

# Considering baseline homophily when generating spatial social networks

Sascha Holzhauser, Friedrich Krebs and Andreas Ernst  
Center for Environmental Systems Research  
University of Kassel  
Germany

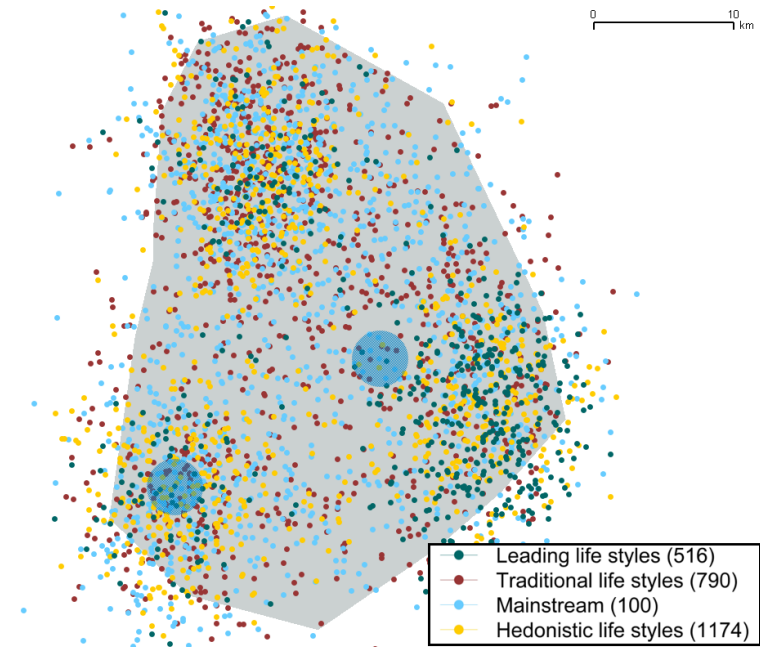


## Motivation

- You join a group that is available and you feel comfortable with
- In many ABM, networking works different:
  - E.g., Yang et.al. 2011: Simulation of Adults' Daily Walking  
„Each individual has 3 to 5 friends who can influence her/his walking attitude, **randomly** selected from the people with the same or similar SES value.“
- The influence of social network on simulation results varies.
- For a factor to prove to be irrelevant we need to consider it.

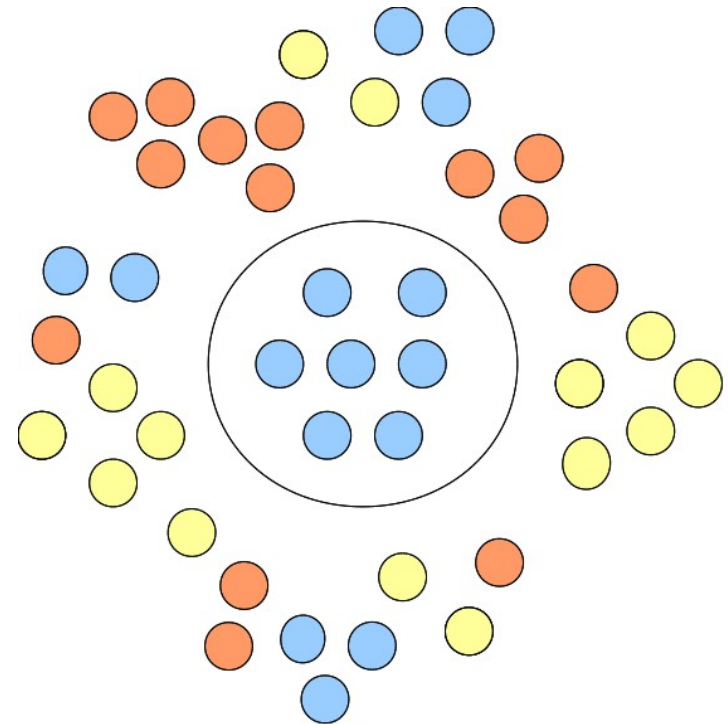
## Our Model Setting

- Modelling for policy consulting
  - planning of campaigns
  - identify area-specific measures
- Agent grouping according to sociological lifestyles
- Dataset of geo-referenced life style composition
- Representative agents are placed in the model region
- Processes of social influence
  - > asymmetrical ties from influencer to influenced



