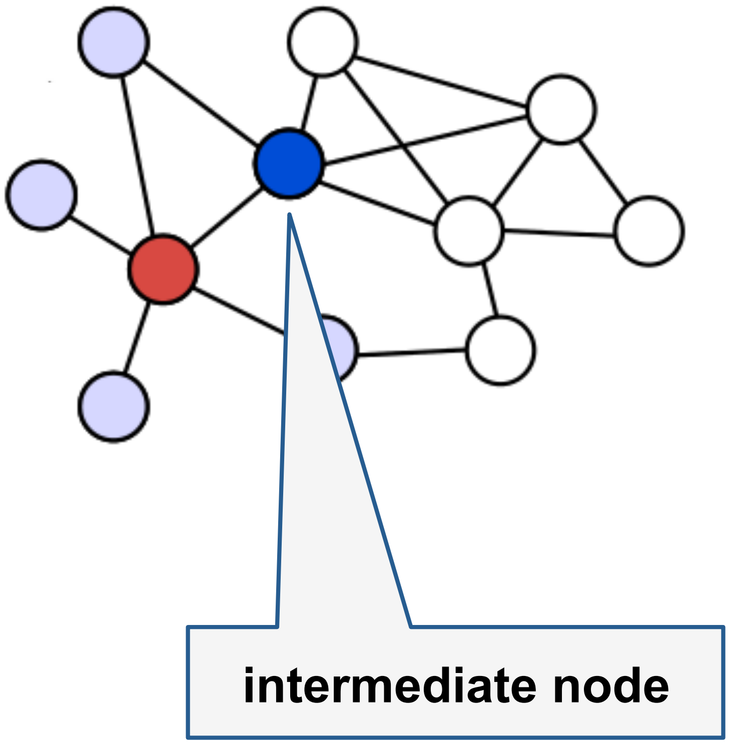


Social attachment

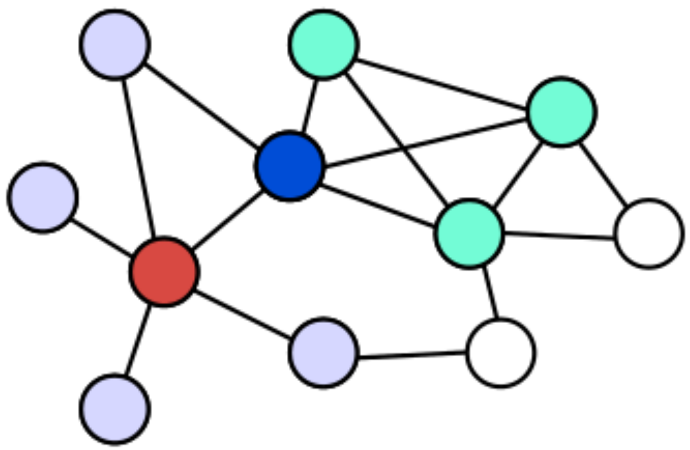


UNIVERSITY OF CAMBRIDGE

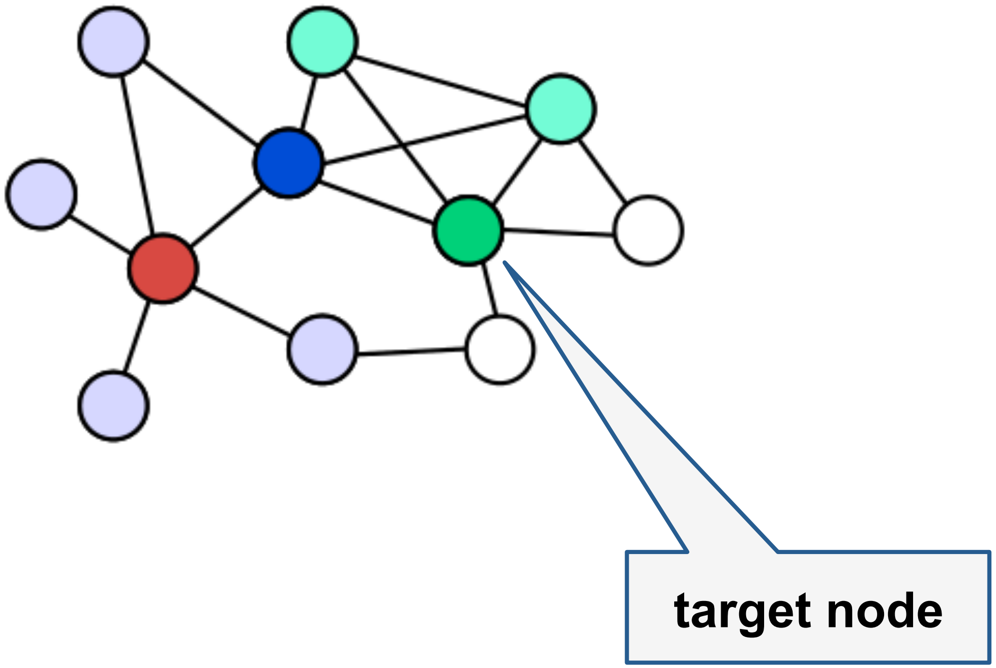
Social attachment



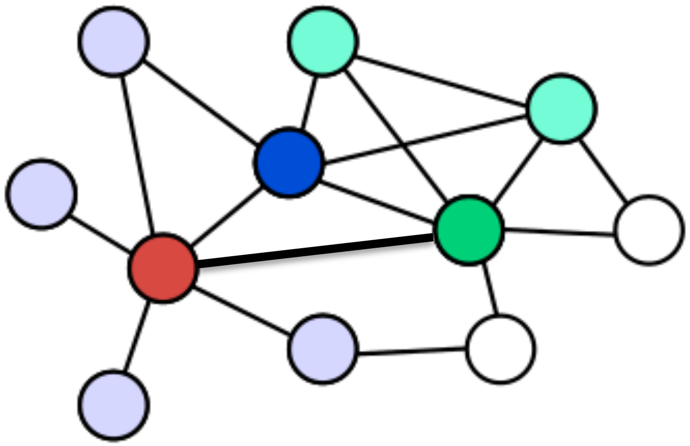
Social attachment



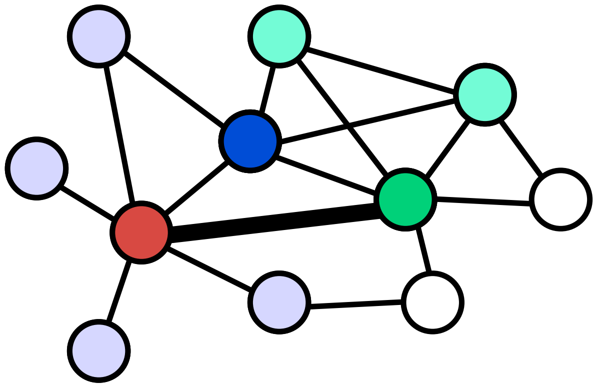
Social attachment



Social attachment



Social attachment models



Intermediate node models

Percentage improvement on random choice (2 hop)

	random	shared	degree	distance	gravity
random	**12.34**	9.48	-3.47	-28.17	-35.26
shared	**14.54**	11.47	-0.95	-24.74	-34.46
degree	7.33	5.16	-6.79	-25.17	-41.98
distance	-0.92	-3.70	-16.94	-39.32	-41.53
gravity	2.71	0.25	-12.11	-33.01	-43.18

Triadic closure is mainly driven by social processes, while geographic distance is not an important factor.

