

# Socio-spatial Properties of Online Location-based Social Networks

---

**Salvatore Scellato**<sup>1</sup>, Anastasios Noulas<sup>1</sup>, Renaud Lambiotte<sup>2</sup>, Cecilia Mascolo<sup>1</sup>  
Computer Laboratory, University of Cambridge<sup>1</sup>  
Department of Mathematics, University of Namur<sup>2</sup>

5th International AAI Conference on Weblogs and Social Media  
**Barcelona, 18th July 2011**





Location-based social networks

More and more people want to share their geographic position with their friends.

## Foursquare Closing In On 7 Million Users

Nearing A Million Users, Gowalla Moves Forward On The “Social Atlas” Path

Location-based app Gowalla partners with Disney

Gowalla Snaps Back: 1 Million Photos ...But Foursquare May Hit That In 2 Weeks

« Official 4sqDay Badge revealed

Today only: save up to 20% at Walgreen's just for checking in »

## Foursquare hits 8 million users

Gowalla shoots for 1 million users with new version of location service

2 Years And 10 Million Users Later, Google Latitude Locates The Check-In

Gowalla Snaps Back: 1 Million Photos ...But Foursquare May Hit That In 2 Weeks

Wow! The foursquare community has over 10,000,000 members!



More and more importance.

Growing levels of popularity, millions of users and the attention of media and investors.



facebook

December 2010

facebook

December 2010

Insight into spatial social behavior

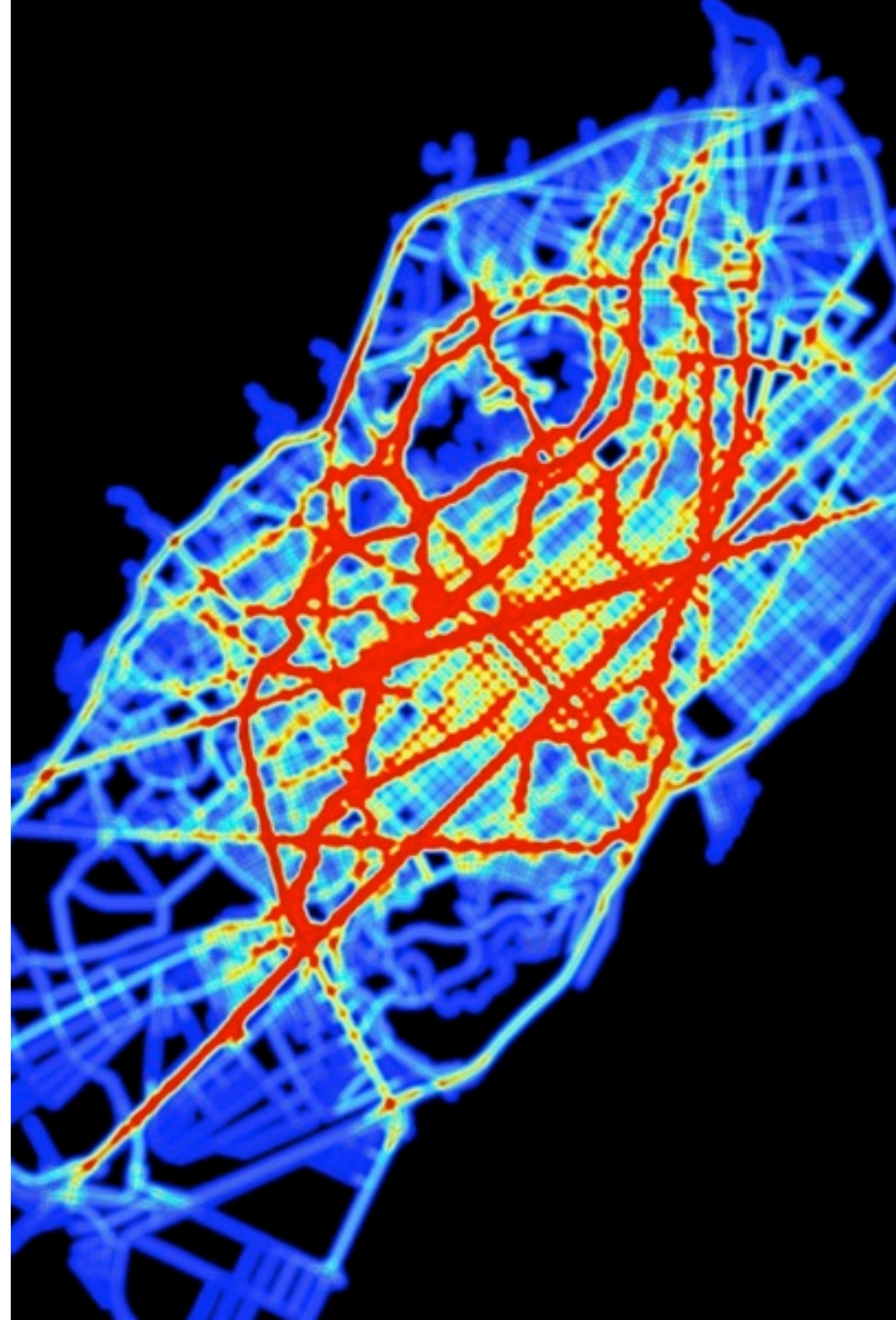
Availability of social and geographic data opens new research opportunities.



# Spatial networks

---

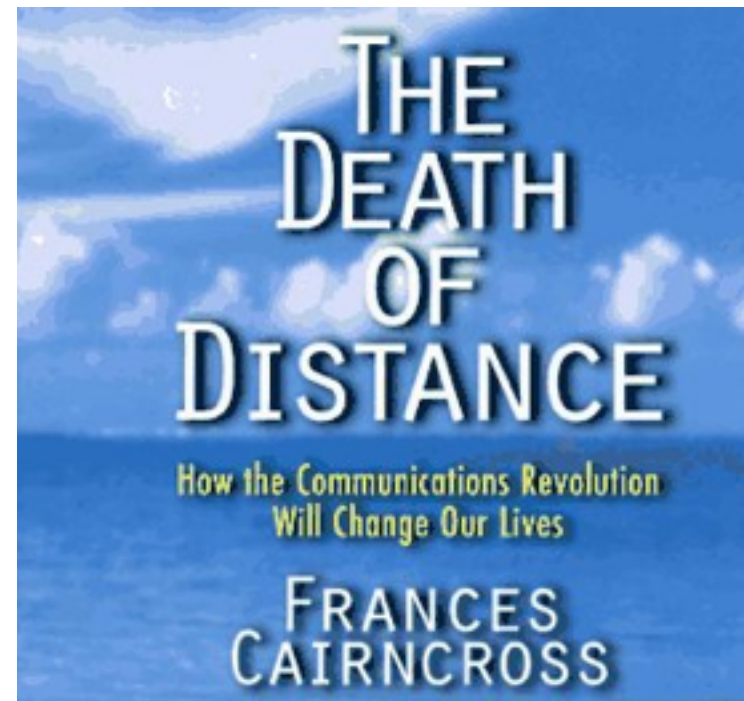
- In a **spatial network** nodes are embedded in a metric space: transportation systems, electric power grids, urban road networks, etc.
- **Metric distance** directly influences the network structure by imposing higher costs on the connections between distant nodes.



