A model for spatially embedded social networks

Johannes Illenberger, Gunnar Flötteröd, Matthias Kowald,
Kai Nagel

Technische Universität Berlin

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE


Image © 2009 Flotron / Perrinjaquet
©2009 Google
Breite 47.372679° Lange 8.536144° Hohe 0 m
Sichthohe 3.39 km

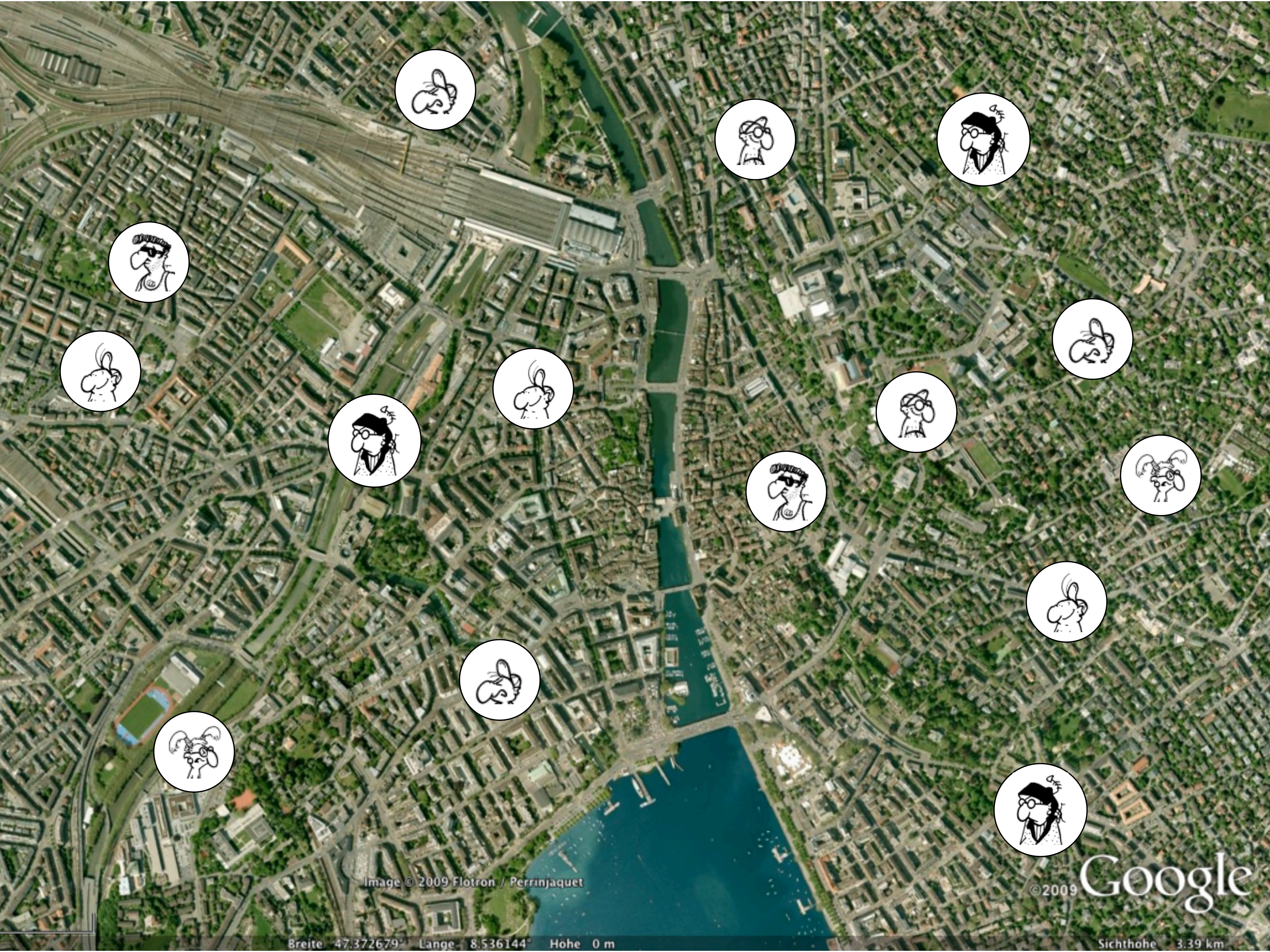


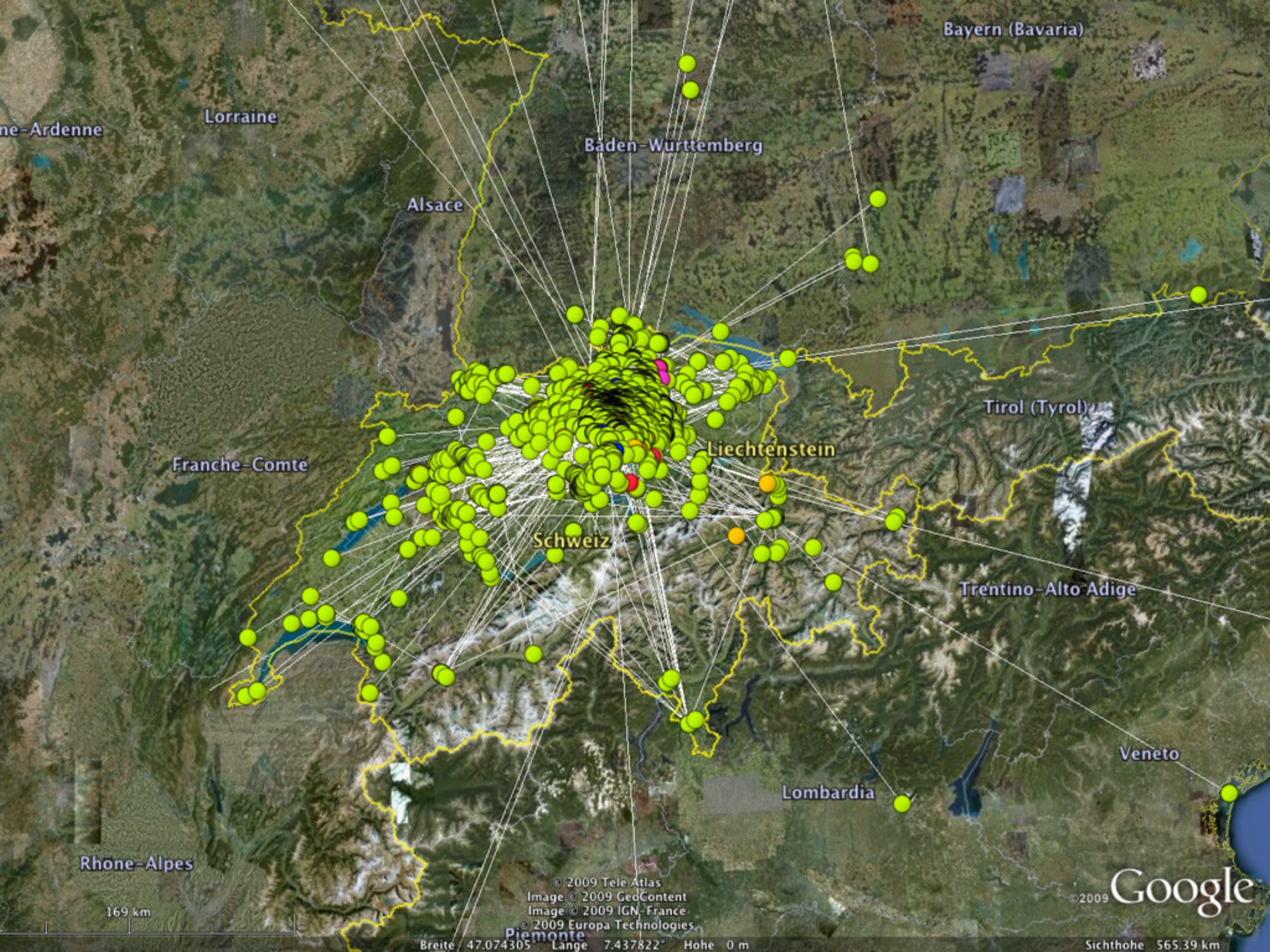
Image © 2009 Flotron / Perrinjaquet
©2009 Google
Breite 47.372679° Länge 8.536144° Höhe 0 m
Sichthöhe 3.39 km