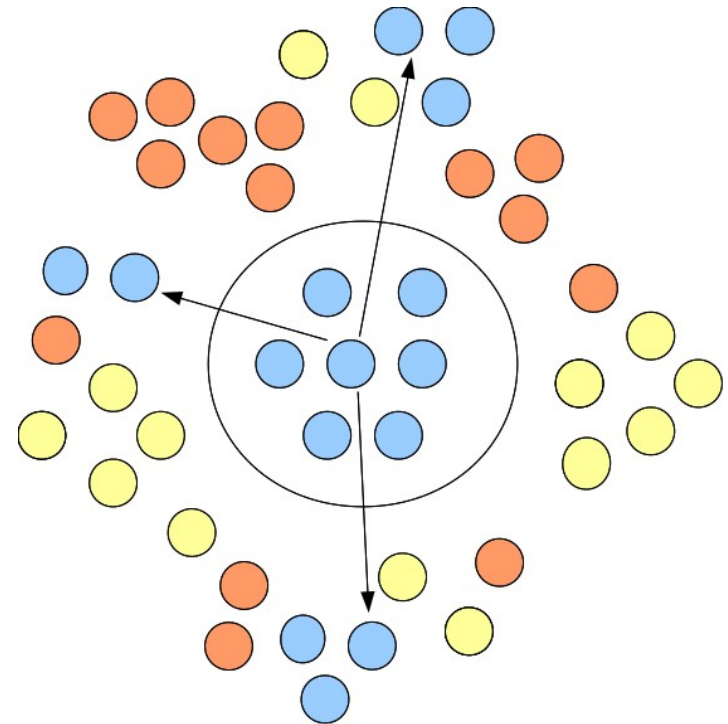
# Homophily

- People are likely to affiliate with others similar in personal or socio-demographic attributes (McPherson et. al. 2001)
- One's social world gets narrowed, which effects access to information, formation of attitudes, etc.
- Inbreeding Homophily
  - Explicitly choose friends that have similar views, income, occupation – above the opportunity set



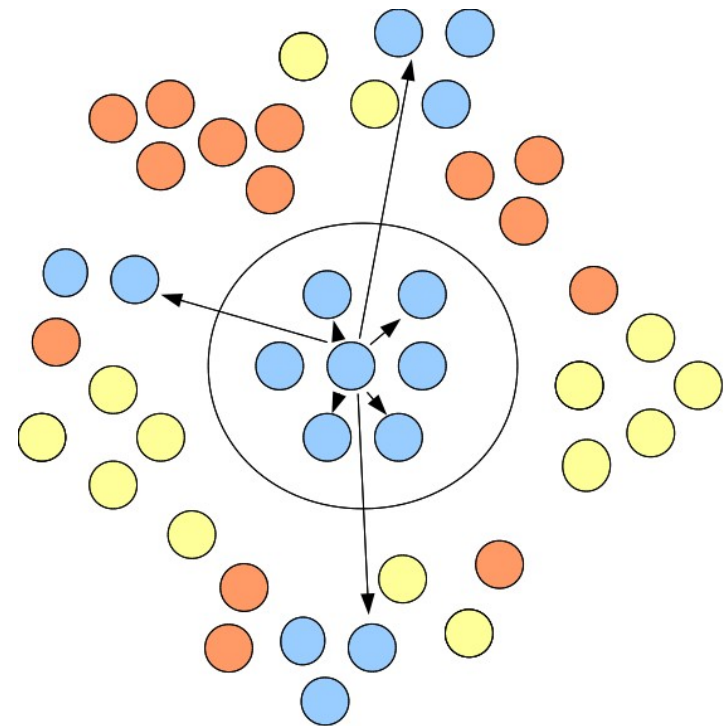
# Homophily

- People are likely to affiliate with others similar in personal or socio-demographic attributes (McPherson et. al. 2001)
- One's social world gets narrowed, which effects access to information, formation of attitudes, etc.
- Inbreeding Homophily
  - Explicitly choose friends that have similar views, income, occupation – above the opportunity set



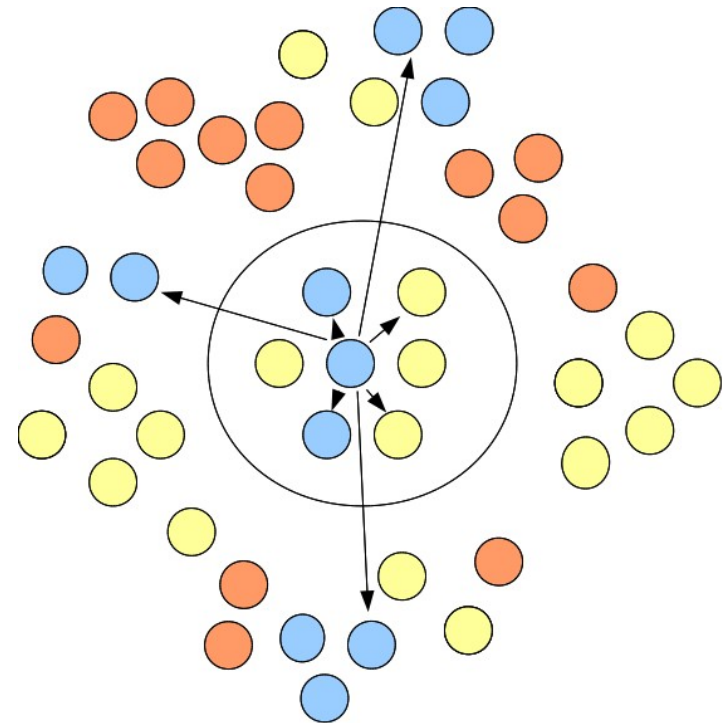
# Homophily

- Baseline Homophily
  - Constraints to the local social world
  - Naturally people work/live/spend leisure time with similar others