# Social ties and geographic distance

---

- A popular assumption is that most individuals try to **minimize the efforts** to maintain a friendship by interacting more with their spatial neighbors.
- The **connection costs** imposed by distance in spatial networks are not as important in social networks.
- Online tools and long-distance travel might result in the **“Death of Distance”**.



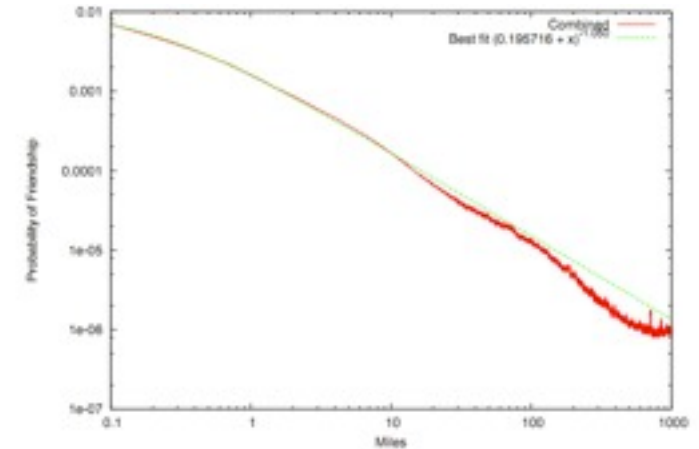
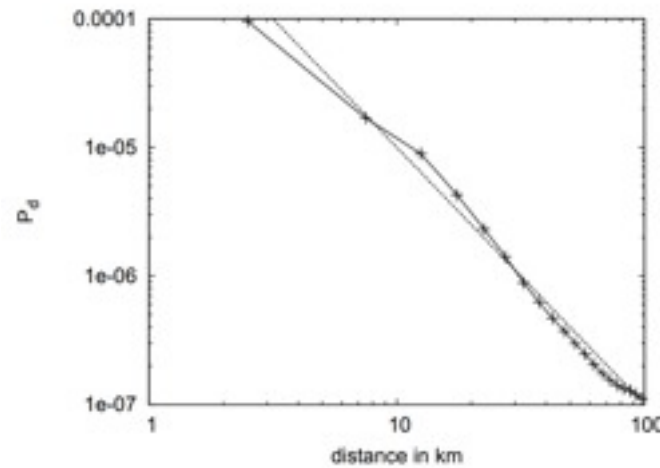
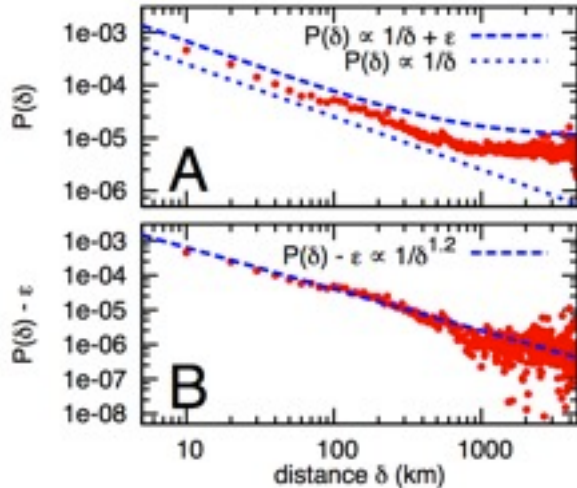
# Effect of distance on social connections

One fundamental spatial property of social networks is the **probability of friendship between two individuals as a function of their geographic distance**.

Livejournal (2005)

Mobile phones (2008)

Facebook (2010)



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

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

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

# Interesting questions...

---

- How is **geographic distance** affecting social ties in online location-based networks?
- Do users exhibit **homogeneous** or **heterogeneous** socio-spatial properties?
- What are the spatial properties of **social triads**?
- How are **spatial and social factors** simultaneously shaping how individuals create their connections?

