

# **Einblick in die Analyse sozialer Netzwerke mit R**

GetTogetherR

Till Hovestadt

Institut für Soziologie, Universität Leipzig

11.05.2023

# Über mich

## Zur Person

- 2016-2019 B.A. Soziologie, Universität Leipzig

## Zur Person

- 2016-2019 B.A. Soziologie, Universität Leipzig
- 2019-2022 M.A. Soziologie, Universität Leipzig

## Zur Person

- 2016-2019 B.A. Soziologie, Universität Leipzig
- 2019-2022 M.A. Soziologie, Universität Leipzig
- 2022– Wissenschaftlicher Mitarbeiter und Promovend im Forschungsprojekt SERIOUS, Universität Leipzig

## Zur Person

- 2016-2019 B.A. Soziologie, Universität Leipzig
- 2019-2022 M.A. Soziologie, Universität Leipzig
- 2022– Wissenschaftlicher Mitarbeiter und Promovend im Forschungsprojekt SERIOUS, Universität Leipzig
- Für weitere Infos/Kontakt: Siehe [Institutswebsite](#), oder [E-Mail](#)

## Forschungsprojekt SERIOUS

- Mitarbeiter des DFG-Forschungsprojektes “Social Embeddedness in Social Networks and the Reproduction of Socioeconomic Inequality in Educational Attainment” (**SERIOUS**), geleitet durch Georg Lorenz

# Forschungsprojekt SERIOUS

- Mitarbeiter des DFG-Forschungsprojektes “Social Embeddedness in Social Networks and the Reproduction of Socioeconomic Inequality in Educational Attainment” (SERIOUS), geleitet durch Georg Lorenz
- Untersuchung der sozialen Einbettung von Schüler:innen (SuS) in Freundschaftsnetzwerke und der Konsequenzen, die das für sozio-ökonomische Bildungsungleichheiten hat:



# Forschungsprojekt SERIOUS

- Mitarbeiter des DFG-Forschungsprojektes “Social Embeddedness in Social Networks and the Reproduction of Socioeconomic Inequality in Educational Attainment” (SERIOUS), geleitet durch Georg Lorenz
- Untersuchung der sozialen Einbettung von Schüler:innen (SuS) in Freundschaftsnetzwerke und der Konsequenzen, die das für sozio-ökonomische Bildungsungleichheiten hat:
  - SES Segregation in Freundschaftsnetzwerken

# Forschungsprojekt SERIOUS

- Mitarbeiter des DFG-Forschungsprojektes “Social Embeddedness in Social Networks and the Reproduction of Socioeconomic Inequality in Educational Attainment” (**SERIOUS**), geleitet durch Georg Lorenz
- Untersuchung der sozialen Einbettung von Schüler:innen (SuS) in Freundschaftsnetzwerke und der Konsequenzen, die das für sozio-ökonomische Bildungsungleichheiten hat:
  - SES Segregation in Freundschaftsnetzwerken
  - Stärke von sozialem Einfluss innerhalb und zwischen sozio-ökonomischen Gruppen

# Forschungsprojekt **SERIOUS**

- Mitarbeiter des DFG-Forschungsprojektes “Social Embeddedness in Social Networks and the Reproduction of Socioeconomic Inequality in Educational Attainment” (**SERIOUS**), geleitet durch Georg Lorenz
- Untersuchung der sozialen Einbettung von Schüler:innen (SuS) in Freundschaftsnetzwerke und der Konsequenzen, die das für sozio-ökonomische Bildungsungleichheiten hat:
  - SES Segregation in Freundschaftsnetzwerken
  - Stärke von sozialem Einfluss innerhalb und zwischen sozio-ökonomischen Gruppen
  - Zusammenspiel von Homophilie, Sozialkapital und sozialem Einfluss in der Produktion von sozio-ökonomischen Bildungsungleichheiten

# Grundlagen der Netzwerkanalyse

## Was ist ein soziales Netzwerk?

- Endliches Set an Akteur:innen (auch Knoten), die durch Beziehungen (auch Kanten) miteinander verbunden sind (Wasserman and Faust 1994)

# Was ist ein soziales Netzwerk?

- Endliches Set an Akteur:innen (auch Knoten), die durch Beziehungen (auch Kanten) miteinander verbunden sind (Wasserman and Faust 1994)
- z.B. Freund:innen in einer Schulklasse; Wissenschaftler:innen, die miteinander publiziert haben

# Was ist ein soziales Netzwerk?

- Endliches Set an Akteur:innen (auch Knoten), die durch Beziehungen (auch Kanten) miteinander verbunden sind (Wasserman and Faust 1994)
- z.B. Freund:innen in einer Schulklasse; Wissenschaftler:innen, die miteinander publiziert haben
- Soziale Netzwerkanalyse untersucht die Auswirkungen von Beziehungen und Netzwerkstruktur

# Forschungsfragen in der Netzwerkforschung

## 1. Selektion:



# Forschungsfragen in der Netzwerkforschung

## 1. Selektion:

- Fokus auf Beziehungen zwischen den Akteur:innen

# Forschungsfragen in der Netzwerkforschung

## 1. Selektion:

- Fokus auf Beziehungen zwischen den Akteur:innen
- Wie und warum entwickeln sich soziale Beziehungen zwischen den Akteur:innen in einem Netzwerk?

# Forschungsfragen in der Netzwerkforschung

## 1. Selektion:

- Fokus auf Beziehungen zwischen den Akteur:innen
- Wie und warum entwickeln sich soziale Beziehungen zwischen den Akteur:innen in einem Netzwerk?

## 2. Einfluss:

# Forschungsfragen in der Netzwerkforschung

## 1. Selektion:

- Fokus auf Beziehungen zwischen den Akteur:innen
- Wie und warum entwickeln sich soziale Beziehungen zwischen den Akteur:innen in einem Netzwerk?

## 2. Einfluss:

- Fokus auf Veränderungen in den Akteur:innen

# Forschungsfragen in der Netzwerkforschung

## 1. Selektion:

- Fokus auf Beziehungen zwischen den Akteur:innen
- Wie und warum entwickeln sich soziale Beziehungen zwischen den Akteur:innen in einem Netzwerk?

## 2. Einfluss:

- Fokus auf Veränderungen in den Akteur:innen
- Wie und warum verändern die Beziehungen zwischen den Akteur:innen die individuellen Attribute der Akteur:innen?

# Klassische Netzwerkstudien

- Segregationsmodell nach Schelling (Schelling 1969, 1971)

# Klassische Netzwerkstudien

- Segregationsmodell nach Schelling (Schelling 1969, 1971)
- Strength of Weak Ties (Granovetter 1973)

# Klassische Netzwerkstudien

- Segregationsmodell nach Schelling (Schelling 1969, 1971)
- Strength of Weak Ties (Granovetter 1973)
- Inequality and Heterogeneity. A primitive theory of social structure (Blau 1977)



# Klassische Netzwerkstudien

- Segregationsmodell nach Schelling (Schelling 1969, 1971)
- Strength of Weak Ties (Granovetter 1973)
- Inequality and Heterogeneity. A primitive theory of social structure (Blau 1977)
- Structural holes: The social structure of competition (Burt 1992)

## Aktuelle Netzwerkstudien

- What drives ethnic homophily? A relational approach on how ethnic identification moderates preferences for same-ethnic friends (Leszczensky and Pink 2019) in *American Sociological Review*

## Aktuelle Netzwerkstudien

- What drives ethnic homophily? A relational approach on how ethnic identification moderates preferences for same-ethnic friends (Leszczensky and Pink 2019) in *American Sociological Review*
- The Trojan-horse mechanism: How networks reduce gender segregation (Arvidsson, Collet, and Hedström 2021) in *Science Advances*

## Aktuelle Netzwerkstudien

- What drives ethnic homophily? A relational approach on how ethnic identification moderates preferences for same-ethnic friends (Leszczensky and Pink 2019) in *American Sociological Review*
- The Trojan-horse mechanism: How networks reduce gender segregation (Arvidsson, Collet, and Hedström 2021) in *Science Advances*
- Ethnic diversity fosters the social integration of refugee students (Boda et al. 2023) in *Nature Human Behavior*

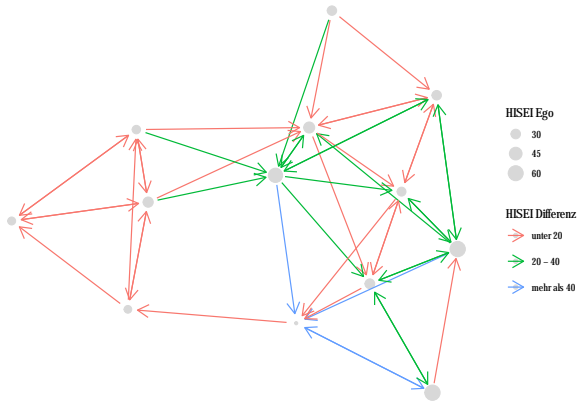
# Einfache Netzwerkanalyse

# Visualisierung

```
firstgraph <- igraph::graph.adjacency(w1) # igraph-object aus Freundschaftsmatrix bilden
class_graph <- ggraph(firstgraph,
  layout = "auto") +
  geom_node_point(mapping = aes(size = (classdata$HISEI / 10)),
    color = "grey",
    alpha = 0.6,
    show.legend = T) +
  geom_edge_link(
    mapping = aes(colour = difference1),
    alpha = 1,
    arrow = arrow(length = unit(4, "mm")),
    end_cap = circle(3, "mm"),
    start_cap = circle(3, "mm")
  )
```

# HISEI Differenzen in Freundschaften von SuS

Welle 1



*Eigene Berechnungen, Daten CILS4EU Deutschland (Kaher et al. 2016)*

# Pfade & Pfadlängen

- Ein Pfad ist die Verbindung zweier Knoten



# Pfade & Pfadlängen

- Ein Pfad ist die Verbindung zweier Knoten
  - Pfadlänge = 1  $\rightarrow$  direkte Verbindung der Knoten

# Pfade & Pfadlängen

- Ein Pfad ist die Verbindung zweier Knoten
  - Pfadlänge = 1  $\rightarrow$  direkte Verbindung der Knoten
  - Pfadlänge  $> 1 \rightarrow$  indirekte Verbindung der Knoten

# Pfade & Pfadlängen

- Ein Pfad ist die Verbindung zweier Knoten
  - Pfadlänge = 1  $\rightarrow$  direkte Verbindung der Knoten
  - Pfadlänge  $> 1 \rightarrow$  indirekte Verbindung der Knoten
- Die Distanz zweier Knoten ist der kürzeste Pfad zwischen ihnen

# Distanz in R

```
# Matrix der Distanzen von Ego (Reihe) zu Alter (Spalte)
dist <- igraph::distances(firstgraph,
                           mode = "in")
dist[c(1:5), c(1:5)]
```

	20200101	20200103	20200104	20200105	20200108
20200101	0	Inf	Inf	Inf	Inf
20200103	3	0	2	2	1
20200104	2	3	0	1	3
20200105	2	4	1	0	2
20200108	2	3	1	1	0

```
# Mittlere Distanz
igraph::mean_distance(firstgraph,
                       directed = TRUE)
```

```
[1] 2.111111
```

```
# Maximale Distanz
max(dist)
```

```
[1] Inf
```

# Zentralität

- Die Zentralität eines Knotens drückt seine Prominenz aus

# Zentralität

- Die Zentralität eines Knotens drückt seine Prominenz aus
- Der Grad (auch **degree**) eines Knotens drückt die Anzahl der Kanten des Knotens aus

# Zentralität

- Die Zentralität eines Knotens drückt seine Prominenz aus
- Der Grad (auch **degree**) eines Knotens drückt die Anzahl der Kanten des Knotens aus
- Die Dichte eines Netzwerkes drückt aus, wie viele Kanten es in Relation zu den (mathematisch) maximal möglichen Kanten gibt

# Degree in R

```
igraph::degree(firstgraph,  
  mode = "in")
```

```
20200101 20200103 20200104 20200105 20200108 20200109 20200113 20200114  
      0      2      5      4      5      5      0      6  
20200116 20200117 20200118 20200120 20200122 20200123  
      5      3      3      3      2      5
```

```
# Mittlerer Grad  
mean(igraph::degree(firstgraph,  
  mode = "in"))
```

```
[1] 3.428571
```

```
# SD  
sqrt(var(igraph::degree(firstgraph,  
  mode = "in")))
```

```
[1] 1.910066
```



## Density in R

```
igraph::graph.density(firstgraph)
```

```
[1] 0.2637363
```

# Homophily

- Tendenz, eher Beziehungen zu ähnlichen Akteur:innen zu schließen  
(McPherson, Smith-Lovin, and Cook 2001)

# Homophily

- Tendenz, eher Beziehungen zu ähnlichen Akteur:innen zu schließen  
(McPherson, Smith-Lovin, and Cook 2001)
- $e_{AB}$  := Anzahl an Kanten zwischen den Gruppen

# Homophily

- Tendenz, eher Beziehungen zu ähnlichen Akteur:innen zu schließen  
(McPherson, Smith-Lovin, and Cook 2001)
- $e_{AB}$  := Anzahl an Kanten zwischen den Gruppen
- $m$  := Gesamtanzahl an Kanten

# Homophily

- Tendenz, eher Beziehungen zu ähnlichen Akteur:innen zu schließen  
(McPherson, Smith-Lovin, and Cook 2001)
- $e_{AB}$  := Anzahl an Kanten zwischen den Gruppen
- $m$  := Gesamtanzahl an Kanten
- Homophily existiert, wenn  $\frac{e_{AB}}{m}$  signifikant geringer als die Wahrscheinlichkeit einer zufälligen Verbindung von zwei Akteur:innen

# Homophily in R i

```
table(factor(  
  classdata$male,  
  levels = c(0:1),  
  labels = c("Female",  
             "Male")  
),  
useNA = "ifany")
```

Female	Male
7	7

# Homophily in R ii

```
# Gender als Attribut hinzufügen
firstgraph <- set_vertex_attr(graph = firstgraph,
                             name = "gender",
                             value = classdata$male)
```

```
# Vektoren mit Männern und Frauen bilden
```

```
male <- V(firstgraph)[gender == 1]
female <- V(firstgraph)[gender == 0]
```

```
# Proportion Male (Gegen-proportion = Female)
(frac_male <- length(male)/vcount(firstgraph))
```

```
[1] 0.5
```

```
# Wahrscheinlichkeit, dass zufällige Kante zwischen zwei Akteur:innen Different-Gender ist
(pmf <- 2 * frac_male * (1 - frac_male))
```

```
[1] 0.5
```

```
# Share an Intergruppenbeziehungen
```

```
emf <- length(E(firstgraph)[male %--% female])
(emf / ecount(firstgraph))
```

```
[1] 0.375
```

# Homophily in R iii

```
# T-Test
prop.test(emf,
          ecount(firstgraph),
          p = pmf,
          alternative="less")
```

1-sample proportions test with continuity correction

```
data:  emf out of ecount(firstgraph), null probability pmf
X-squared = 2.5208, df = 1, p-value = 0.05618
alternative hypothesis: true p is less than 0.5
95 percent confidence interval:
 0.0000000 0.5041197
sample estimates:
      p
0.375
```



# Kohäsion

- Z.B. mit Clustering-Koeffizienten

# Kohäsion

- Z.B. mit Clustering-Koeffizienten
- Zeigt Anteil an  $i$ 's Kontakten, die untereinander verbunden sind

# Kohäsion

- Z.B. mit Clustering-Koeffizienten
- Zeigt Anteil an i's Kontakten, die untereinander verbunden sind
- 

$$C_l(i) = \frac{2 \times E(i)}{k(i) \times (k(i) - 1)}$$

# Kohäsion

- Z.B. mit Clustering-Koeffizienten
- Zeigt Anteil an i's Kontakten, die untereinander verbunden sind

- 

$$C_l(i) = \frac{2 \times E(i)}{k(i) \times (k(i) - 1)}$$

- $E(i) :=$  Kanten unter i's Kontakten

# Kohäsion

- Z.B. mit Clustering-Koeffizienten
- Zeigt Anteil an i's Kontakten, die untereinander verbunden sind
- 

$$C_l(i) = \frac{2 \times E(i)}{k(i) \times (k(i) - 1)}$$

- $E(i)$  := Kanten unter i's Kontakten
- $k(i)$  := Anzahl Kanten von i

# Kohäsion in R

```
c_l <- igraph::transitivity(firstgraph,  
                             type = "local")
```

```
table(c_l)
```

```
c_l  
0.428571428571429 0.464285714285714 0.466666666666667 0.5  
      1          1          1          1  
0.6 0.666666666666667 0.7 0.733333333333333  
      3          1          1          1  
      1  
      3
```

```
mean(c_l,  
      na.rm = T)
```

```
[1] 0.6738095
```

# Fortgeschrittene Netzwerkanalyse

# SAOMs

- **Stochastic Actor Oriented Models (SAOMs)** (Koskinen and Snijders 2023; Ripley et al. 2021; Snijders 2011)



# SAOMs

- Stochastic Actor Oriented Models (SAOMs) (Koskinen and Snijders 2023; Ripley et al. 2021; Snijders 2011)
  - Arbeiten mit längsschnittliche Netzwerkdaten

# SAOMs

- Stochastic Actor Oriented Models (SAOMs) (Koskinen and Snijders 2023; Ripley et al. 2021; Snijders 2011)
  - Arbeiten mit längsschnittliche Netzwerkdaten
  - Adressieren das Problem der Trennung von Selektion und Einfluss

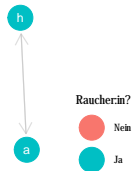
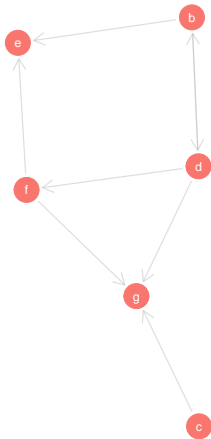
- Stochastic Actor Oriented Models (SAOMs) (Koskinen and Snijders 2023; Ripley et al. 2021; Snijders 2011)
  - Arbeiten mit längsschnittliche Netzwerkdaten
  - Adressieren das Problem der Trennung von Selektion und Einfluss
    - Gleichzeitige Entwicklung von Netzwerkstruktur (Selektion) und Verhalten (Einfluss) modelliert

- Stochastic Actor Oriented Models (SAOMs) (Koskinen and Snijders 2023; Ripley et al. 2021; Snijders 2011)
  - Arbeiten mit längsschnittliche Netzwerkdaten
  - Adressieren das Problem der Trennung von Selektion und Einfluss
    - Gleichzeitige Entwicklung von Netzwerkstruktur (Selektion) und Verhalten (Einfluss) modelliert
  - Kombiniert Logik aus Regressionsmodellen mit Agentenbasierten Modellen

- Stochastic Actor Oriented Models (SAOMs) (Koskinen and Snijders 2023; Ripley et al. 2021; Snijders 2011)
  - Arbeiten mit längsschnittliche Netzwerkdaten
  - Adressieren das Problem der Trennung von Selektion und Einfluss
    - Gleichzeitige Entwicklung von Netzwerkstruktur (Selektion) und Verhalten (Einfluss) modelliert
  - Kombiniert Logik aus Regressionsmodellen mit Agentenbasierten Modellen
    - Zentrale Logik: Annahme kontinuierlicher Zeit, Simulation der Entwicklung von  $t_1$  zu  $t_2$  mit *Micro Steps*

# Selektion vs Einfluss

Freundschaften in einer Seminargruppe (T1)



Raucher:in?

● Nein

● Ja

*Simuliertes Netzwerk*

# Selektion vs Einfluss - Selektion

Freundschaften in einer Seminargruppe (T0)

h

Freundschaften in einer Seminargruppe (T1)

h

a

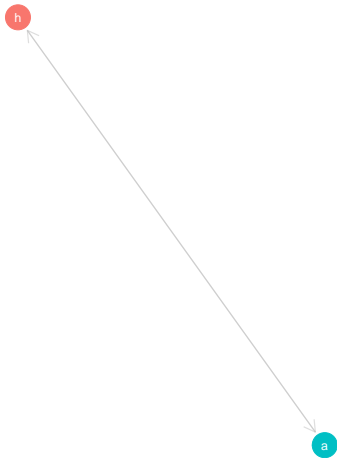
a

Raucher:in?

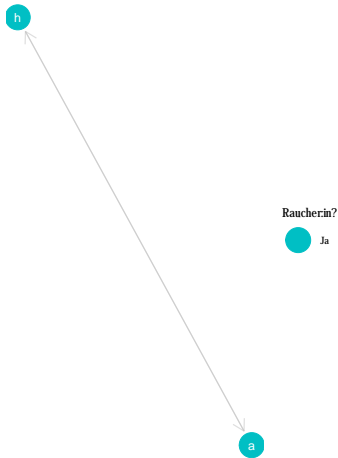
Ja

# Selektion vs Einfluss - Einfluss

Freundschaften in einer Seminargruppe (T0)



Freundschaften in einer Seminargruppe (T1)





# SAOM Micro-Steps i

S. Pink et al.

Journal of Choice Modelling 34 (2020) 100202

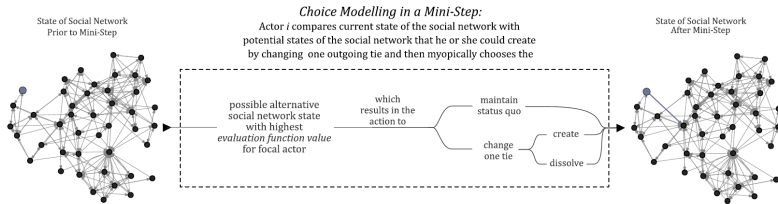


Fig. 1. Illustration of modelling network choices (selection) using the stochastic actor-oriented model.

(Pink, Kretschmer, and Leszczensky 2020: S. 5)

## SAOM Micro-Steps ii

- Micro-Steps, Beispiel Selektion (Einfluss analog modelliert) (Koskinen and Snijders 2021; Ripley et al. 2021; Snijders 2011)

## SAOM Micro-Steps ii

- Micro-Steps, Beispiel Selektion (Einfluss analog modelliert) (Koskinen and Snijders 2021; Ripley et al. 2021; Snijders 2011)
- *Zufällige:r* Akteur:in wird gewählt und kann

## SAOM Micro-Steps ii

- Micro-Steps, Beispiel Selektion (Einfluss analog modelliert) (Koskinen and Snijders 2021; Ripley et al. 2021; Snijders 2011)
- *Zufällige:r* Akteur:in wird gewählt und kann
  - neue Kante schließen,

## SAOM Micro-Steps ii

- Micro-Steps, Beispiel Selektion (Einfluss analog modelliert) (Koskinen and Snijders 2021; Ripley et al. 2021; Snijders 2011)
- *Zufällige:r* Akteur:in wird gewählt und kann
  - neue Kante schließen,
  - bestehende Kante auflösen oder

## SAOM Micro-Steps ii

- Micro-Steps, Beispiel Selektion (Einfluss analog modelliert) (Koskinen and Snijders 2021; Ripley et al. 2021; Snijders 2011)
- *Zufällige:r* Akteur:in wird gewählt und kann
  - neue Kante schließen,
  - bestehende Kante auflösen oder
  - Status Quo beibehalten

## SAOM Micro-Steps ii

- Micro-Steps, Beispiel Selektion (Einfluss analog modelliert) (Koskinen and Snijders 2021; Ripley et al. 2021; Snijders 2011)
- Zufällige:r Akteur:in wird gewählt und kann
  - neue Kante schließen,
  - bestehende Kante auflösen oder
  - Status Quo beibehalten
- Bewertung der Alternative erfolgt anhand der *objective function* (~ Regressionsmodell, spezifiziert durch Nutzer:in)

## SAOM Micro-Steps ii

- Micro-Steps, Beispiel Selektion (Einfluss analog modelliert) (Koskinen and Snijders 2021; Ripley et al. 2021; Snijders 2011)
- Zufällige:r Akteur:in wird gewählt und kann
  - neue Kante schließen,
  - bestehende Kante auflösen oder
  - Status Quo beibehalten
- Bewertung der Alternative erfolgt anhand der *objective function* (~ Regressionsmodell, spezifiziert durch Nutzer:in)
- Beginn bei  $t_1$ , Wiederholung des Vorganges, bis Netzwerk wie zu  $t_2$  erreicht ist



# Iteration of Micro-Steps

S. Pink et al.

Journal of Choice Modelling 34 (2020) 100202

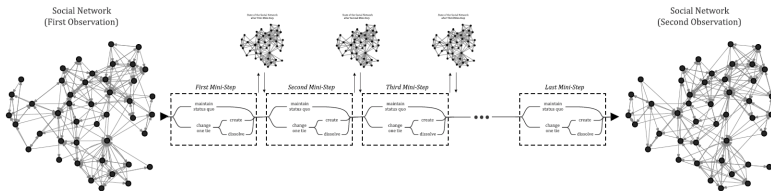


Fig. 2. Illustration of modelling interdependent individual choices using SAOMs.

(Pink et al. 2020: S. 7)

- Viele Iterationen (ca. 10, 000) dieses Vorganges der Netzwerkevolution

- Viele Iterationen (ca. 10, 000) dieses Vorganges der Netzwerkevolution
  - Wichtig für Konvergenz (Unverzerrtheit) und Präzision der Parameter

# Visuell Simulation der Netzwerkentwicklung

Visualisierung von SAOM Micro Steps (Selektion) in der Praxis (Adams and Schaefer 2018)

- Daten

## Visualisierung von SAOM Micro Steps (Selektion) in der Praxis (Adams and Schaefer 2018)

- Daten
  - Subsample der Teenage Friends and Lifestyle Study (West and Sweeting 1995)

## Visualisierung von SAOM Micro Steps (Selektion) in der Praxis (Adams and Schaefer 2018)

- Daten
  - Subsample der Teenage Friends and Lifestyle Study (West and Sweeting 1995)
  - Rauchverhalten und Alkoholkonsum unter schottischen Jugendlichen

# SAOM in R

# Results

saom.model.results

Estimates, standard errors and convergence t-ratios

	Estimate	Standard Error	Convergence t-ratio
1. rate constant friends rate (period 1)	15.0270	( 1.7663 )	-0.3205
2. rate constant friends rate (period 2)	12.8003	( 20.9791 )	-0.5005
3. rate constant friends rate (period 3)	11.9277	( 0.7374 )	0.1116
4. rate constant friends rate (period 4)	9.7823	( 2.0472 )	-0.3248
5. rate constant friends rate (period 5)	10.2772	( 0.5303 )	-0.1586
6. eval outdegree (density)	-2.7746	( 0.1816 )	-0.0670
7. eval reciprocity	2.7406	( 0.1972 )	-0.5113
8. eval GWESP I -> K -> J (69)	1.9082	( 0.0910 )	0.0935
9. eval indegree-popularity	-0.0535	( 0.0134 )	-0.2772
10. eval outdegree-activity	0.0711	( 0.0096 )	-0.0617
11. eval indegree-activity	-0.1994	( 0.0249 )	-0.4017
12. eval gender alter	-0.3366	( 0.1413 )	-0.1123
13. eval gender ego	-0.4257	( 0.1534 )	-0.3078
14. eval same class	0.5835	( 0.0919 )	-0.0122
15. eval native alter	-0.1126	( 0.0817 )	-0.0463
16. eval native ego	-0.0150	( 0.0761 )	0.0686
17. eval same native	0.0918	( 0.0519 )	0.1332
18. eval int. reciprocity x GWESP I -> K -> J (69)	-1.1469	( 0.1125 )	-0.4620
19. eval int. gender ego x gender alter	0.5789	( 0.2389 )	-0.1308

Overall maximum convergence ratio: 4.3808

Degrees constrained to maximum values:  
friends : 10

Total of 586 iteration steps.



# Nutzung des URZ Rechenclusters

- Das URZ der Uni Leipzig (UL) hat ein Scientific Computing (SC) Rechencluster, das sehr viele Ressourcen bereitstellt:

# Scientific Computing an der UL

- Das URZ der Uni Leipzig (UL) hat ein Scientific Computing (SC) Rechencluster, das sehr viele Ressourcen bereitstellt:
  - [URZ SC Website](#)

# Scientific Computing an der UL

- Das URZ der Uni Leipzig (UL) hat ein Scientific Computing (SC) Rechencluster, das sehr viele Ressourcen bereitstellt:
  - [URZ SC Website](#)
- Hier können remote Analysen auf leistungsstarken Rechenclustern durchgeführt werden

## Workflow

- Grundsätzlich sollten Analysen zunächst auf dem eigenen Rechner oder der online RStudio Instanz des URZ ([Zugriff zu RStudio](#)) getestet werden

## Workflow

- Grundsätzlich sollten Analysen zunächst auf dem eigenen Rechner oder der online RStudio Instanz des URZ ([Zugriff zu RStudio](#)) getestet werden
- Auf den Clustern können keine interaktiven Jobs laufen. **!Bei Fehlern wird das R-Skript abgebrochen!**

## Workflow

- Grundsätzlich sollten Analysen zunächst auf dem eigenen Rechner oder der online RStudio Instanz des URZ ([Zugriff zu RStudio](#)) getestet werden
- Auf den Clustern können keine interaktiven Jobs laufen. **!Bei Fehlern wird das R-Skript abgebrochen!**
- Per [SSH](#) kann man den Server erreichen und Befehle geben

## Workflow

- Grundsätzlich sollten Analysen zunächst auf dem eigenen Rechner oder der online RStudio Instanz des URZ ([Zugriff zu RStudio](#)) getestet werden
- Auf den Clustern können keine interaktiven Jobs laufen. **!Bei Fehlern wird das R-Skript abgebrochen!**
- Per [SSH](#) kann man den Server erreichen und Befehle geben
- Per [SCP](#) kann man Dateien zwischen dem lokalen (eigenen) und remote (SC Server) austauschen



# Slurm Workload Manager

- **Slurm** ist die Software mittels derer Rechenzeit auf dem Cluster “beantragt” wird

# Slurm Workload Manager

- **Slurm** ist die Software mittels derer Rechenzeit auf dem Cluster “beantragt” wird
- Das SC stellt **ausführliche Anleitungen** (siehe auch die sehr gute **Kurzanleitung**) zu Slurm bereit

# Beispiel Slurm Job Datei

```
#!/bin/bash

## Resource Request: What do you need, '##' is a comment '#' not
#SBATCH --time=14-00:00:00      ## time: days-hours:min:sec
#SBATCH --mem=32G               ## RAM
#SBATCH --ntasks=1              ## tasks (for us 1)
#SBATCH --cpus-per-task=24      ## cpus as specified in the R script
#SBATCH --job-name="jobname"    ## job name (to find it later in the job queue)
#SBATCH --partition=galaxy-job   ## server partition -- here: galaxy

#SBATCH --mail-type=BEGIN       ## e-mail at starts
#SBATCH --mail-type=END         ## e-mail at end
#SBATCH --mail-type=FAIL        ## e-mail if fail
#SBATCH --mail-type=TIME_LIMIT_80 ## e-mail at 80% time limit

## Dependencies: load R into the system
module load R

## Job steps: which Scripts are to be run in which order?
Rscript name-of-rscript_1.R
Rscript name-of-rscript_2.R
Rscript name-of-rscript_3.R
```

# Gute Weiterführende Literatur

# Netzwerkanalyse im Allgemeinen

- Wasserman, Stanley & Faust, Katherine (1994). Social Network Analysis. Methods and Applications. Cambridge: Cambridge University Press

# Netzwerkanalyse im Allgemeinen

- Wasserman, Stanley & Faust, Katherine (1994). Social Network Analysis. Methods and Applications. Cambridge: Cambridge University Press
- Easley, David & Kleinberg, Jon (2010). Networks, Crowds, and Markets. Reasoning About a Highly Connected World. Cambridge: Cambridge University Press

## SAOM Literatur

- Koskinen, Johan, and Tom A. B. Snijders. 2023. "Multilevel longitudinal analysis of social networks". Journal of the Royal Statistical Society Series A: Statistics in Society 00:1–25. doi: 10.1093/jrsssa/qnac009.

## SAOM Literatur

- Koskinen, Johan, and Tom A. B. Snijders. 2023. "Multilevel longitudinal analysis of social networks". Journal of the Royal Statistical Society Series A: Statistics in Society 00:1–25. doi: 10.1093/jrsssa/qnac009.
- Pink, Sebastian, David Kretschmer, and Lars Leszczensky. 2020. "Choice Modelling in Social Networks Using Stochastic Actor-Oriented Models." Journal of Choice Modelling 34:100202. doi: 10.1016/j.jocm.2020.100202.



## SAOM Literatur

- Koskinen, Johan, and Tom A. B. Snijders. 2023. "Multilevel longitudinal analysis of social networks". *Journal of the Royal Statistical Society Series A: Statistics in Society* 00:1–25. doi: 10.1093/jrsssa/qnac009.
- Pink, Sebastian, David Kretschmer, and Lars Leszczensky. 2020. "Choice Modelling in Social Networks Using Stochastic Actor-Oriented Models." *Journal of Choice Modelling* 34:100202. doi: 10.1016/j.jocm.2020.100202.
- Ripley, Ruth M., Tom A. B. Snijders, Zsófia Boda, András Vörös, and Paulina Preciado. 2021. *Manual for RSIENA (Version April 24, 2021)*. Oxford: University of Oxford, Department of Statistics, Nuffield College; University of Groningen, Department of Sociology.

## SAOM Literatur

- Koskinen, Johan, and Tom A. B. Snijders. 2023. "Multilevel longitudinal analysis of social networks". *Journal of the Royal Statistical Society Series A: Statistics in Society* 00:1–25. doi: 10.1093/jrssa/qnac009.
- Pink, Sebastian, David Kretschmer, and Lars Leszczensky. 2020. "Choice Modelling in Social Networks Using Stochastic Actor-Oriented Models." *Journal of Choice Modelling* 34:100202. doi: 10.1016/j.jocm.2020.100202.
- Ripley, Ruth M., Tom A. B. Snijders, Zsófia Boda, András Vörös, and Paulina Preciado. 2021. *Manual for RSIENA (Version April 24, 2021)*. Oxford: University of Oxford, Department of Statistics, Nuffield College; University of Groningen, Department of Sociology.
- Snijders, Tom AB. 2011. "Statistical Models for Social Networks." *Annual Review of Sociology* 37:131–53.

# Hilfreiche Webseiten für SAOMs

- [RSiena Website](#)

# Hilfreiche Webseiten für SAOMs

- RSiena Website
- RSiena Mailingliste

# Danke für die Aufmerksamkeit!

[till.hovestadt@uni-leipzig.de](mailto:till.hovestadt@uni-leipzig.de)

# Bibliographie

## Bibliographie I

- Adams, Jimi, and David R. Schaefer. 2018. "Visualizing Stochastic Actor-Based Model Microsteps." *Socius: Sociological Research for a Dynamic World* 4:237802311881654. doi: 10.1177/2378023118816545.
- Arvidsson, Martin, F. Collet, and Peter Hedström. 2021. "The Trojan-Horse Mechanism: How Networks Reduce Gender Segregation." *Science Advances* 7(16):1–7. doi: 10.1126/sciadv.abf6730.
- Blau, Peter Michael. 1977. *Inequality and Heterogeneity. A Primitive Theory of Social Structure*. New York: Free Press.
- Boda, Zsófia, Georg Lorenz, Malte Jansen, Petra Stanat, and Aileen Edele. 2023. "Ethnic Diversity Fosters the Social Integration of Refugee Students." *Nature Human Behaviour*. doi: 10.1038/s41562-023-01577-x.
- Burt, Ronald S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.
- Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78(6):1360–80.

## Bibliographie II

- Koskinen, Johan, and Tom A. B. Snijders. 2021. "Random-Effect Multilevel Stochastic Actor-Oriented Models." *In Preparation*.
- Koskinen, Johan, and Tom A. B. Snijders. 2023. "Multilevel Longitudinal Analysis of Social Networks." *Journal of the Royal Statistical Society Series A: Statistics in Society* qnac009. doi: 10.1093/jrssa/qnac009.
- Leszczensky, Lars, and Sebastian Pink. 2019. "What Drives Ethnic Homophily? A Relational Approach on How Ethnic Identification Moderates Preferences for Same-Ethnic Friends." *American Sociological Review* 84(3):394–419. doi: 10.1177/0003122419846849.
- McPherson, J. Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44. doi: 10.1146/annurev.soc.27.1.415.
- Pink, Sebastian, David Kretschmer, and Lars Leszczensky. 2020. "Choice Modelling in Social Networks Using Stochastic Actor-Oriented Models." *Journal of Choice Modelling* 34:100202. doi: 10.1016/j.jocm.2020.100202.



## Bibliographie III

- Ripley, Ruth M., Tom A. B. Snijders, Zsófia Boda, András Vörös, and Paulina Preciado. 2021. *Manual for RSIENA (Version April 24, 2021)*. Oxford: University of Oxford, Department of Statistics, Nuffield College; University of Groningen, Department of Sociology.
- Schelling, Thomas C. 1969. "Models of Segregation." *The American Economic Review* 59(2):488–93.
- Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1:143–86.
- Snijders, Tom AB. 2011. "Statistical Models for Social Networks." *Annual Review of Sociology* 37:131–53. doi: 10.1146/annurev.soc.012809.102709.
- Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.
- West, P., and H. Sweeting. 1995. *Background Rationale and Design of the West of Scotland 11-16 Study. Working Paper*. 52. Glasgow: MRC Medical Sociology Unit Glasgow.