UNIVERSITY OF CAMBRIDGE

What about geography?

30% of new edges are established between users that share at least one **common place**.

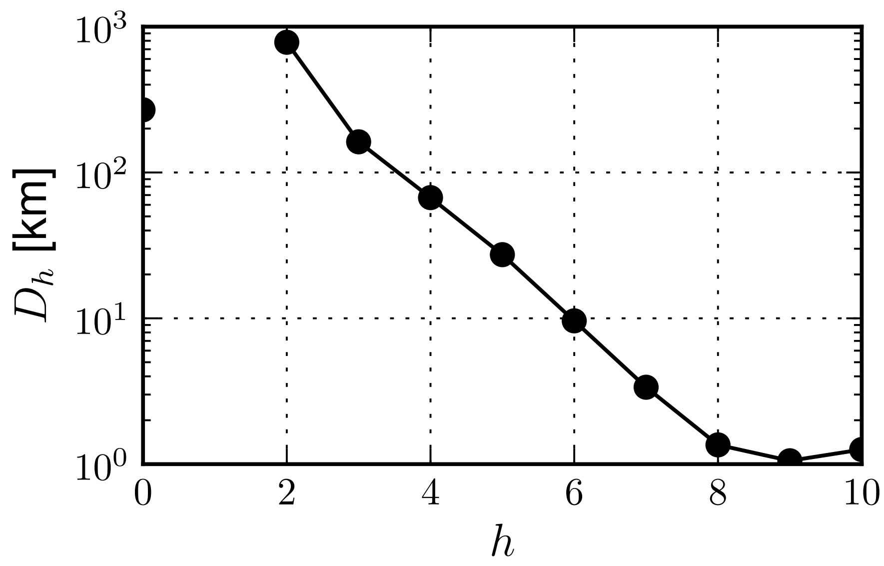
10% of new links are created between users that do share **common places**, but *no* common friends

A **social only** model would **fail** to reproduce that users create new social connections beyond their 2-hop neighborhood.

UNIVERSITY OF CAMBRIDGE

Effect of distance

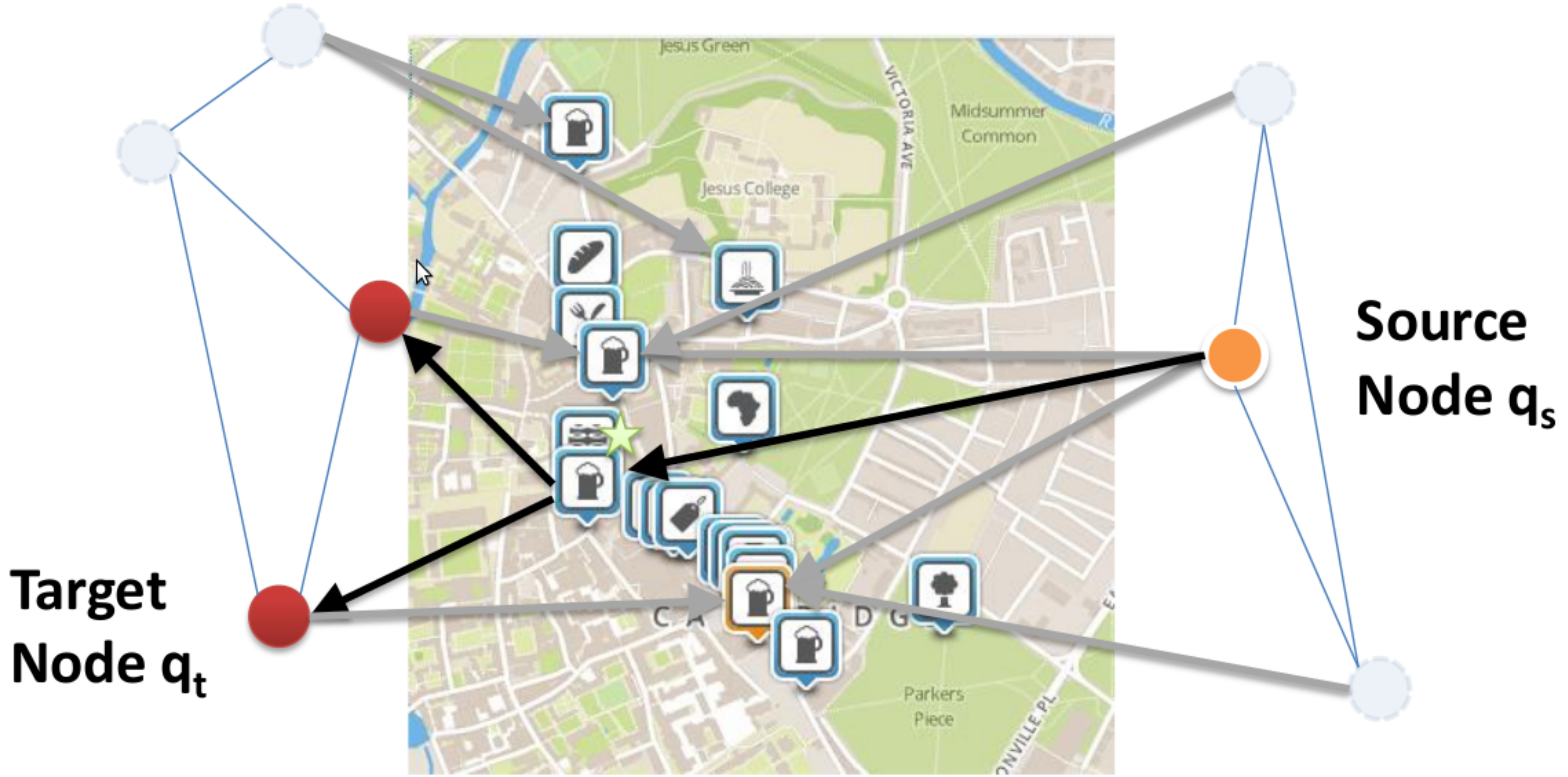
Distance of new links



Geographic proximity appears complementary to social closeness: **being close in space connects people at large social distances.**

UNIVERSITY OF CAMBRIDGE

Mobility-driven attachment



Choosing friends geographically

Select users who visit the same places

Select a visited place:

- Visited by many friends
- Visited by the user many times
- Very popular
- Close

