Complex Systems Engineering

Graphen und Netzwerke

Prof. Dr. Christin Seifert

22. November 2017

University of Passau, WS 2017/2018

Überblick

1. Einführung
2. Grundlagen
3. Kleine-Welt Netzwerke
4. Skalenfreie Netzwerke
5. Zusammenfassung

Einführung

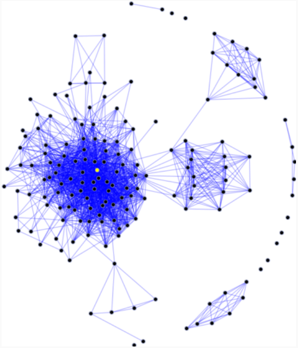


Abbildung 1: Visualisierung eines soziales Netzwerkes mit Verbindungen zwischen Personen (CC-SA 3.0, DarwinPeacock, via Wikimedia Commons)

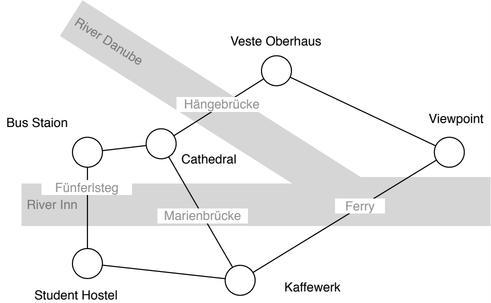


Abbildung 2: Skizze wichtiger Orter in Passau und ihrer Verbindungen

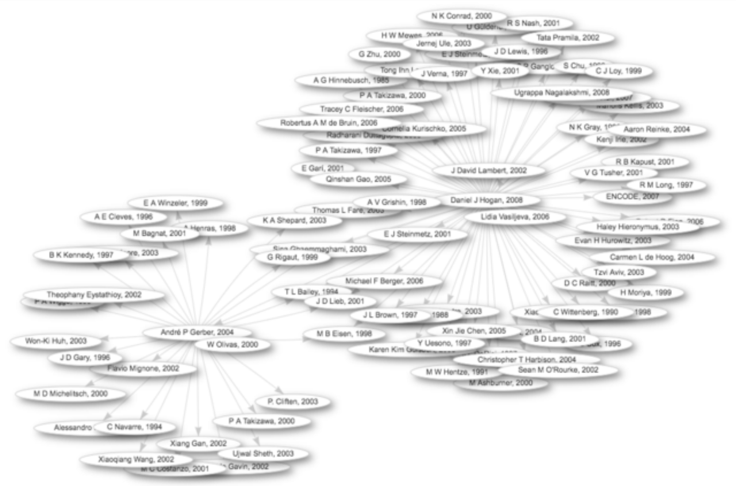


Abbildung 3: Zitationsgraph des Artikels „A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes “(CC-0 Public Domain Dedication, Daniel Mietchen, via Wikimedia Commons)

Netzwerke

Abbildung 4: Strommasten des Stromnetzwerkes in Österreich (CC-SA 3.0, DM.Dufek, via Wikimedia Commons)

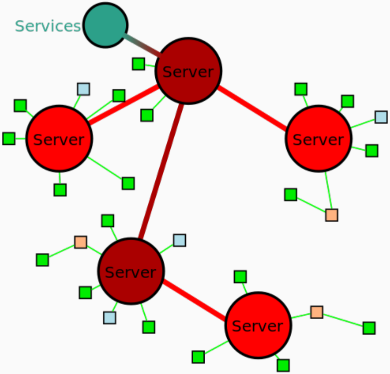


Abbildung 5: Internet Relay Chat (IRC) Netzwerk (Public Domain, via Wikimedia Commons)

Weitere Beispiele:

- Das gesamte Internet
- Wikipedia
- Aktienanteile von Firmen an anderen Firmen
- Schauspieler*innen, die in den gleichen Filmen spielen
- Telefonnetzwerke (mobil oder Festnetz)
- Schienennetz der Bahn, das Städte verbindet
- Reaktionen zwischen Chemikalien

Viele Phänomene der realen Welt haben die Struktur von Netzwerken. Graphentheoretische Ansätze helfen, diese zu modellieren und zu verstehen.

Grundlagen

Definition Graph

Ein Graph ist ein geordnetes Paar $G = (V, E)$ mit der Knotenmenge und V der Kantenmenge $E \subseteq V \times V$.

Definition gerichteter Graph

Ein gerichteter Graph ist ein Graph mit Kantenmenge $E = \Big\{(u, v) | u, v \in V, u \neq v\Big\}$.

Definition ungerichteter Graph

Ein unggerichteter Graph ist ein Graph mit Kantenmenge $E = \Big\{\{u, v\} | u, v \in V, u \neq v\Big\}$.

Bemerkung: Diese Definitionen erlauben keine Kanten von einem Knoten zu sich selbst (engl: self-loops, self-edges).

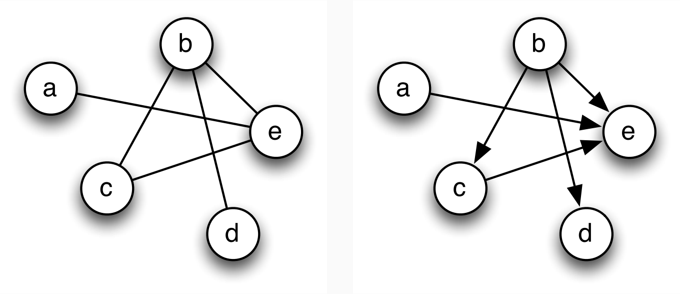


Abbildung 6: Ungerichteter und gerichteter Graphen mit jeweils 5 Knoten

Grad eines Knotens

Grad eines Knotens

In einem ungerichteten Graphen ist der Grad eines Knotens die Anzahl der Kanten zu (von) diesem Knoten.

Eingangsgrad eines Knotens

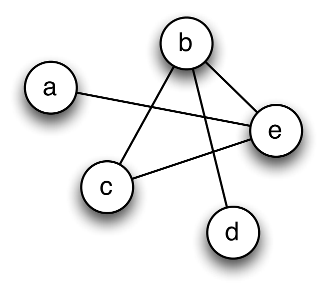
In einem gerichteten Graphen ist der Eingangsgrad eines Knotens die Anzahl der Kanten, die zu diesem Knoten führen.

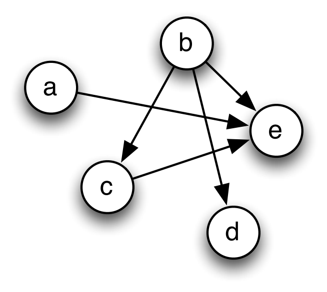
Ausgangsgrad eines Knotens

In einem gerichteten Graphen ist der Ausgangsgrad eines Knotens die Anzahl der Kanten, die von diesem Knoten weg führen.