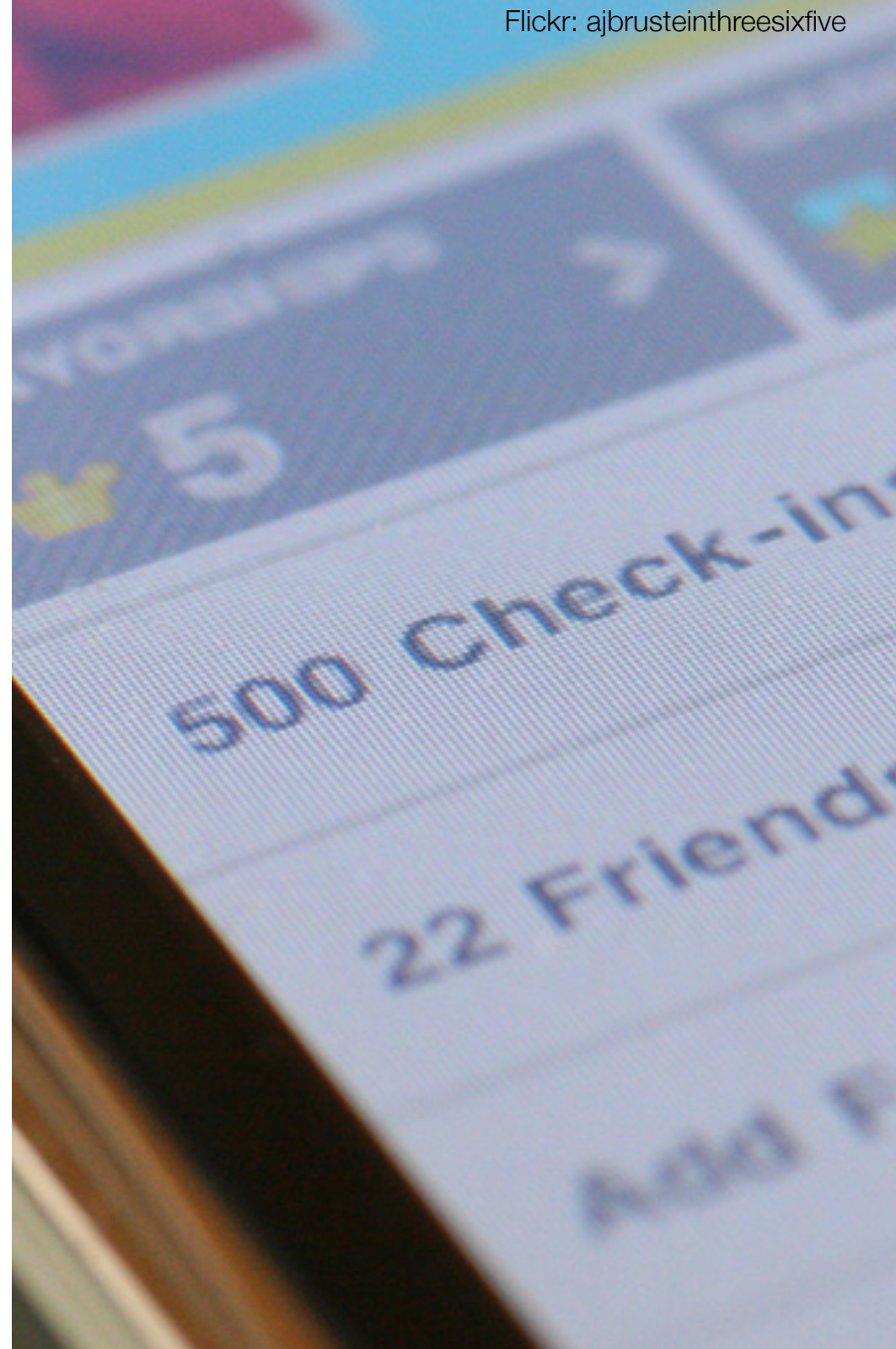







# Approach

---

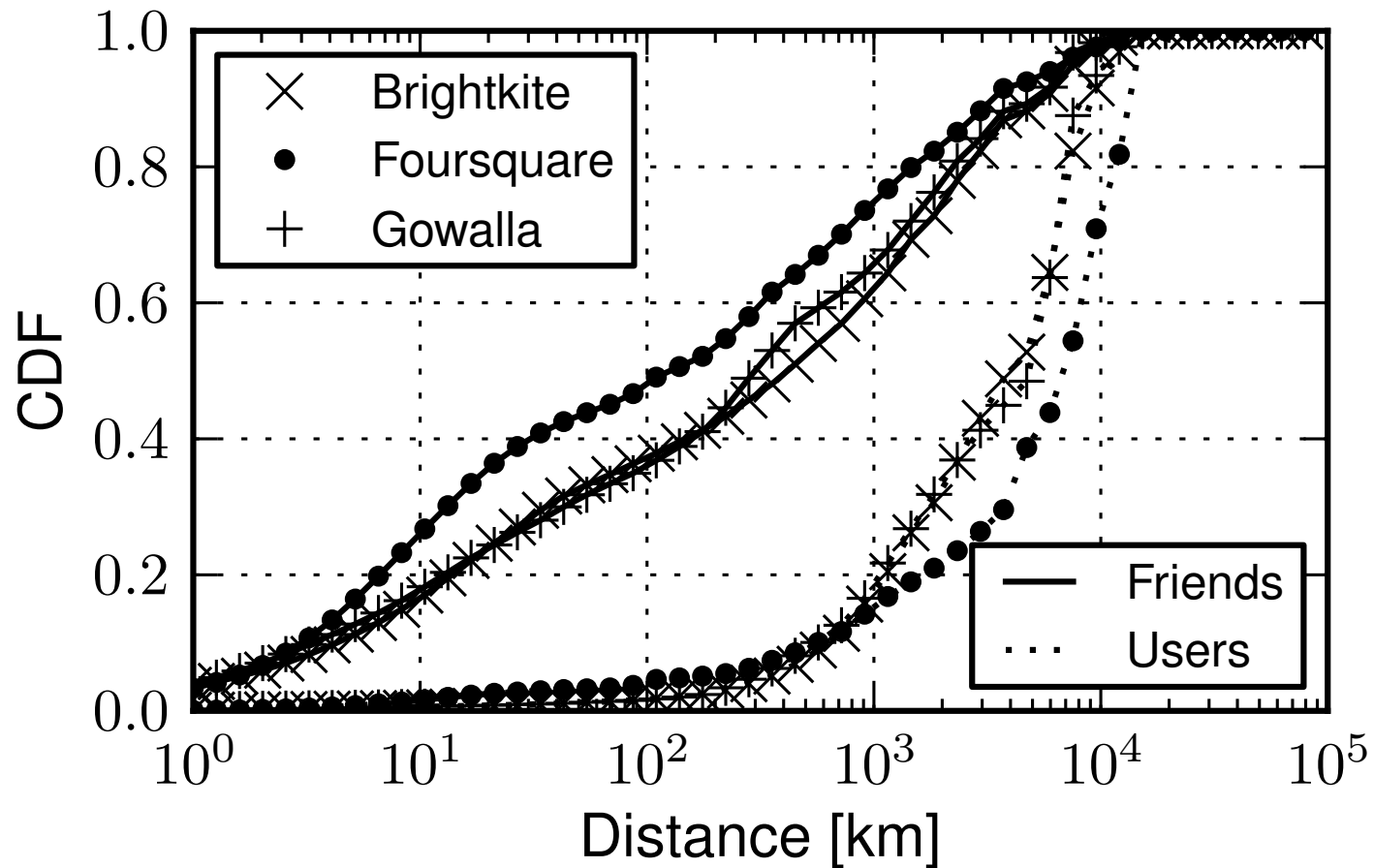
- We have acquired data about the socio-spatial network of **3 real-world location-based services**
- We design **two randomized models** of a socio-spatial network to better understand which factors shape the real networks.
- We study how individual users create their **social links** and their **social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