# Homophily

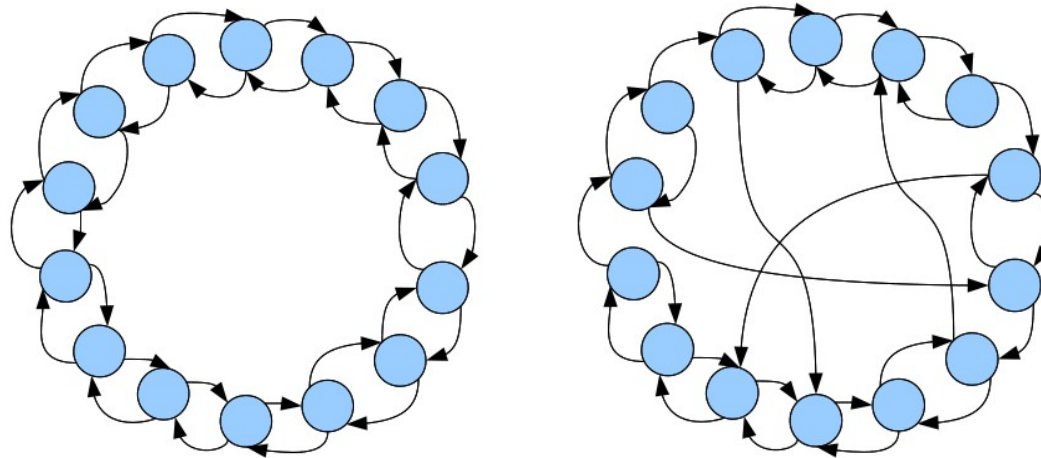
- Baseline Homophily
  - Constraints to the local social world
  - Naturally people work/live/spend leisure time with similar others
  - But: One incidentally meets people that do not match affiliation preference set





## Watts-Strogatz' Small-World Generator (SW)

- Agents are randomly put in line and connected to  $k$  neighbours
- Rewiring according to lifestyle preferences
- Does not consider agent's geographical position
  - Requires regular structure with equal number of in and out links for all agents



## Social distance attachment (SD)

- Link agents with probability based on social distance in each of a set of dimensions (Boguna et.al. 2004)

- Here: Two dimensions

- Geographical distance

- Life style preferences

$$p_{ij} = \left( w_i^g (b^g)^{-1} c_g (h_i^g h_j^g) \right) + \left( w_i^l (b^l)^{-1} c_l (h_i^l h_j^l) \right) k_i$$

$$b_g = \sum_{k,l=1}^{k,l=n} c_k (h_k^g h_l^g)$$

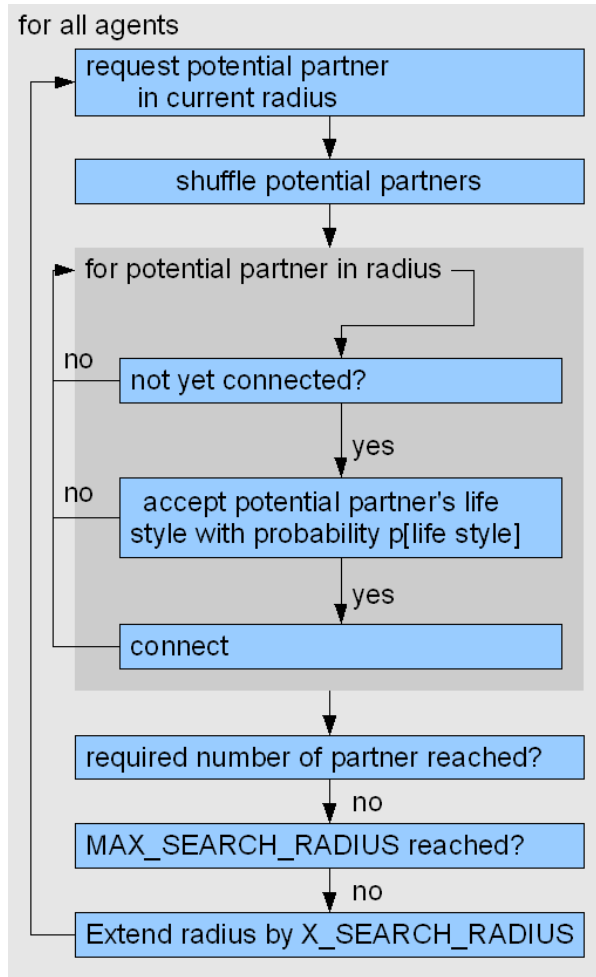
w=dim. weight

b= length normalisation

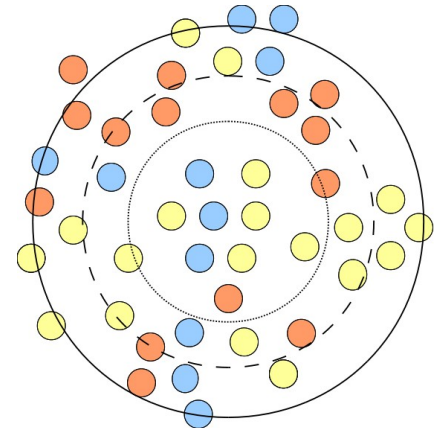
k= degree preference

- Considers local circumstances in supporting near-by partners (inbreeding homophily)
- Drawback: requires sort of global knowledge
- Asymmetrical relationships: Define an individual's position in the social space for both in-going and out-going links