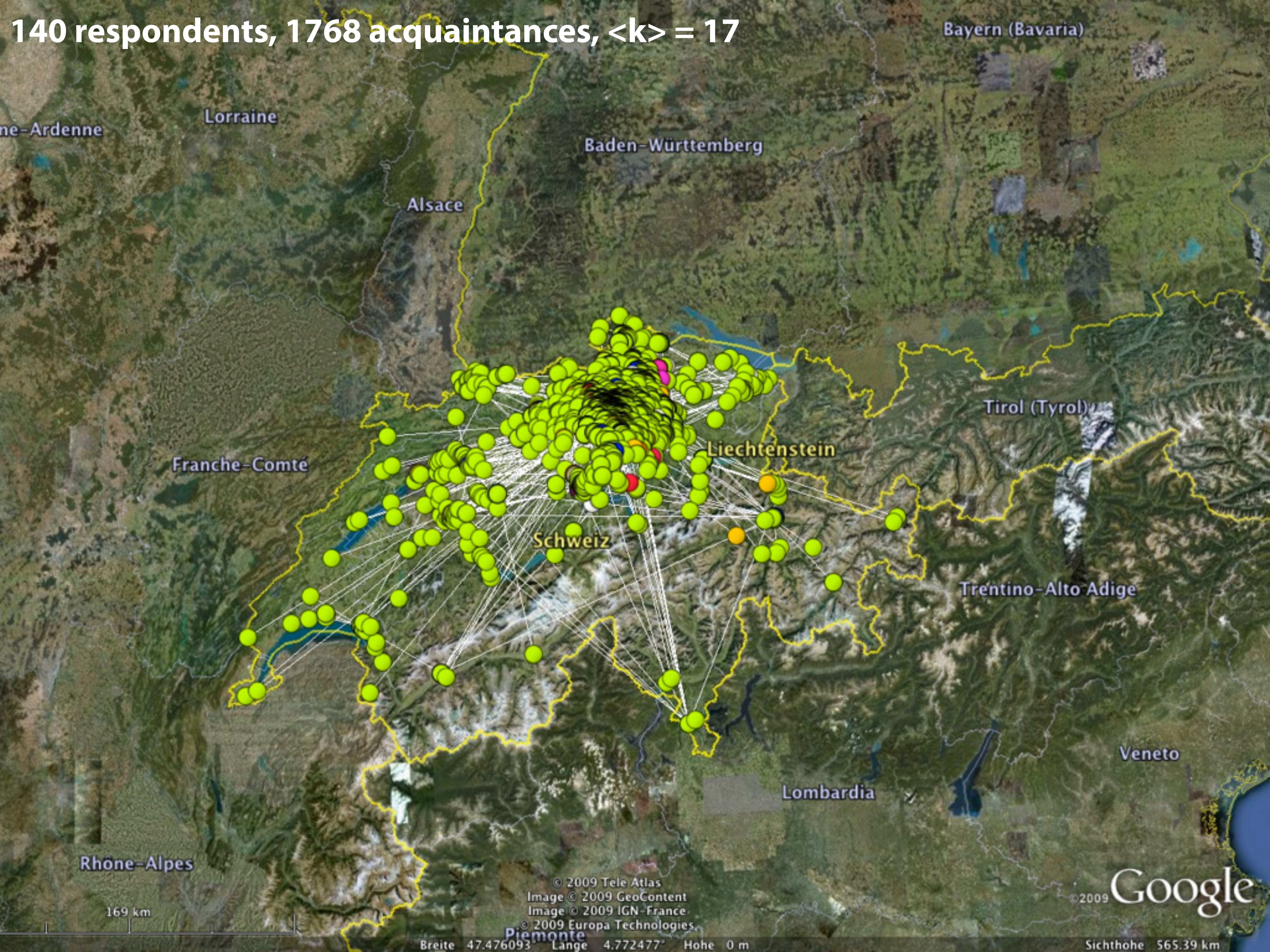

Image © 2009 Flotron / Perrinjaquet
©2009 Google
Breite 47.372679° Länge 8.536144° Höhe 0 m
Sichthöhe 3.39 km

Motivation

- Can we reproduce a social network with a model that uses easy to measure indicators
  - e.g. the spatial distribution of individuals?