# Datasets

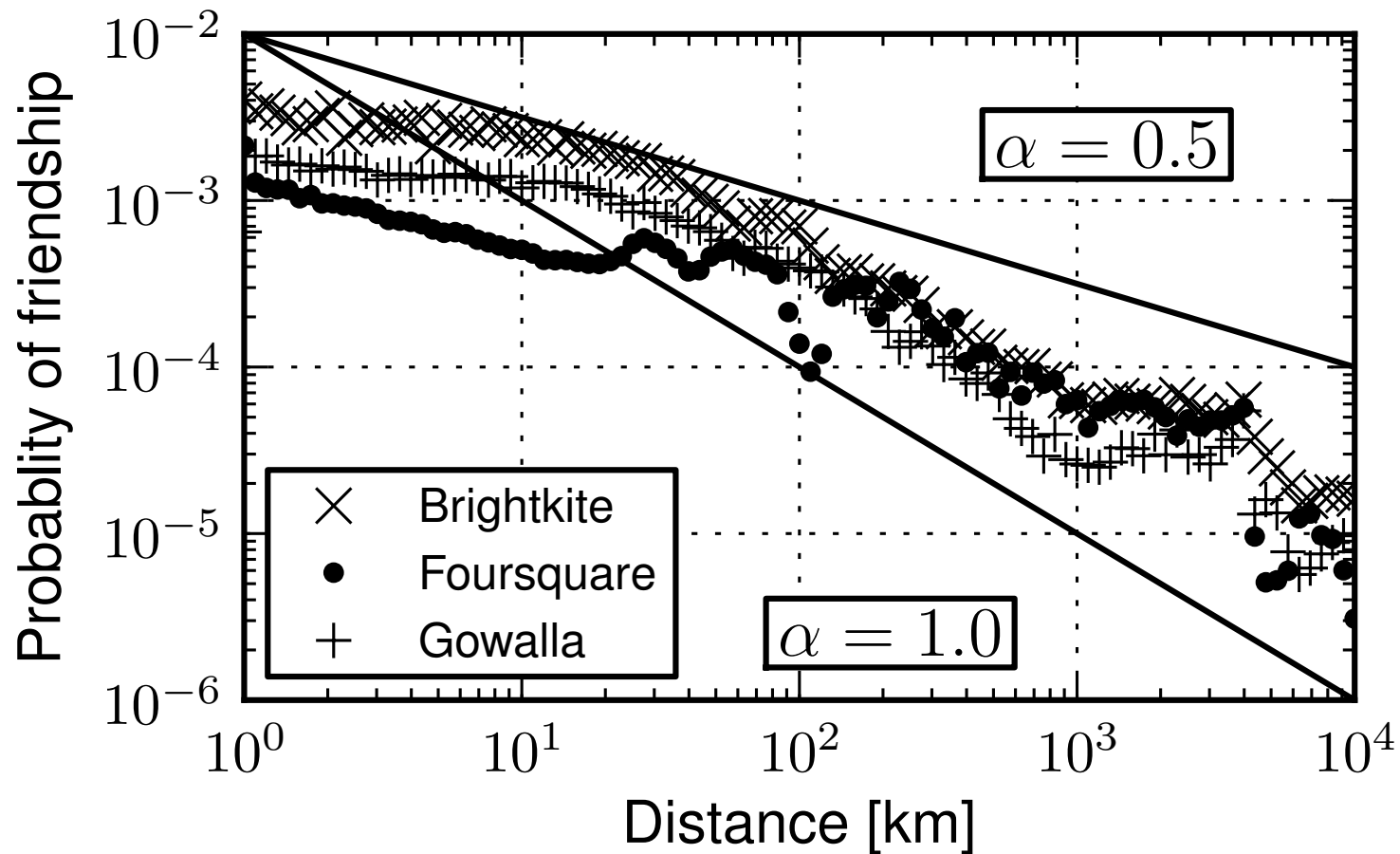
Service			
Nodes	54,190	258,706	122,414
Social links	213,668	2,854,957	580,446
Average degree	7.88	22.07	9.48
Average clustering coefficient	0.181	0.191	0.254
Average distance between friends [km]	2,041	1,442	1,792
Average distance between users [km]	5,651	8,494	5,663

# Distance between users and 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.

# Probability of friendship vs. geographic distance



The decay is less sharp than in other systems: **location-based services appear affected by distance in a weaker way.**

How are individual users affected by geographic distance when they create social links?

# Network randomization

---

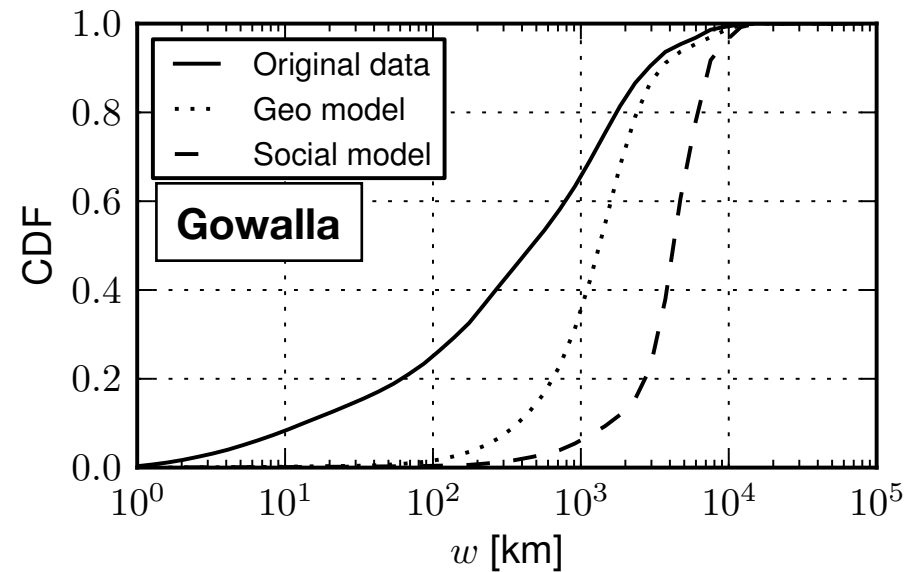
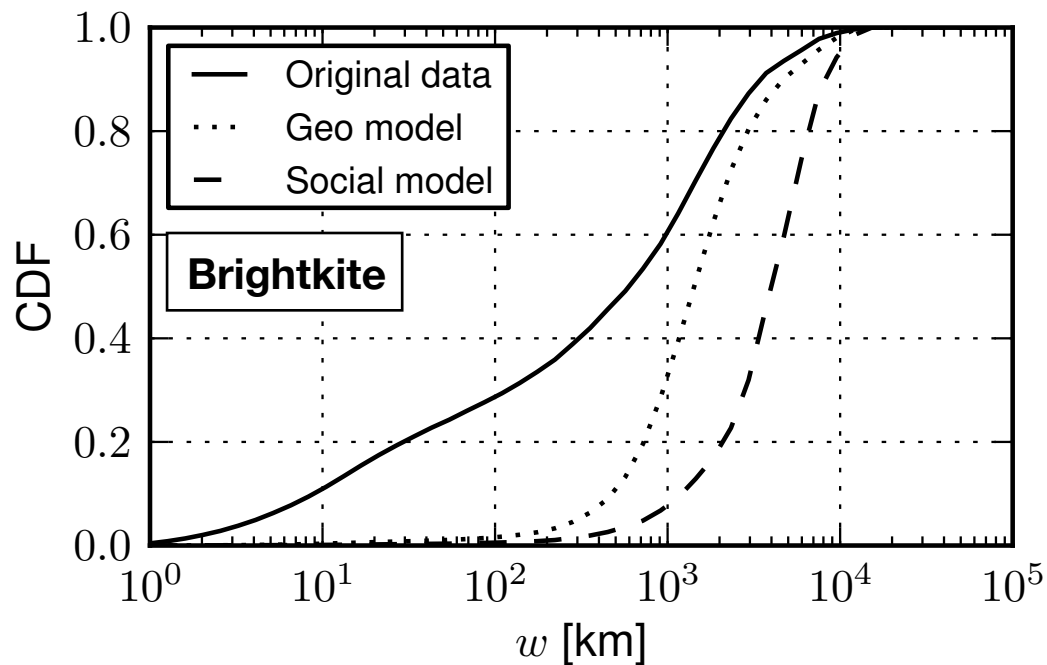
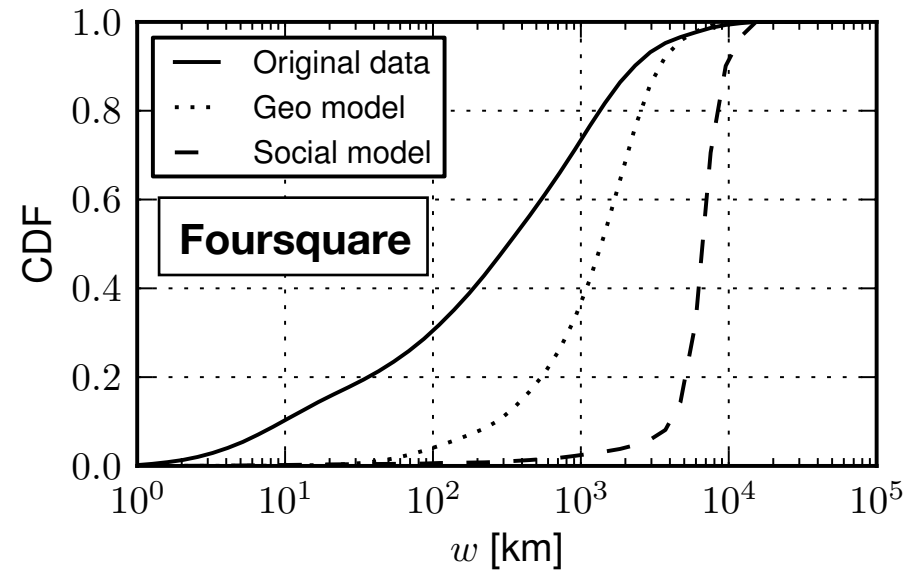
**Two randomized models**, which capture either the geographic or the social properties of the original social networks and randomize everything else.