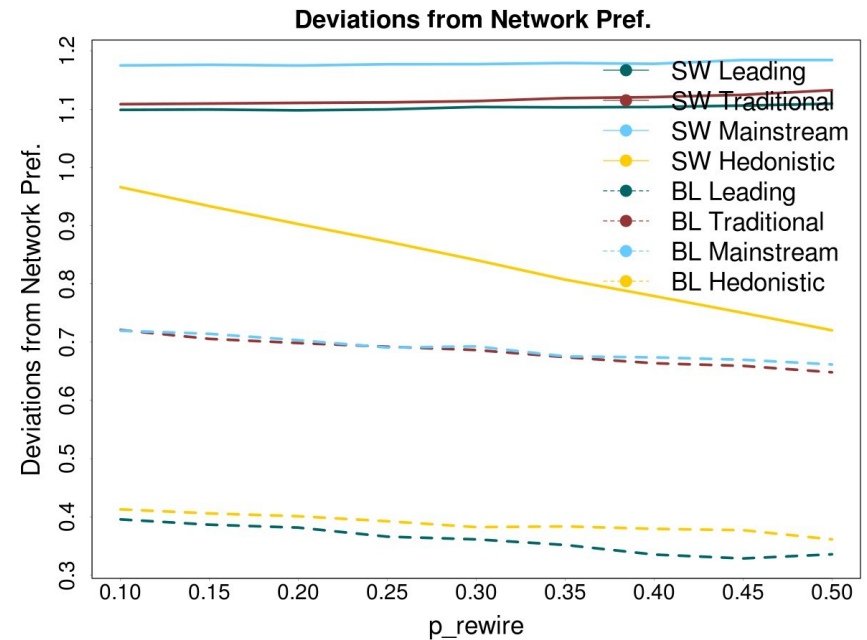
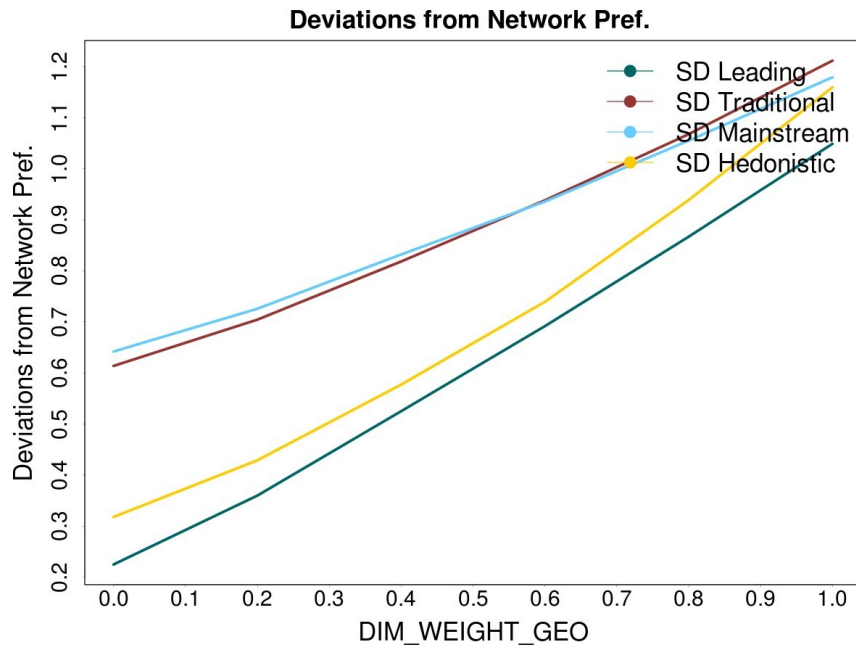
## Baseline homophily considering network generator (BL)



- Considers local circumstances: Actors may connect to those others who are available within the boundaries they are agitating
- Distant links are random  
> no global knowledge required
- Max-Search-Radius and Extending-Search-Radius are potentially life style specific