Empirical Data

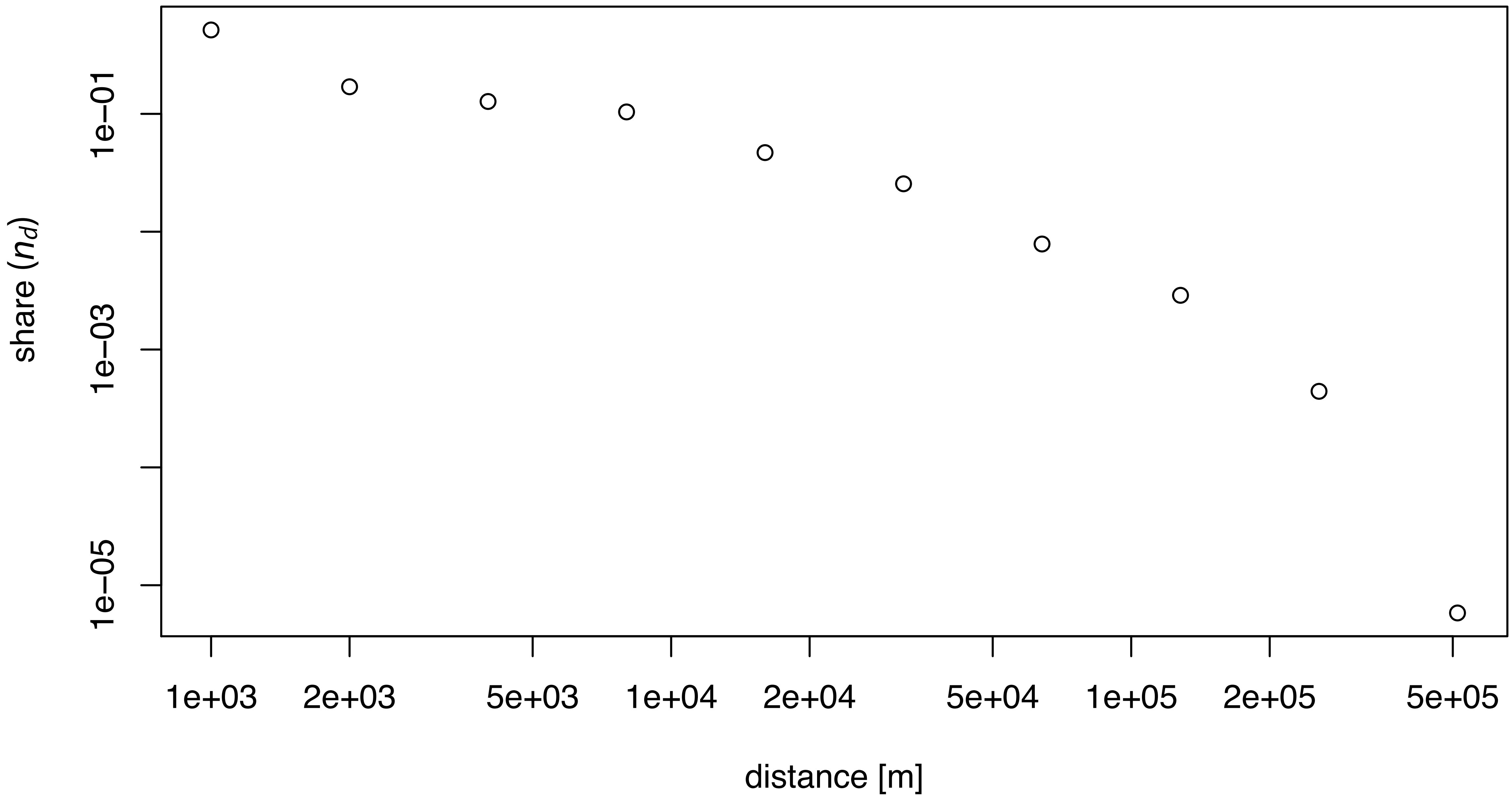

Bayern (Bavaria)
Lorraine
ne-Ardenne
Baden-Württemberg
Alsace
Tirol (Tyrol)
Liechtenstein
Franche-Comté
Schweiz
Trentino-Alto Adige
Veneto
Lombardia
Rhône-Alpes
Piemonte
169 km
© 2009 Tele Atlas
Image © 2009 GeoContent
Image © 2009 IGN-France
© 2009 Europa Technologies
©2009 Google
Breite 47.074305° Länge 7.437822° Höhe 0 m
Sichthöhe 565.39 km



140 respondents, 1768 acquaintances, <k> = 17
Bayern (Bavaria)
Lorraine
Baden-Württemberg
Alsace
Franche-Comté
Liechtenstein
Tirol (Tyrol)
Schweiz
Trentino-Alto Adige
Veneto
Lombardia
Rhône-Alpes
Piemonte
169 km
© 2009 Tele Atlas
Image © 2009 GeoContent
Image © 2009 IGN-France
© 2009 Europa Technologies
©2009 Google
Breite 47.476093° Länge 4.772477° Höhe 0 m
Sichthöhe 565.39 km

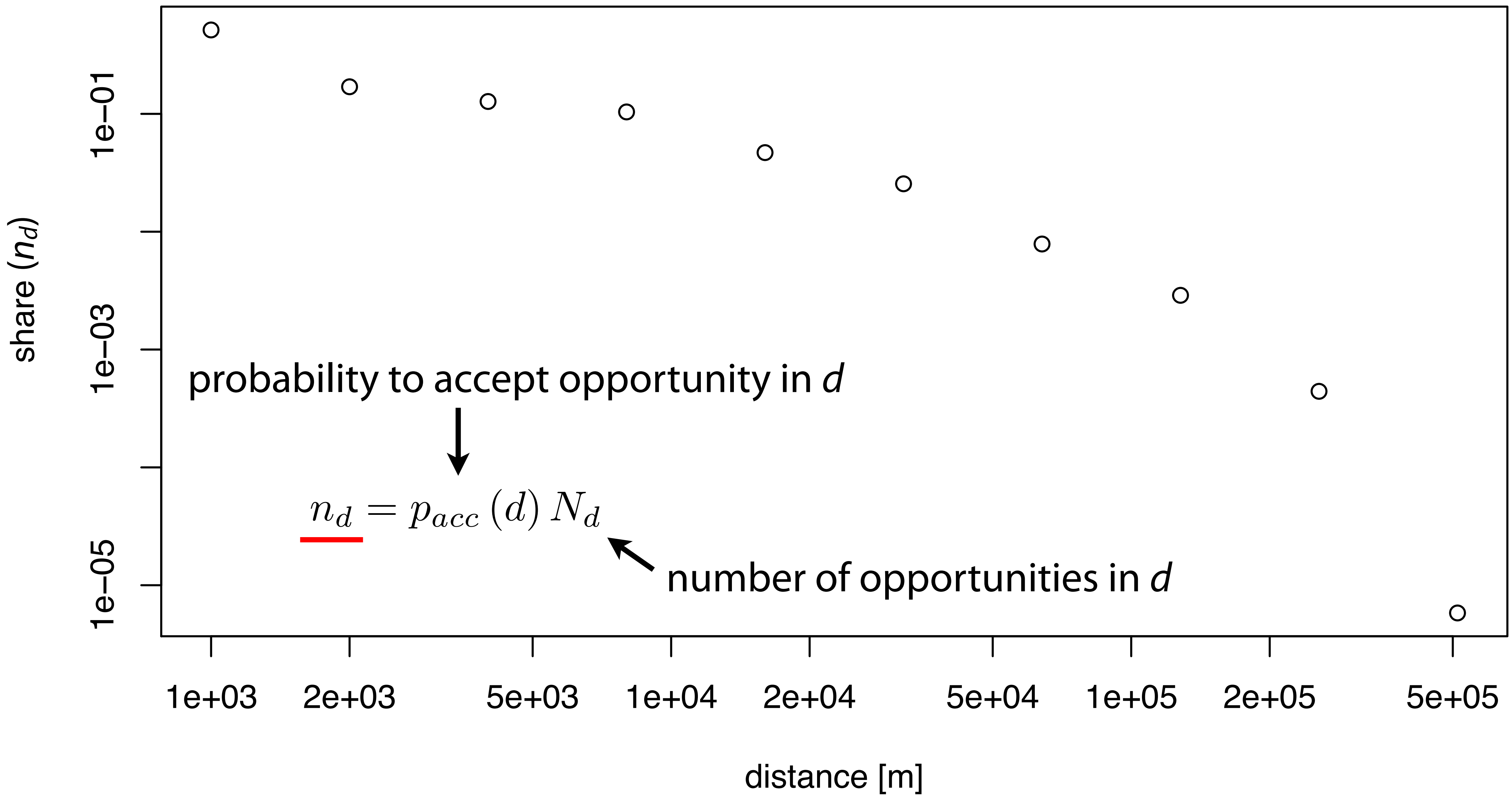
Empirical Data

- Edge length distribution (edge length = beeline distance)