	Description	Social properties	Spatial properties
<b>Original data</b>	No modification.	✓	✓
<b>Geo model</b>	Fix node locations and reassign all links according to probability $P(d)$ .	✗	✓
<b>Social model</b>	Fix links and shuffle all node locations.	✓	✗

# Average friend distance

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

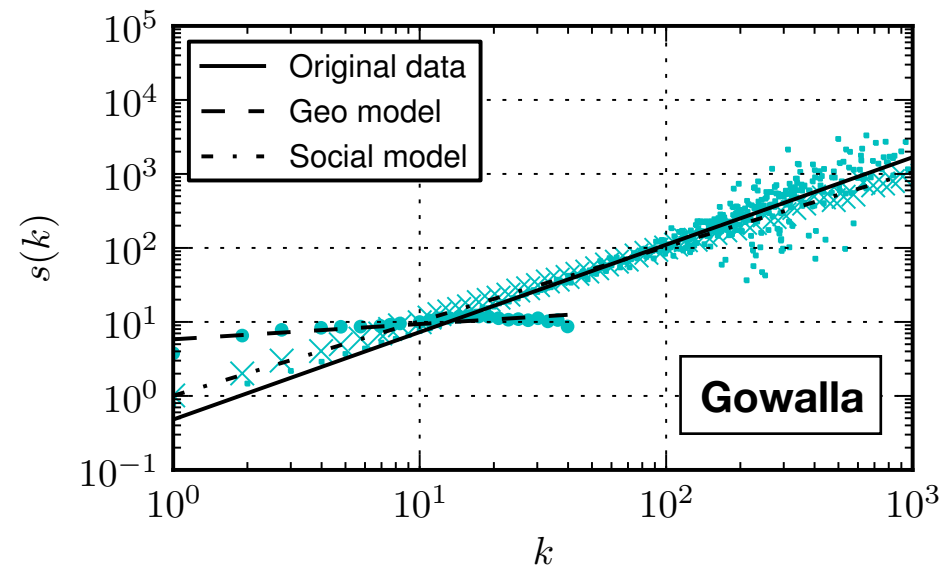
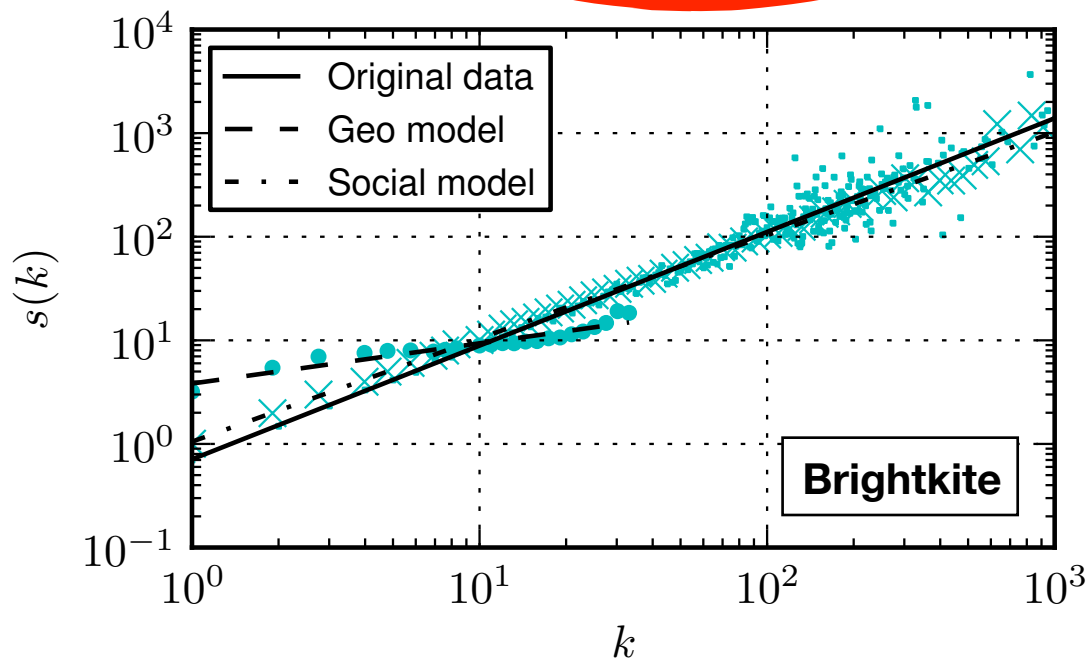
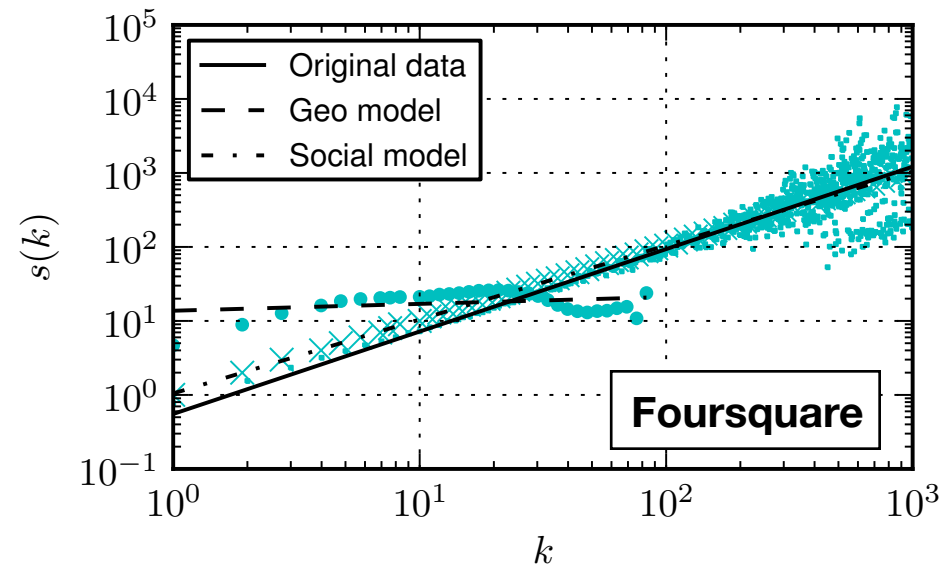
Node degree  $k_i$       Link length  $l_{ij}$       Node neighborhood  $\Gamma_i$