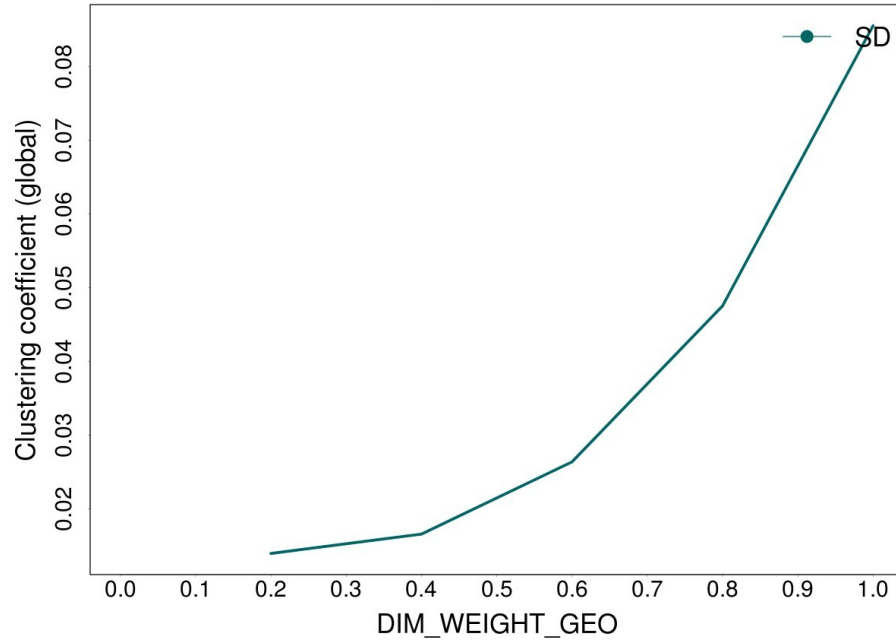


# Deviations from partner life style preferences

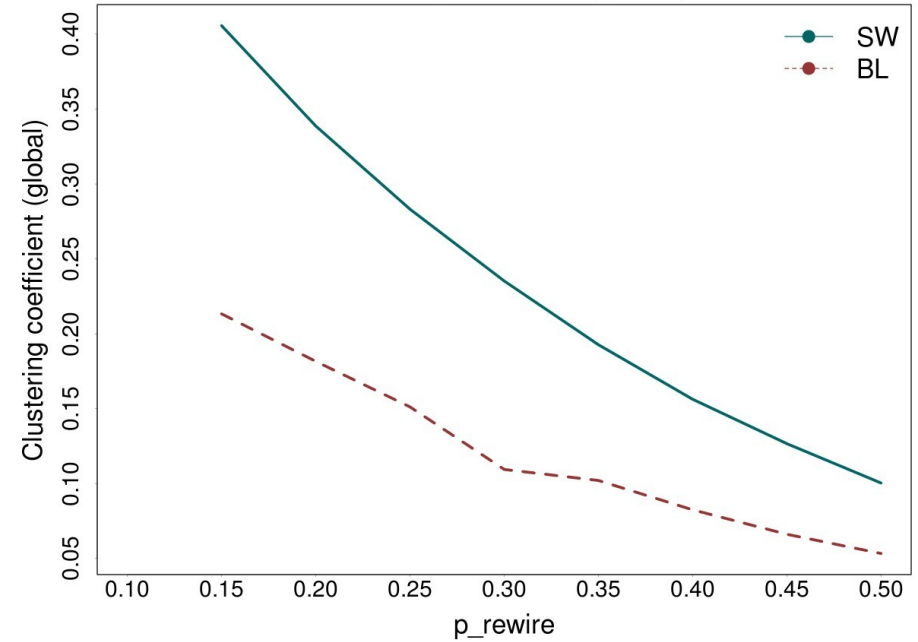