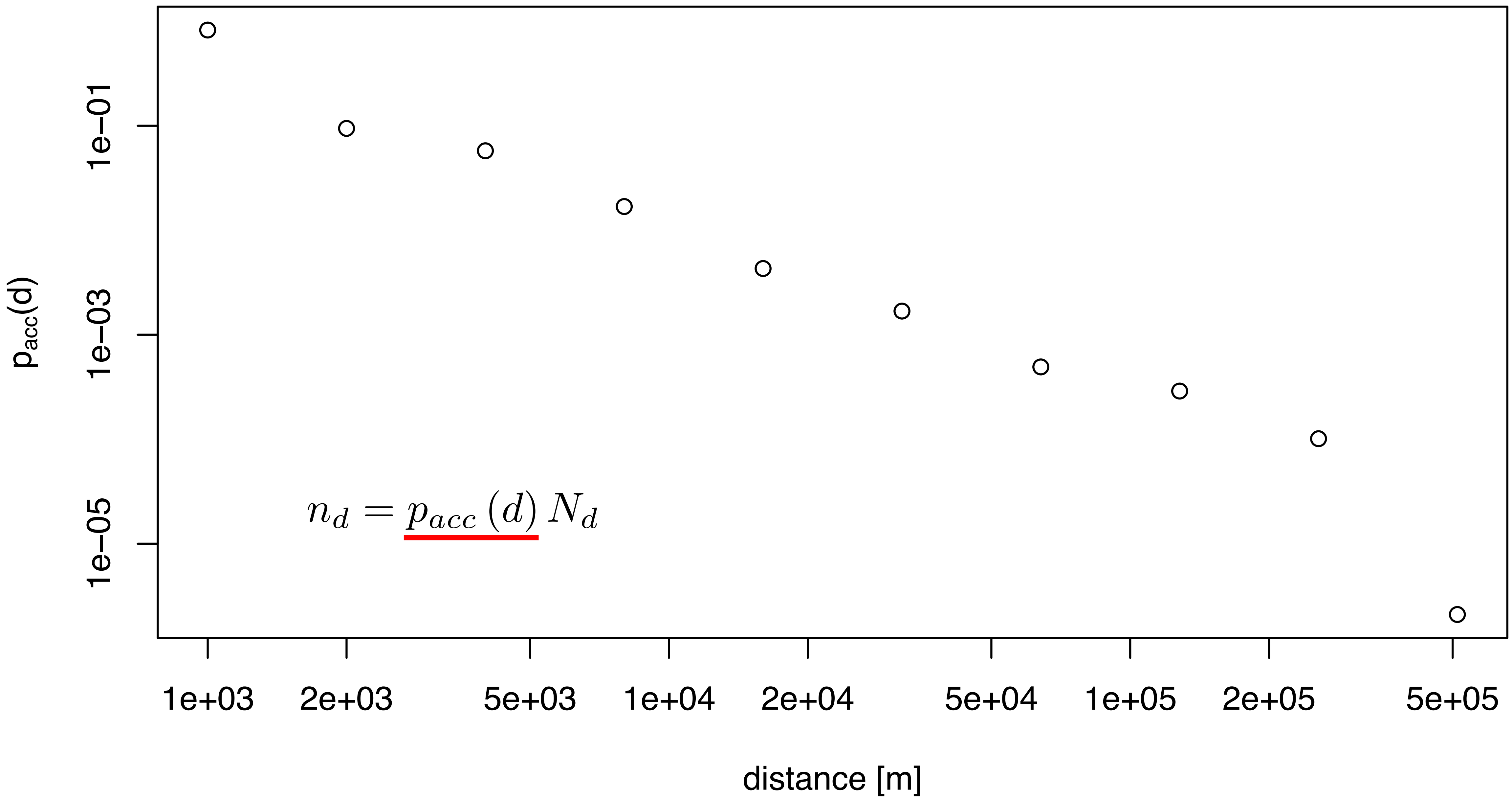
Empirical Data

- Edge length distribution (edge length = beeline distance)



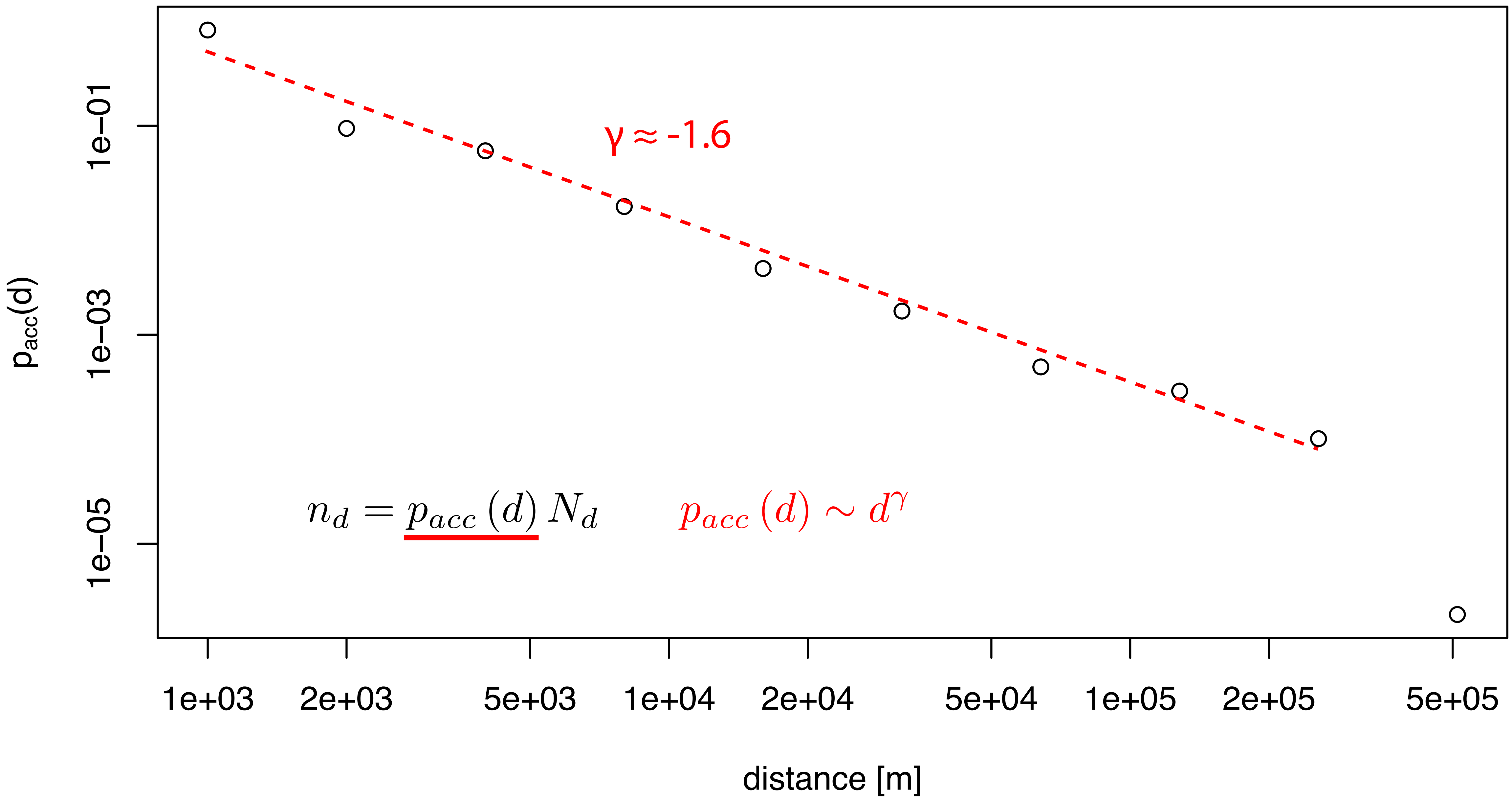
Empirical Data

- Acceptance probability (N_d from land use data)



Empirical Data

- Acceptance probability (N_d from land use data)



Methodology

Methodology

- Gravity model $$T_{ij} = K \frac{W_i W_j}{e^{c_{ij}}}$$
- Translated to a graph $$p_{ij} \sim e^{-c_{ij}}$$
- Empirical data $$p_{ij} \sim d_{ij}^{-\gamma}$$
- Costs of an edge $$c_{ij} = \gamma \ln d_{ij} + const$$

Methodology

- Assign each vertex a fixed budget C^*
- Create a graph with the constraint that
  - for each vertex i with k neighbours