Select a user:

- Popular
- Active
- Close to the place

Place-user choice

Choice of Place

Choice of user

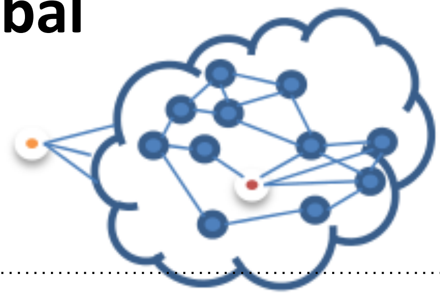
	random	degree	deg-diffusion	user-checkins	tot-checkins	inv-tot-checkins	distance	gravity
random	0.28	6.88	9.24	0.16	-17.02	-4.51	-19.36	-7.04
friends	4.70	**11.60**	**13.63**	4.74	-10.63	-1.56	-14.88	-1.71
user-checkins	0.05	6.59	8.94	-0.03	-17.27	-4.80	-19.69	-7.41
tot-checkins	6.09	**13.13**	**15.18**	6.14	-9.29	0.04	-13.15	-0.02
tot-users	5.10	**12.33**	**14.33**	5.16	-9.96	-1.08	-14.19	-0.84
place-distance	-23.41	-15.57	-13.21	-23.56	-40.82	-28.27	-43.67	-30.17
place-gravity	0.37	7.22	9.46	0.32	-16.26	-5.29	-19.60	-6.81

proportional to user's degree and inversely proportional the logarithm of user's total number of visited places;

UNIVERSITY OF CAMBRIDGE

Putting it all together

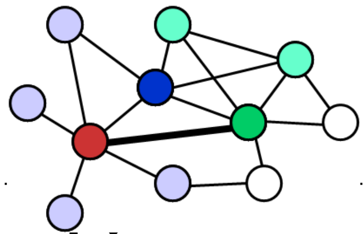
Global



gravity model

Distance & degree are important on a global level

Local Social



random-random model

Triangle closure not affected by distance

Mobility-driven

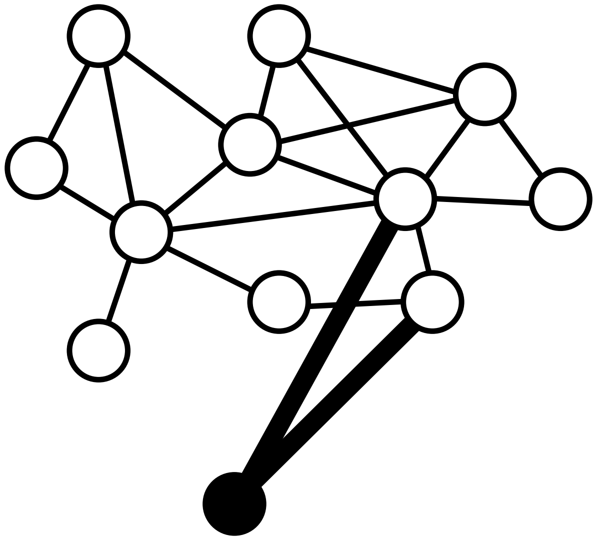
tot-checkins - degree model

Popular places are important

Small-scale preferential attachment

UNIVERSITY OF CAMBRIDGE

Putting all pieces together: a new growth model



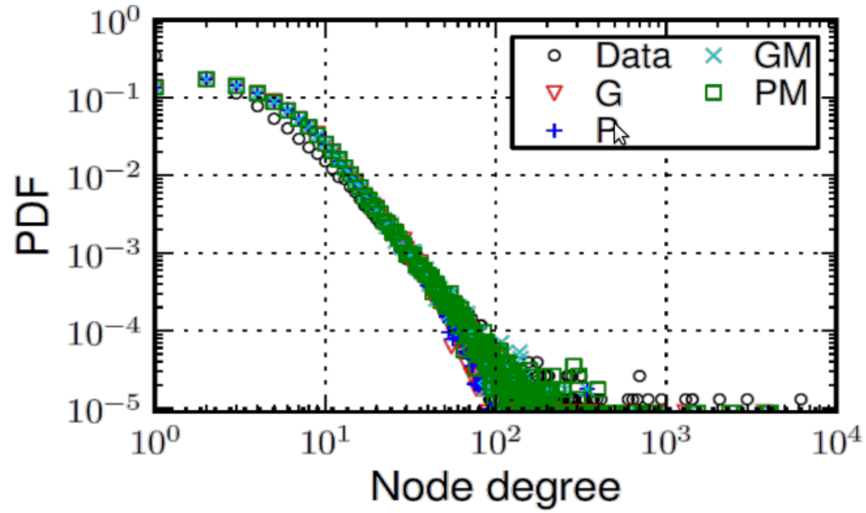
1. A **new node** joins the network and positions itself over the space;
2. It samples its **lifetime** from an exponential distribution;
3. The new node adds its first edge according to a **preferential attachment** or **gravity model**;
4. The node samples a **time gap** from the degree-dependent distribution and then goes to sleep for that time gap;
5. When a node wakes up, if its lifetime has not expired yet with probability p the node uses the random−random social triangle-closing model, otherwise it uses the tot-checkins − degree mobility- based closure.
6. The node repeats step 4.