Bemerkung: Beachte, dass die Definitionen spezifisch für Typen von Graphen sind (gerichtet/ungerichtet).

Aufgabe





- Bestimmen Sie die Knotengrade in den Graphen.

Grad eines Knotens

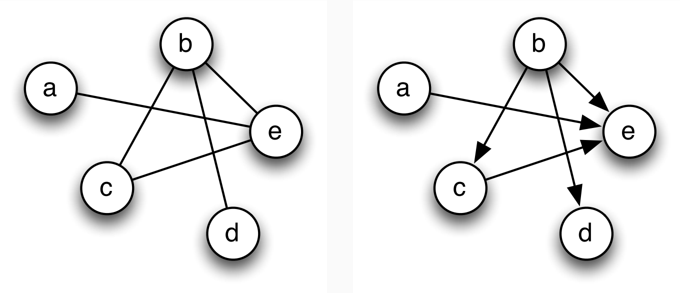


Abbildung 7: Ungerichteter und gerichteter Graphen mit jeweils 5 Knoten

- Grade im ungerichteten Graphen. Knoten a: 1; Knoten b: 3 usw.
- Grade im gerichteten Graphen. Knoten c: Eingangsgrad 1, Ausgangsgrad 1; Knoten b: Eingangsgrad 0, Ausgangsgrad 3

Dichte von Graphen

- Die Dichte eines Graphen wird über das Verhältnis der Anzahl existierender Kanten zu möglichen Kanten definiert.
- Anzahl möglicher Kanten in einem ungerichteten Graphen mit n Knoten kann über den Binomialkoeffizienten (Ziehen ohne Zurücklegen) berechnet werden. Ziehen von $k = 2$ Knoten aus n Knoten.

$$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}$$

Dichte eines ungerichteten Graphen

Die Dichte ρ eines ungerichteten Graphen $G = (V, E)$ mit n Knoten und m Kanten ist definiert als $\rho = \frac{2m}{n(n-1)}$.

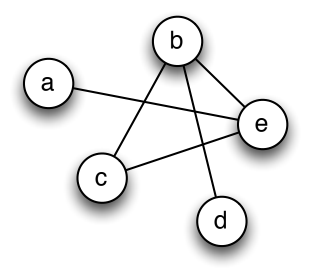
Dichte von Graphen

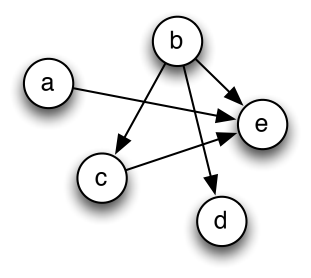
- Anzahl möglicher Kanten in einem gerichteten Graphen mit n Knoten kann ergibt sich als $n(n-1)$, wenn ein Knoten keine Kante zu sich selbst haben darf.
- Idee: Alle Knoten befinden sich in einer Urne. Zufälliges Ziehen von 2 Knoten. Beim Ziehen des ersten Knotens gibt es n Möglichkeiten, beim Ziehen des zweiten nur noch $n-1$.

Dichte eines gerichteten Graphen

Die Dichte ρ eines gerichteten Graphen $G=(V,E)$ mit n Knoten und m Kanten ist definiert als $\rho = \frac{m}{n(n-1)}$.

AUFGABE





- Bestimmen Sie die Dichte der beiden Graphen.

Dichte von Graphen

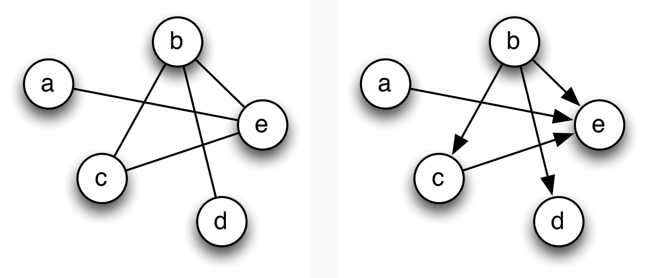


Abbildung 8: Ungerichteter und gerichteter Graphen mit jeweils 5 Knoten

- Dichte des ungerichteten Graphen: $\frac{2 \cdot 5}{5 \cdot 4} = 0.5$. D.h., der Graph hat 50% der Kanten des vollständigen ungerichteten Graphen mit der gleichen Knotenmenge.
- Dichte des gerichteten Graphen: $\frac{5}{5 \cdot 4} = 0.25$. D.h., der Graph hat 25% der Kanten des vollständigen gerichteten Graphen mit der gleichen Knotenmenge.

Pfad[1]

Sei $G = (V, E)$ ein Graph. Ein Pfad $P_{u,v}$ von Knoten u zu Knoten v ist eine (geordnete) Liste von Kanten, sodass jeweils der Endknoten der vorherigen Kante der Anfangsknoten der nachfolgenden Kante ist. Außerdem gilt, dass der Anfangsknoten der ersten Kante u und der Endknoten der letzten Kante v ist.

Bemerkungen:

- Die Definition ist ähnlich für gerichtete und ungerichtete Graphen. Der Unterschied besteht darin, dass die Kanten einmal geordnete Paare sind (und es Anfangs-und Endknoten gibt) und einmal Mengen (mit 2 Knoten).
- In dieser Definition sind Zyklen erlaubt und es ist auch nicht gefordert, dass $u \neq v$.

[1] Alternativ kann ein Pfad auch als Liste von Knoten definiert werden.

Zyklus

Ein Zyklus ist ein Pfad bei dem Anfangs- und Endknoten gleich sind.

Länge eines Pfades

Die Länge eines Pfades ist die Anzahl der Kanten des Pfades.

Distanz zwischen Knoten

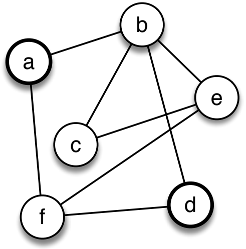
Die Distanz $dist(u, v)$ zwischen zwei Knoten u, v ist die Länge des kürzesten Pfades zwischen ihnen.

Bemerkungen:

- Ein Pfad kann Zyklen enthalten, kürzeste Pfade sind azyklisch.
- Es muss kein kürzester Pfad zwischen 2 Knoten existieren. In dem Fall ist die Distanz zwischen ihnen ∞.
- Es kann mehrere kürzeste Pfade zwischen 2 Knoten geben.
- $dist(u, u) = 0$[2]

[2] Es sind keine Kanten von einem Knoten zu sich selbst erlaubt.

Aufgabe



- Wieviele Pfade gibt es zwischen Knoten a und d?
- Bestimmen Sie den kürzesten Pfades zwischen Knoten a und d.
- Nennen Sie einen Zyklus (wenn vorhanden).

