

Generierung von sozialen Netzwerken

Steffen Brauer
WiSe 2011/12
HAW Hamburg



Agenda

- Motivation
- Soziale Netzwerke
- Modelle
- Metriken
- Forschungsumfeld
- Ausblick

Motivation

- Wo gibt es Netzwerke?
 - Computernetzwerke
 - Kommunikationsnetzwerke
- Forschungsansätze
 - Existierende Netzwerke analysieren
 - Modelle finden, die reale Netzwerke nachbilden
- Probleme der Forschung
 - Netzwerke sind verteilt
 - Datenschutz

Motivation

Projekt Mindstone

Ziel:

- **Content-centric social Network**
- **Fokus auf eLearning**
- **Find "Learning-Friends" in Online Social Networks**



[w8]

Motivation

Woher Daten nehmen?



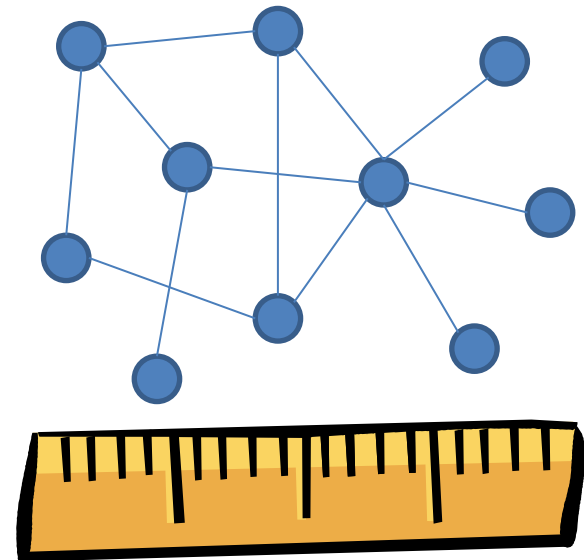
- Extrahieren
 - via APIs aus Netzwerken
- Generieren
 - und in ein modifiziertes Open Source Netzwerk laden

Soziale Netzwerke

- Soziale Netzwerke
 - Verbindung zwischen Personen
 - The Small-World Problem von Milgram (1967)^[5]
 - Alle Menschen in den USA kennen sich über 5 Freunde
 - Thema der Soziologie^[6]
 - Soziales Kapital
 - Interorganisations-Netzwerke
 - Interpersonelle Netzwerke
 - Auch mit Informatik: Wie entstehen Netzwerke?

Soziale Netzwerke

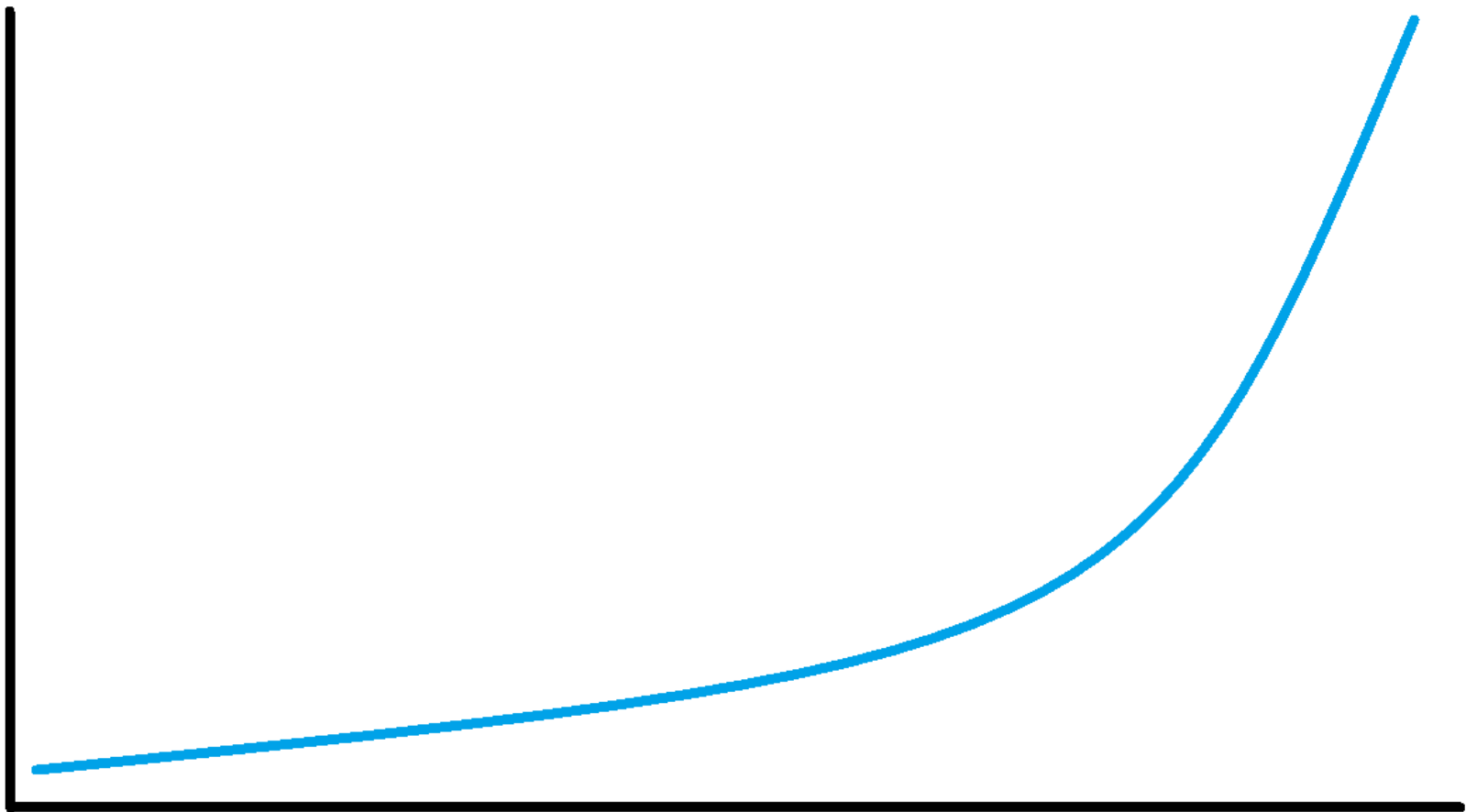
- Benutzer -> Knoten
- Freundschaftsbeziehungen
-> Kanten
- Eigenschaften
 - Degree Power Law
 - Densification Power Law
 - Small Diameter
 - Shrinking Diameter