# Clustering Coefficient

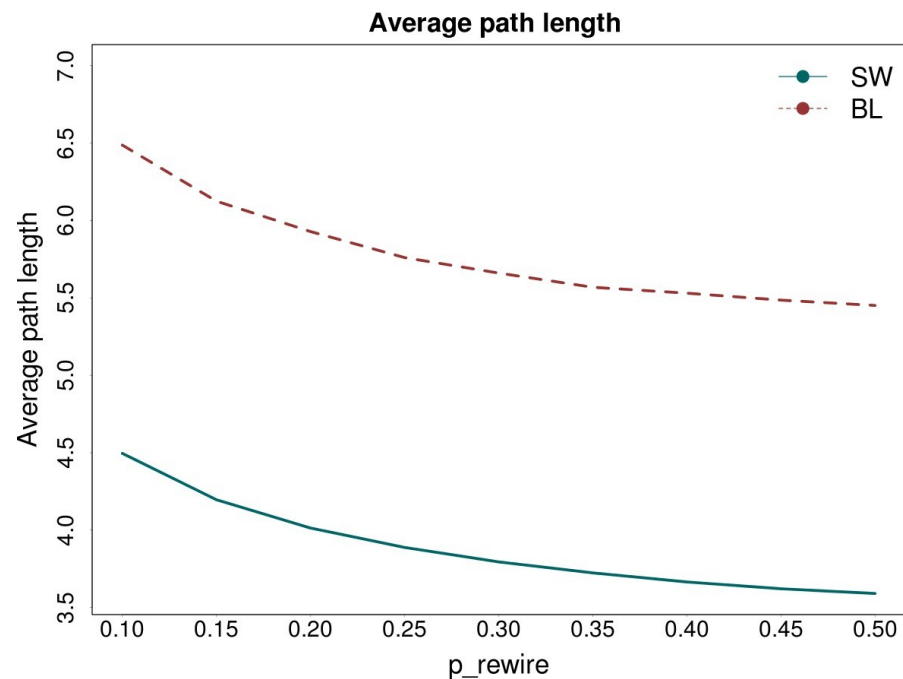
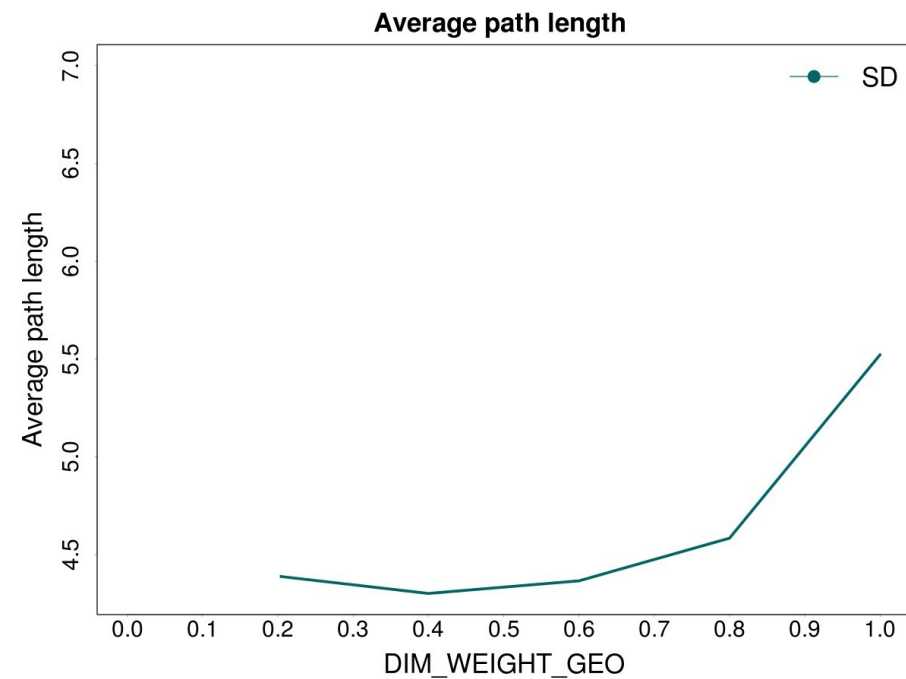
Clustering coefficient (global)



Clustering coefficient (global)