Pfade, Zyklen, Distanz

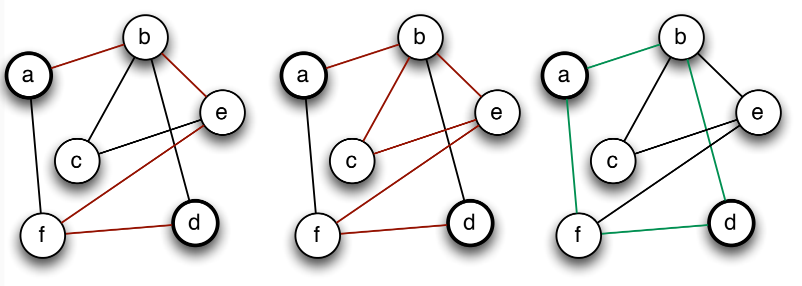
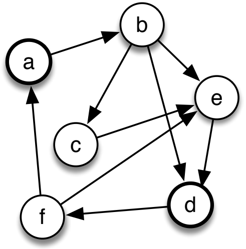


Abbildung 9: Ungerichteter Graph. Rot - Pfade zwischen Knoten a und d. Grün: kürzeste Pfade zwischen Knoten a und d.

- Länge des kürzesten Pfades zwischen Knoten a und d: 2
- Distanz zwischen Knoten a und d: 2
- Graph enthält Zyklen: z.B. zwischen Knoten b,c,e



- Wieviele Pfade gibt es zwischen Knoten a und d?
- Bestimmen Sie den kürzesten Pfades zwischen Knoten a und d.
- Nennen Sie einen Zyklus (wenn vorhanden).

Pfade, Zyklen, Distanz

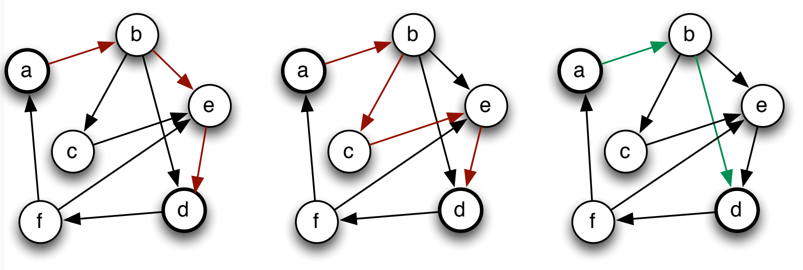


Abbildung 10: Ungerichteter Graph. Rot - Pfade zwischen Knoten a und d. Grün: kürzester Pfad zwischen Knoten a und d.

- Länge des kürzesten Pfades zwischen Knoten a und d: 2
- Distanz zwischen Knoten a und d: 2
- Graph enthält Zyklen: z.B. zwischen Knoten auf dem Pfad a,b,d,f,a

Durchmesser und Zusammenhang

Zusammenhang eines ungerichteten Graphen

Ein ungerichteter Graph $G = (V, E)$ heißt zusammenhängend, wenn es für alle $\{u, v\} \in V$ ein Pfad existiert.

Zusammenhang eines gerichteter Graphen

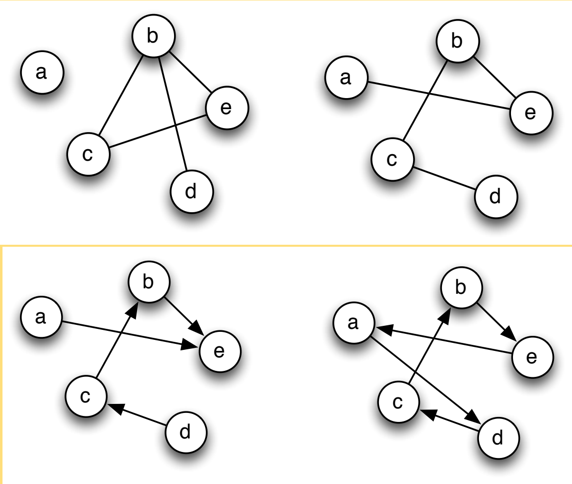
Ein gerichteter Graph $G = (V, E)$ heißt stark zusammenhängend, wenn es für alle $\{u, v\} \in V$ ein Pfad existiert. Ein gerichteter Graph $G = (V, E)$ heißt schwach zusammenhängend, wenn der zugehörige ungerichtete Graph zusammenhängend ist.

Durchmesser eines Graphen

Der Durchmesser D eines Graphen G ist definiert als die Länge des längsten kürzesten Pfades, d.h. entspricht der maximalen Distanz von Knoten im Graphen.

$$D(G) = \max_{u,v \in V} \big(dist(u, v) \big)$$

Aufgabe



- Sind die dargestellten Graphen zusammenhängend? Begründen Sie.

DURCHMESSER UND ZUSAMMENHANG

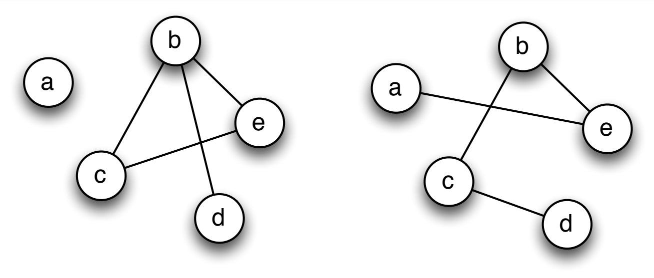


Abbildung 11: Zusammenhang ungerichteter Graphen

- Der linke Graph ist nicht zusammenhängend (es gibt keinen Pfad von Knoten a zu irgendeinem anderen Knoten).
- Der rechte Graph ist zusammenhängend (jeder Knoten ist von jedem anderen aus erreichbar).
- Die Anzahl Kanten in beiden Graphen ist gleich.

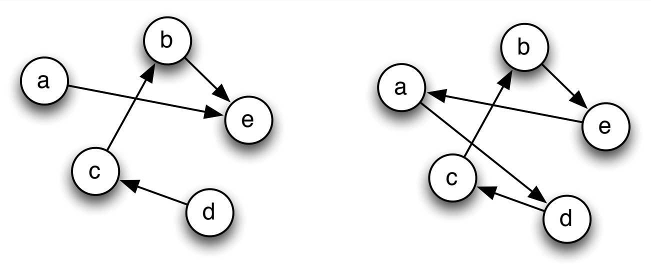
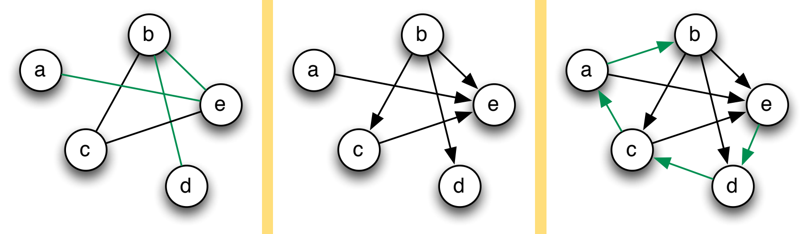


Abbildung 12: Zusammenhang gerichteter Graphen

- Der linke Graph ist schwach zusammenhängend (es gibt z.B keinen Pfad von Knoten e zu Knoten a), im ungerichteten Graphen jedoch schon (siehe Folie 23).
- Der rechte Graph ist stark zusammenhängend (jeder Knoten ist von jedem anderen aus erreichbar).

AUFGABE



- Bestimmen Sie die Durchmesser der abgebildeten Graphen.

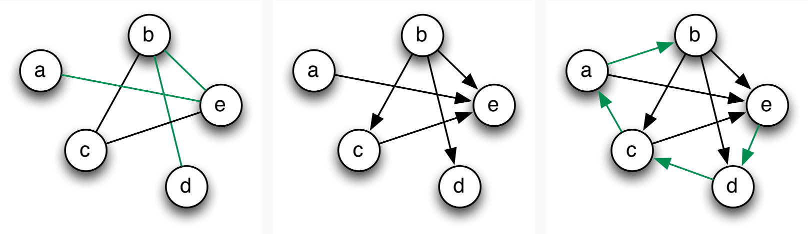


Abbildung 13: Längster kürzester Pfad im gerichteten und ungerichteten Graphen

- Der Durchmesser des ungerichteten Graphen ist 3 (es gibt mehrere Knotenpaare mit Distanz 3)
- Der Durchmesser des mittleren gerichteten Graphen ist ∞ (es gibt z.B. keinen Pfad zwischen Knoten e und d)
- Der Durchmesser des rechten gerichteten Graphen ist 4 (der längste kürzeste Pfad durchläuft alle Knoten)

- Der Clusterkoeffizient[3] beschreibt die Dichte der Verbindung der Nachbarn eines Knotens v.
- Dazu wird das Verhältnis der tatsächlichen Verbindungen zwischen den Nachbarn (ohne v) ins Verhältnis zur Anzahl der theoretisch möglichen Verbindungen zwischen den Nachbarn gesetzt.

Lokaler Clusterkoeffizient

Der lokale Clusterkoeffizient eines Knotens v ist definiert als

$$C_v = \frac{2m}{k_v \cdot (k_v - 1)}$$

wobei m die Anzahl der Kanten zwischen den Nachbarn ist und k_v die Anzahl der Nachbarn von v: $k_v = \Big|\big\{u \in V | \{u, v\} \in E, u \neq v\big\}\Big|$.

[3] Der Einfachheit halber werden nur ungerichtete Graphen betrachtet.

Der globale Clusterkoeffizient eines Graphen[4] ist der mittlere Clusterkoeffizient seiner Knoten.

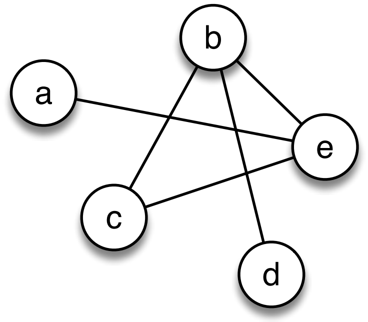
Globaler Clusterkoeffizient

Der globale Clusterkoeffizient eines Graphen $G = (V, E)$ ist definiert als

$$C(G) = \frac{1}{|V|} \sum_{v \in V} C_v$$

[4] Nach Watts & Strogatz [3]

Aufgabe



- Bestimmen Sie die Clusterkoeffizienten aller Knoten, sowie den Clusterkoeffizienten des Graphen.

Clusterkoeffizient

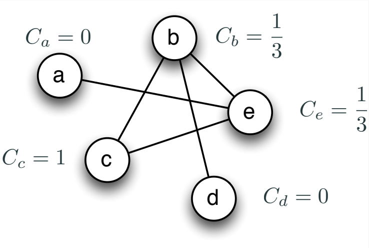


Abbildung 14: Clusterkoeffizienten

- Randfall: Knoten a und d haben jeweils nur einen Nachbarn, ihr Clusterkoeffizient ist definiert als 0
- Knoten c hat Nachbarn b und e, die miteinander verbunden sind
- Knoten b hat 3 Nachbarn c, d, e; von den möglichen drei Verbindungen zwischen ihnen ist eine vorhanden
- $C(G) = \frac{1}{5}(0 + 0 + 1 + \frac{1}{3} + \frac{1}{3}) = \frac{1}{3}$

Kleine-Welt Netzwerke

Small World

Die Welt ist ein Dorf. (dt. Redewendung)

- Hypothese: Zwei zufällige Menschen sind über Freund*innen und Freund*innen von Freund*innen usw. auf relativ kurzem Wege verbunden
- Die kurzen Pfade zwischen zwei beliebigen Akteuren in Netzwerken nennt man Kleine-Welt-Phänomen (Small-World Phenomenon)
- 1967 von Stanley Milgram, mehrere Experimente zur Untersuchung des Kleine-Welt-Phänomens

Milgram Experiment[5]

Milgram Experiment Beschreibung

- An 296 zufällig ausgewählte Startpersonen in den USA (aus Omaha, Nebraska, and Wichita, Kansas) wurden Briefe verschickt
- Die Briefe enthielten Anweisungen, sie an eine Zielperson zu schicken, die in Boston lebt (Aktienhändler in Boston)
- Briefe enthielten die Adresse der Zielperson und eine Beschreibung
- Personen sollten Briefe an Personen weiterschicken, mit denen sie per Du sind und von denen sie glauben, dass über sie der Brief das Ziel so schnell wie möglich erreicht
- In den Briefen wurde jeweils mitgeschrieben, wer sie wann an wen weitergeschickt hat

[5] Nicht zu verwechseln mit dem berühmten Experiment zu Autoritätsgehorsam

Milgram Experiment

- 64 der 296 (22%) Briefe erreichten ihr Ziel, von dem Rest ist unbekannt, wo sie verblieben sind
- Von den 64 Briefen war die Länge des Reisepfades bekannt (jedem Brief lag eine Liste des Reiseverlaufes bei)
- Der Mittelwert der Reisepfadlängen betrug 5,5, der Median 6
- Maximale Pfadlänge 11, minimale 1
- Schlussfolgerung: Alle Menschen in den USA sind über eine Distanz von 6 miteinander verbunden
- Kritik an den Experimenten
  - Niedrige Response-Rate
  - Daraus, dass eine Person einen Brief einen anderen weiter leitet kann man nicht notwendigerweise eine Schlussfolgerung über die Enge der Beziehung ziehen (würde diese Person auch bei der Jobsuche helfen?)

Six Degrees of Separation

Beschreibt die Annahme, dass alle Menschen in der Welt maximal 6 Schritte in Beziehungsnetzwerken voneinander entfernt ist.

- Dies ist in späteren Experimenten in verschiedenen Online- und Offline Netzwerken empirisch belegt worden.
- Teilweise wurden auch höhere mittlere Pfadlängen beobachtet (z.B. 8,11), aber insgesamt waren die Pfadlängen überraschend kurz.

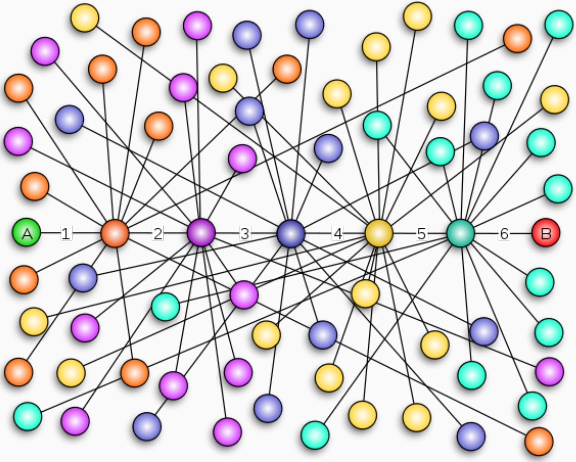


Abbildung 15: Beispielnetzwerk zu Six Degrees of Separation (CC-SA-30, Dannie-walker via Wikimedia Commons)

Eigenschaften von Kleine-Welt-Netzwerken

- Transitivität
    - Zwei Knoten, die über einen dritten Knoten verbunden sind, sind auch mit hoher Wahrscheinlichkeit untereinander verbunden (d.h. der Graph besteht aus vielen Dreiecken)
    - Grad der Transitivität über Clusterkoeffizienten messbar
- Geringer Durchmesser
    - Der längste kürzeste Pfad ist tendenziell sehr kurz (vlg. six degrees of separation auf Folie 34).
    - In Informationsnetzwerken heißt dies, dass die Informationen sehr schnell von einem Knoten zu einem beliebigen anderen übertragen werden können.

Small World Networks

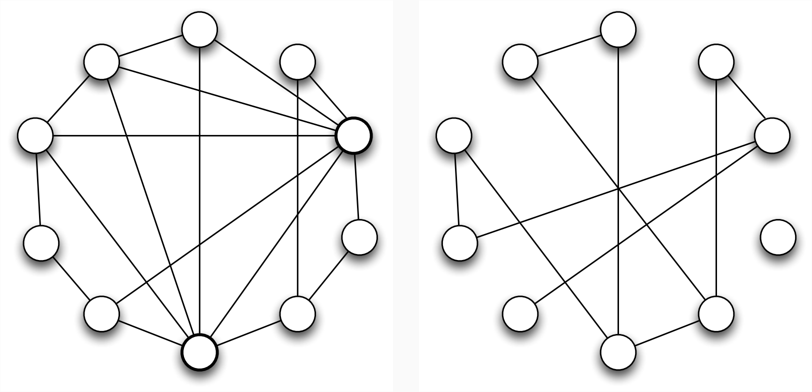


Abbildung 16: Einfaches Small World Network mit zwei stark verbundenen Knoten (sog. Hubs, dicker Umriss) und zufälliger Graph (rechts)

Beispiel Erdős-Zahl

- Beobachtung des Kleine-Welt-Phänomens ins professionellen Netzwerken
- Der Mathematiker Paul Erdős hat in seiner Karriere 1500 Artikel veröffentlicht
- Die Erdős-Zahl einer Person gibt die Distanz dieser im Co-Autor*innen-Netzwerk an

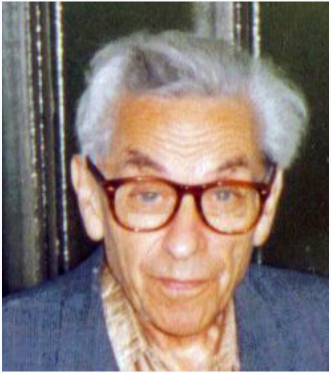


Abbildung 17: Paul Erdős

Beispiel Erdős-Zahl

- Co-Autor*innen-Netzwerke sind ungerichtete Graphen
- Knoten repräsentieren Personen
- Semantik von Kanten: „hat publiziert mit“
- Die Erdős-Zahl einer Person ist die Distanz dieser Person zu Paul Erdős

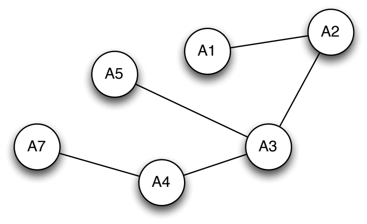


Abbildung 18: Einfaches Co-Autor*innen-Netzwerk

Abbildung 19: Illustrations zur Erdős-Zahl (CC-SA-30, H2g2bob via Wikimedia Commons)

Beispiel Erdős-Zahl

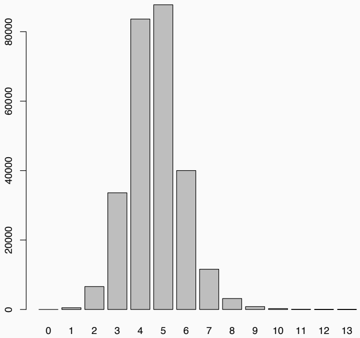


Abbildung 20: Balkendiagramm zur Erdős-Zahl (x-Achse) und ihrer Häufigkeit (y-Achse)

- Es gibt 268015 Personen, die eine endliche Erdős-Zahl haben
- Durchschnittliche Erdős-Zahl 4,65
- Beispiel einer Kleinen Welt
- In der Physik gibt es eine ähnliche Zahl, die Pauli-Zahl (nach dem Physiker Wolfgang Pauli)

Bestimmung der Erdös Number:

`https://mathscinet.ams.org/mathscinet/searchauthors.html`, dann Collaborator Distance auf dem Profil

Beispiel Kevin-Bacon-Zahl

- Wie sind Schauspieler*innen über Filme miteinander verbunden?
- Netzwerk mit 2 verschiedenen Knotenmengen
- Semantik der Kanten „hat gespielt in“
- Verbindung zwischen Person A2 und A4 über Filme M1 und M2, Distanz ist 2.

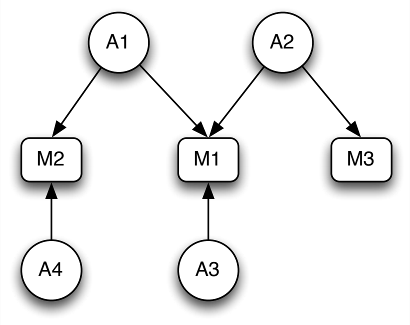


Abbildung 21: Einfaches Schauspieler*innen-Filme Netzwerk

Beispiel Kevin-Bacon-Zahl

- Die Kevin-Bacon-Zahl einer Person gibt an, über wie viele Filme die Person mit Kevin-Bacon verbunden ist
- Netzwerk mit 2.301.643 Schauspieler*innen, durchschnittliche Kevin-Bacon-Zahl $\approx$ 3.026 (August 2017)

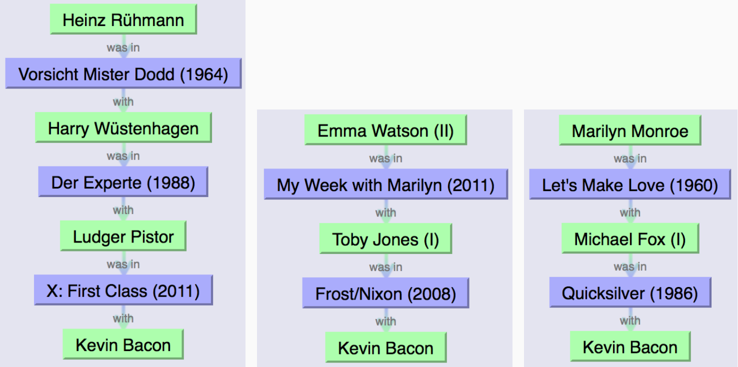


Abbildung 22: Kevin-Bacon-Zahl (`http://www.oracleofbacon.org`)

Beispiel Facebook Gruppen

Untersuchung von Communities im Facebook Netzwerk [4].

- Konstruktion von mehreren Graphen, die jeweils Gruppen zu gleichen Themen enthalten
- Knoten sind (öffentliche) Gruppen
- Kanten werden zwischen Gruppen eingefügt, wenn sie mindestens ein Mitglied gemeinsam haben

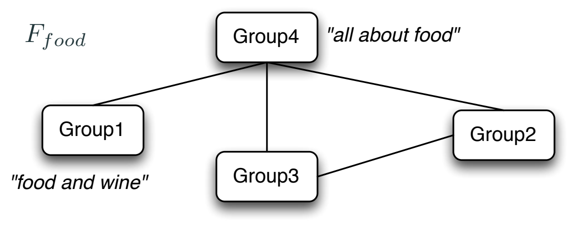


Abbildung 23: Schema des Graphen für das Wort *food*

Beispiel Facebook Gruppen

Beobachtungen:

- Hoher Clusterkoeffizient
- Niedrige mittlere Länge der kürzesten Pfade (≈ 2)
- Beispielbilder der Netzwerke in: Jason Wohlgemuth und Mihaela Teodora Matache. "Small-world properties of facebook group networks". In: *Complex Systems* 23.3 (2014), S. 197–225. ISSN: 0891-2513. URL: http://www.complex-systems.com/pdf/23-3-1.pdf

Gruppe	$\lvert V\rvert$	$\lvert E\rvert$	**C**	**l**
F_{anime}	276	9770	0,708	1,837
F_{bieber}	117	1714	0,646	1,903
F_{muslim}	233	4326	0,564	2,121

(C - Clusteringkoeffizient des Graphen, l mittlere Länge der kürzesten Pfade)

Beispiel Nervenzellen C. Elegans

Beispiel. Nervennetzwerk des Wurmes C.elegans [3].
$|V| = 306, |E| = 2345, C = 0.164$, Durchmesser 14, durschnittliche Pfadlänge 4.

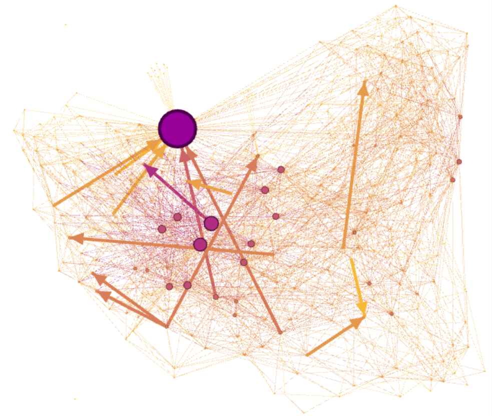


Abbildung 24: Visualisierung Nervennetzwerk. (Erstellt mit `https://gephi.org`)

Watts-Strogatz Modell für Kleine Welten

Wie kann man einen Graphen generieren, der sowohl Transitivität als auch geringen Durchmesser zeigt?

Idee Konstruktion nach Watts-Strogatz

- Starte mit 2-D Grid-Layout ohne Verbindungen (in jeder Grid-Zelle befindet sich ein Knoten).
- Verbinde alle Knoten, die sich auf dem Grid in Entfernung r zum Ursprungskonten befinden, mit Kanten. Dies simuliert die engmaschigen, lokalen Verbindungen in einem Netzwerk.[6]
- Mit Wahrscheinlichkeit p füge zufällig Verbindungen zwischen zwei beliebigen Knoten hinzu. Dies simuliert die Zufallsbekanntschaften über lokale Nachbarschaften hinweg.[7]

[6] Der Fachbegriff dafür ist Homophilie-Verbindungen.
[7] Der Fachbegriff dafür ist weak ties.

Watts-Strogatz Modell für Kleine Welten

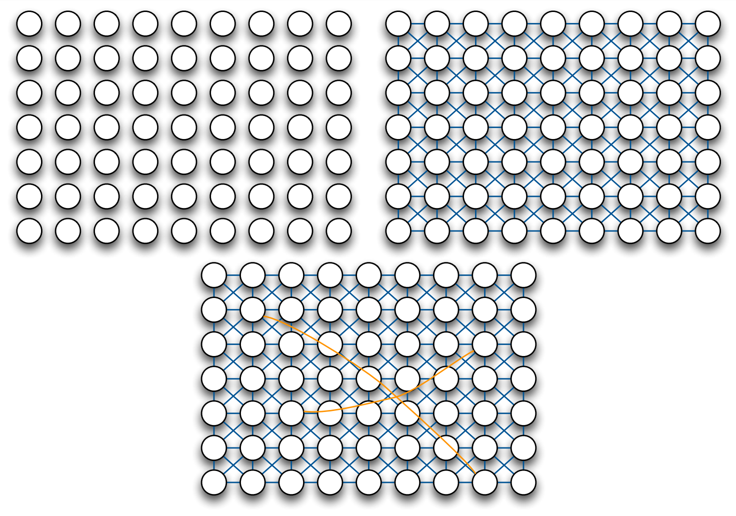


Abbildung 25: Konstruktion des Watts-Strogatz-Modells in 3 Schritten. Lokale, dichte Verbindungen (blau) und zufällige Verbindungen (orange).

Watts-Strogatz Modell für Kleine Welten

- Netzwerk hat hohen Clusterkoeffizienten (hohe Transitivität durch das lokale Grid).
- Zwei Konten *u* und *v*, die im Grid weit entfernt voneinander sind, können über die zufälligen Kanten, über wenige Schritte erreicht werden (kleiner Durchmesser).
- Beobachtung: Es reicht ein relativ kleiner Anteil zufälliger Kanten aus, um eine Kleine Welt zu schaffen.

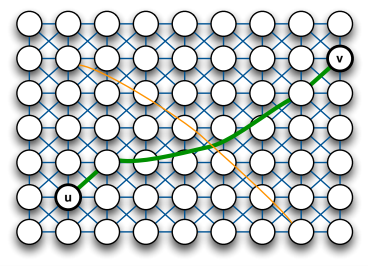


Abbildung 26: Distanz im Watts-Strogatz-Modell

Zusammenfassung Small-World-Netzwerke

- Small-World-Netzwerke haben sowohl Charakteristika chaotischer Systeme (zufällige Netzwerke) als auch Eigenschaften hochstrukturierter System (vlg. Watts-Strogatz-Modell auf Folie 47)
- Anwendung in der Simulation der Übertragung von Infektionskrankheiten (Desease-Contact-Networks)
- Anwendung in der Simulation der Übertragung von Computerviren
- Nutzung der Eigenschaften für Protokolle in peer-to-peer Netzwerken
- Auftreten in biologischen (z.B. das neuronale Netz des Fadenwurms Caenorhabditis elegans) und sozialen Systemen (z.B. soziale Netzwerke)

Skalenfreie Netzwerke

Skalenfreies Netzwerk

Ein skalenfreies Netzwerk ist ein Graph, dessen Verteilung der Knotengrade k einem Potenzgesetz folgt, d.h.

$$P(k) \propto k^{-\gamma}$$

- Empirisch wurde für viele Netzwerke $\gamma \in \{2, 3\}$ gefunden
- Skalenfreie Netzwerke haben sehr viele Knoten mit kleinen Graden (wenigen Verbindungen) und einige wenige, die viele Verbindungen haben
- Knoten mit vielen Verbindungen heißen Hubs
- Viele Knoten mit wenigen Verbindungen werden durch einige Hubs zusammengehalten
- Skalenfreie Netzwerke sind robust gegen zufälligen Ausfall von Knoten und sensitiv gegen Angriffe auf Hubs

Skalenfreie Netzwerke

Potenzverteilung für verschiedene $\gamma \in \{2, 3\}$. Im Log-Log-Plot ergeben sich Geraden mit Anstieg γ. x-Achse: Grad der Knoten aufsteigend sortiert, y-Achse: Anzahl der Knoten mit dem jeweiligen Grad.

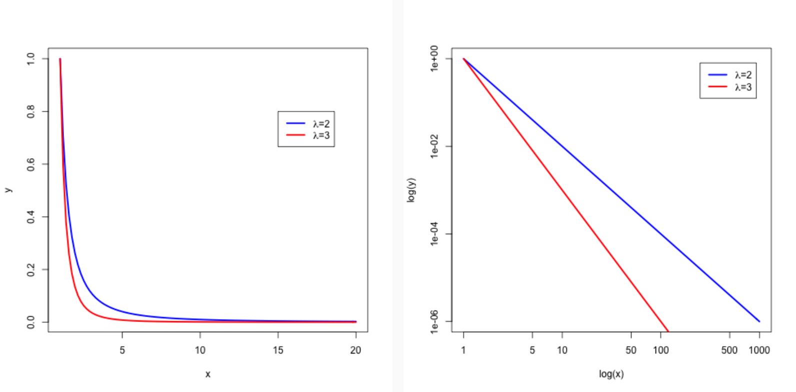


Abbildung 27: Funktion $y = k^{-\gamma}$. Rechts: log-log skalierte Achsen

Beispiel Jazz-Netzwerk [2]. Knoten repräsentieren Musiker*innen, Kanten repräsentieren Kollaborationen.

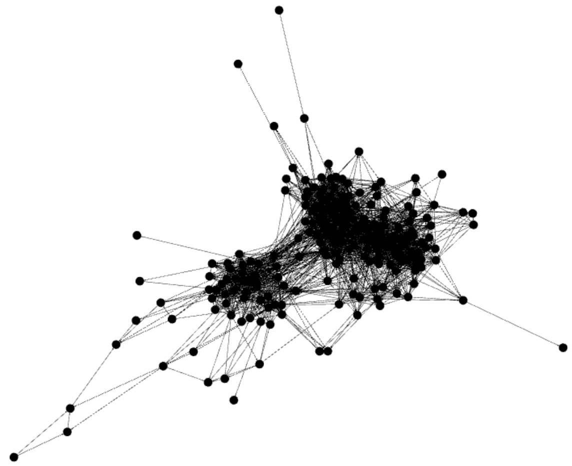


Abbildung 28: Visualisierung Jazz-Netzwerk. (Erstellt mit `https://gephi.org`)

Beispiel Jazz-Netzwerk

Anzahl Knoten	198
Anzahl Kanten	2742
Durchmesser	9
⌀ Pfadlänge	2.23
Clusterkoeff.	0.309

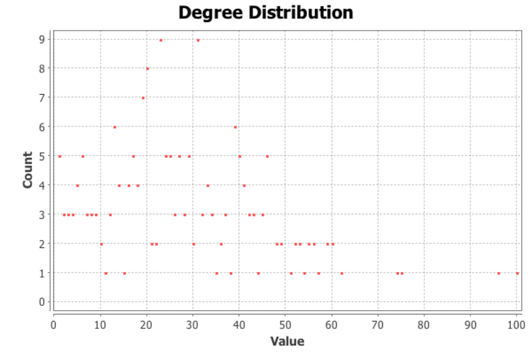


Abbildung 29: Verteilung der Knotengrade

- Verteilung der Knotengrade entspricht keiner Potenzverteilung.
- Kein skalenfreies Netz!

Beispiel Energie-Netzwerk

Beispiel. Energienetzwerk im Westen der USA [3]. Kanten sind Energieversorgungskabel. Knoten repräsentieren Generatoren, Transformatoren oder Speicherstationen.

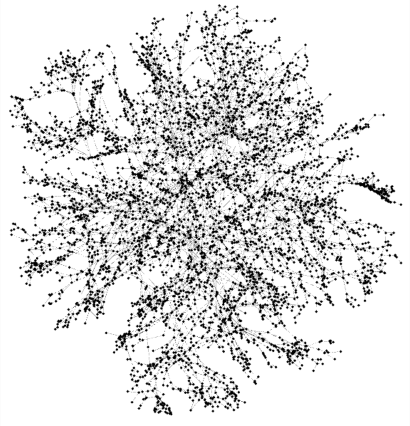


Abbildung 30: Visualisierung Energie-Netzwerk. (Erstellt mit `https://gephi.org`)

Beispiel Energie-Netzwerk

Anzahl Knoten	4941
Anzahl Kanten	6594
Durchmesser	46
⌀ Pfadlänge	18
Clusterkoeff.	

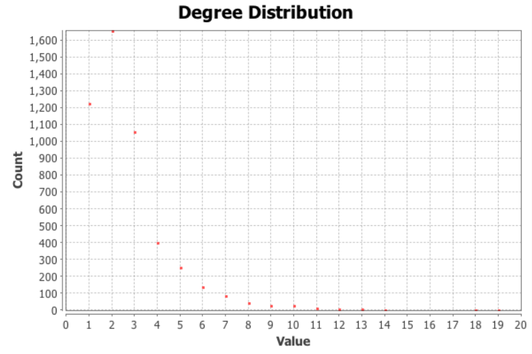


Abbildung 31: Verteilung der Knotengrade

- Verteilung ähnlich einer Potenzverteilung
- Netzwerk ist ein skalenfreies Netzwerk

Barabási-Albert Modell

- Modell zur Konstruktion von skalenfreien Netzen, um die Existenz skalenfreier Netzwerke (z.B. Internet und soziale Netzwerke) zu erklären
- Kernkonzepte der Konstruktion sind Wachstum und Preferential Attachment
- Wachstum: beginnend mit einigen Knoten wird das Netz erweitert und neue Knoten und Kanten hinzugefügt
- Preferential Attachment[8]: Knoten, die bereits viele Kanten haben, werden bei der Auswahl für neue Verbindungen bevorzugt („Reiche-werden-reicher-Phänomen“)

[8] Ungefähr übersetzbar mit „Bevorzugung der Bevorzugten “

Konstruktion nach Barabási-Albert

1. Starte mit einer kleinen Anzahl m_0 von Knoten, die vollständig miteinander verbunden sind
2. Füge einen neuen Knoten v ein
3. Wähle $m \leq m_0$ weitere Knoten u_i aus den existierenden aus, wobei die Wahrscheinlichkeit p_j, dass Knoten u_j gewählt wird, proportional zum Grad des Knotens $deg(u_j)$ ist

$$p_j = \frac{deg(u_j)}{\sum_{u_k \in V} deg(u_k)}$$

4. Füge m Kanten ein, jeweils zwischen Knoten v und u_i.

Wiederhole Schritte 2 bis 4

- Erstellen Sie einen Graphen nach dem Barabási-Albert-Modell. Starten Sie mit $m_0 = 3$ Knoten.

Barabási-Albert Modell

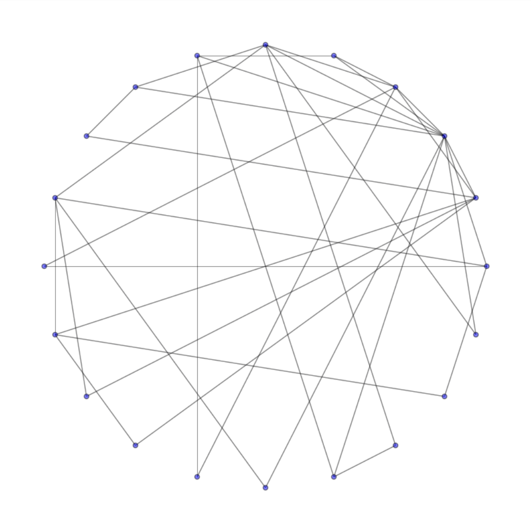


Abbildung 32: Barabási-Albert Modell mit $m = m_0 = 2$ (CC-SA 3.0, Horváth Árpád, via Wikimedia Commons)

Zusammenfassung

Zusammenfassung

- Graphen können Beziehungen zwischen Objekten modellieren
- Phänomene der Realwelt können über Charakteristiken von Graphen erklärt werden
  - Six degrees of separation (Durchmesser und Clusterkoeffizient)
    $\rightarrow$ Kleine-Welt Netzwerke
  - Ausfallsicherheit zufälliger Knoten (Potenzverteilung der Knotengrade)
    $\rightarrow$ Skalenfreie Netzwerke
- Kleine-Welt Netzwerke sind nicht notwendigerweise skalenfrei und umgekehrt
- Theoretische Modelle erlauben die Konstruktion von Netzwerken mit bestimmten Eigenschaften und damit Erklärung bestimmter Phänomene
  $\rightarrow$ Watts-Strogatz-Modell und Barabási-Albert Modell

Zusammenfassung

Wichtige Konzepte

- Graphen (Pfade, Zyklen, Distanz, Knotengrad, Dichte, Durchmesser, Zusammenhang)
- Clusterkoeffizient
- Kleine-Welt Netzwerke
- Skalenfreie Netzwerke
- Six degrees of separation
- Bacon-Zahl, Erdős-Zahl
- Watts-Strogatz-Modell für Kleine-Welt Netzwerke
- Barabási-Albert Modell für skalenfreie Netzwerke

Acknowledgement

- Teile der Folien sind angelehnt an Inhalte aus der englischen Veranstaltung Soziale Netzwerkanalyse von Michael Granitzer, Universität Passau
- Datensätze von Gephi `http://gephi.org` und aus der Koblenz Network Collection `http://konect.uni-koblenz.de`.

Weiterführende Literatur

- Networks, Crowds, and Markets: Reasoning About a Highly Connected World (PDF Version online verfügbar)
  David Easley und Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. New York, NY, USA: Cambridge University Press, 2010. ISBN: 0521195330, 9780521195331. URL: https://www.cs.cornell.edu/home/kleinber/networks-book/networks-book.pdf

Literatur

[1] David Easley und Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. New York, NY, USA: Cambridge University Press, 2010. ISBN: 0521195330, 9780521195331. URL: https://www.cs.cornell.edu/home/kleinber/networks-book/networks-book.pdf.

[2] Pablo M. Gleiser und Leon Danon. "Community Structure in Jazz". In: *Advances in Complex Systems* 6.4 (2003), S. 565–573.

[3] Duncan J. Watts und Steven H. Strogatz. "Collective dynamics of small-world networks". In: *Nature* 393.6684 (Juni 1998), S. 440–442. ISSN: 00280836. URL: http://worrydream.com/refs/Watts-CollectiveDynamicsOfSmallWorldNetworks.pdf.

[4] Jason Wohlgemuth und Mihaela Teodora Matache. "Small-world properties of facebook group networks". In: *Complex Systems* 23.3 (2014), S. 197–225. ISSN: 0891-2513. URL: http://www.complex-systems.com/pdf/23-3-1.pdf.