# Distance strength and correlation with degree

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

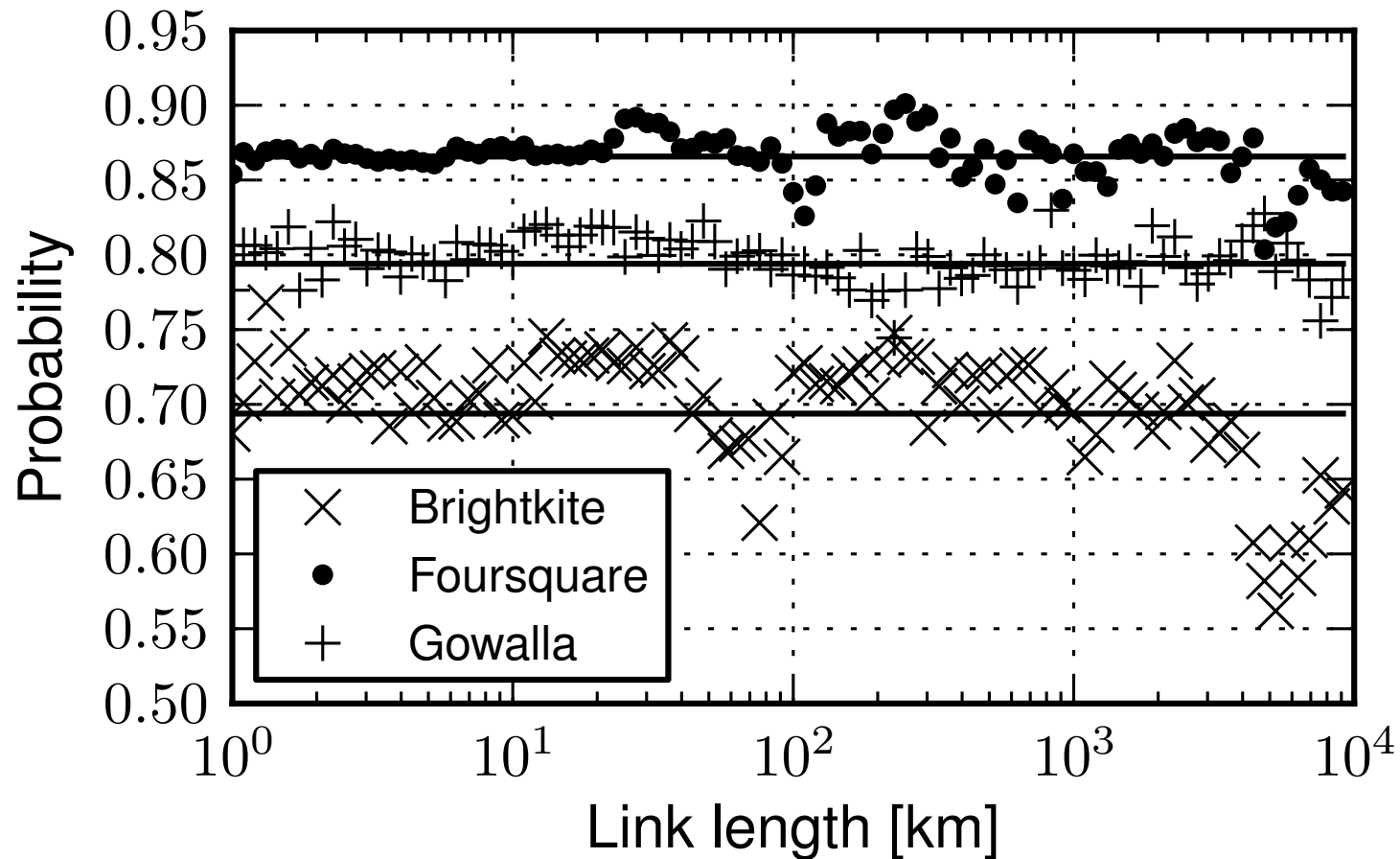
$$s(k) \propto k^\beta \quad \beta \in [1.10, 1.18]$$





What are the spatial properties of the social triangles users belong to?