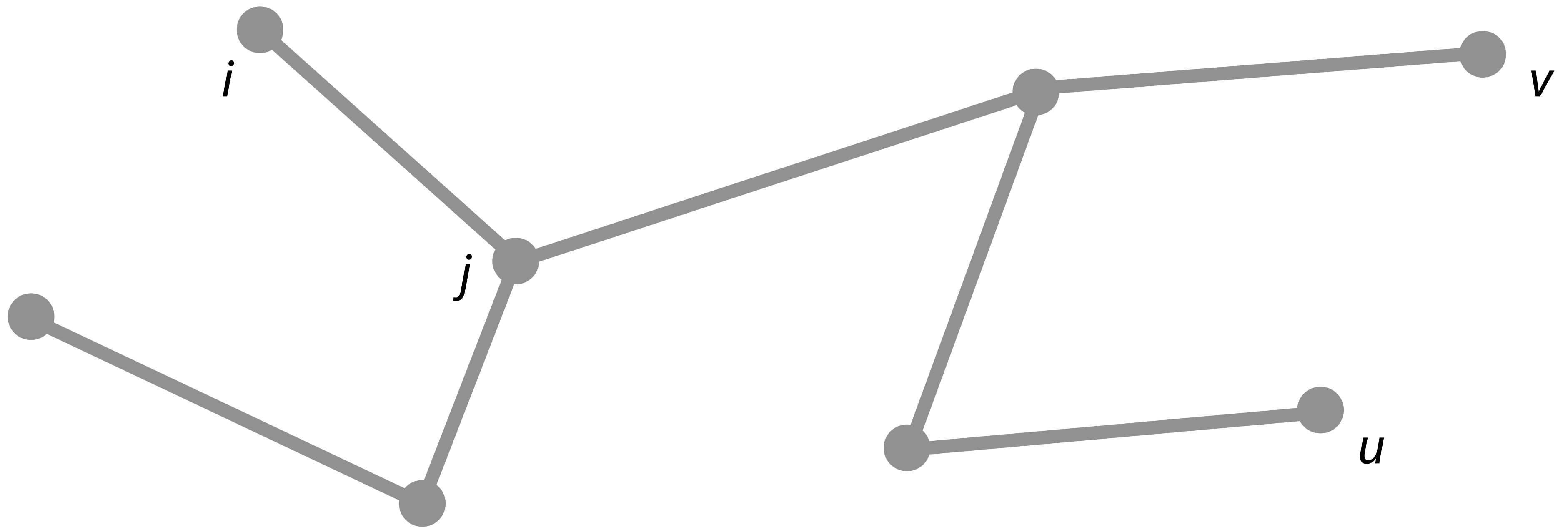
$$\sum_{j}^{k} c_{ij} = C_i \overset{!}{=} C^*$$

- Consider $P(Y)$ the probability distribution of all realisable graphs
- Choose $P(Y)$ so that the values of C_i meet as good as possible the desired value C^*
  - Exponential Random Graph Model

Methodology

- Generate a random graph with the desired number of edges
- Re-order edges until steady state distribution (Gibbs sampling)

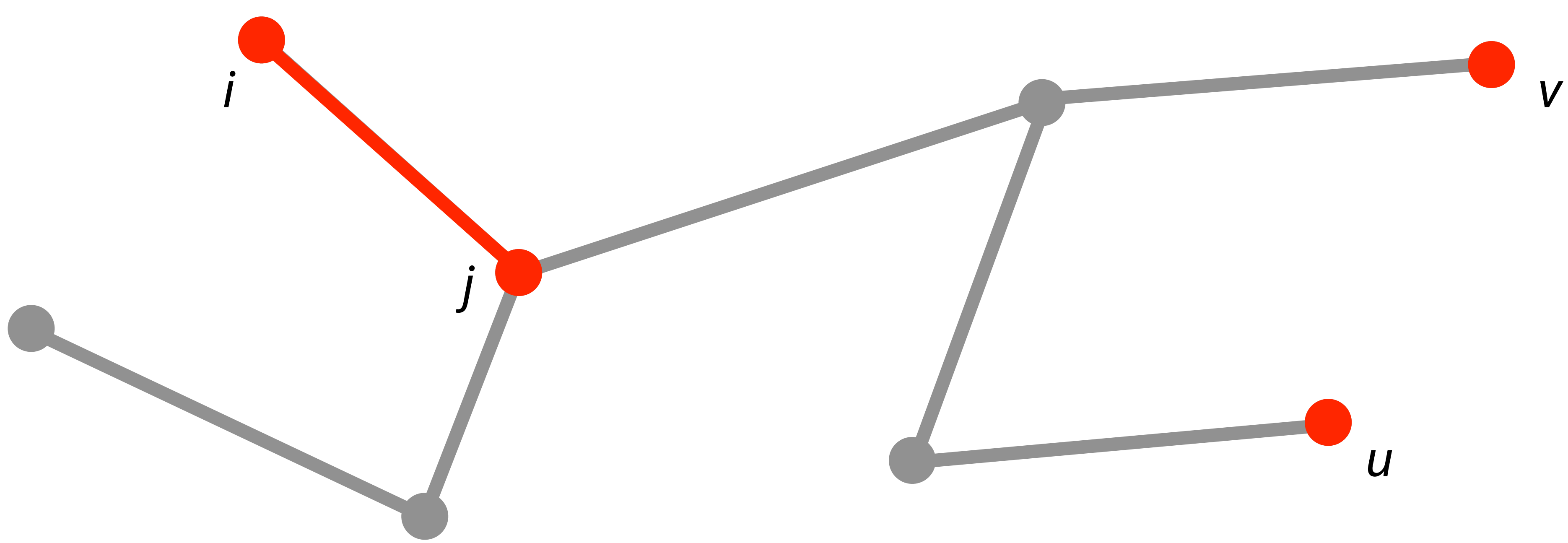
$$P_{accept} = \frac{P\left(Y|y_{ij}=0, y_{uv}=1\right)}{P\left(Y|y_{ij}=0, y_{uv}=1\right) + P\left(Y|y_{ij}=1, y_{uv}=0\right)}$$



Methodology

- Generate a random graph with the desired number of edges
- Re-order edges until steady state distribution (Gibbs sampling)

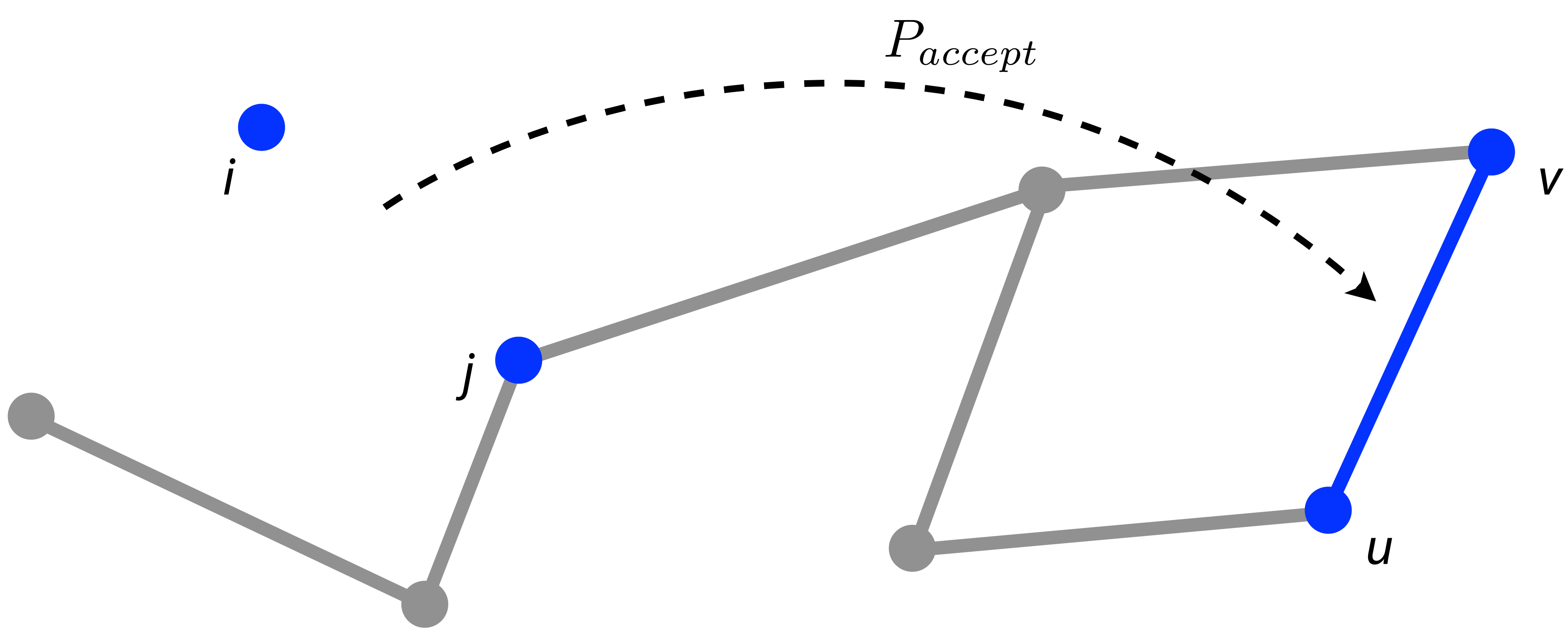
$$P_{accept} = \frac{P\left(Y|y_{ij}=0, y_{uv}=1\right)}{P\left(Y|y_{ij}=0, y_{uv}=1\right) + \underline{P\left(Y|y_{ij}=1, y_{uv}=0\right)}}$$