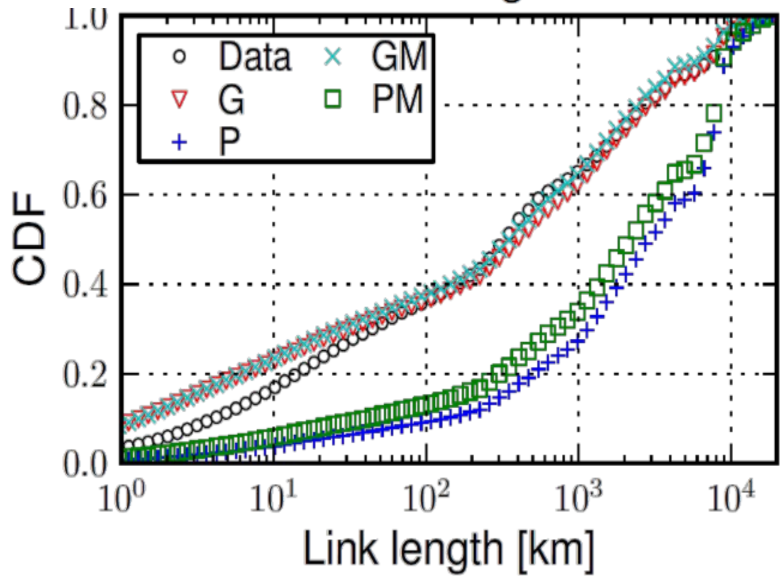
UNIVERSITY OF CAMBRIDGE

Degree and link length distribution

Degree distribution is generally similar for all models

Link length shows the precision of gravity based models

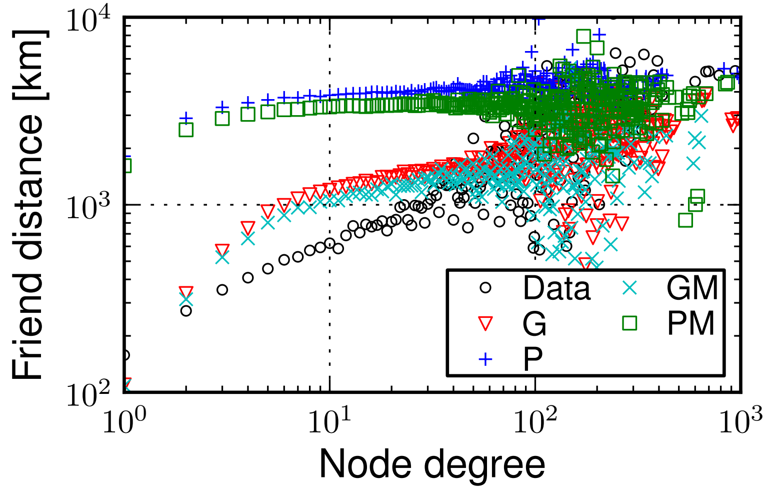


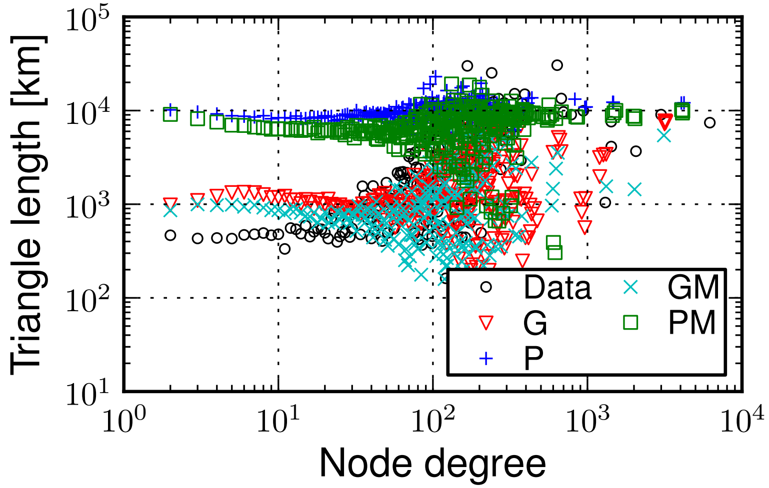


UNIVERSITY OF CAMBRIDGE

Friend distance and triangle length

Friend distance and triangle length of gravity based models correlation with degree are matching the data





Understanding community evolution and role of places

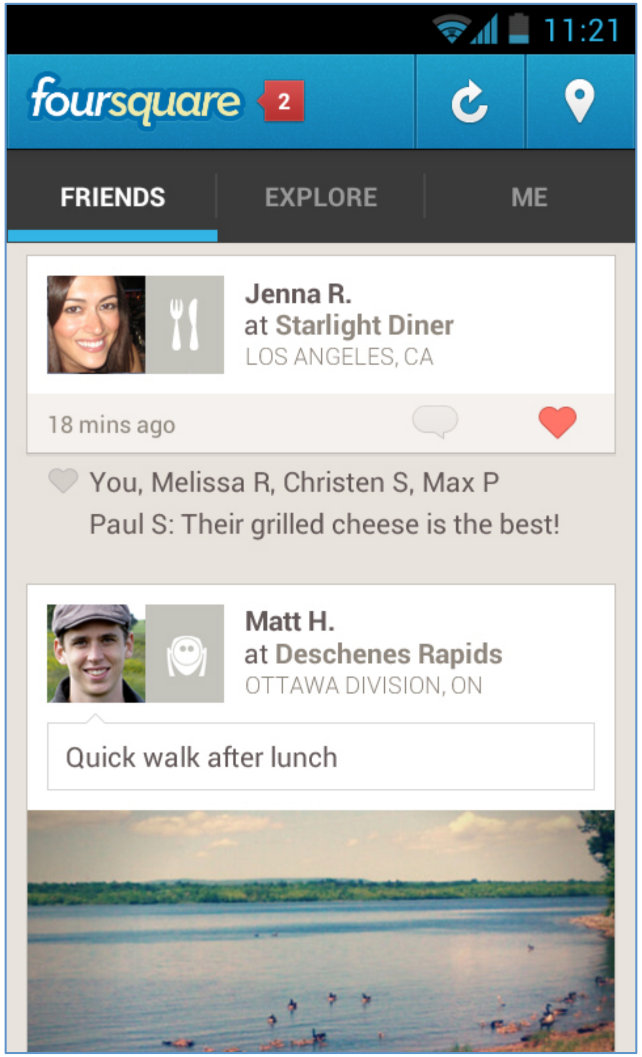
UNIVERSITY OF CAMBRIDGE

A Place-focused Model for Social Network Formation in Cities. Chloë Brown, Anastasios Noulas, Cecilia Mascolo, Vincent Blondel. NetMob 2013. Boston, MA. May 2013.