[1][7][2]

Soziale Netzwerke

Power Law



Modelle

- Warum Modelle?
 - Beobachtete Effekte nachbilden
 - Extrapolation von realen Netzwerken
 - Experimente: Kleine Netzwerke vergrößern

Feature-driven

- Bilden Eigenschaften der realen Graphen nach

Intent-driven

- Simulieren den Erstellungsprozess eines Netzwerks

Structure-driven

- Bilden Strukturen eines vorgegebenen Graphen nach

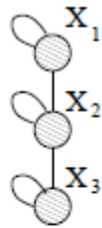
[4]

Forest Fire Modell

- Forward burning probability p
- Backward burning ratio r
- v erstellt Verbindungen
 - v wählt ambassador node w
 - Zufällig x, y mit geometrischer Verteilung $(1-p)^{-1}$ und $(1-rp)^{-1}$
 - x ausgehende Kanten
 - y eingehende Kanten
- Rekursion mit allen neuen w

[1]

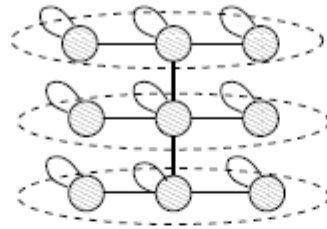
Kronecker Graphs



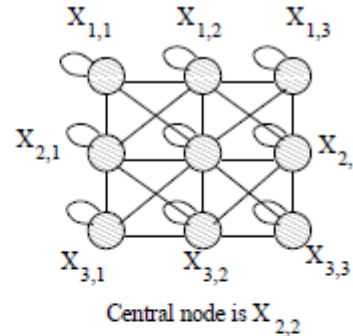
(a) Graph G_1

1	1	0
1	1	1
0	1	1

(d) Adjacency matrix of G_1



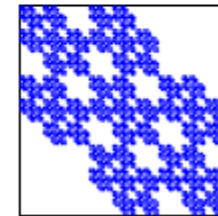
(b) Intermediate stage



(c) Graph $G_2 = G_1 \otimes G_1$

G_1	G_1	0
G_1	G_1	G_1
0	G_1	G_1

(e) Adjacency matrix of $G_2 = G_1 \otimes G_1$



(f) Plot of G_4

$\text{Kante}(X_{ij}, X_{kl}) \in G \otimes H$ iff $(X_i, X_k) \in G$ und $(X_j, X_l) \in H$

[3][7]

Background & Behavior Modell 1/2

- Background
 - Gruppen zu denen eine Person gehört
- Behaviour
 - Wie schnell bilde ich Beziehungen zu anderen
- Erste Phase (Generierung der Welt)
 - Welt wird in Cliques eingeteilt
 - Knoten werden in zufällige Gruppen aufgeteilt

[2]