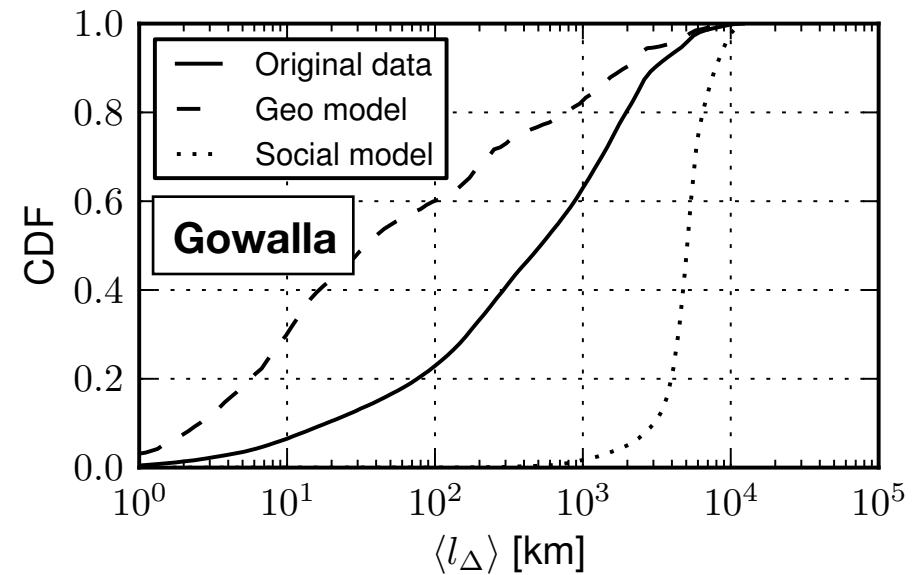
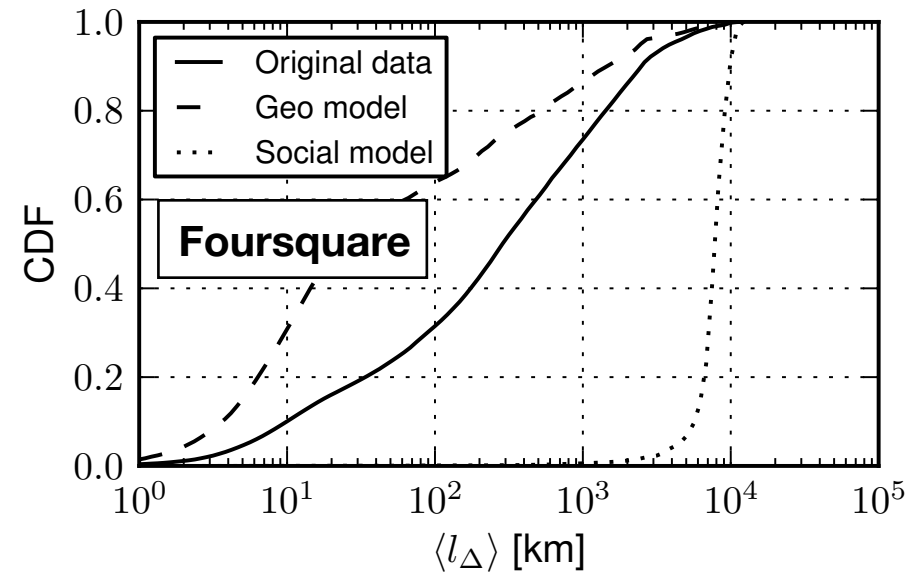
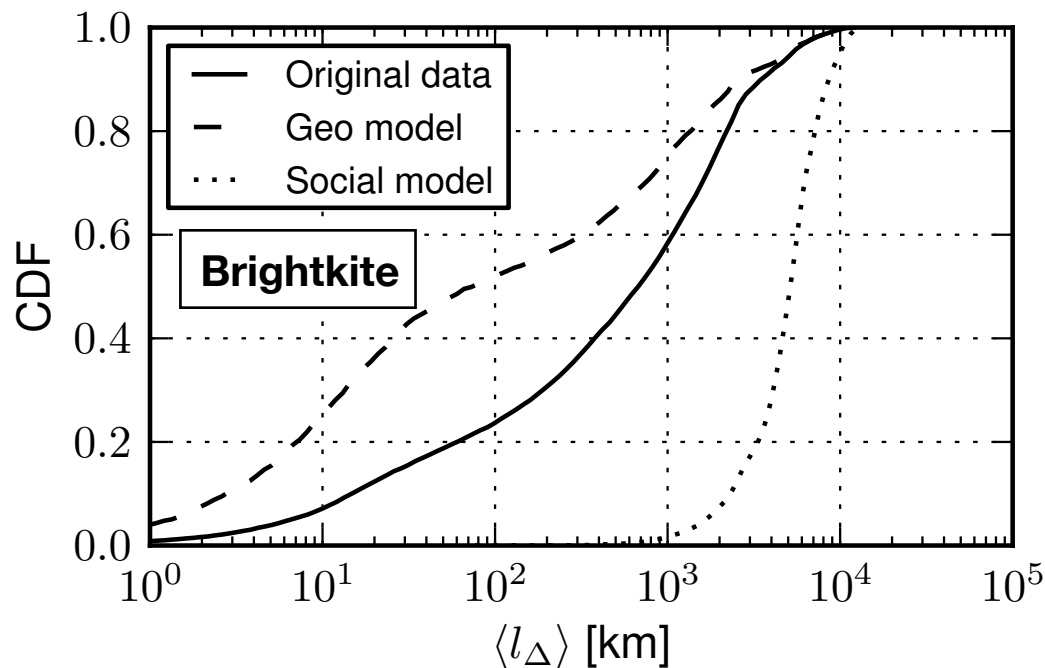
# Social links in social triangles



A link is equally likely to belong to a social triangle regardless of its geographic length.

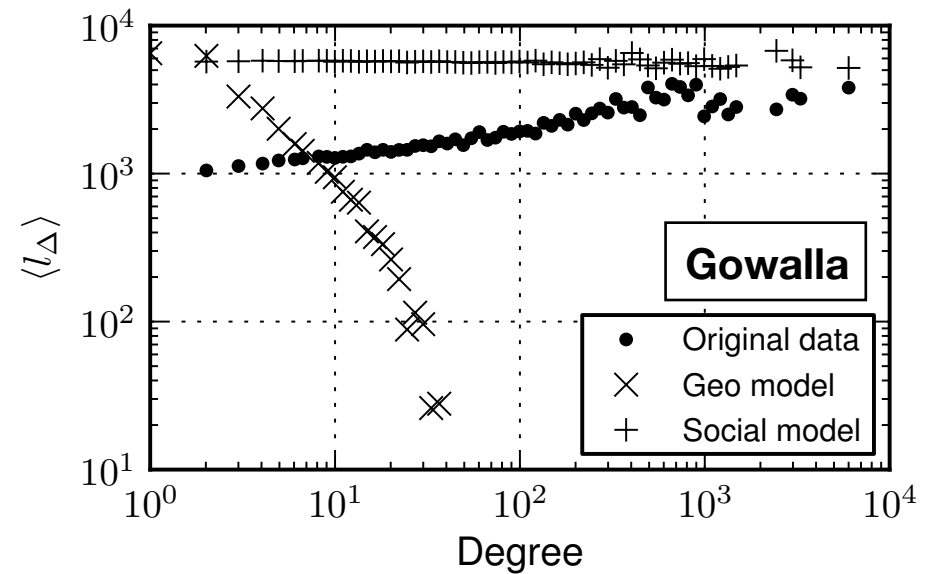
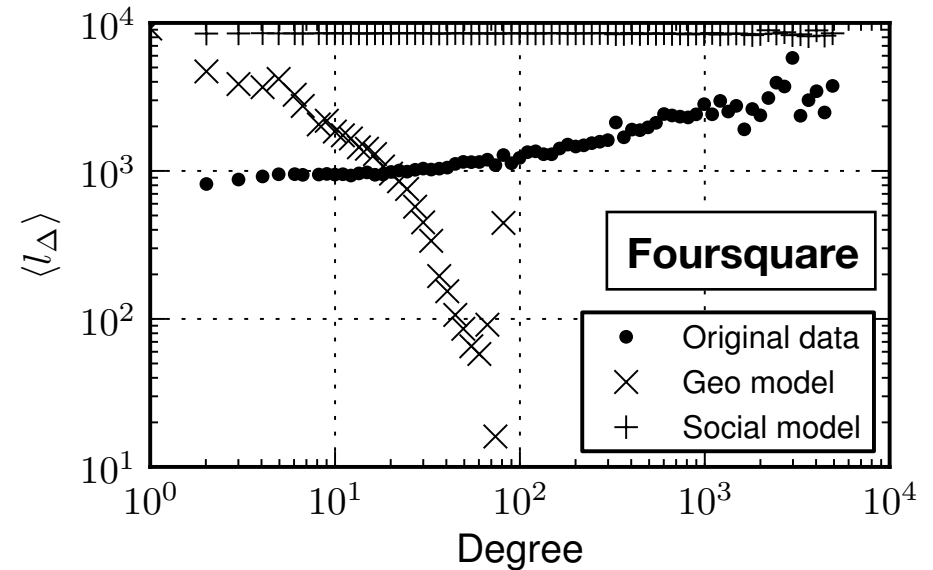
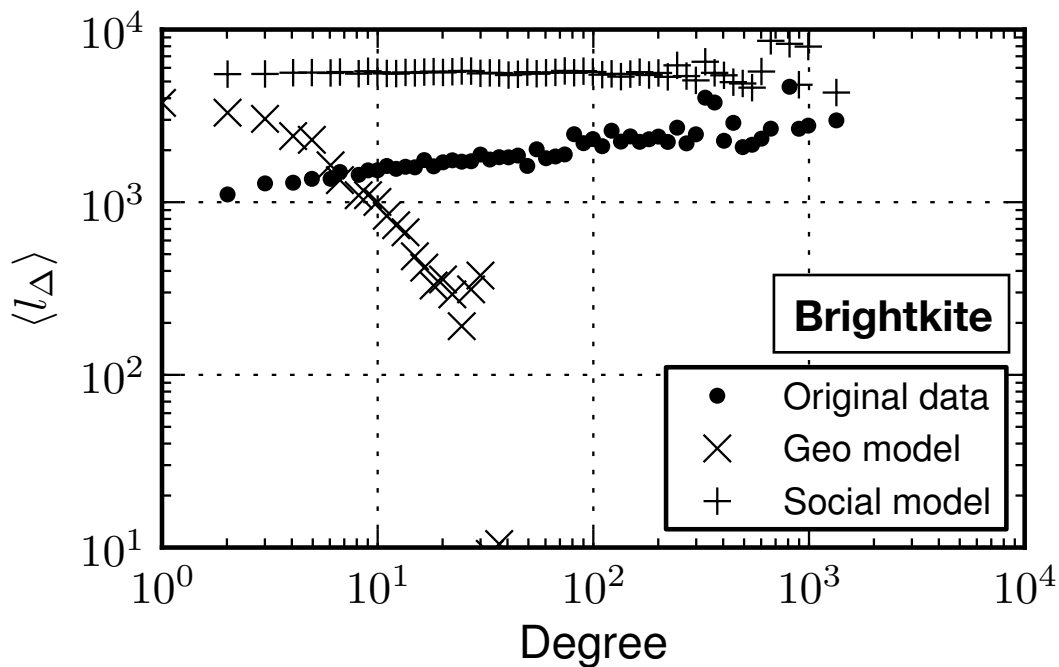
# Average triangle geographic length