Place-friend vs social networks

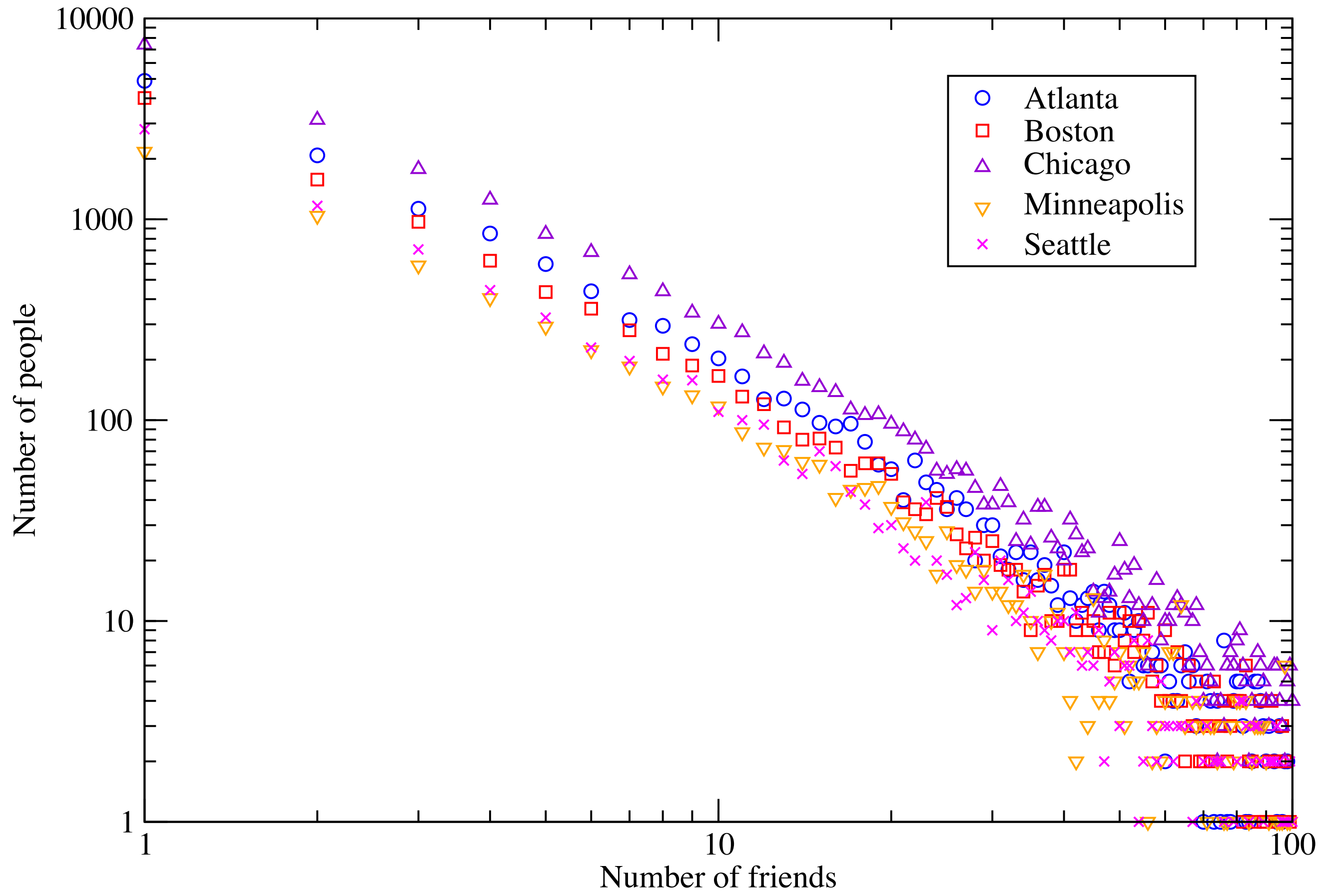
- People who have checked in at a place in a given city, and their friends **who have also checked in** at those places.

What do these place-friend networks look like?



Place-friend networks

- **Degree**: power-law distribution

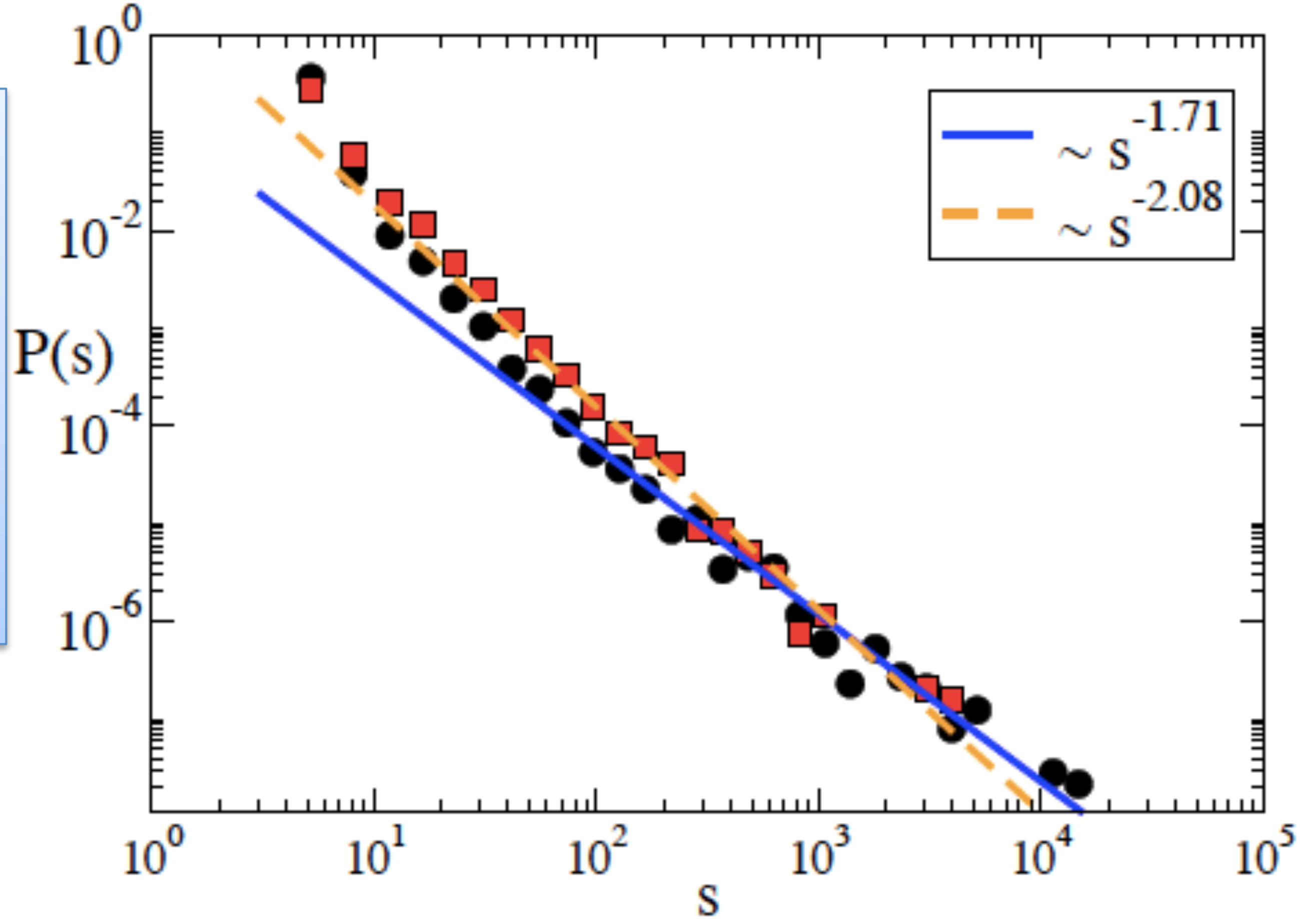


UNIVERSITY OF CAMBRIDGE

Distribution of community sizes

Black: social network communities.
Red: place-friend (local) communities.

Size of local communities is consistently smaller.



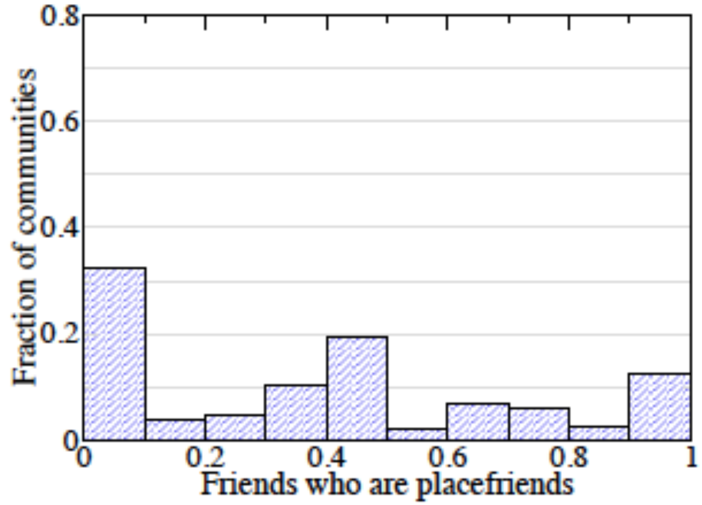
Intra-community links

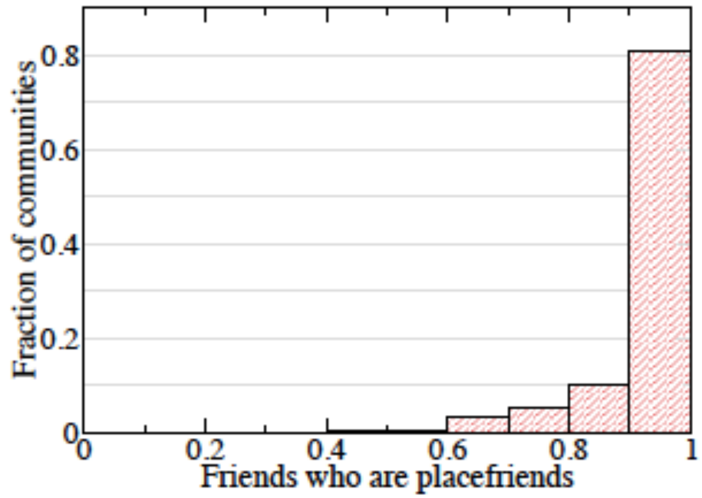
Proportion of communities having intra-community ties which are place friends.

More than 30% of social communities have less than 10% placefriends.

More than 80% of local communities have more than 90% placefriends.

Community Detection on the social graph might not capture local communities.

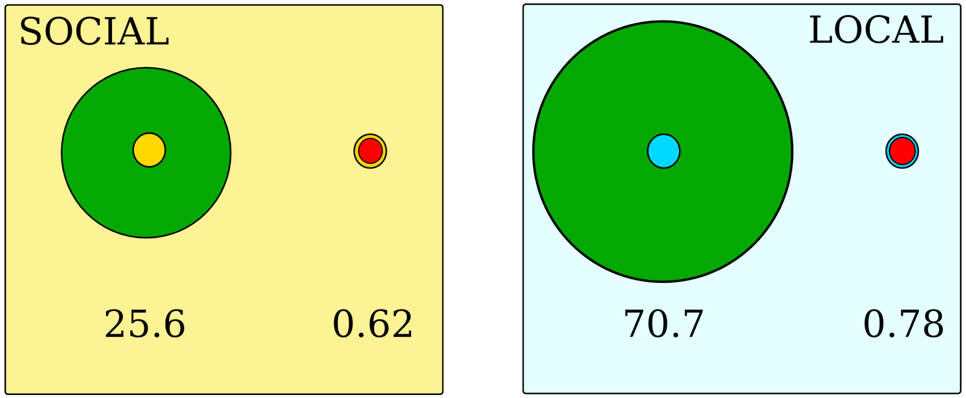




UNIVERSITY OF CAMBRIDGE

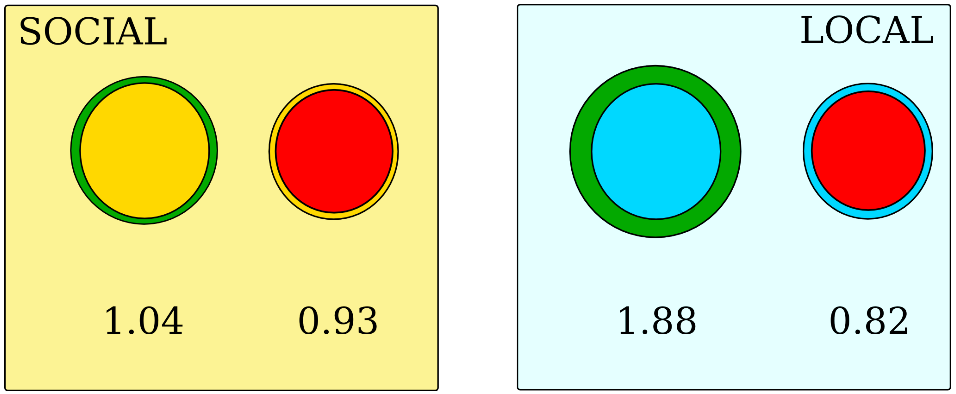
Temporal community evolution

FORMED EDGES



Number of appearing links inside communities wrt to random appearance of links:
Local communities could be very good predictors.

FRIENDS -> PLACEFIRENDS

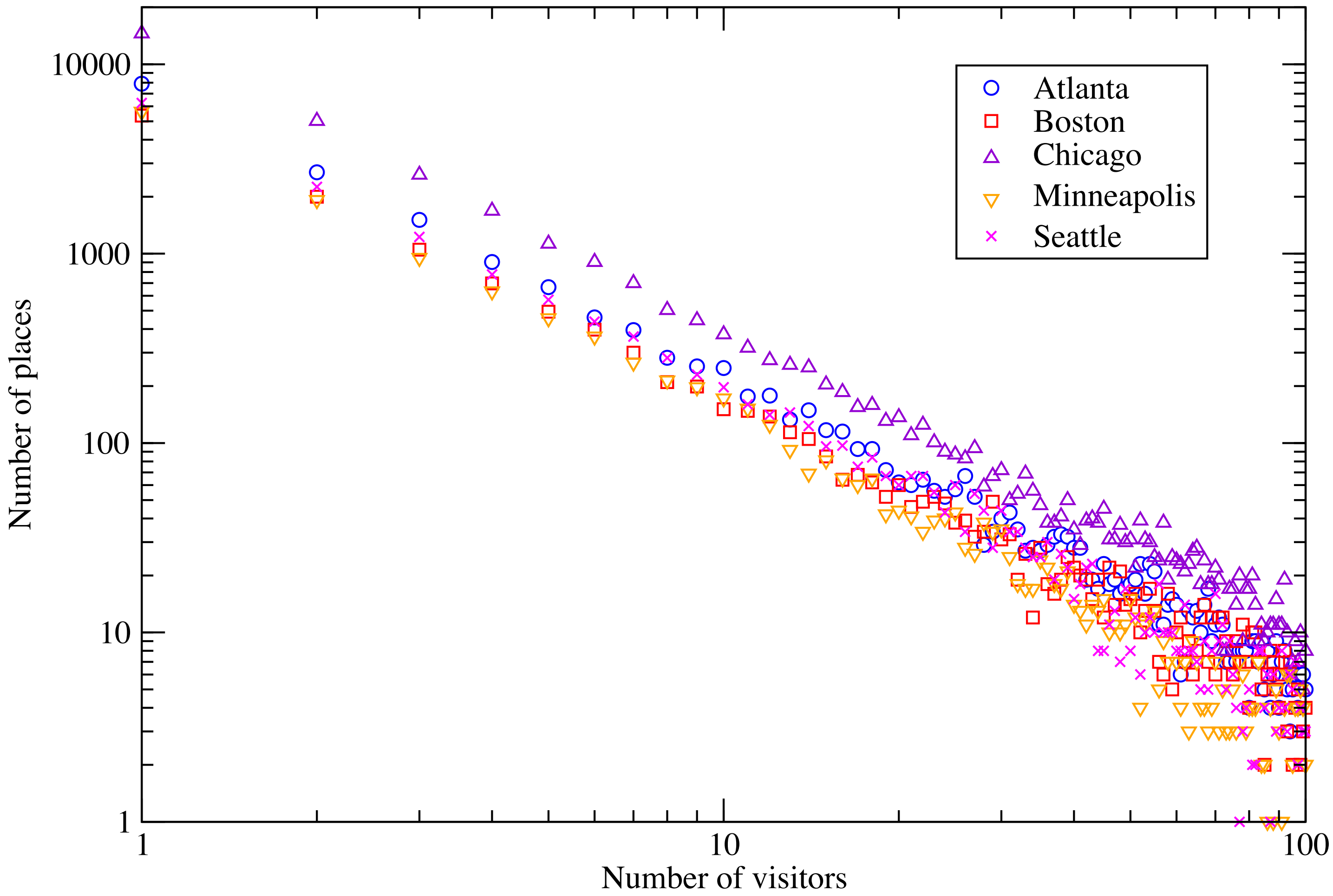


Number of placefriends appearing**: local communities are again good predictors.**

UNIVERSITY OF CAMBRIDGE

The role of places

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



UNIVERSITY OF CAMBRIDGE

Places vital for tie formation

>70% of triangles have one place shared between all three people.

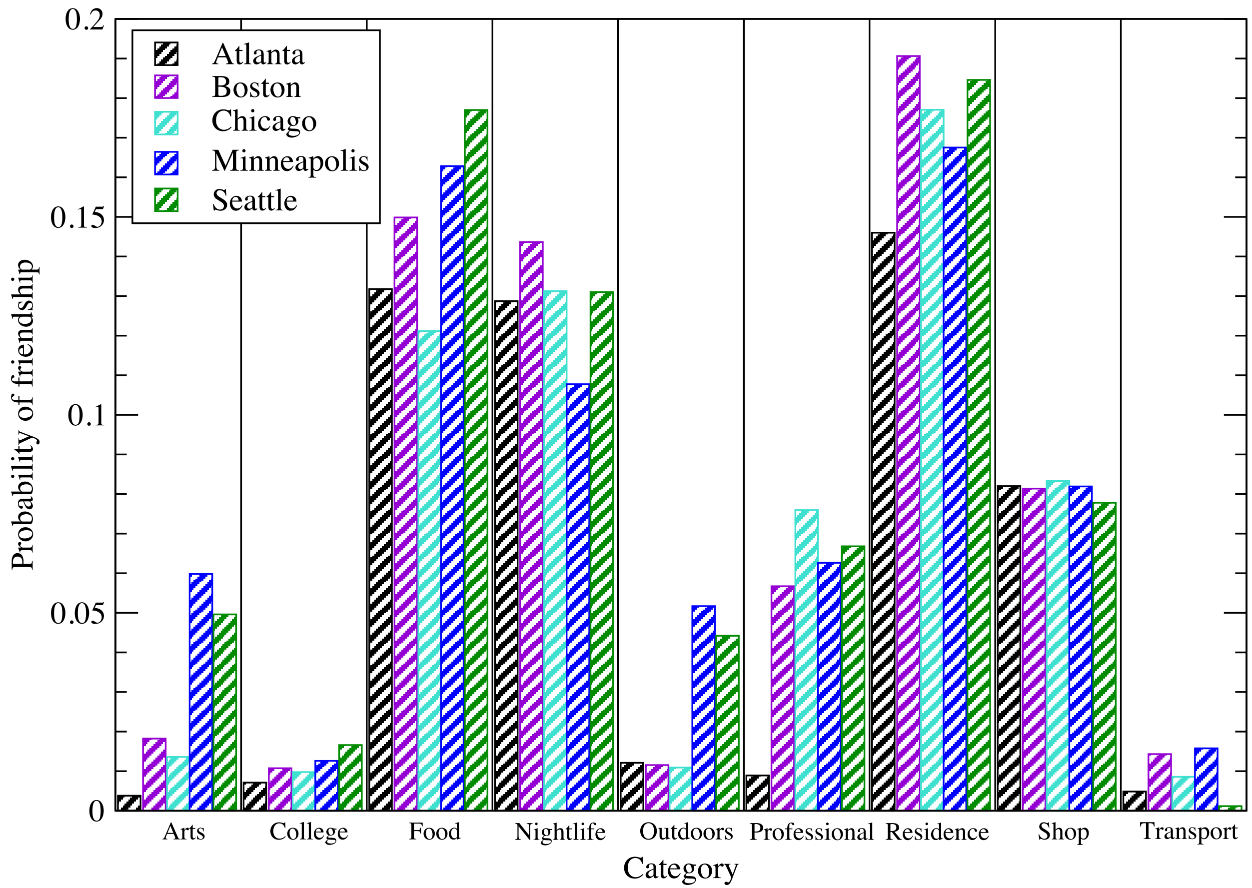
Clustering around certain places

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.

Friend 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.

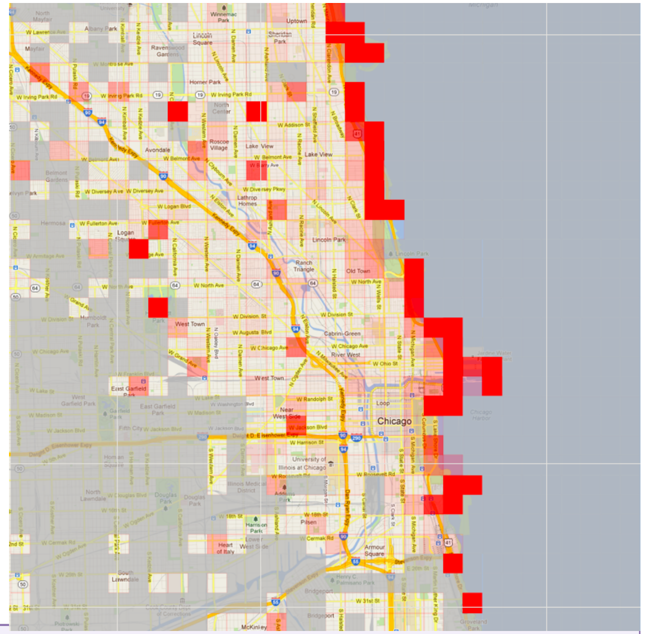
The Importance of Being Placefriends: Discovering Location-focused Online Communities. Chloë Brown, Vincenzo Nicosia, Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo. In ACM SIGCOMM Workshop on Online Social Networks (WOSN 2012). Helsinki, Finland. August 2012.

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.

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.

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). Full Paper. Dublin, 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). Helsinki, Finland. August 2012.

References

- **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. doi:10.1371/journal.pone.0037027
- **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.
- **A Place-focused Model for Social Network Formation in Cities**. Chloë Brown, Anastasios Noulas, Cecilia Mascolo, Vincent Blondel. NetMob 2013. Boston, MA. May 2013.
- **Social and place-focused communities in location-based online social networks**. Chloë Brown, Vincenzo Nicosia, Salvatore Scellato, Anastasios Noulas, Cecilia Mascolo. To appear in European Physical Journal B.

UNIVERSITY OF CAMBRIDGE

Acks..of course!

Militiadis Allamanis

Vincent Blondel

Chloë Brown

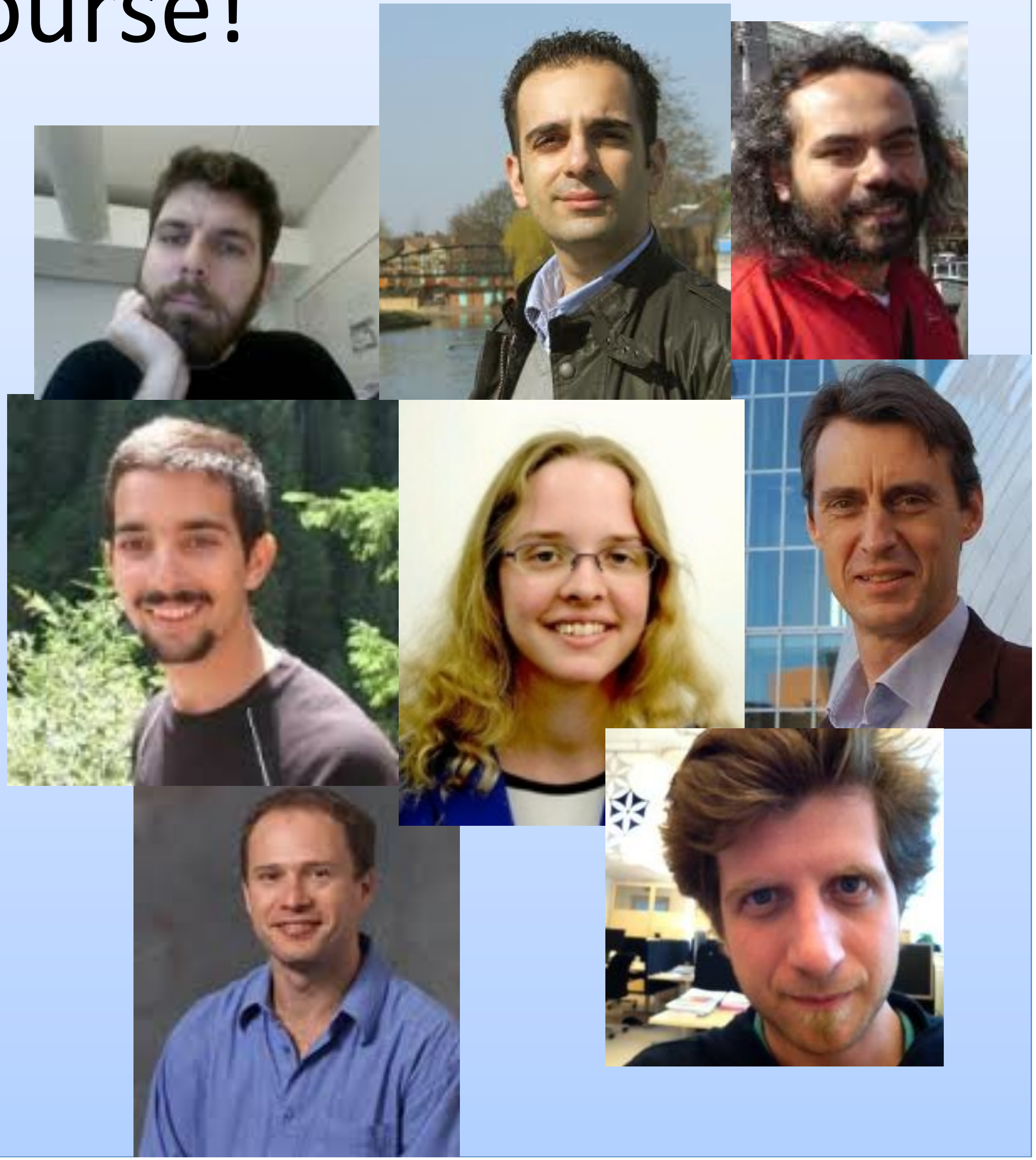
Renaud Lambiotte

Vincenzo Nicosia

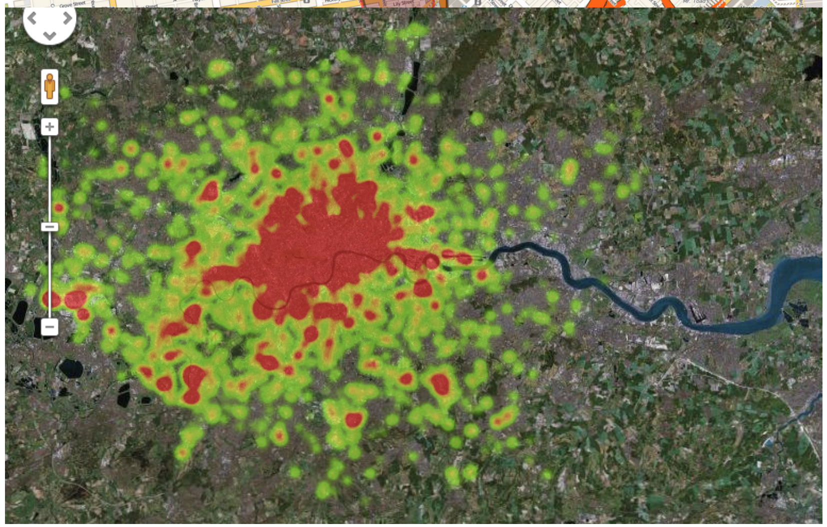
Anastasios Noulas

Massi Pontil

Salvatore Scellato



Thanks! Questions?



Cecilia Mascolo

@cecim
cecilia.mascolo@cl.cam.ac.uk
www.cl.cam.ac.uk/users/cm542

UNIVERSITY OF CAMBRIDGE