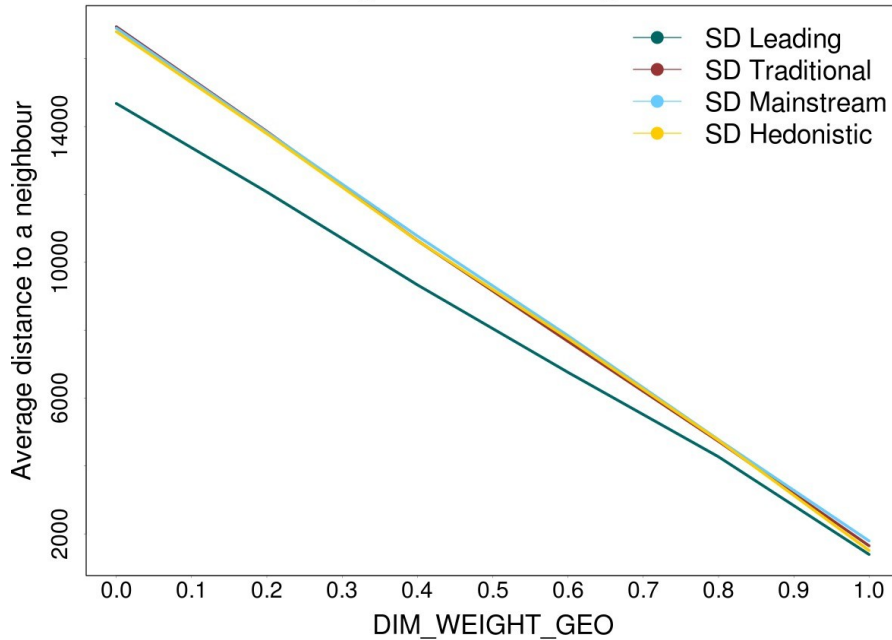


# Average path length

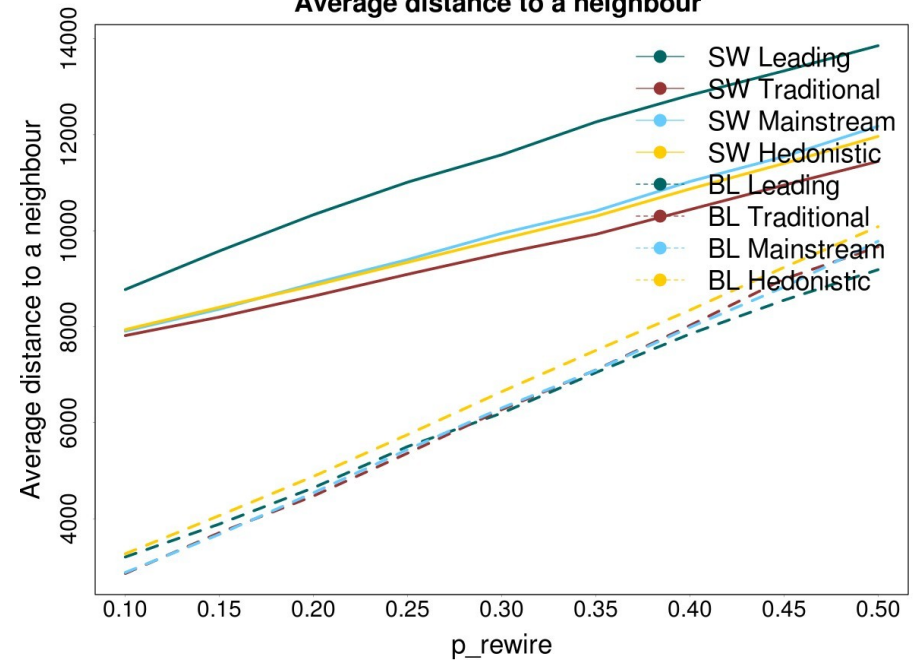


# Average distance to nearest neighbours

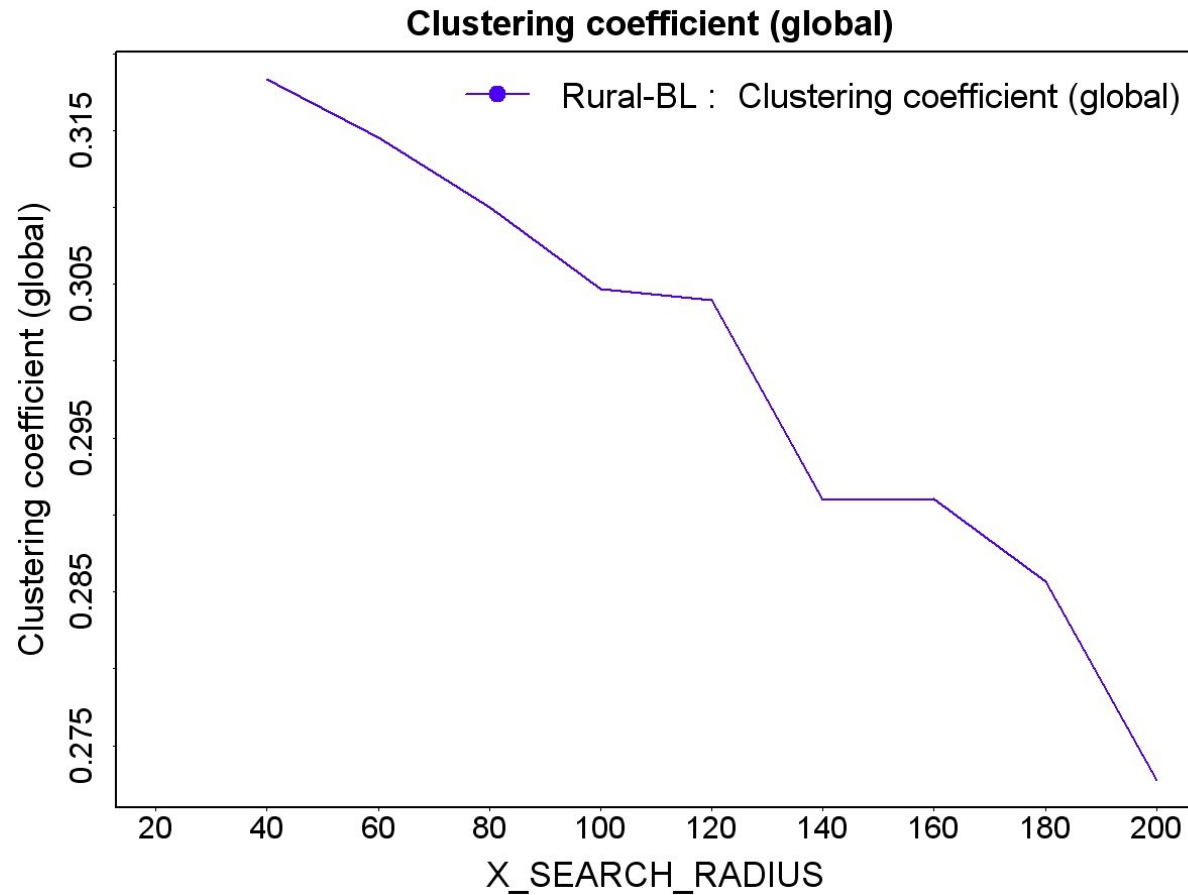
Average distance to a neighbour



Average distance to a neighbour



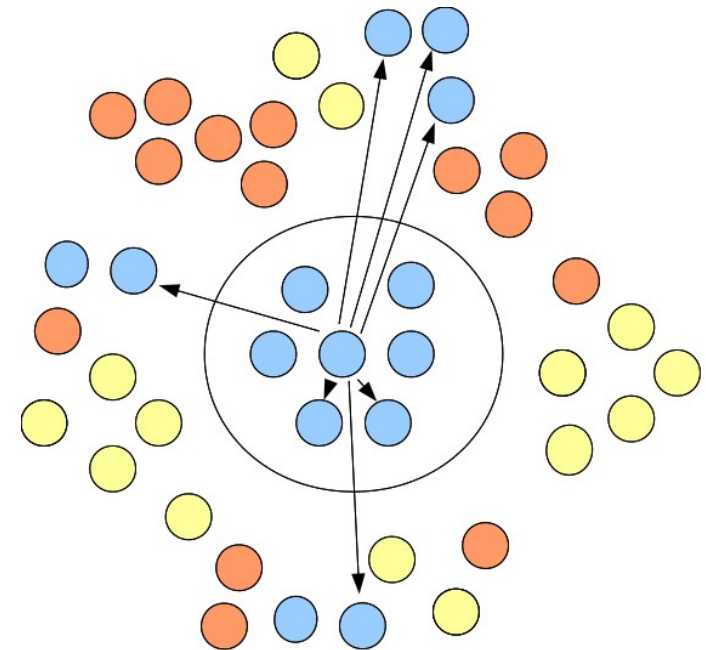
# Maximum search radius





## Challenges

- Collection of empirical data
  - Network size
  - Geographical position
  - Preferences regarding network partner's life style
- Several distant links to a cluster:  
If Agent A knows Agent B far away, it is likely to also know Agent C that is acquainted to B



## Discussion

- Our baseline homophily generator:
  - Simple
  - Plausible
  - Adjustable
- Social Distance Attachement
- Outlook
  - Studie interplay of parameters
  - challenges

# Appendix

## Lifestyle network preferences

	Leading	Traditional	Mainstream	Hedonistic
In-degree	15	5	5	10
p_rewire	0.2	0.05	0.1	0.2
p_links to....				
Leading	0.8	0.0	0.0	0.2
Traditional	0.6	0.3	0.1	0.0
Main-stream	0.6	0.1	0.3	0.0
Hedonistic	0.5	0.0	0.0	0.5

## Comparison: Boguna 2004 ↔ SpatialMilieuDistanceGen.

- Boguna:
  - Explicit homophily factor (alpha)
  - Conversion distance > prob in formula
  - Degree dist: length factor
- SpatialMilieu
  - Agent function: distance → 0..1
  - Explicit k

$$\sum_{n=1}^{n=dim} w_n \frac{1}{1 + b_n^{-1} d_n (h_i^n, h_j^n)^{a_n}}$$

$$p_{ij} = w_i^g (b^g)^{-1} d_g (h_i^g, h_j^g) + w_i^l (b^l)^{-1} d_l (h_i^l, h_j^l) k_i$$

$$b_g = \sum_{k,l=1}^{k,l=n} d_k (h_k^g, h_l^g)$$

# Netzwerkgeneratoren

	inbreed-ing homo-phily	baseline homo-phily	geograph-ical location	local clustering	re-wiring
<b>BaselineDhhRadius- NetworkBuilder</b>	1	1	1	1	1
<b>IdealDhhRadius- NetworkBuilder</b>	1	0	1	1	1
<b>SmallWorld-NetworkBuilder</b>	0	0	0	0	1

# Stufen der Netzwerkmodellierung

	<i>Einfachheit</i> ←	→ <i>Realistische Abbildung</i>		
	<b>First-Guess Modellierung</b>	<b>Homogene Modellierung</b>	<b>Heterogene Modellierung</b>	<b>Dynamische Modellierung</b>
<b>Netzwerkstruktur</b>	Gitter (Clustering)	Small-World (kurze Pfadlängen)	Soziale Kreise ( $L_2$ )	Social Distance Attachment ( $L_1$ )
<b>Netzwerkdynamik</b>	Statisches Netzwerk	Entstehung und Auflösung von Beziehungen	Berücksichtigung von Homophilie, heterogene Beziehungsstärke	Dynamische Entwicklung der Beziehungsstärke
<b>Kontaktdynamik</b>	Kontakte zu jedem Simulationszeitschritt	Längere Kontaktabstände randomisiert	Kontaktabstände basierend auf Agenteneigenschaften	Dynamische Kontaktabstände
<b>Soziales Modell des Agenten</b>	Wahrnehmung des gesamten Netzwerkes	Homogene, eingeschränkte Sichtweite	Heterogene Sichtweiten	Individuelles soziales Modell

# Kriterien für die Agenten-basierte Modellierung

- Effizienz der Generierung
- Verfügbarkeit der Parameter
- Transparenz der Prozesse
- mögliche Abschätzung von Unsicherheiten
- Angemessenheit bezüglich des Modellzwecks
- **Kompatibilität mit Lebensstilkonzept**