Methodology

- Generate a random graph with the desired number of edges
- Re-order edges until steady state distribution (Gibbs sampling)

$$P_{accept} = \frac{P\left(Y|y_{ij}=0, y_{uv}=1\right)}{P\left(Y|y_{ij}=0, y_{uv}=1\right) + P\left(Y|y_{ij}=1, y_{uv}=0\right)}$$



- probability of move is larger with larger *P(Y)* of new *Y*

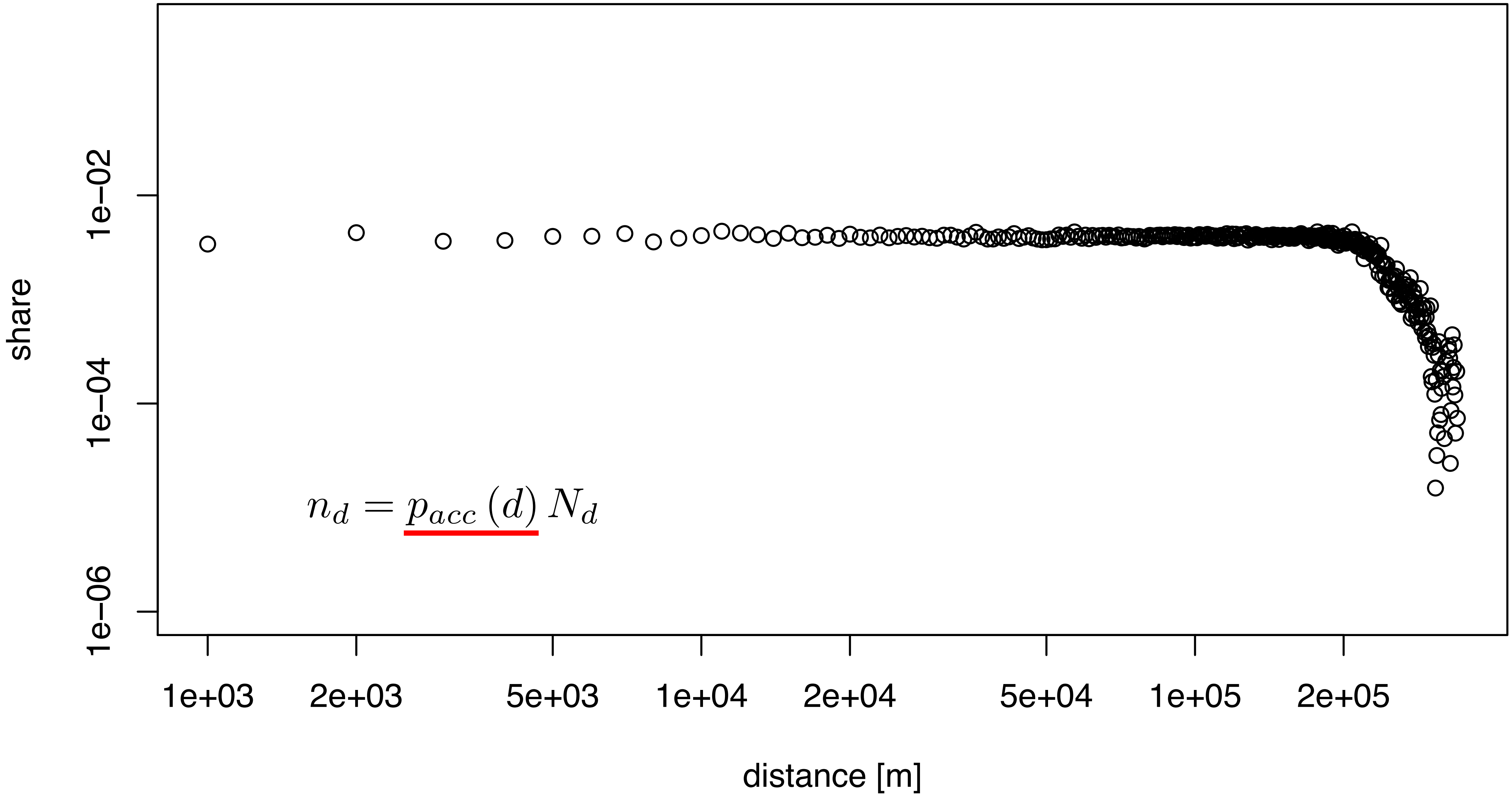
Simulation Scenario

Simulation Scenario

- Starting point: synthetic population of Switzerland (37k persons $\hat{\approx}$ 1% sample)
  - synthetic persons with residential locations (gender, age, income...)
  - i.e., fixed set of vertices with geographic coordinates
- Create a random graph with $\langle k \rangle = 17$
- Assign each vertex a budget C^*
- Reorder edges for $\sim 10^9$ iterations