Users exhibit **heterogeneity**, as there are users with smaller triangles and users with wider ones.



# Correlation with degree

Users with many friends belong to triangles with longer links.

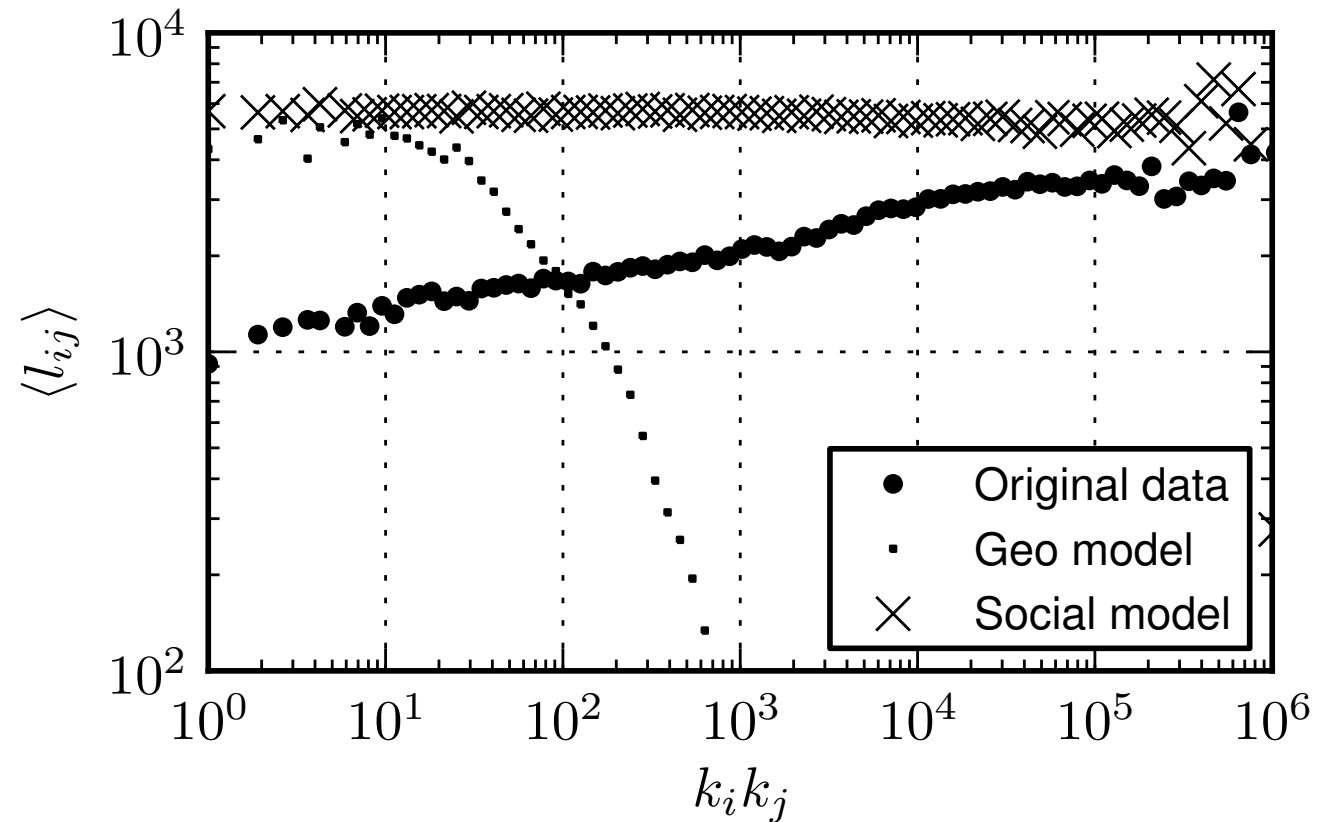


How can we combine spatial and social factors?

# Gravity models for spatial social networks

Gravity model

$$P(i, j) \propto \frac{N_i N_j}{f(d_{ij})}$$



Links connecting popular users tend to be much longer, **while a user might connect to an unimportant one only when they are close to each other.**

# Implications

---

- **Robust and universal socio-spatial features** observed in different datasets.
- **High degree of user heterogeneity** in their socio-spatial properties, often correlated with node degree.
- Social ties appear with different probabilities at different distances, but **spatial factors are not enough to describe real networks.**
- **Geographic distance can not be neglected** when modeling or studying these social networks.



# RETHINK

Flickr: vanagas

Take-away message

Users have heterogeneous socio-spatial properties potentially compatible with gravity models.



Thanks!

Questions?

---

**Salvatore Scellato**

Email: [salvatore.scellato@cl.cam.ac.uk](mailto:salvatore.scellato@cl.cam.ac.uk)

Web: <http://www.cl.cam.ac.uk/~ss824/>

Twitter: [@thetarro](https://twitter.com/thetarro)

Thanks to Flickr authors for CC-licensed images.



Flickr: sean dreilinger