Background & Behavior Modell 2/2

- Zweite Phase
 - Knoten betritt Netzwerk und hat einen Wert für
 - Extravertiertheit
 - Qualität
 - Verbindung bisherigen Knoten mit min/max
 - Knoten startet mit einer Wahrscheinlichkeit Walk durch Netzwerk
 - und verbindet sich mit Extravertiertheit*Qualität
- [2]

Metriken 1/3

- Wie bewertet man Modelle und Netzwerke?
 - **Metriken!**
- Degree Distribution
 - wird mittels Regression ermittelt
- Diameter
 - aufwendige Berechnung
- Clustering
 - $ij \wedge jk \rightarrow ik$

Metriken 2/3

- Betweenness Centrality
 - Relevanz für die Verbindung anderer Knoten
 - $P_i(kj)$
 - Assortativity Coefficient
 - Korrelation zwischen dem Grad der Nachbarn
- [2]

Metriken 3/3

	Princeton	Georgetown	Oklahoma	UNC	B & B	Forest Fire
Nodes	6596	9414	17425	18163	9000	9000
Edges	293320	425638	892528	766800	394512	300130
Avg. Grad	88,93	90,42	102,442	84,43	87,66	66,69
Grad γ	-1,13	-1,26	-1,40	-1,46	-1,19	-0,99
AC	0,091	0,075	0,073	0,0007	0,066	-0,34
Clustering	0,244	0,231	0,235	0,206	0,21	0,561
Diameter	9	11	9	7	7	10
Betweenness	2,525E-4	1,856E-4	1,01E-4	9,913E-5	1,031E-4	1,856E-4

[2]

Forschungsumfeld



[w9]

- **Stanford Network Analysis Project**
 - Jure Leskovec [w1]
- ACM Transactions on Knowledge Discovery from Data [w2]
- International Workshop on **Hot Topics in Peer-to-peer computing and Online Social networking** – HotPOST [w3]
- International Conference on **Knowledge Discovery and Data Mining** – SIGKDD [w4]
- International Conference on **Web Search and Data Mining** – WSDM [w5]
- Workshop on Algorithms and Models for the Web Graph – WAW [w6]
- **Workshop on Online Social Networks** – WOSN [w7]

Ausblick

- Testnetzwerk für Mindstone Projekt schaffen
 - Umgebung aufbauen
 - Mit Modellen experimentieren (erweitern/ändern)
 - Konkret: Gewichtete Kanten
 - Benutzerverhalten durch empirische Daten erweitern
- Risiken
 - Schwierige Evaluierung
 - Nur die Effekte werden erkannt, die auch modelliert wurden

Quellen

- [1] Leskovec, J., Kleinberg, J., Faloutsos, C. ; Graphs over time: Densification laws, shrinking diameters and possible explanations. In Proc. of ACM KDD(2005).
- [2] Foudalis, I., Jain, K., Papadimitriou, C., Sideri, M.; Modeling Social Networks through User Background and Behavior, Springer (2011).
- [3] Leskovec, J., Faloutsos, C.; Scalable modeling of real graphs using kronecker multiplication. In Proc. Of ICML (2007).
- [4] Alessandra Sala, Lili Cao, Christo Wilson, Robert Zablit, Haitao Zheng, and Ben Y. Zhao. 2010. Measurement-calibrated graph models for social network experiments. In *Proceedings of the 19th international conference on World wide web (WWW '10)*.

Quellen

- [5] Milgram, S. ; The Small-World Problem ; Psychology Today 1 (1967)
- [6] Weyer, J. ; Soziale Netzwerke: Konzepte und Methoden der sozialwissenschaftlichen Netzwerkforschung ; Oldenbourg Wissenschaftsverlag (2011)
- [7] Leskovec J. , Chakrabarti D., Kleinberg J., Faloutsos C.; Realistic, Mathematically Tractable Graph Generation and Evolution, Using Kronecker Multiplication in Knowledge Discovery in Databases PKDD 2005 (2005)

Quellen

- [w1] SNAP - Stanford Network Analysis Project - <http://snap.stanford.edu/index.html>
- [w2] Publications: TKDD - <http://dl.acm.org/pub.cfm?id=J1054>
- [w3] HotPOST 2012 - <http://cseweb.ucsd.edu/~tixu/HotPost/>
- [w4] ACM SIGKDD: Home Page - <http://www.kdd.org/>
- [w5] WSDM 2009 - <http://www.wsdm2009.org/>
- [w6] WAW 2012 - <http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=19029©ownerid=2>
- [w7] WOSN 2010 - <http://www.usenix.org/events/wosn10/index.html>
- [w8] Mindstone - <http://inet.cpt.haw-hamburg.de/projects/mindstone>
- [w9] <http://cs.stanford.edu/people/jure/images/jure-6.jpg>