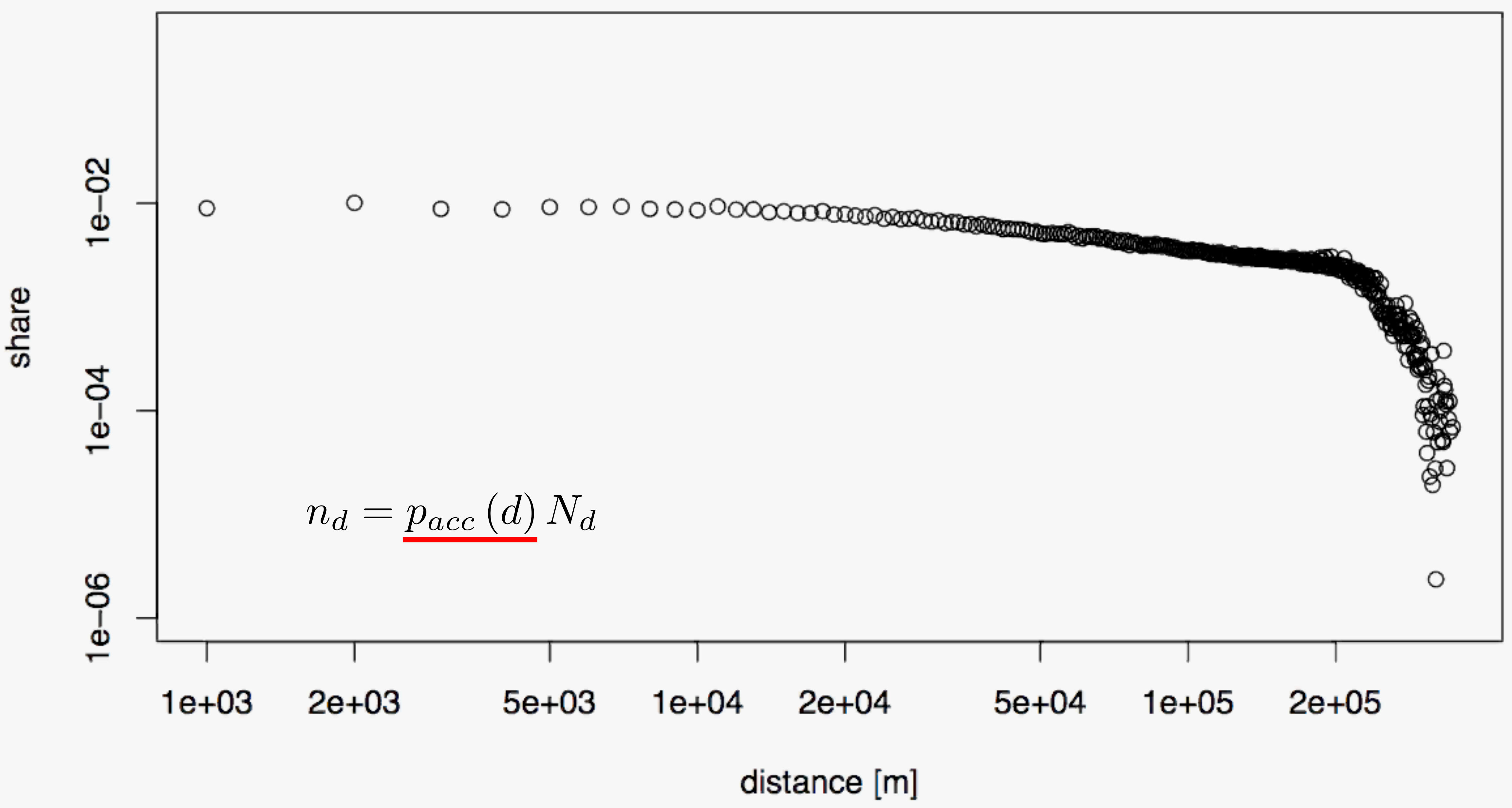
Simulation Results

- Acceptance probability



Simulation Results

- Acceptance probability



$$n_d = p_{acc}(d)\, N_d$$

Simulation Results

- How to control γ?

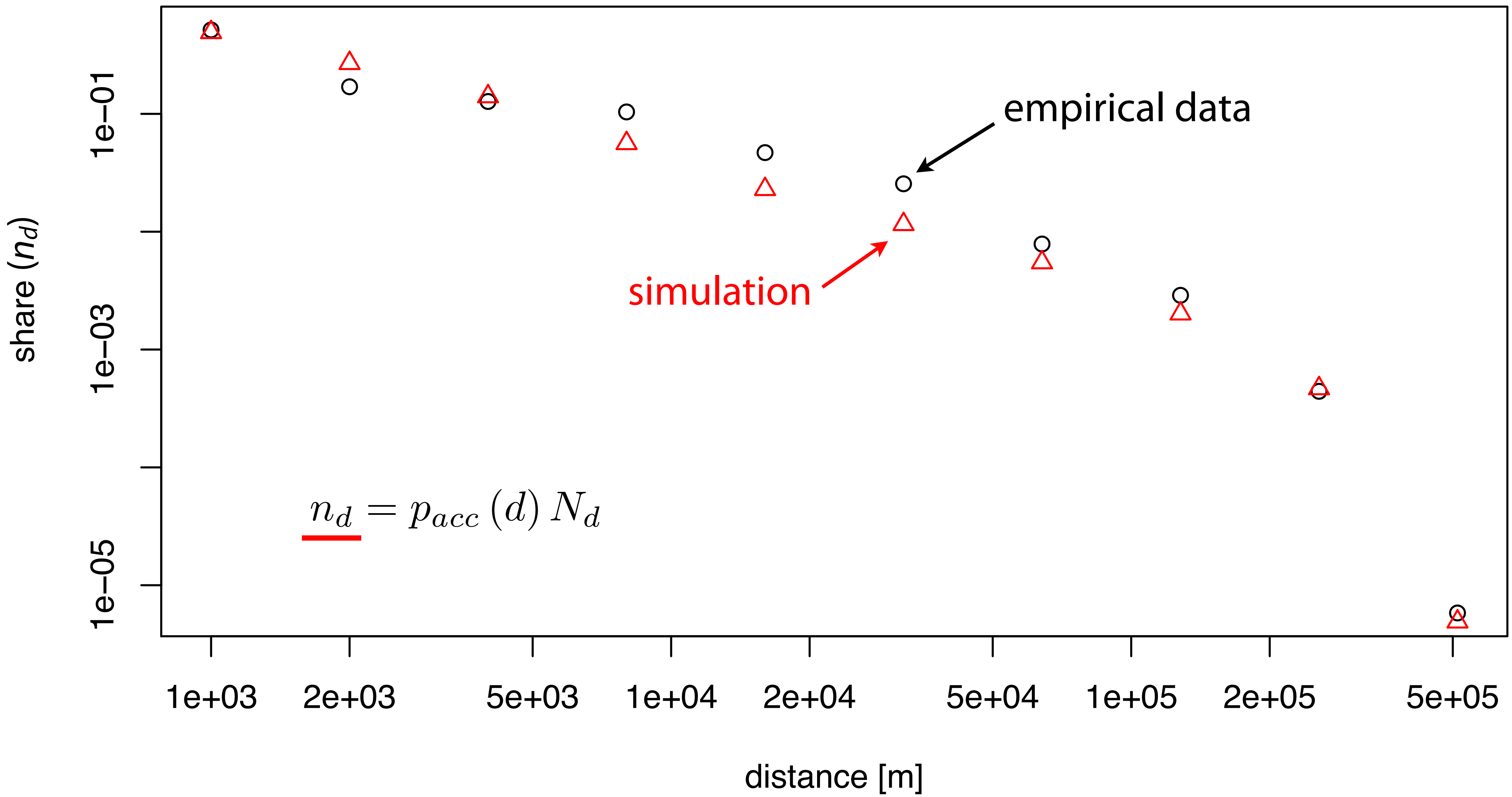
$$\sum_{j}^{k} c_{ij} = C_i \stackrel{!}{=} C^* \quad \text{where} \quad c_{ij} = \gamma \ln d_{ij} + const$$

C^* ↑	longer edges	γ ↓	flatter slope
C^* ↓	shorter edges	γ ↑	steeper slope

- by trial and error...
  - for <*k*> = 17 one obtains with $C^* = 25 \rightarrow \gamma = 1.6$
    (i.e. with <k> given from empirical data, adjust C* until *γ=1.6 from empirical data is achieved)*

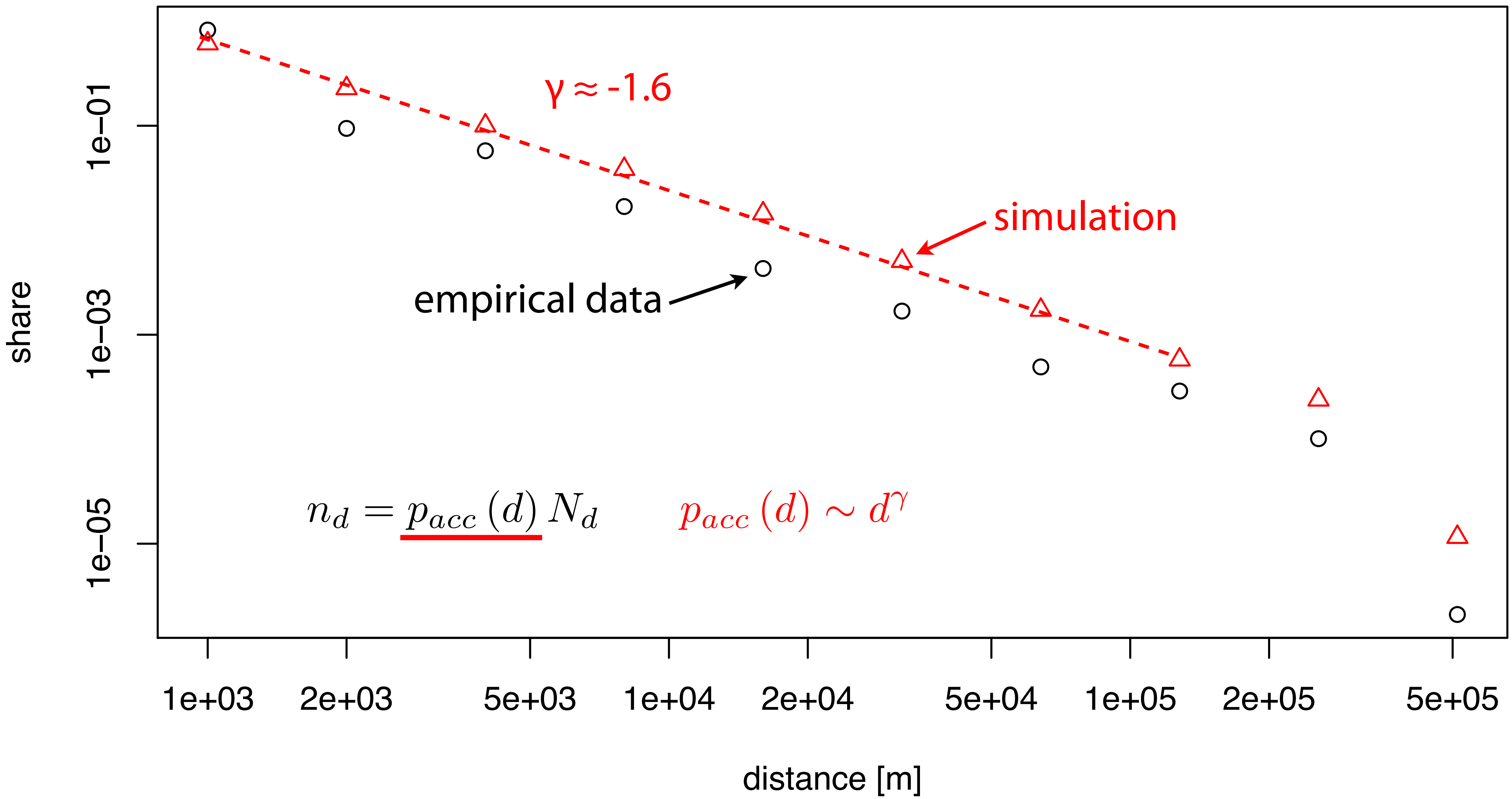
Simulation Results

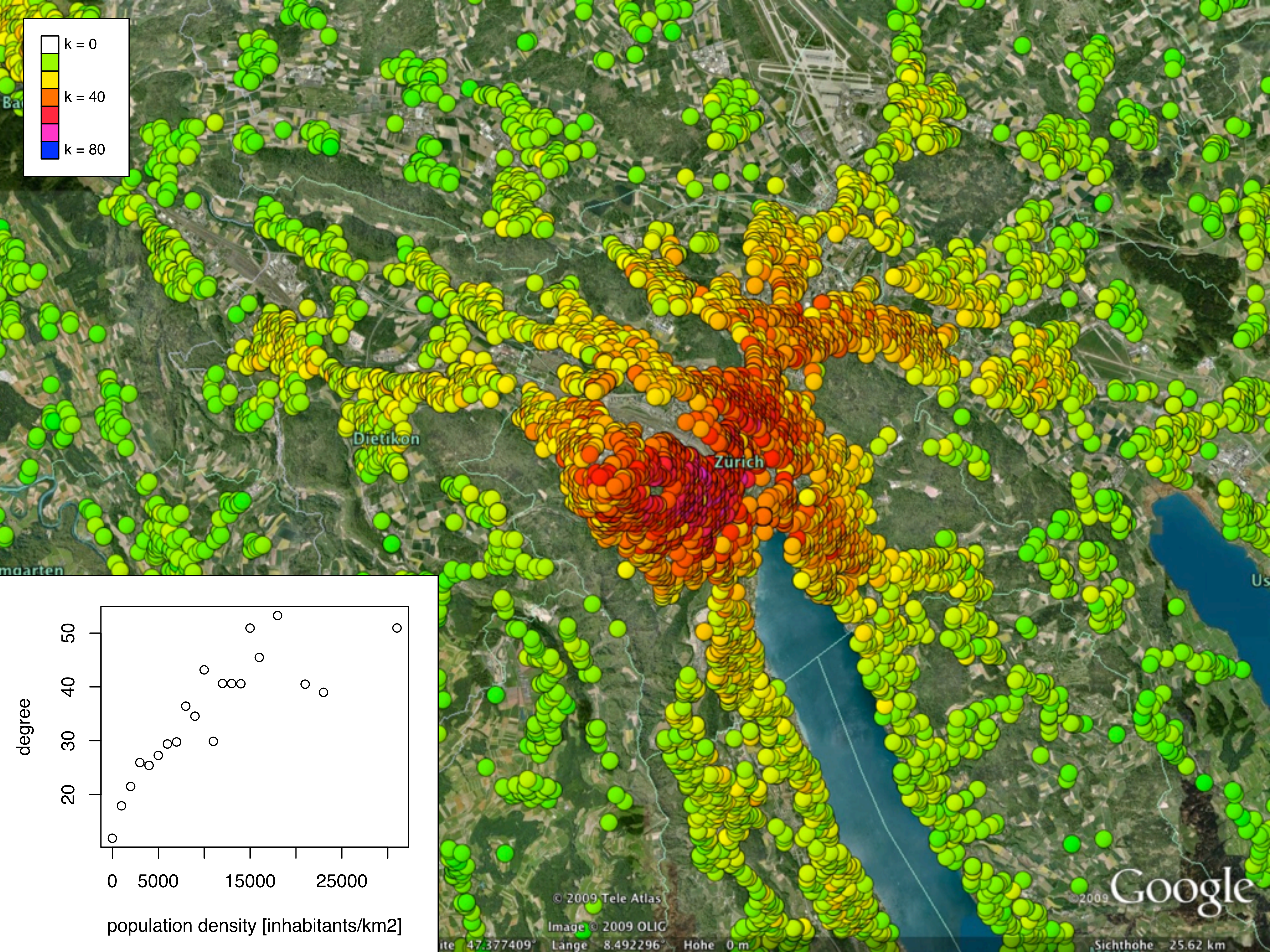
- Edge length distribution



Simulation Results

- Acceptance probability





k = 0
k = 40
k = 80
Dietikon
Zürich
degree
population density [inhabitants/km2]
0
5000
15000
25000
20
30
40
50
© 2009 Tele Atlas
Image © 2009 OLIG
©2009 Google

Conclusion

- A model that uses only the spatial location of vertices as an explanatory variable
- Reproduces the edge length distribution of empirical data
- Spatial distribution of costs controls the spatial distribution of degrees
- No clustering. Small-World effects?

- Outlook
  - Calibrate cost function
  - Include socio-demographic attributes

Thanks for your attention!

Questions? Comments?

The following people contributed to the work presented in this talk:

- Gunnar Flötteröd
- Matthias Kowald
- Kai Nagel
- Kay W. Axhausen

Technische Universität Berlin

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE