

Physik der sozio-ökonomischen Systeme *mit dem Computer*

*JOHANN WOLFGANG GOETHE UNIVERSITÄT
11.12.2020*

MATTHIAS HANAUSKE

*FRANKFURT INSTITUTE FOR ADVANCED STUDIES
JOHANN WOLFGANG GOETHE UNIVERSITÄT
INSTITUT FÜR THEORETISCHE PHYSIK
ARBEITSGRUPPE RELATIVISTISCHE ASTROPHYSIK
D-60438 FRANKFURT AM MAIN
GERMANY*

6. Vorlesung

Aufgrund der Corona Krise findet die Vorlesung und die freiwilligen Übungstermine in diesem Semester nur Online statt.

Plan für die heutige Vorlesung

- Einführung in die Theorie der komplexen Netzwerke
 - Warum evolutionäre Spieltheorie auf komplexen Netzwerken?
 - Mathematische Beschreibung von komplexen Netzwerken
 - Definition unterschiedlicher Netzwerk-Typen
 - Komplexe Netzwerke in der Realität
 - Definition wichtiger Netzwerk-charakterisierender Größen
 - Definition unterschiedlicher Netzwerk-Klassen
 - Die Klasse der zufälligen Netzwerke (random networks)
 - Die Klasse der „Kleine Welt“ Netzwerke (small world networks)
 - Komplexe Netzwerke analysieren mit Python (NetworkX)

Einführung in die Theorie der komplexen Netzwerke

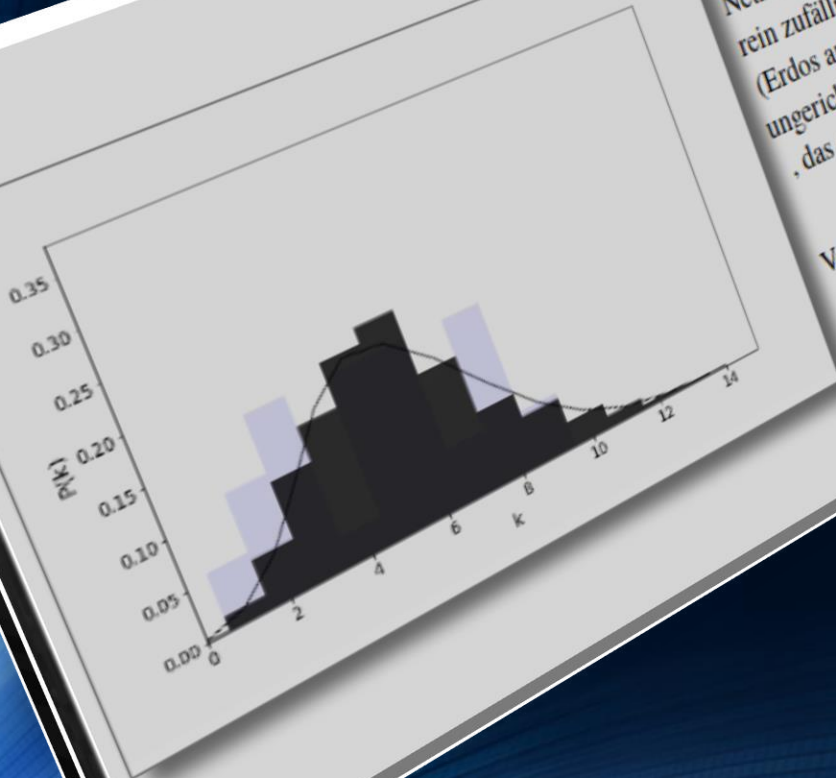
Internetseite der Vorlesung

Zufällige Netzwerke (random networks)

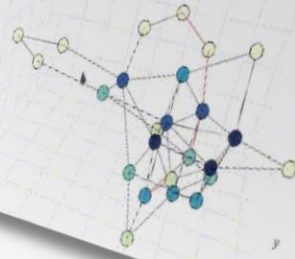
Im Folgenden betrachten wir die Zufälligen Netzwerke. Die einzelnen Knoten werden von den Kanten rein zufällig Muster aus ungerichteten Kanten, das ein Knoten

$p =$
Verteilung

$$P(k) = N^{-1} \sum_{i=1}^N k_i$$



Einführung in die Theorie der komplexen Netzwerke



Ein Netzwerk wird formal mittels graphentheoretischer Größen beschrieben. Wir nehmen im Folgenden zunächst ein ungerichtetes, ungewichtetes Netzwerk bestehend aus einer Knotenart an, wobei N die Anzahl der Knoten und L die gesamte Anzahl der Links (Kanten) des Netzwerks ist. Jeder einzelne Knoten i ($i \in \{1, 2, 3, \dots, N\}$) besitzt eine gewisse Anzahl von Verbindungskanten zu anderen Knoten und die Anzahl dieser Kanten bezeichnet man als den Knotengrad k_i des Knotens. Der durchschnittliche Knotengrad $\langle k \rangle$ eines ungerichteten Netzwerks ist $\langle k \rangle = \frac{2L}{N}$ und die maximale Anzahl möglicher Kanten (L_{max} (ein sogenannter vollständiger Graph) ist $L_{max} = \frac{N(N-1)}{2}$). Eine der wichtigsten Netzwerk-charakterisierenden Größen ist die Verteilungsfunktion der Knotengrade $P(k) = N(k)/N$, wobei $N(k)$ die Anzahl der Knoten mit Knotengrad k beschreibt. Weitere wichtige graphentheoretische Größen werden in dem Jupyter Notebook: Einführung in die Theorie der komplexen Netzwerke (siehe auch Einführung in die Theorie der komplexen Netzwerke und Albert-Laszlo Barabasi, Network science, Chapter 2 Graph Theory).

Einführung Teil II

Physik der sozio-ökonomischer Systeme mit dem Computer von Dr.phil.nat.Dr.rer.pol. Matthias Hanauske

[Home](#) [Research](#) [Contact](#)

[Einführung](#)

[Teil I](#)

[Teil II](#)

[Teil III](#)

[E-Learning](#)

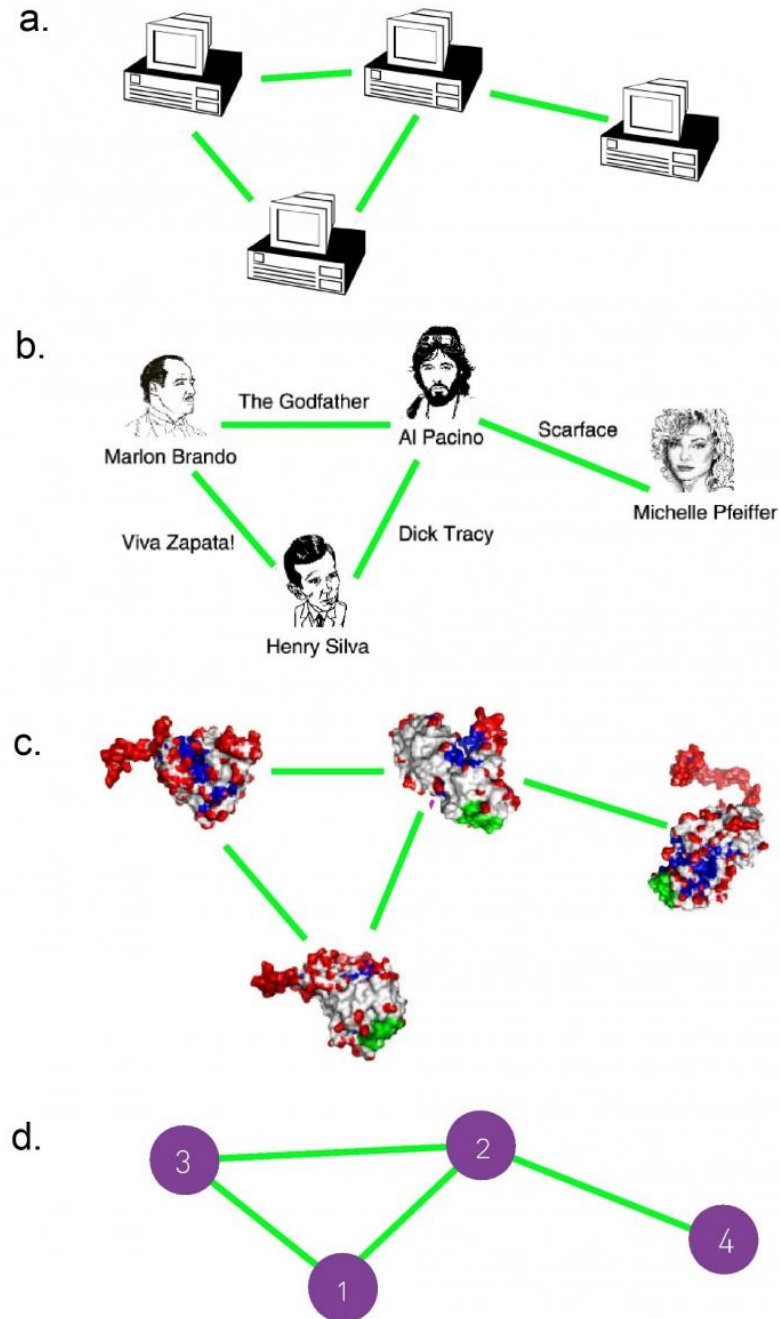
II.1 Theorie der sozio-ökonomischen Netzwerke

Eine bedeutende Einschränkung der deterministischen, evolutionären Spieltheorie ist deren zugrundeliegende Netzwerkstruktur (Netzwerktopologie). Die jeweiligen Spieler der betrachteten Population suchen in jeder Spielperiode einen neuen Spielpartner, wobei sie hierbei zufällig vorgehen (zufälliges Netzwerk) und vom Prinzip her mit jedem Spieler innerhalb der Population potentiell das zugrundeliegende Spiel spielen können (vollständig verbundenes Netzwerk). In Bimatrix Spielen suchen sich die Spieler der Teilpopulation A einen zufälligen Spielpartner aus Gruppe B (bzw. umgekehrt). Betrachtet man sich jedoch real existierende sozio-ökonomische Netzwerke, so zeigt sich, dass diese Annahme oft nicht erfüllt ist. Personen kennen oft nur eine Teilmenge von Spielern innerhalb der Population (kein vollständig verbundenes Netzwerk) und die Wahl der potentiellen Spielpartner erfolgt oft auch nicht nach zufälligen Mustern.

II.2 Einführung in die Theorie der komplexe Netzwerke

Die Theorie der komplexen Netzwerke bildet die Grundlage zur Beschreibung einer Vielzahl von unterschiedlichen biologischen und sozio-ökonomischen Systemen. Die mathematische Beschreibung komplexer Netzwerke ist in mehreren Übersichtsartikel (siehe z.B. <http://barabasi.com/f/103.pdf>, <https://arxiv.org/abs/cond-mat/0106144>, <http://epubs.siam.org/doi/pdf/10.1137/S003614450342480>) zusammengefasst (siehe auch [Claudius Gros, "Complex and Adaptive Dynamical Systems, a Primer"](#)), wobei im

Komplexe Netzwerke Knoten und Kanten



Different Networks, Same Graph

The figure shows a small subset of (a) the Internet, where routers (specialized computers) are connected to each other; (b) the Hollywood actor network, where two actors are connected if they played in the same movie; (c) a protein-protein interaction network, where two proteins are connected if there is experimental evidence that they can bind to each other in the cell. While the nature of the nodes and the links differs, these networks have the same graph representation, consisting of $N = 4$ nodes and $L = 4$ links, shown in (d).

Abbildung entnommen von:

Network Science by Albert-László Barabási

<http://networksciencebook.com/>

Viele der im Teil 2 behandelten Themen sind in diesem Buch ausführlich behandelt.

Netzwerke in der Realität

Netzwerke finden sich in den unterschiedlichsten sozialen, physikalischen und biologischen Systemen

- **Biologische Netzwerke**
 - Protein- und Gennetzwerke
- **Soziale Netzwerke**
 - Beziehungs- und Freundschaftsnetzwerke
 - Netzwerke von Geschäftsbeziehungen und Firmenbeteiligungen
 - Internetbasierte, soziale Web2.0 Netzwerke
- **Technologische Netzwerke**
 - Transportnetzwerke (Flug-, Zugrouten)
 - Internetverbindungen zwischen Computerservern
- **Informationsnetzwerke**
 - Wissensnetzwerke, Verlinkungen von Internetseiten
 - Zitationsnetzwerke von wissenschaftlichen Artikeln
 - Linguistische Netzwerke

Theorie der komplexen Netzwerke (II)

Komplexe Netzwerke lassen sich wie folgt untergliedern:

- Handelt es sich nur um eine Knotenart (Spielergruppe), oder besteht das Netzwerk aus mehreren Knotenarten (z.B. Bi-Matrix Spiele).
- Sind die Kanten (Verbindungslinien zwischen den Knoten) gerichtet oder ungerichtet.
- Besitzen die Kanten zahlenmäßige Gewichtungen oder geben sie einfach an ob ein Knoten mit einem anderen verbunden oder nicht verbunden ist.
- Gibt es zeitliche Veränderungen des Netzwerks; ist die Anzahl der Knoten konstant oder wächst bzw. schrumpft sie im Laufe der Zeit.

Theorie der komplexen Netzwerke (III)

(Beispiele unterschiedlicher komplexer Netzwerke)

- a) Nicht gerichtetes und ungewichtetes Netzwerk einer einzigen Knotenart.
- b) Nicht gerichtetes und ungewichtetes Netzwerk dreier verschiedener Knotenarten, wobei zusätzlich drei verschiedene Kantenarten existieren.
- c) Nicht gerichtetes aber gewichtetes Netzwerk. Sowohl die Knoten als auch die Kanten des Netzwerks besitzen zahlenmäßige Gewichtungen.
- d) Gerichtetes aber nicht gewichtetes Netzwerk. Es existiert nur eine Knoten- und gerichtete Kantenart.

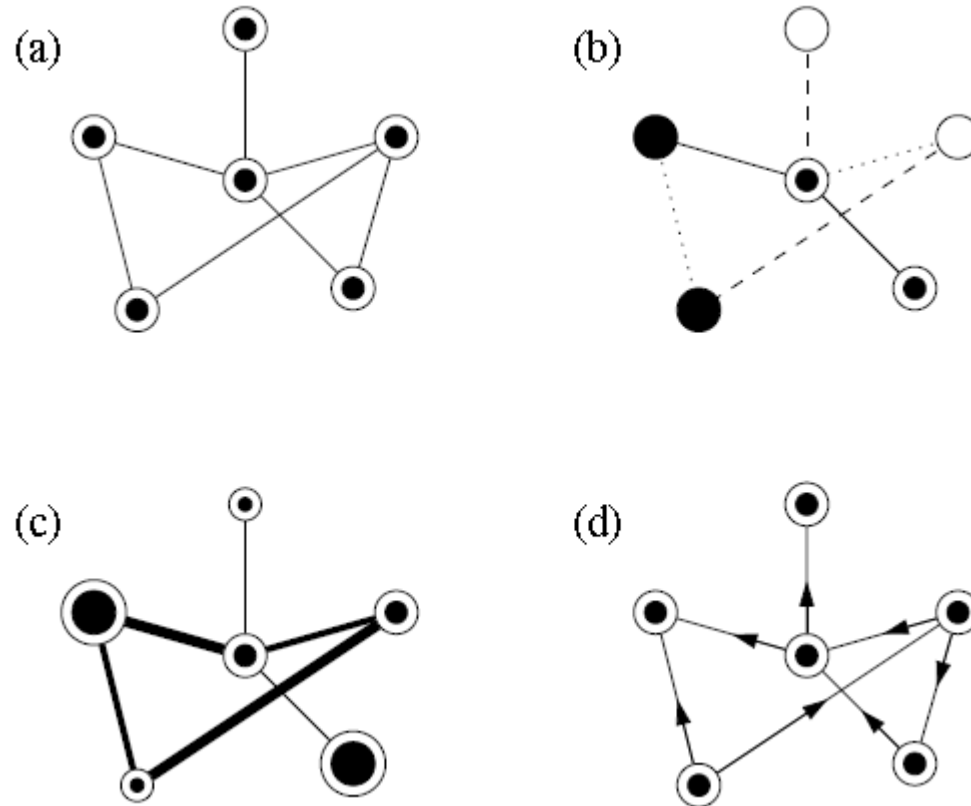


Abbildung: Unterschiedliche Netzwerktypen
Die Abbildung ist dem folgenden Artikel entnommen:
M. E. J. Newman,
„The structure and function of complex networks”

Theorie der komplexen Netzwerke (IV)

(Größen die ein Netzwerk charakterisieren)

- **Der Knotengrad k_i**

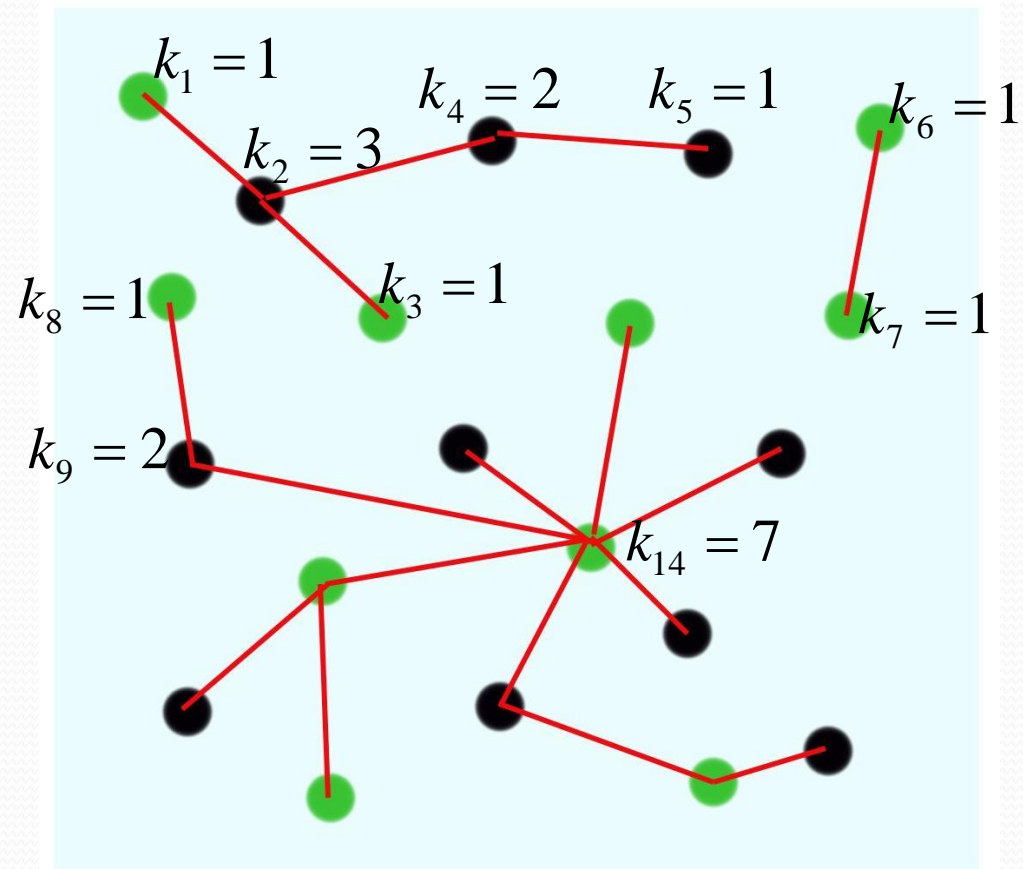
Der Knotengrad des Knotens i ist gleich der Anzahl der Kanten die der Knoten i besitzt. Bei gerichteten Netzwerken unterscheidet man zwischen dem eingehenden und ausgehenden Knotengrad. Bei gewichteten Netzwerken summiert man über die Zahlenfaktoren der gewichteten Kanten.

- **Der Clusterkoeffizient C_i**

Der Clusterkoeffizient gibt die Wahrscheinlichkeit an, dass zwei nächste Nachbarn eines Knotens ebenfalls nächste Nachbarn untereinander sind. Der globale Wert C des Netzwerks stellt demnach eine Art von „Cliques“-Nachbarschafts-Eigenschaft des Netzwerks dar

- **Der Durchmesser des Netzwerks**

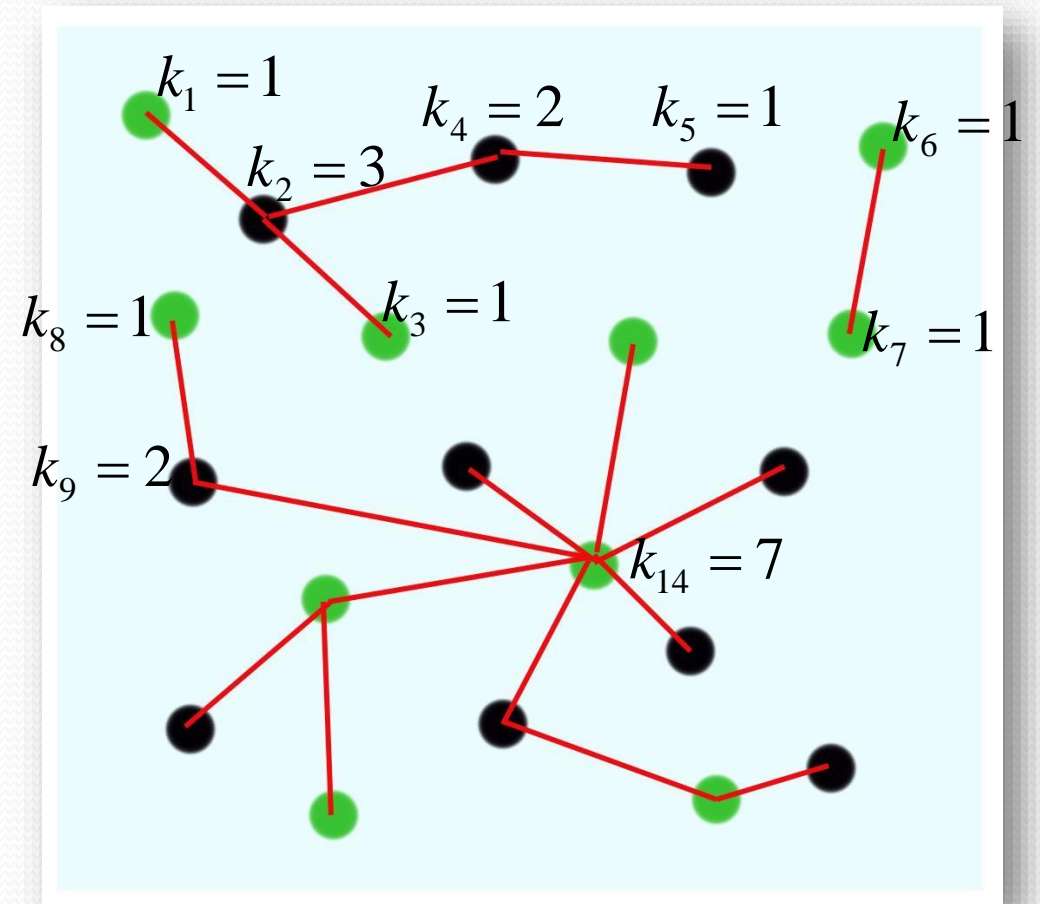
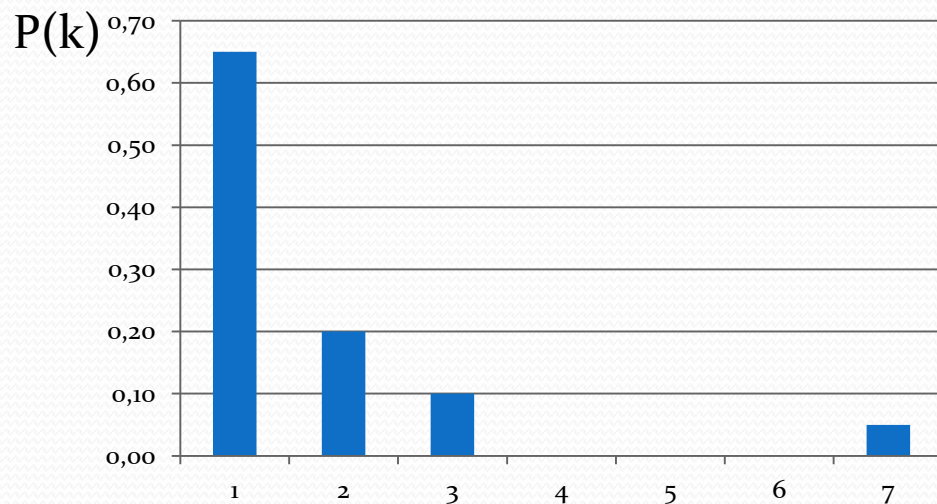
Der Durchmesser des Netzwerks gibt die maximale kürzeste Kantenlänge zwischen zwei beliebigen Knoten des Netzwerkes an.



Theorie der komplexen Netzwerke (V)

(Die Verteilungsfunktion der Knotengrade)

Die Verteilungsfunktion der Knotengrade $P(k)$ (bzw. $N(k)$) ist eine wichtige das Netzwerk charakterisierende Größe. Sie gibt an, wie groß der Anteil an Netzwerkknoten mit Knotengrad k ist. Bei realen (endlichen) Netzwerken ist diese Funktion keine kontinuierliche, sondern eine diskrete Funktion. In dem rechten Beispiel besitzt die Verteilungsfunktion das folgende Aussehen:



k

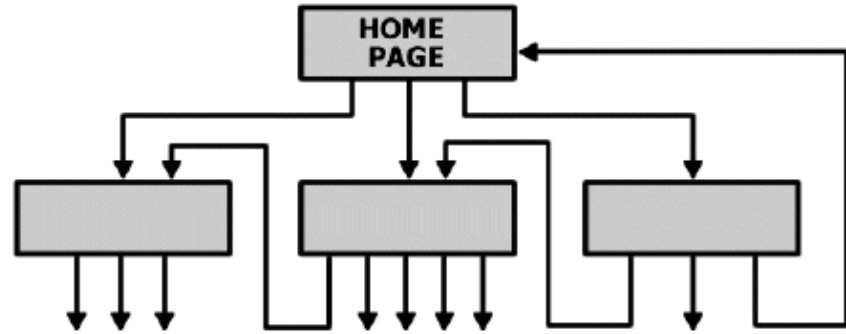
The Structure and Function of Complex Networks*

M. E. J. Newman[†]

Netzwerkstrukturen in unterschiedlichsten Systemen

	Network	Type	n	m
Social	film actors	undirected	449 913	25 516 482
	company directors	undirected	7 673	55 392
	math coauthorship	undirected	253 339	496 489
	physics coauthorship	undirected	52 909	245 300
	biology coauthorship	undirected	1 520 251	11 803 064
	telephone call graph	undirected	47 000 000	80 000 000
	email messages	directed	59 912	86 300
	email address books	directed	16 881	57 029
	student relationships	undirected	573	477
	sexual contacts	undirected	2 810	
Information	WWW nd.edu	directed	269 504	1 497 135
	WWW Altavista	directed	203 549 046	2 130 000 000
	citation network	directed	783 339	6 716 198
	Roget's Thesaurus	directed	1 022	5 103
	word co-occurrence	undirected	460 902	17 000 000
Technological	Internet	undirected	10 697	31 992
	power grid	undirected	4 941	6 594
	train routes	undirected	587	19 603
	software packages	directed	1 439	1 723
	software classes	directed	1 377	2 213
	electronic circuits	undirected	24 097	53 248
	peer-to-peer network	undirected	880	1 296
Biological	metabolic network	undirected	765	3 686
	protein interactions	undirected	2 115	2 240
	marine food web	directed	135	598
	freshwater food web	directed	92	997
	neural network	directed	307	2 359

WORLD-WIDE WEB



INTERNET

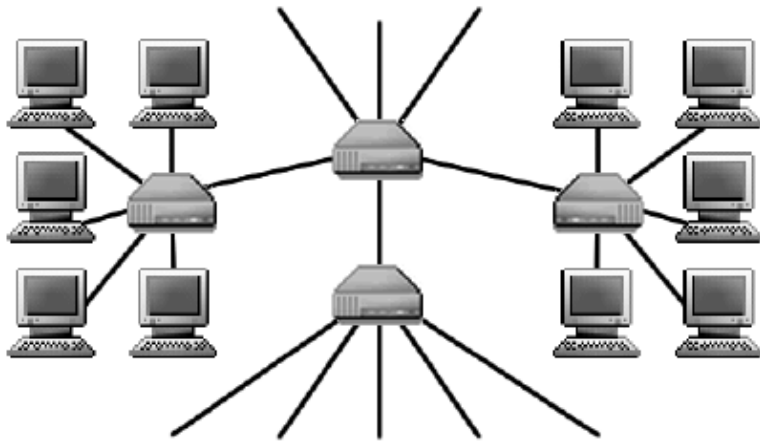
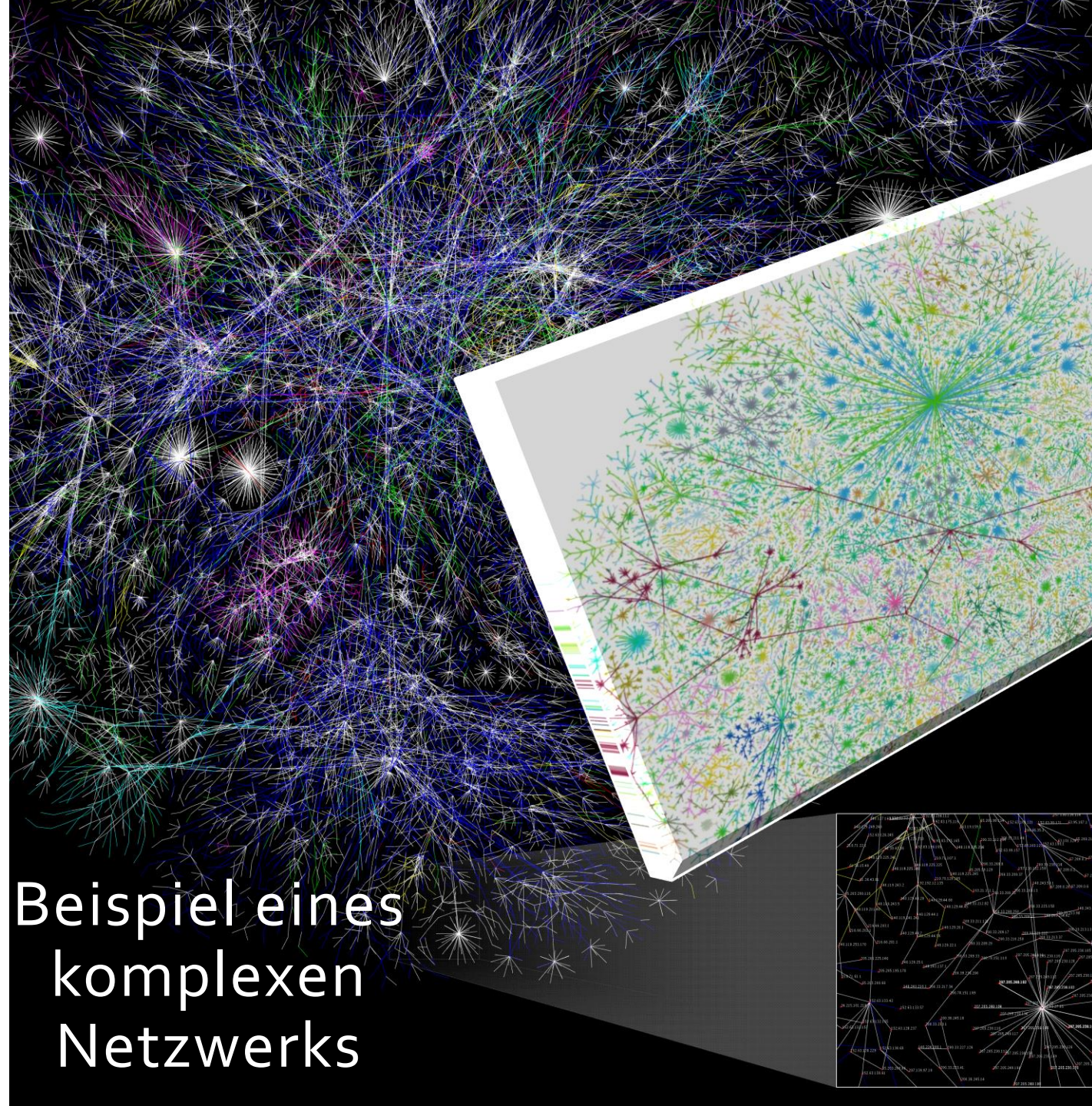


FIG. 1. Network structure of the World-Wide Web and the Internet. Upper panel: the nodes of the World-Wide Web are web documents, connected with directed hyperlinks (URLs). Lower panel: on the Internet the nodes are the routers and computers, the edges are the wires and cables that physically connect them. Figure courtesy of István Albert.

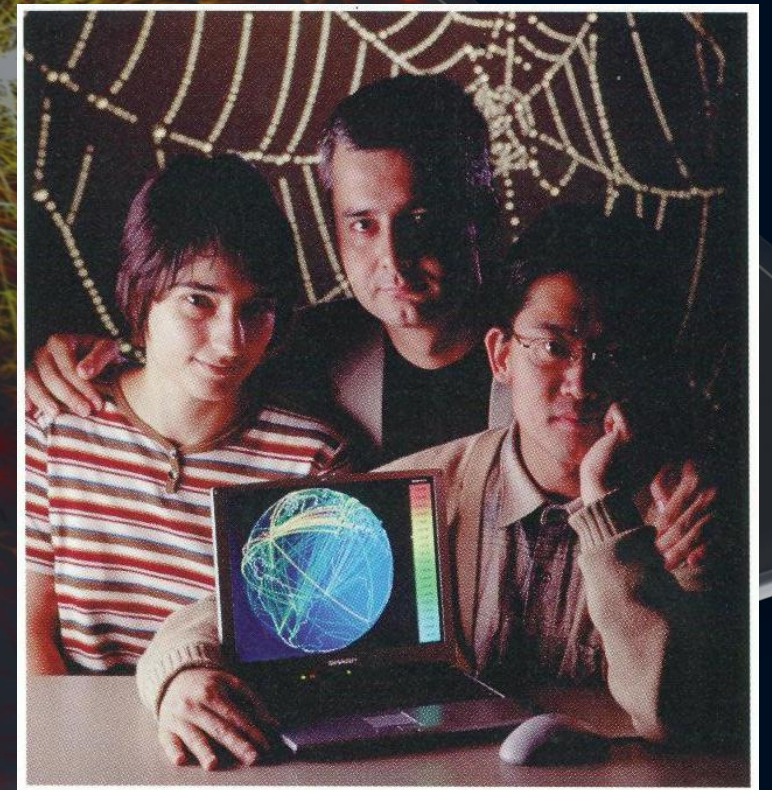
Beispiel eines
komplexen
Netzwerks



Wechselwirkungen und Struktur im Internet

Im Jahre 1999 untersuchten Albert-Laszlo Barabasi und Mitarbeiter die topologische Struktur des Internets (WWW)

A photo taken for Business 2.0 magazine in 2000, showing Reka Albert, Hawoong Jeong and Albert-Laszlo Barabasi, soon after our publication of the paper on the topology of the WWW.
(see <http://networksciencebook.com/>)



Viele der folgenden Abbildungen sind aus dem frei zugänglichen Buch „Network Science“ von Albert-Laszlo Barabasi entnommen.
<http://networksciencebook.com/>

Netzwerk-Klassen

Aufgrund ihrer unterschiedlichen Eigenschaften unterscheidet man die folgenden Netzwerk-Klassen:

i. Zufällige Netzwerke

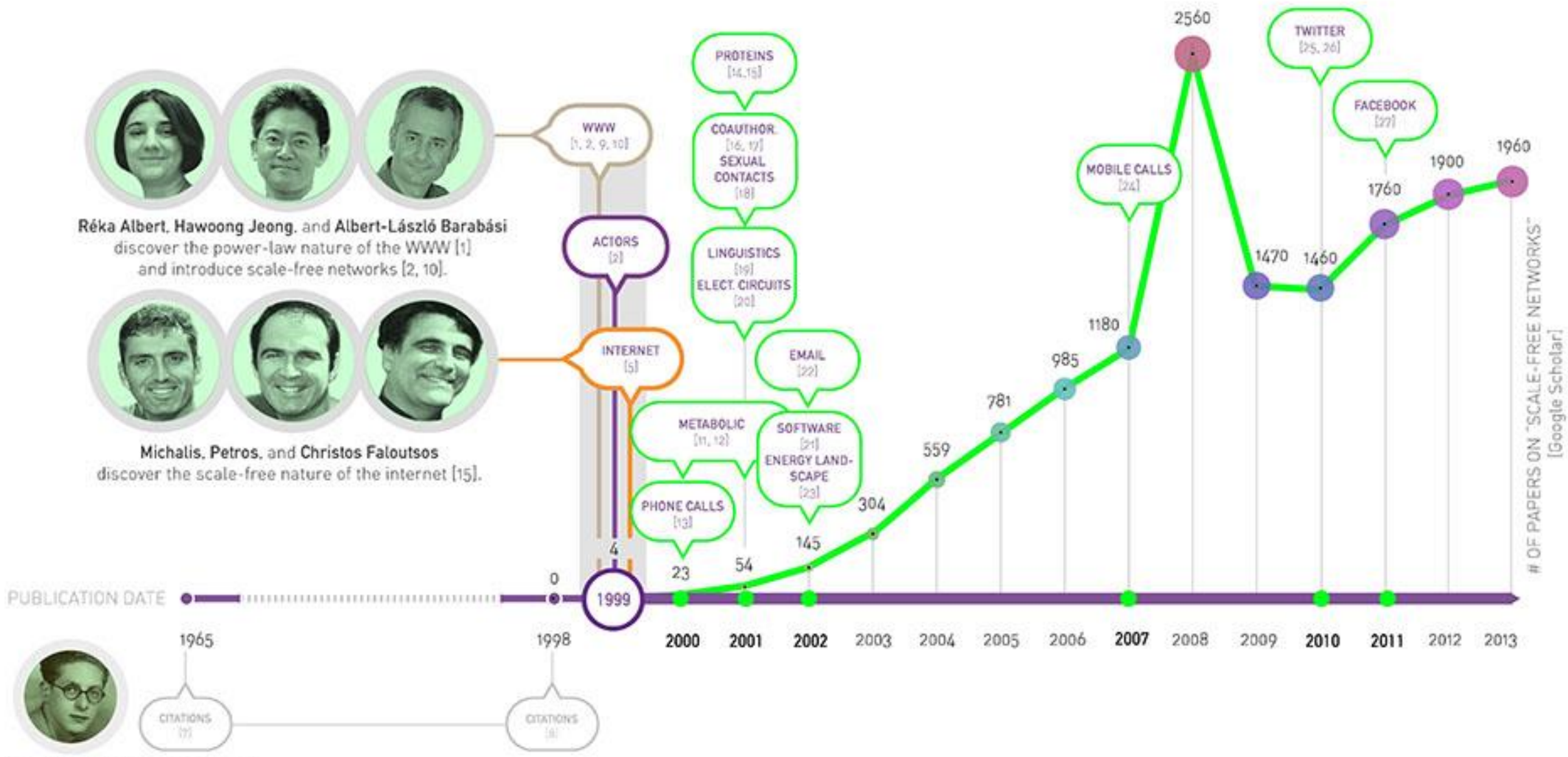
Die einzelnen Kanten bei *zufälligen Netzwerke* werden von den Knoten (Spielern) nach einem rein zufälligen Muster ausgewählt.

ii. „Kleine Welt“-Netzwerke (small-world networks)

- i. „Kleine Welt“-Netzwerke zeichnen sich durch einen kleinen Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten des Netzwerkes und einem großen Wert des Clusterkoeffizienten aus.

iii. Exponentielle Netzwerke

iv. Skalenfreie Netzwerke



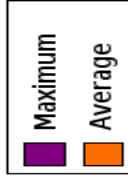
Derek de Solla Price [1922 - 1983]
 discovers that citations follow a power-law distribution [7], a finding later attributed to the scale-free nature of the citation network [2].

Skalenfreie Netzwerke auch in vielen weiteren sozio-ökonomischen und biologischen Systemen

Top 80 Internet exchange points

Gbit/s

500
450
400
350
300
250
200
150
100
50
0



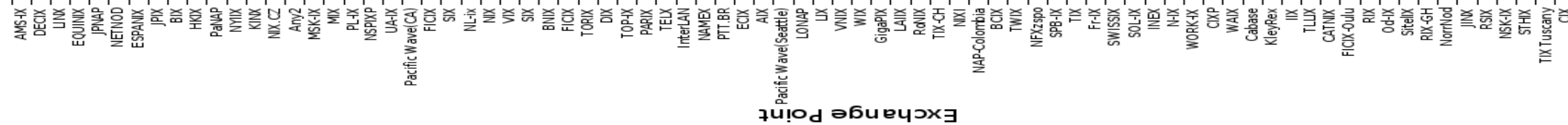
Gbit/s

500
450
400
350
300
250
200
150
100
50
0

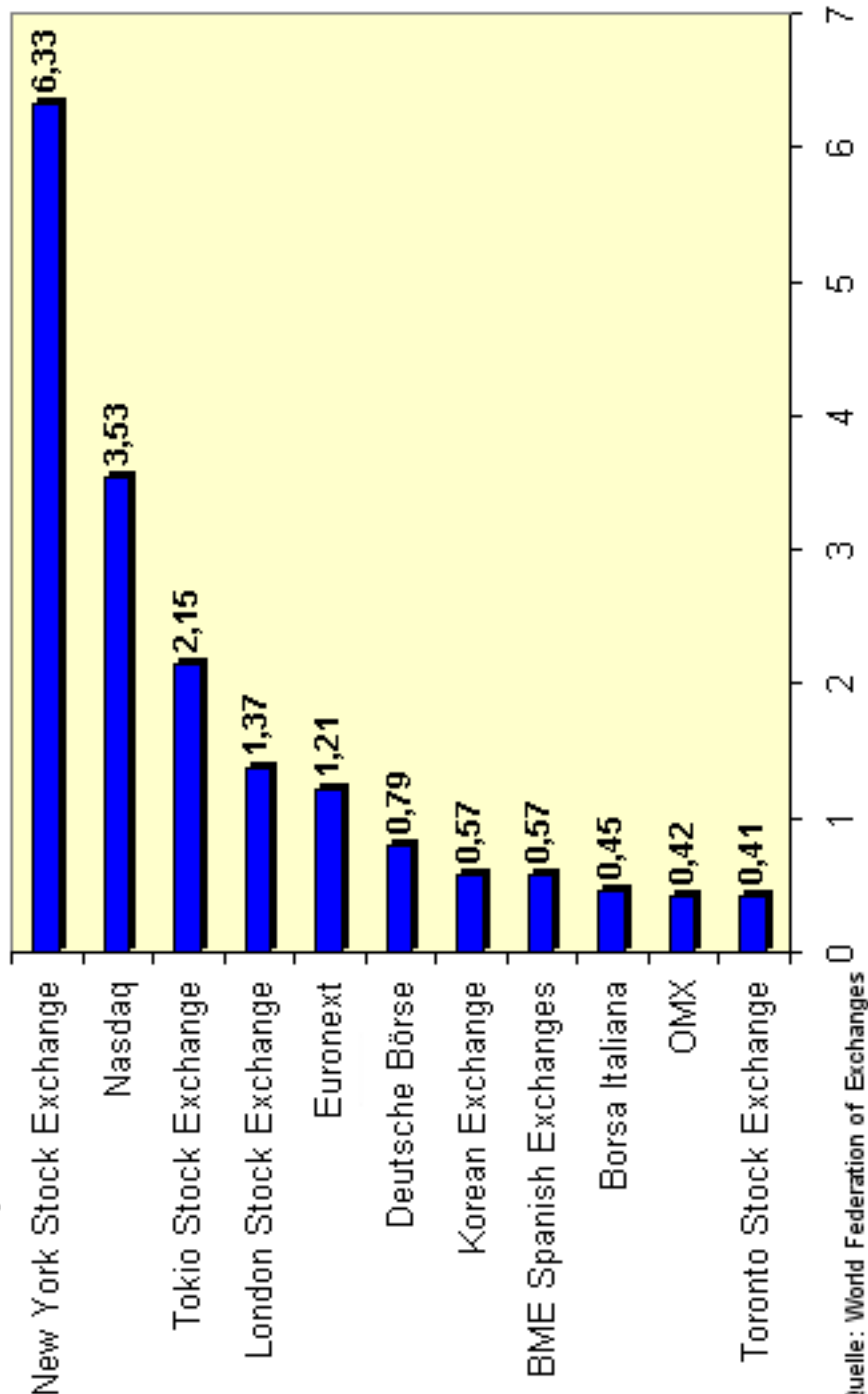
Frankfurt als mächtiger Knoten der Informationsströme Internet-Knoten

Das Internet stellt ein skalenfreies Netzwerk. Trägt man die 80 wichtigsten Internet-Knoten sortiert auf, so erhält man den typischen Verlauf eines skalenfreien Netzwerks. Frankfurt am Main mit seinem Internet-Knoten *DE-CIX* ist in der veralteten Grafik an Stelle 2, laut Wikipedia ist er jetzt sogar auf Rang 1

www.alpro.de



Die weltweit größten Wertpapierbörsen
nach Handelsvolumen (notierte inländische Aktien) in Billionen US-Dollar
Januar bis April 2006



Die Frankfurter Wertpapierbörse (FWB)

ist die bedeutendste deutsche Börse mit Sitz in Frankfurt am Main. Betreiberin und Träger ist die Deutsche Börse AG. Im Jahr 2000 wurde die Neue Börse im Industriehof in Frankfurt am Main in einem neuen Gebäude bezogen. Im Jahr 2005 wurden an den deutschen Börsen rund 3,8 Billionen Euro umgesetzt. Dabei entfielen vom Gesamtumsatz rund 3,2 Billionen Euro auf Aktien, Optionsscheine und börsengehandelte Fonds und rund 615 Milliarden Euro auf Anleihen. Der Aktienumsatz betrug 1,3 Billionen Euro, bei deutschen Aktien entfallen rund 98 % des Handels auf die Frankfurter Wertpapierbörse und Xetra, das elektronische Handelssystem der Deutschen Börse. Im Oktober 2008 entfielen 97 % der Umsätze in deutschen Aktien auf Xetra und die Frankfurter Parkettbörse. Bei ausländischen Aktien liefen über 86 % des Umsatzes über Xetra und den Präsenzhandel.'



NetworkX

Network Analysis in Python

Contact

[Mailing list](#)

[Issue tracker](#)

[Source](#)

Releases

Stable ([notes](#))

2.5 – August 2020

[download](#) | [doc](#) | [pdf](#)

Latest ([notes](#))

2.6 development

[github](#) | [doc](#) | [pdf](#)

[Archive](#)

<https://networkx.org/>

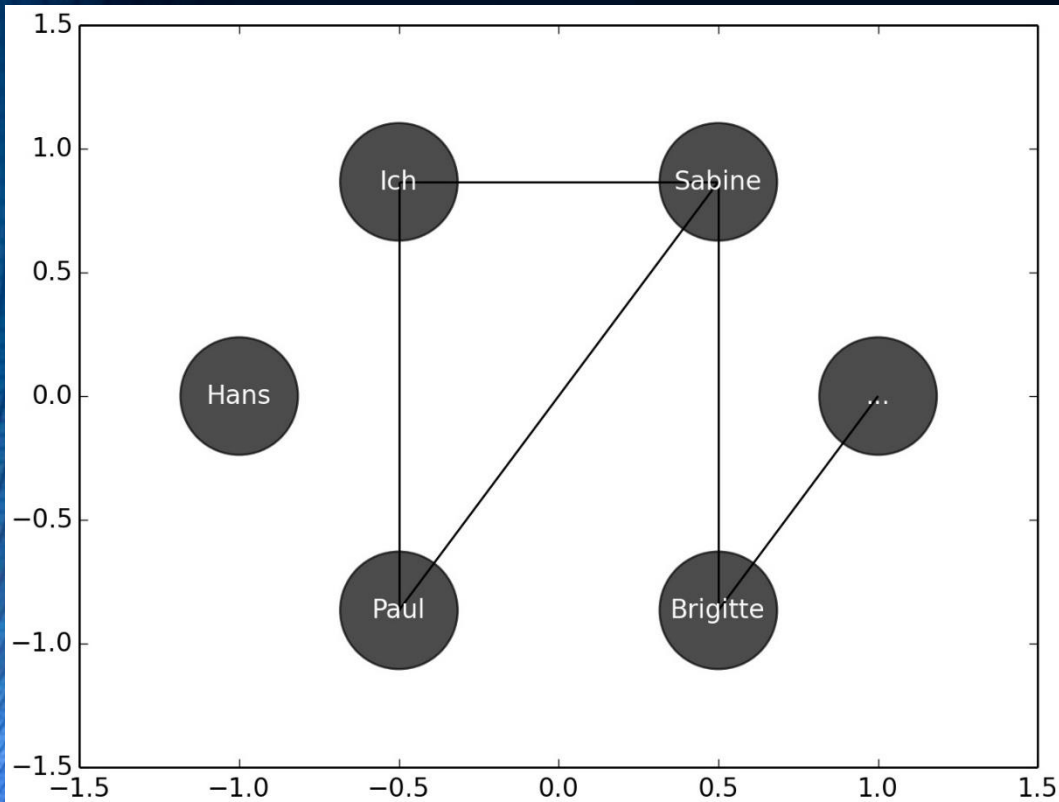
NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks.



Software for complex networks

- Data structures for graphs, digraphs, and multigraphs
- Many standard graph algorithms
- Network structure and analysis measures
- Generators for classic graphs, random graphs, and synthetic networks
- Nodes can be "anything" (e.g., text, images, XML records)
- Edges can hold arbitrary data (e.g., weights, time-series)
- Open source [3-clause BSD license](#)
- Well tested with over 90% code coverage
- Additional benefits from Python include fast prototyping, easy to teach, and multi-platform

Simulation und Darstellung von komplexen Netzwerken mit Python (Teil II)



```
import networkx as nx
import matplotlib.pyplot as plt
```

```
G=nx.Graph()
```

```
#Hinzufuegen von Knoten zum Netzwerk
```

```
G.add_nodes_from(["Ich", "Paul", "Brigitte", "Sabine", "Hans", "..."])
```

```
G.add_edge("Ich", "Paul")
```

```
G.add_edge("Ich", "Sabine")
```

```
G.add_edge("Sabine", "Brigitte")
```

```
G.add_edge("Sabine", "Paul")
```

```
G.add_edge("Paul", "Brigitte")
```

```
G.add_edge("Brigitte", "...")
```

```
#Erzeugung des Netzwerk-Bildes
```

```
graph_pos=nx.shell_layout(G)
```

```
#graph_pos=nx.shell_layout(G)
```

```
#graph_pos=nx.spring_layout(G)
```

```
nx.draw_networkx_nodes(G, graph_pos, node_size=3000, alpha=0.7, node_color="black")
```

```
nx.draw_networkx_edges(G, graph_pos)
```

```
nx.draw_networkx_labels(G, graph_pos, font_color="white")
```

```
#Speicherung des Bildes als .jpg
```

```
saveFig="./Netzwerk.jpg"
```

```
plt.savefig(saveFig, dpi=200, bbox_inches="tight", pad_inches=0.05, format="jpg")
```

```
plt.show()
```

Jupyter Notebook: Network1.ipynb

Physik der sozio-ökonomischen Systeme mit dem Computer (Physics of Socio-Economic Systems with the Computer)

Vorlesung gehalten an der J.W.Goethe-Universität in Frankfurt am Main

(Wintersemester 2020/21)

von Dr.phil.nat. Dr.rer.pol. Matthias Hanauske

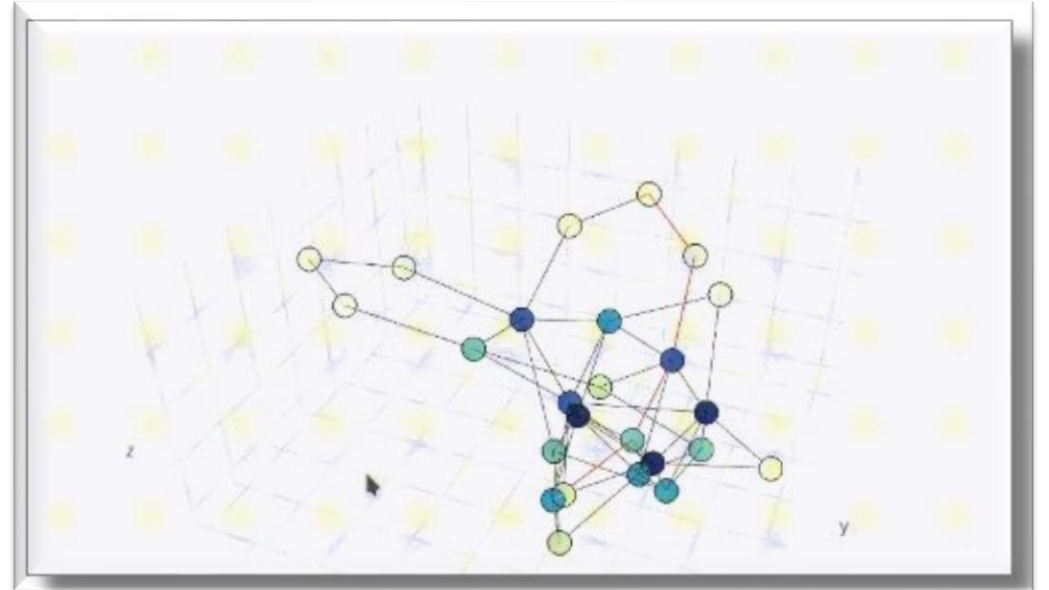
Frankfurt am Main 02.11.2020

Zweiter Vorlesungsteil:

Einführung In die Theorie der komplexen Netzwerke

Einführung

Eine bedeutende Einschränkung der deterministischen, evolutionären Spieltheorie ist deren zugrundeliegende Netzwerkstruktur (Netzwerktopologie). Die jeweiligen Spieler der betrachteten Population suchen in jeder Spielperiode einen neuen Spielpartner, wobei sie hierbei zufällig vorgehen (zufälliges Netzwerk) und vom Prinzip her mit jedem Spieler innerhalb der Population potentiell das zugrundeliegende Spiel spielen können (vollständig verbundenes



Wichtige Netzwerk charakterisierende Größen

Teil II.2 Einführung in die Theorie der komplexe Netzwerke - Mozilla Firefox

Teil II.2 Einführung in die T x +

https://itp.uni-frankfurt.de/~hاناuske/VPSOC/T2/II-2.html

170%

☆ 📄 🌐 ☰

Meistbesucht Getting Started

- **Der Knotengrad k_i**

Der Knotengrad des Knotens i ist gleich der Anzahl der Kanten die der Knoten i besitzt. Bei gerichteten Netzwerken unterscheidet man zwischen dem eingehenden k_i^{in} und ausgehenden Knotengrad k_i^{out} . Bei gewichteten Netzwerken summiert man über die Zahlenfaktoren der gewichteten Kanten.

- **Der durchschnittliche Knotengrad $\langle k \rangle$**

Der durchschnittliche Knotengrad eines ungerichteten Netzwerks $\langle k \rangle$ ist gegeben durch

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2L}{N}, \text{ wobei } L = \text{die gesamte Anzahl der Links (Kanten) des Netzwerks ist.}$$

- **Die Nachbarschaftsmatrix (Adjazenzmatrix) A_{ij}**

Die Nachbarschaftsmatrix des Netzwerkes beschreibt, welche Knoten des Graphen durch eine Kante verbunden sind. Zeigt eine Kante von Knoten j auf den Knoten i , dann ist der zugehörige Matrixwert $A_{ij} = 1$, wohingegen ein Eintrag $A_{ij} = 0$ bedeutet, dass keine Kante von j nach i zeigt. Bei ungerichteten Netzwerken ist die Nachbarschaftsmatrix symmetrisch. Der Knotengrad k_i des Knotens i

(bzw. k_i^{in} und k_i^{out} bei gerichteten Graphen) lässt sich mittels der Nachbarschaftsmatrix wie folgt bestimmen: $k_i = \sum_{j=1}^N A_{ij}$ ($k_i^{\text{in}} = \sum_{j=1}^N A_{ij}$ und $k_i^{\text{out}} = \sum_{j=1}^N A_{ji}$ bei gerichteten Graphen), wobei

die gesamte Anzahl der Kanten sich wie folgt bestimmen lässt: $L = \sum_{i,j=1}^N A_{ij}/2$ und die Größe N die Anzahl der Knoten des Netzwerkes bezeichnet. In gewichteten Netzwerken sind die Einträge der Nachbarschaftsmatrix nicht mehr auf Null und Eins beschränkt $A_{ij} = w_{ij}$.

- **Die Verteilungsfunktion der Knotengrade $P(k) = N(k)/N$**

Nachbarschaftsmatrix nicht mehr auf Null und Eins beschränkt $A_{ij} = w_{ij}$.

- **Die Verteilungsfunktion der Knotengrade** $P(k) = N(k)/N$

Die Verteilungsfunktion der Knotengrade $P(k)$ (bzw. $N(k)$) ist eine wichtige das Netzwerk charakterisierende Größe. Sie gibt an, wie groß der Anteil an Netzwerkknoten mit Knotengrad k ist. Bei realen (endlichen) Netzwerken ist diese Funktion keine kontinuierliche, sondern eine diskrete Funktion $P(k) \approx P_k$. Bezeichnen wir mit N_k die Anzahl der Knoten mit Knotengrad k , so gilt $P_k = \frac{N_k}{N}$. Es gilt $\sum_{k=0}^{\infty} P_k = 1$ und der durchschnittliche Knotengrad berechnet sich wie folgt $\langle k \rangle = \sum_{k=0}^{\infty} k P_k$.

- **Die maximale Anzahl möglicher Kanten** L_{max}

Die maximale Anzahl möglicher Kanten in einem ungerichteten Netzwerk (ein sogenannter 'complete graph') besitzt $L_{max} = \frac{N(N-1)}{2}$ Kanten. Viele real existierende Netzwerke sind dünnbesetzt ($L \ll L_{max}$).

- **Die kürzeste Verbindungstrecke zwischen zwei Knoten** d_{ij}

Die Anzahl der Kanten die in einem Verbindungsweg von Knoten i zum Knoten j durchlaufen wird hängt vom gewählten Pfad ab. Die kürzeste Verbindungstrecke d_{ij} kann hierbei im allgemeinen auf unterschiedlichen Wegen realisiert sein, wobei der Pfad keine Schleifen enthalten darf und sich nicht kreuzen darf. Die Anzahl der kürzeste Verbindungswege N_{ij} lässt sich mittels der Nachbarschaftsmatrix berechnen. Existieren z.B. Verbindungswege vom Knoten i zum Knoten j mit $d_{ij} = 2$, so muss die Anzahl der kürzeste Verbindungswege $N_{ij} = \sum_{l=1}^N A_{il} A_{jl} \geq 1$ sein.

- **Der Clusterkoeffizient** C_i bzw. C

Der Clusterkoeffizient gibt die Wahrscheinlichkeit an, dass zwei nächste Nachbarn eines Knotens

berechnen. Existieren z.B. Verbindungswege vom Knoten i zum Knoten j mit $w_{ij} = 2$, so muss die Anzahl der kürzeste Verbindungswege $N_{ij} = \sum_{l=1}^N A_{il}A_{jl} \geq 1$ sein.

- **Der Clusterkoeffizient C_i bzw. C**

Der Clusterkoeffizient gibt die Wahrscheinlichkeit an, dass zwei nächste Nachbarn eines Knotens ebenfalls nächste Nachbarn untereinander sind. Für einen speziellen Knoten i berechnet er sich mittels:

$C_i = \frac{2L_i}{k_i(k_i-1)}$, wobei L_i die Anzahl der Kanten darstellt, die die nächsten Nachbarn des Knoten i

miteinander verbinden. Der globale Wert C des Clusterkoeffizienten ist der Mittelwert aller C_i 's ($C = \frac{1}{N} \sum_{i=1}^N C_i$) und stellt demnach eine Art von der Enge der Nachbarschaftsbeziehungen des

Netzwerks dar.

- **Durchschnittliche Anzahl der m -nächsten Nachbarn z_m**

z_1 stellt hierbei den Wert der mittleren Knotenzahl des Netzwerkes dar ($z_1 = \langle k \rangle$) und z_2 die mittlere Anzahl zweiter-nächster Nachbarn ($z_2 = \langle k^2 \rangle - \langle k \rangle$), siehe z.B. [Claudius Gros, "Complex and Adaptive Dynamical Systems, a Primer", S:18](#)).

- **Durchmesser des Netzwerks l**

Der Durchmesser des Netzwerks gibt die maximale kürzeste Kantenlänge zwischen zwei beliebigen Knoten des Netzwerkes an: $l = \log(N/z_1)/\log(z_2/z_1) + 1$, wobei N die Anzahl der Knoten des Netzwerkes darstellt (siehe z.B. [Claudius Gros, "Complex and Adaptive Dynamical Systems, a Primer", S:20](#)). Der Wert l wird auch in einigen Lehrbüchern als d_{\max} bezeichnet (siehe [Albert-Laszlo Barabasi, Network science, Chapter 2 Graph Theory](#)).

- **Größter verbundener Knotenclusters N_G (Giant component, Hub)**

Die Anzahl der Knoten im größten verbundener Knotencluster des Netzwerkes ist mit N

miteinander verbinden. Der globale Wert C des Clusterkoeffizienten ist der Mittelwert aller C_i 's ($C = \frac{1}{N} \sum_{i=1}^N C_i$) und stellt demnach eine Art von der Enge der Nachbarschaftsbeziehungen des Netzwerks dar.

- **Durchschnittliche Anzahl der m -nächsten Nachbarn z_m**

z_1 stellt hierbei den Wert der mittleren Knotenzahl des Netzwerkes dar ($z_1 = \langle k \rangle$) und z_2 die mittlere Anzahl zweiter-nächster Nachbarn ($z_2 = \langle k^2 \rangle - \langle k \rangle$, siehe z.B. [Claudius Gros, "Complex and Adaptive Dynamical Systems, a Primer", S:18](#)).

- **Durchmesser des Netzwerks l**

Der Durchmesser des Netzwerks gibt die maximale kürzeste Kantenlänge zwischen zwei beliebigen Knoten des Netzwerkes an: $l = \log(N/z_1)/\log(z_2/z_1) + 1$, wobei N die Anzahl der Knoten des Netzwerkes darstellt (siehe z.B. [Claudius Gros, "Complex and Adaptive Dynamical Systems, a Primer", S:20](#)). Der Wert l wird auch in einigen Lehrbüchern als d_{\max} bezeichnet (siehe [Albert-Laszlo Barabasi, Network science, Chapter 2 Graph Theory](#)).

- **Größter verbundener Knotenclusters N_G (Giant component, Hub)**

Die Anzahl der Knoten im größten verbundenen Knotenclusters des Netzwerkes wird mit N_G bezeichnet.

Diese und weitere graphentheoretischen Größen (Number of Shortest Paths Between Two Nodes, Shortest Path, Average Path Length, ...) sind in [Albert-Laszlo Barabasi, Network science, Chapter 2 Graph Theory](#) zu finden.



Zufällige Netzwerke

Verteilungsfunktion
der Knotengrade

$$P(k)$$

bzw. $N(k) := N * P(k)$

VI. CLASSICAL RANDOM GRAPHS, THE ERDÖS-RÉNYI MODEL

The simplest and most studied network with undirected edges was introduced by Erdős and Rényi (ER model) [77,78]. In this network:

- (i) the total number of vertices, N , is fixed;
- (ii) the probability that two arbitrary vertices are connected equals p .

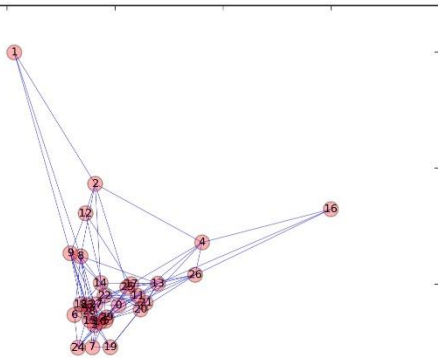
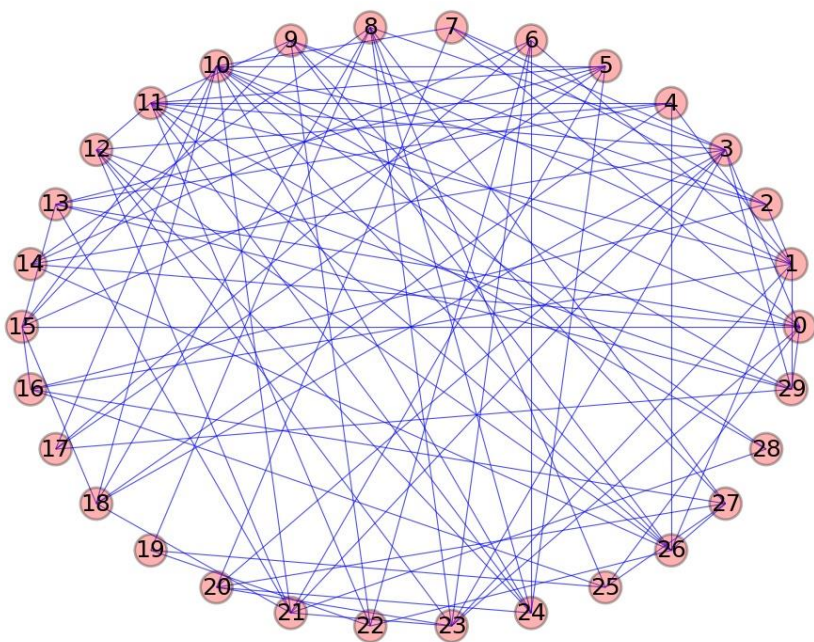
One sees that, on average, the network contains $pN(N-1)/2$ edges. The degree distribution is binomial,

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}, \quad (4)$$

so the average degree is $\bar{k} = p(N-1)$. For large N , the distribution, Eq. (4) takes the Poisson form,

$$P(k) = e^{-\bar{k}} \bar{k}^k / k!. \quad (5)$$

Python (Version 2)



```
import networkx as nx
import matplotlib.pyplot as plt
from random import randint
from random import uniform

G=nx.Graph()

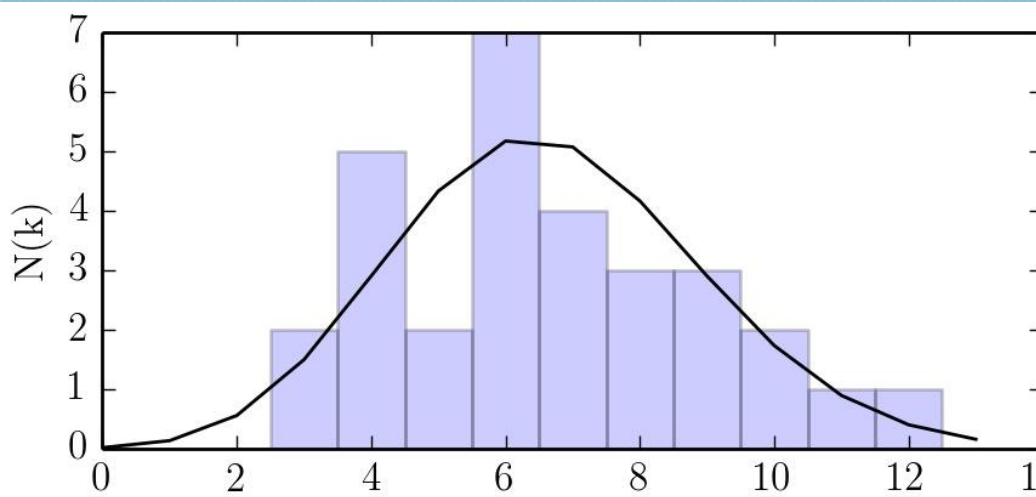
#Allgemeine Festlegungen des Netzwerks
NKn =30 #Anzahl der Knoten (vertices)
anzedges=100 #Gesamte Anzahl der im Netzwerk bestehenden Kanten (links, connections)

#Hinzufuegen der Knoten zum Netzwerk
G.add_nodes_from(range(0,NKn,1))

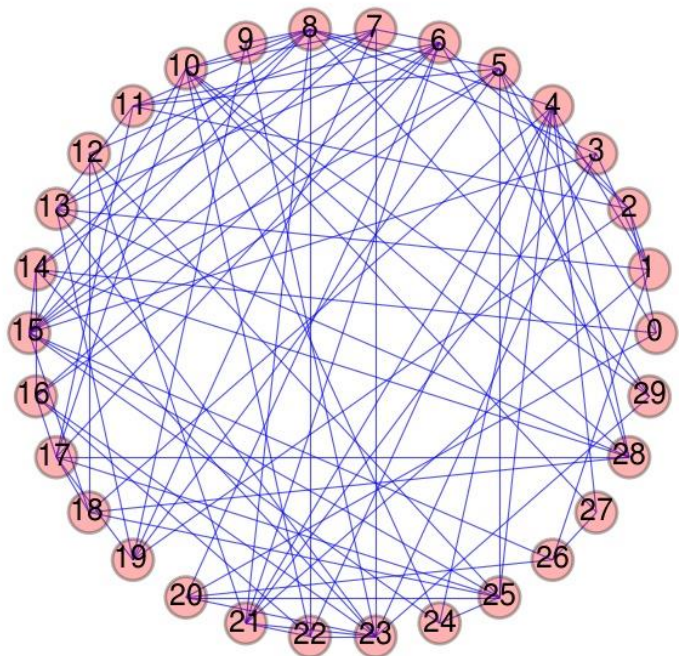
#Erzeugung der Kanten des Netzwerkes (zufaellig anzedges-Kanten zwischen den NKn-Knoten erzeugen)
links=0
while links < anzedges:
    KnA = randint(0, NKn-1)
    KnB = randint(0, NKn-1)
    if KnA != KnB and list(G.edges()).count((KnA,KnB))==0 and list(G.edges()).count((KnB,KnA))==0:
        G.add_edge(KnA,KnB)
        links=links+1

#Erzeugung des Netzwerk-Bildes
node_size=150|
node_alpha=0.3
node_color="red"
edge_tickness=0.4
edge_alpha=0.7
edge_color="blue"
node_text_size=9
text_font="sans-serif"
graph_pos=nx.shell_layout(G)
#graph_pos=nx.spectral_layout(G)
#graph_pos=nx.spring_layout(G)
nx.draw_networkx_nodes(G,graph_pos,node_size=node_size,alpha=node_alpha,node_color=node_color)
nx.draw_networkx_edges(G,graph_pos,width=edge_tickness,alpha=edge_alpha,edge_color=edge_color)
nx.draw_networkx_labels(G,graph_pos,font_size=node_text_size,font_family=text_font)

#Speicherung des Bildes als .jpg und .pdf Datei
saveFig="./Netzwerk.jpg"
```



$$N(k) = N \binom{N-1}{k} p^k (1-p)^{N-1-k}$$



$$N = 30, p = 0.2299, m = 100$$

```

import networkx as nx
import matplotlib.pyplot as plt
from random import randint
from random import uniform
import numpy as np
import matplotlib
import matplotlib.gridspec as gridspec
from matplotlib.ticker import NullFormatter

nullfmt = NullFormatter() # Keine Labels im Netzwerkplot

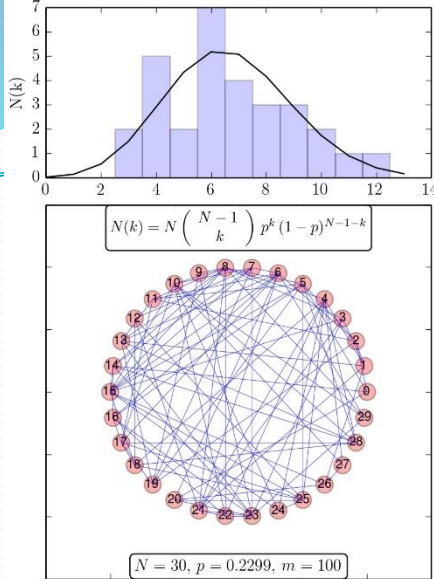
#Von https://de.wikibooks.org/wiki/Algorithmensammlung:_Statistik:_Binomialkoeffizient
def binomialkoeffizient(n, k):
    if k == 0: return 1
    if 2*k > n:
        ergebnis = binomialkoeffizient(n, n-k)
    else:
        ergebnis = n-k+1
        for i in range(2, k+1): # i in [2; k]
            ergebnis *= (n-k+i) # Selbstmultiplikation
            ergebnis /= i # Achtung: Ergebnis ist eine Kommazahl!
    return int(ergebnis)]

#Analytische Verteilungsfunktion eines zufaelligen Netzwerkes
def P(n, kmin, kmax, p):
    ergebnis=[]
    for k in range(kmin, kmax, 1):
        ergebnis.append(binomialkoeffizient(n-1, k)*p**k*(1-p)**(n-1-k))
    return ergebnis

#plot settings
params = {
    'figure.figsize' : [5, 7.2],
    'text.usetex' : True,
}
matplotlib.rcParams.update(params)

#Grid
plt.figure(0)
gs = gridspec.GridSpec(2, 1, height_ratios=[1,2.2], hspace=0.1)
ax1 = plt.subplot(gs[0])
ax2 = plt.subplot(gs[1])

```



```
#####  
#Beginn des eigentlichen Python Programms "Zufaelliches Netzwerk"  
G=nx.Graph()  
#Allgemeine Festlegungen des Netzwerks  
NKn =30 #Anzahl der Knoten (vertices)  
anzedges=100 #Gesamte Anzahl der im Netzwerk bestehenden Kanten (links, connections)  
  
#Hinzufuegen der Knoten zum Netzwerk  
G.add_nodes_from(range(0,NKn,1))  
  
#Erzeugung der Kanten des Netzwerkes (zufaellich anzedges-Kanten zwischen den NKn-Knoten erzeugen)  
links=0  
while links < anzedges:  
    KnA = randint(0, NKn-1)  
    KnB = randint(0, NKn-1)  
    if KnA != KnB and list(G.edges()).count((KnA,KnB))==0 and list(G.edges()).count((KnB,KnA))==0:  
        G.add_edge(KnA,KnB)  
        links=links+1  
  
#Liste der Knotengrade  
degree_sequence=sorted([d for n,d in G.degree()],reverse=True) # degree sequence  
maxk=np.max(degree_sequence)  
  
#Berechnung der Wahrscheinlichkeit das Knoten KnA mit KnB verbunden ist  
p=(2*anzedges/float(NKn*(NKn-1)))  
  
#Erzeugung des Bildes der Verteilungsfunktion der Knotengrade N(k) (analytisch,simulativ)  
ax1.plot(range(0,int(maxk+2),1),np.array(P(NKn,0,int(maxk+2),p))*NKn, linewidth=1, linestyle='-', c="black")  
ax1.hist(degree_sequence,bins=range(0,int(maxk+2),1), align="left", histtype='bar', color="blue", alpha=0.2)  
  
#Achsenbeschriftung  
ax1.set_ylabel(r'$\rm N(k)$')  
ax2.yaxis.set_major_formatter(nullfmt)  
ax2.xaxis.set_major_formatter(nullfmt)
```

```
#Erzeugung des Netzwerk-Bildes
```

```
node_size=150  
node_alpha=0.3  
node_color="red"  
edge_tickness=0.4  
edge_alpha=0.7  
edge_color="blue"  
node_text_size=9  
text_font="sans-serif"  
graph_pos=nx.shell_layout(G)  
nx.draw_networkx_nodes(G,graph_pos,node_size=node_size,alpha=node_alpha, node_color=node_color)  
nx.draw_networkx_edges(G,graph_pos,width=edge_tickness, alpha=edge_alpha,edge_color=edge_color)  
nx.draw_networkx_labels(G, graph_pos,font_size=node_text_size,font_family=text_font)
```

```
# Plotten der Netzwerkeigenschaften in das Bild
```

```
roundp="%.4f"%p  
textstr1=r'$N(k) = N \left( \begin{array}{c} N-1 \\ k \end{array} \right) p^k (1-p)^{N-1-k}$'  
textstr2=r'$N='+str(NKn)+' , p='+roundp+' , m='+str(anzedges)+'$'  
props = dict(boxstyle='round', facecolor='white', alpha=0.92)  
plt.text(0, 1.46, textstr1, fontsize=11, verticalalignment='top', horizontalalignment='center', bbox=props)  
plt.text(0, -1.46, textstr2, fontsize=12, verticalalignment='bottom', horizontalalignment='center', bbox=props)
```

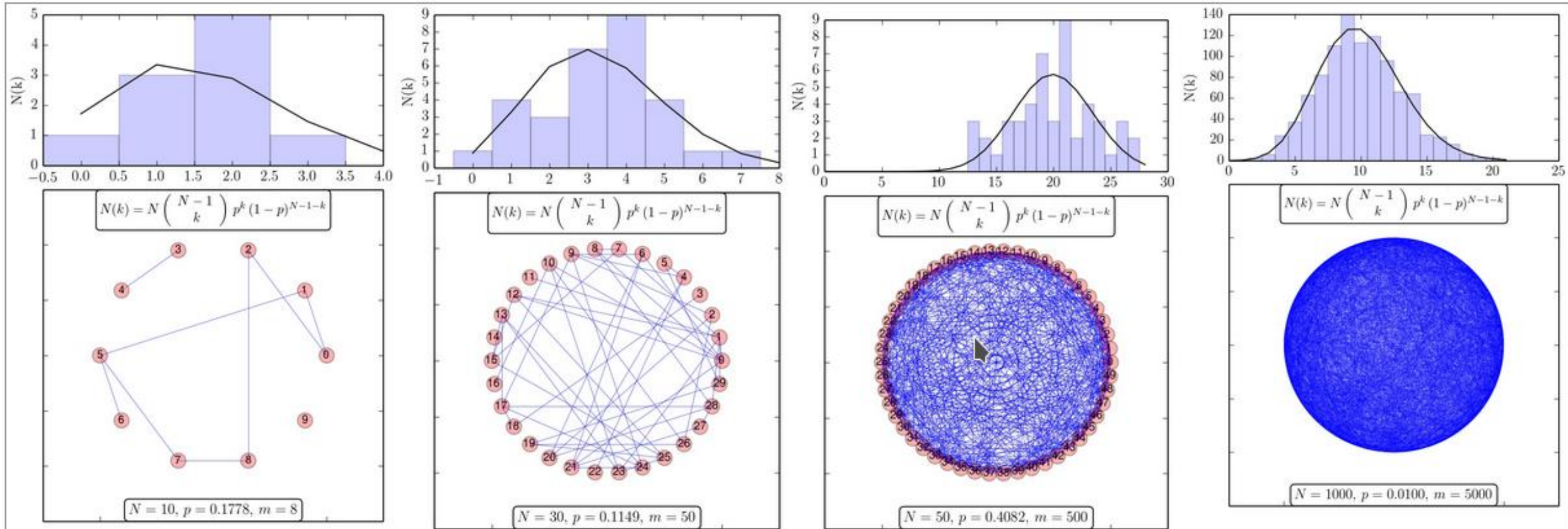
```
#Speicherung des Bildes als .jpg und .pdf Datei
```

```
saveFig="./Netzwerk.jpg"  
plt.savefig(saveFig, dpi=200,bbox_inches="tight",pad_inches=0.05,format="jpg")  
saveFig="./Netzwerk.pdf"  
plt.savefig(saveFig,bbox_inches="tight",pad_inches=0.05,format="pdf")  
plt.show()
```

Python (Version 3)

Python Programm *RandomNetwork.py*

II.2.1 Zufällige Netzwerke (Random Networks)



Ergebnisse des Python Skriptes RandomNetwork.py: Zufallsgraph und die zugehörige Verteilungsfunktion der Knotengrade (schwarze Kurve: analytische Verteilung).

Python Programm VPSOC-RandomNetwork_evol.py

Die einzelnen Kanten bei zufälligen Netzwerken werden von den Knoten (Spielern) nach einem rein zufälligen Muster ausgewählt. Im Erdos-Renyi Modell (Erdos and Renyi, 1959) werden N Knoten zufällig mit m ungerichteten Kanten verbunden. Die Wahrscheinlichkeit p , dass ein Knoten mit dem anderen verbunden ist, demnach $p = 2m/(N(N - 1))$. Die Verteilungsfunktion der Knotengrade $P(k)$ ist binomialverteilt:

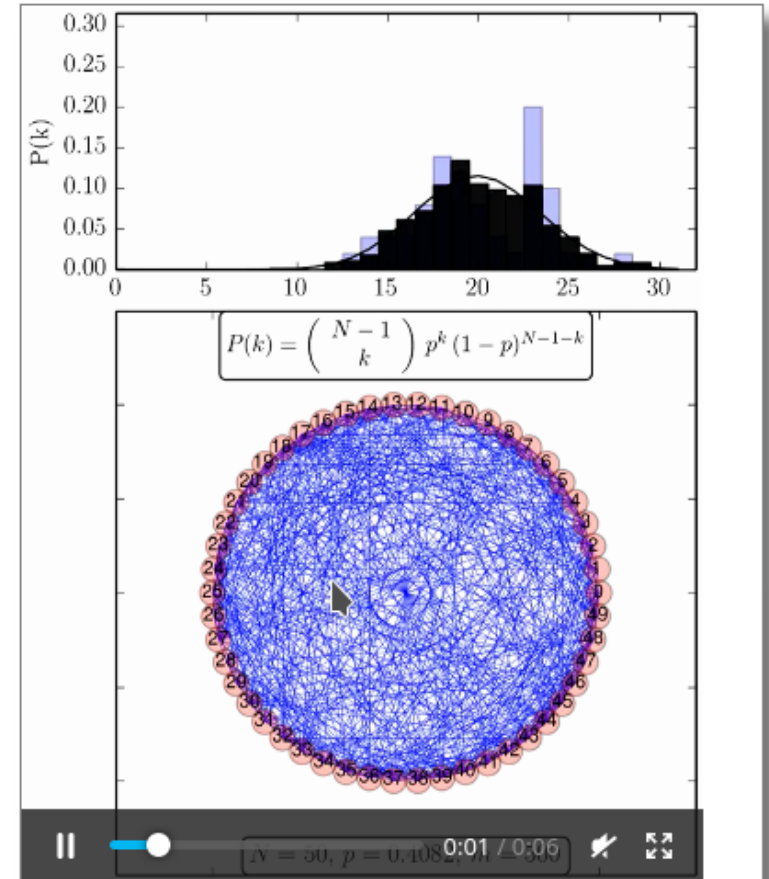
$$P(k) = N(k)/N = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$

Für große N geht diese Verteilung in die folgende Poisson Verteilungsfunktion über

$$P(k) = \frac{e^{-\bar{k}} \bar{k}^k}{k!},$$

wobei $\bar{k} = \langle k \rangle = p(N - 1)$ der mittlere Knotengrad im Netzwerk ist (siehe Abb. 3.4 in [Chapter 3: Albert-Laszlo Barabasi, Network Science](#)). Die oberen Abbildungen zeigen Ergebnisse des Python Skriptes ([Python Skript: RandomNetwork.py](#)) welches einen Zufallsgraphen mit N -Knoten und m -Kanten und die zugehörige Verteilungsfunktion $N(k) = N P(k)$ der Knotengrade darstellt. Die einzelnen blauen Kästchen des Balkendiagramms stellen die spezifische, zufällig erzeugte Realisierung des Zufallsgraphen dar, wohingegen die schwarze Kurve die analytische Binomialverteilung zeigt. Mittelt man über die einzelnen Verteilungsfunktionen $P(k)$ mehrerer zufälliger Netzwerke, erzeugt man ein Ensemble-Mittelwert der möglichen Realisierungen des Zufallsgraphen. In der nebenstehenden rechten Abbildung ist eine solche Mittelung anhand einer Animation veranschaulicht. Man erkennt deutlich, dass die über mehrere zufällige Netzwerke gemittelte Verteilung (dunkel graue Kästchen) sich der analytischen Verteilung (schwarze Kurve) immer mehr annähert.

Eine weitere wichtige Eigenschaft des Netzwerkes ist die relative Größe des größten verbundenen Knotenclusters (Giant component, Hub) N_G/N . Zufällige Netzwerke mit kleinem mittlerem Knotengrad $\langle k \rangle < 1$ besitzen keine



Mittelt man über die einzelnen Verteilungsfunktionen $P(k)$ mehrerer zufälliger Netzwerke so ergeben sich die analytisch ermittelten Verteilungsfunktionen (Python Skript [VPSOC-RandomNetwork_evol.py](#)).

Jupyter Notebook: RandomNetworks.ipynb

Physik der sozio-ökonomischen Systeme mit dem Computer

(Physics of Socio-Economic Systems with the Computer)

Vorlesung gehalten an der J.W.Goethe-Universität in Frankfurt am Main

(Wintersemester 2020/21)

von Dr.phil.nat. Dr.rer.pol. Matthias Hanauske

Frankfurt am Main 02.11.2020

Zweiter Vorlesungsteil:

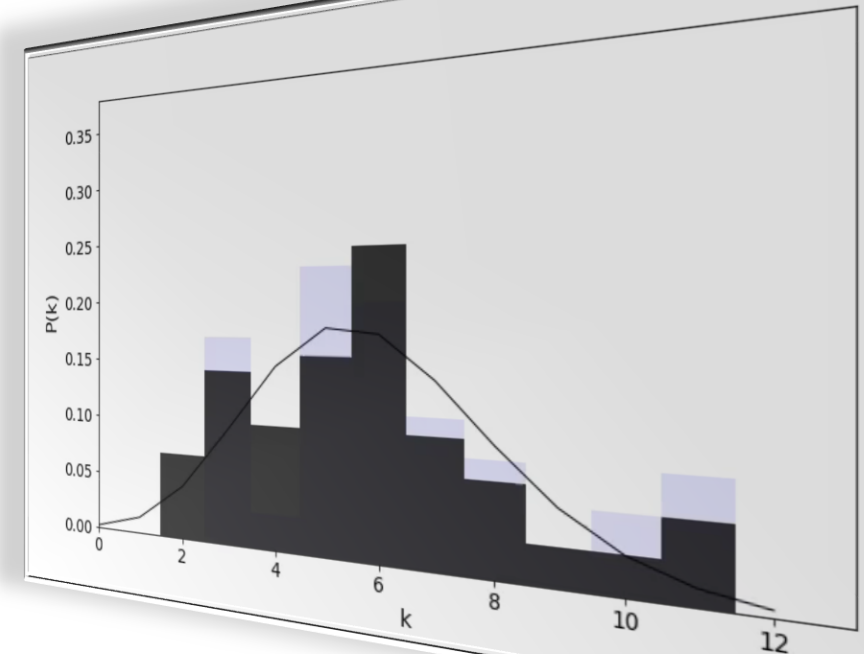
Zufällige komplexe Netzwerke (random networks)

Einführung

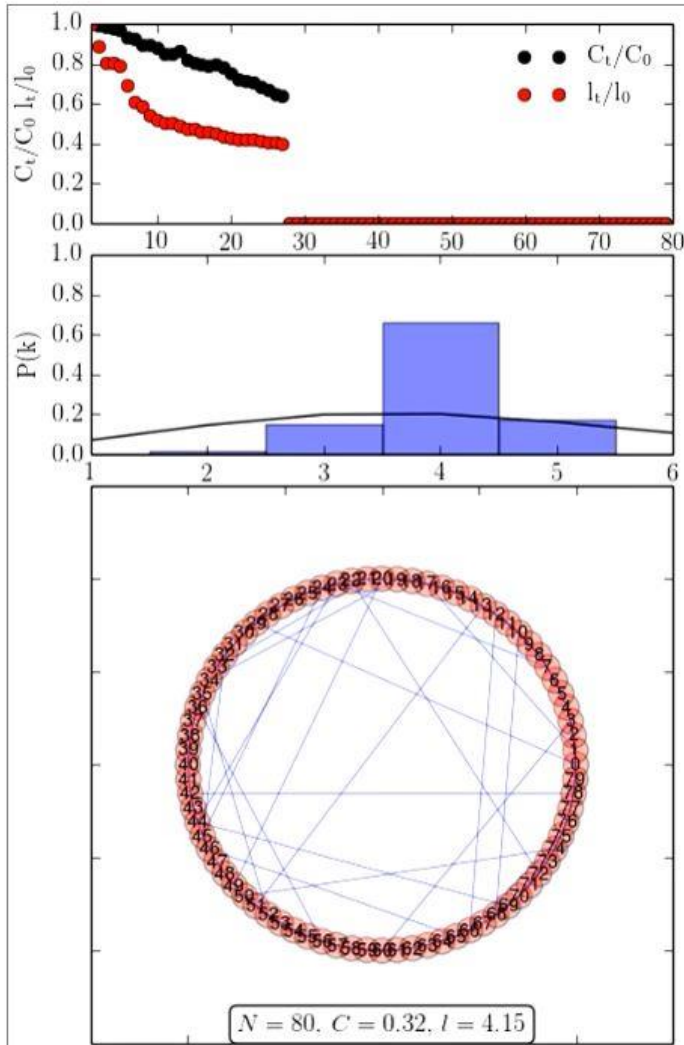
Aufgrund ihrer unterschiedlichen Eigenschaften unterscheidet man die folgenden Netzwerk-Klassen: Zufällige Netzwerke (Random Networks: Die einzelnen Kanten bei zufälligen Netzwerke werden von den Knoten (Spielern) nach einem rein zufälligen Muster ausgewählt), Kleine Welt-Netzwerke (Small World Networks, Kleine Welt-Netzwerke zeichnen sich durch einen kleinen Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten des Netzwerkes und einem großen Wert des Clusterkoeffizienten aus), Exponentielle Netzwerke (Exponential Networks) und Skalenfreie Netzwerke (Scale-Free Networks).

Bei einigen Modellnetzwerken können analytische Ergebnisse gewonnen werden. Im folgenden betrachten wir die Klasse der zufällige Netzwerke. Die einzelnen Kanten bei zufälligen Netzwerken werden von den Knoten (Spielern) nach einem rein zufälligen Muster ausgewählt. Im Erdos-Renyi Modell (Erdos and Renyi, 1959) werden N Knoten zufällig mit L ungerichteten Kanten verbunden. Die Wahrscheinlichkeit p , das ein Knoten mit dem anderen verbunden ist demnach $p = \frac{2L}{N(N-1)}$. Die Verteilungsfunktion der Knotengrade $P(k)$ ist binomialverteilt:

$$P(k) = N(k)/N = \binom{N-1}{k} p^k (1-p)^{N-1-k}$$



II.2.2 Kleine Welt Netzwerke (Small-World Networks)



Übergang zwischen einem regulären Netzwerk, einem *Kleine Welt* Netzwerk und einem zufälligen Netzwerk (Python Skript [VPSOC-SmallWorld-Network_evol.py](#)).

Kleine Welt-Netzwerke zeichnen sich durch einen kleinen Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten des Netzwerkes und einem großen Wert des Clusterkoeffizienten aus. Ein einfaches Modell welches den Übergang von einem Netzwerk mit regulärer Struktur (Gitter-ähnlicher Struktur) über ein *kleines Welt* Netzwerk hin zu einem zufälligen Netzwerk veranschaulicht, wurde von Watts und Strogatz im Jahre 1998 vorgestellt: Im einfachsten Fall startet man hierbei mit einem eindimensionalen Gitter-Netzwerk mit N Knoten, wobei jeder Knoten mit seinen K -nächsten Nachbarn (hier speziell $K=2$) verbunden ist. Nun löscht man mit der Wahrscheinlichkeit p jede der existierenden Verbindungen (Kanten) und stellt eine neue Kante im Netzwerk in zufälliger Weise her. Für $p=0$ bleibt die ursprüngliche Gitterstruktur erhalten und für $p=1$ erzeugt man ein vollständig zufälliges Netzwerk. Watts und Strogatz konnten in ihrem Modell zeigen, dass man schon für kleine p ($0 < p \ll 1$) ein *kleines Welt*-Netzwerk erzeugt, das durch einen kleinen Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten und einem großen Wert des Clusterkoeffizienten gekennzeichnet ist.

Die linke, nebenstehende Abbildung zeigt eine Animation eines leicht abgeänderten Watts-Strogatz Modells. Zum Zeitpunkt $t=0$ startet man wiederum mit einer regulären Gitterstruktur (hier $N=80, K=2$). Zu jedem folgenden Zeitpunkt wird nun zufällig eine Kante im Netzwerk ausgewählt, die in zufälliger Weise neu angeordnet wird. Im Laufe der Zeitentwicklung erkennt man, dass sich die Verteilungsfunktion der Knotengrade immer mehr an die binomialverteilte Funktion des zufälligen Netzwerkes annähert ($P(k)$, siehe schwarze Kurve im mittleren Diagramm in der linken Animation). Im oberen Diagramm der linken Animation ist hingegen der normierte Wert des Clusterkoeffizienten (C_t/C_0 , schwarze Punkte) und der normierte Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten des Netzwerkes (l_t/l_0 , rote Punkte) als Funktion des Iterationszeitpunktes t dargestellt. Man erkennt, dass sich l_t schon nach wenigen Iterationsschritten schnell verkleinert - im Bereich $5 < t < 20$ hat das Netzwerk eine klare *Kleine Welt*-Struktur. Für große Zeiten nähert sich das Netzwerk immer weiter einem zufälligen Netzwerk an.

Neben den hier dargestellten Netzwerkeigenschaften von zufälligen und kleinen Welt Netzwerken findet sich eine ausführliche Darstellung in [Chapter 3: Albert-Laszlo Barabasi, Network Science](#). In der Section 3

Jupyter Notebook: SmallWorldNetworks.ipynb

Physik der sozio-ökonomischen Systeme mit dem Computer

(Physics of Socio-Economic Systems with the Computer)

Vorlesung gehalten an der J.W.Goethe-Universität in Frankfurt am Main

(Wintersemester 2020/21)

von Dr.phil.nat. Dr.rer.pol. Matthias Hanauske

Frankfurt am Main 02.11.2020

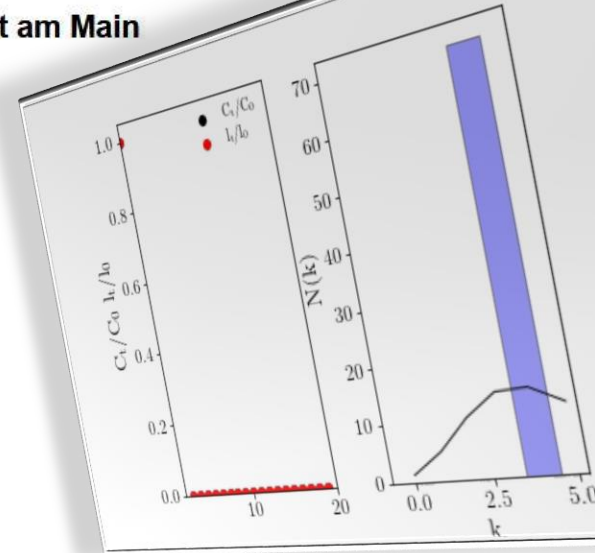
Zweiter Vorlesungsteil:

Kleine Welt Netzwerke (small world networks)

Einführung

Aufgrund ihrer unterschiedlichen Eigenschaften unterscheidet man die folgenden Netzwerk-Klassen: Zufällige Netzwerke (Random Networks: Die einzelnen Kanten bei zufälligen Netzwerke werden von den Knoten (Spielern) nach einem rein zufälligen Muster ausgewählt), Kleine Welt-Netzwerke (Small World Networks, Kleine Welt-Netzwerke zeichnen sich durch einen kleinen Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten des Netzwerkes und einem großen Wert des Clusterkoeffizienten aus), Exponentielle Netzwerke (Exponential Networks) und Skalenfreie Netzwerke (Scale-Free Networks). Bei einigen Modellnetzwerken können analytische Ergebnisse gewonnen werden. Im folgenden betrachten wir die Klasse der kleinen Welt Netzwerke.

"Kleine Welt"-Netzwerke zeichnen sich durch einen kleinen Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten des Netzwerkes und einem großen Wert des Clusterkoeffizienten aus. Ein einfaches Modell welches den Übergang von einem Netzwerk mit regulärer Struktur (Gitter-ähnlicher Struktur) über ein "kleines Welt"-Netzwerk hin zu einem zufälligen Netzwerk veranschaulicht, wurde von Watts und Strogatz im Jahre 1998 vorgestellt. Im einfachsten Fall startet man hierbei mit einem eindimensionalen Gitter-Netzwerk mit N Knoten, wobei jeder Knoten mit seinen K -nächsten Nachbarn (im folgenden Beispiel speziell $K=2$) verbunden ist. Nun löscht man mit der Wahrscheinlichkeit p jede der existierenden Verbindungen (Kanten) und stellt eine neue Kante im Netzwerk in zufälliger Weise her. Für $p=0$ bleibt die ursprüngliche Gitterstruktur erhalten und für $p=1$ erzeugt man ein vollständig verbundenes Netzwerk. Watts und Strogatz konnten in ihrem Modell zeigen, dass man schon für kleine p ($0 < p \ll 1$) ein "kleines Welt"-Netzwerk erzeugt, das durch einen kleinen Wert der durchschnittlichen kürzesten Verbindung zwischen den Knoten und einem großen Wert des Clusterkoeffizienten gekennzeichnet ist. Im folgenden



Network Science

by Albert-László Barabási

Personal Introduction

1. Introduction

2. Graph Theory

3. Random Networks

4. The Scale-Free Property

5. The Barabási-Albert Model

6. Evolving Networks

7. Degree Correlations

8. Network Robustness

9. Communities

10. Spreading Phenomena

Preface

<http://networksciencebook.com/>

Start Reading

English

Русский

Magyar

فارسی

日本語

Section 3.2

The Random Network Model

Wie erzeugt man mittels eines mathematischen Algorithmus ein zufälliges Netzwerk (siehe Box 3.1)

Section 3.4

Degree Distribution

Wie sieht die Verteilungsfunktion der Knotengrade in zufällige Netzwerken aus (siehe Image 3.4 Binomial vs. Poisson Degree Distribution)

Section 3.5

Real Networks are Not Poisson

Vergleich: Real existierende Netzwerke \leftrightarrow Zufällige Netzwerke (siehe Image 3.6 Degree Distribution of Real Networks)

Section 3.6

The Evolution of a Random Network

Relativen Größe des Hubs (grösster verbundener Knotencluster) hängt von dem durchschnittlichen Knotengrad des Netzwerkes ab.
Definition von unterschiedlichen Regimen in zufälligen Netzwerken (subcritical, supercritical, fully connected) (siehe Image 3.7 Evolution of a Random Network)

Section 3.7

Real Networks are Supercritical

Sind real existierende Netzwerke subcritical, supercritical oder fully connected? (siehe Table 3.1 Are Real Networks Connected? und Image 3.9 Most Real Networks are Supercritical)

Section 3.8

Small Worlds

Definition der kleinen Welt Eigenschaft in komplexen Netzwerken "In the language of network science the small world phenomenon implies that the distance between two randomly chosen nodes in a network is short." Mittlerer Abstand zwischen zwei Knoten im Netzwerk $\langle d \rangle$ bestimmt die Eigenschaft von kleinen Welt Netzwerken (siehe Image 3.10 Six Degree of Separation and Image 3.11 Why are Small Worlds Surprising? und Table 3.2 Six Degrees of Separation)

Section 3.9

Clustering Coefficient

Der Clusterkoeffizient in real existierenden und zufälligen Netzwerken (siehe Image 3.13 Clustering in Real Networks und Box 3.9 Watts-Strogatz Model)

Random Networks: a Brief History

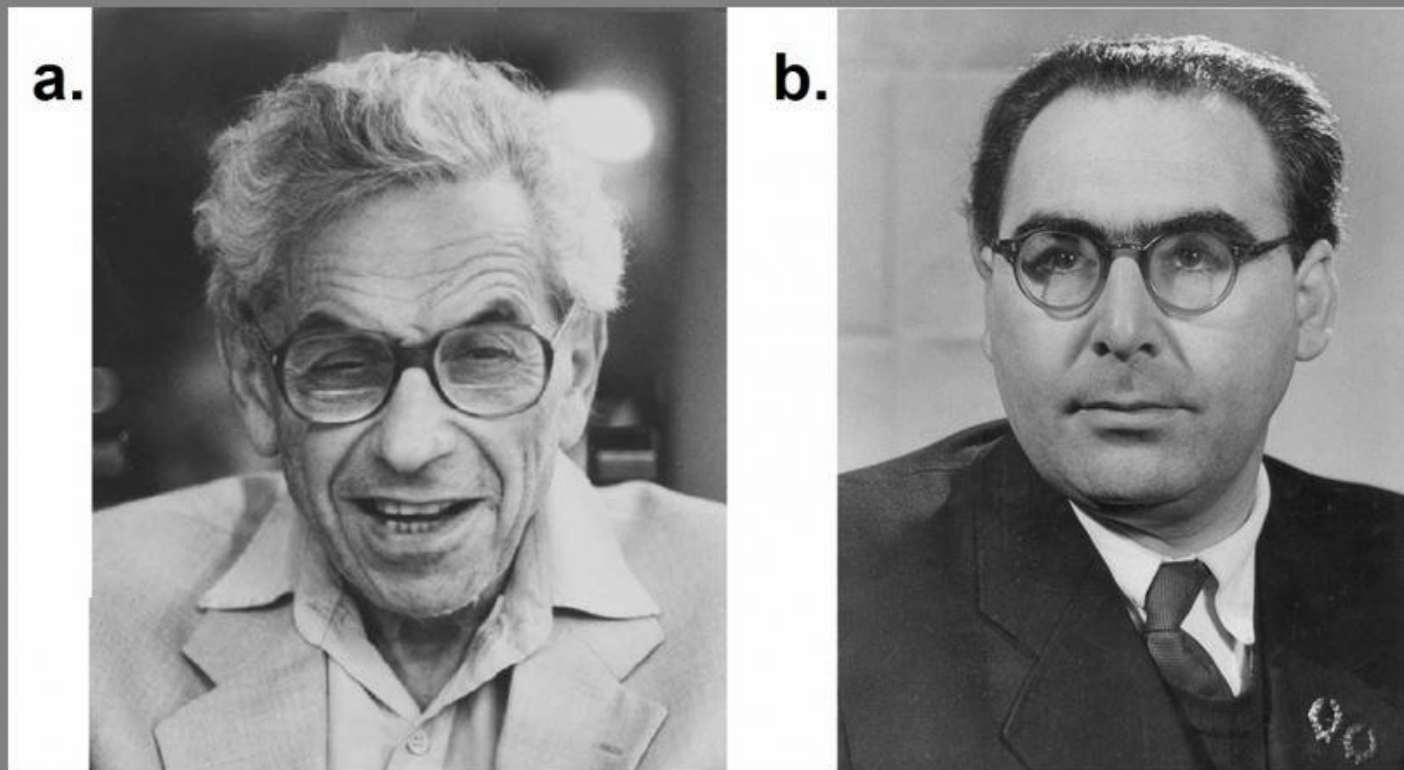


Image 3.2

(a) Pál Erdős (1913-1996)

Hungarian mathematician known for both his exceptional scientific output and eccentricity. Indeed, Erdős published more papers than any other mathematician in the history of mathematics. He co-authored papers with over five hundred mathematicians, inspiring the concept of *Erdős number*. His legendary personality and profound professional impact has inspired two biographies [12, 13] and a documentary [14] ([Video 3.1](#)).

(b) Alfréd Rényi (1921-1970)

Hungarian mathematician with fundamental contributions to combinatorics, graph theory, and number theory. His impact goes beyond mathematics: The Rényi entropy is widely used in chaos theory and the random network theory he co-developed is at the heart of network science. He is remembered through the hotbed of Hungarian mathematics, the Alfréd Rényi Institute of Mathematics in Budapest.

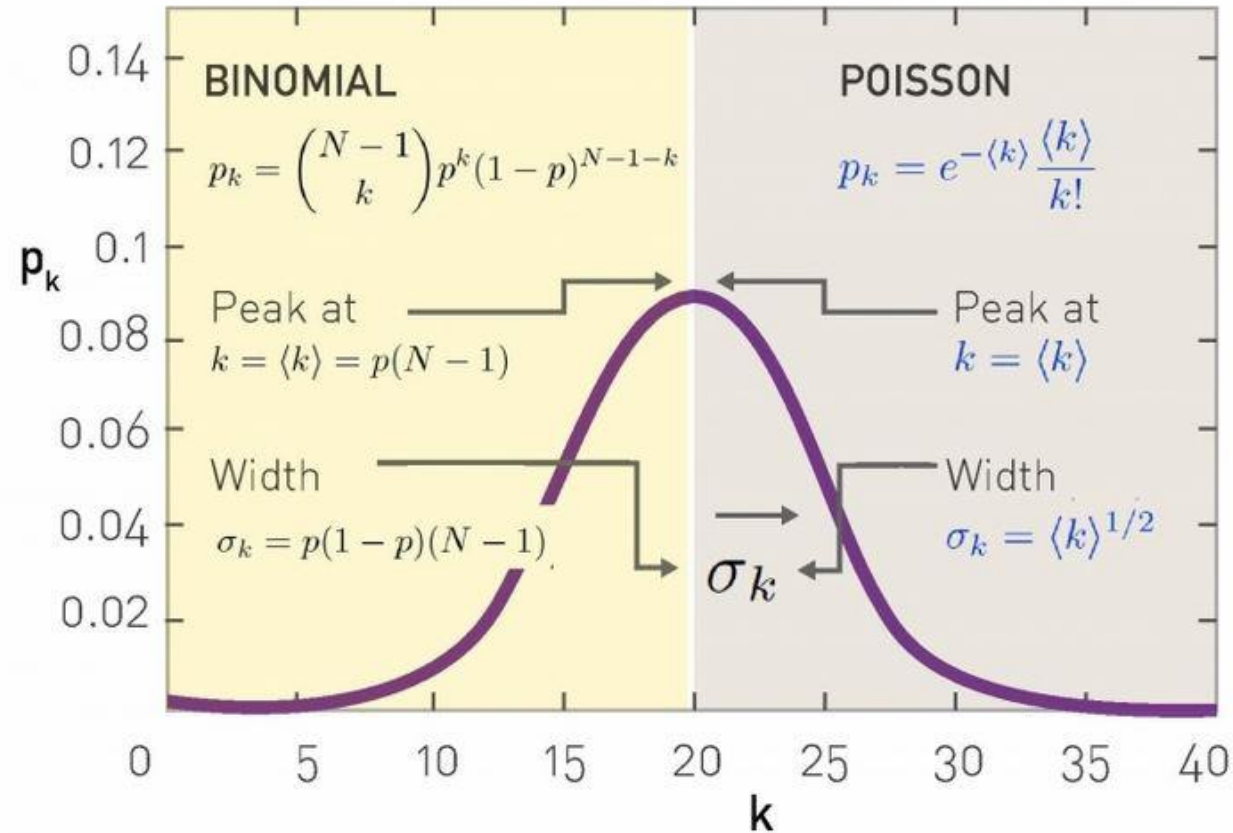


Image 3.4

Binomial vs. Poisson Degree Distribution

The exact form of the degree distribution of a random network is the binomial distribution (left half). For $N \gg \langle k \rangle$ the binomial is well approximated by a Poisson distribution (right half). As both formulas describe the same distribution, they have the identical properties, but they are expressed in terms of different parameters: The binomial distribution depends on p and N , while the Poisson distribution has only one parameter, $\langle k \rangle$. It is this simplicity that makes the Poisson form preferred in calculations.

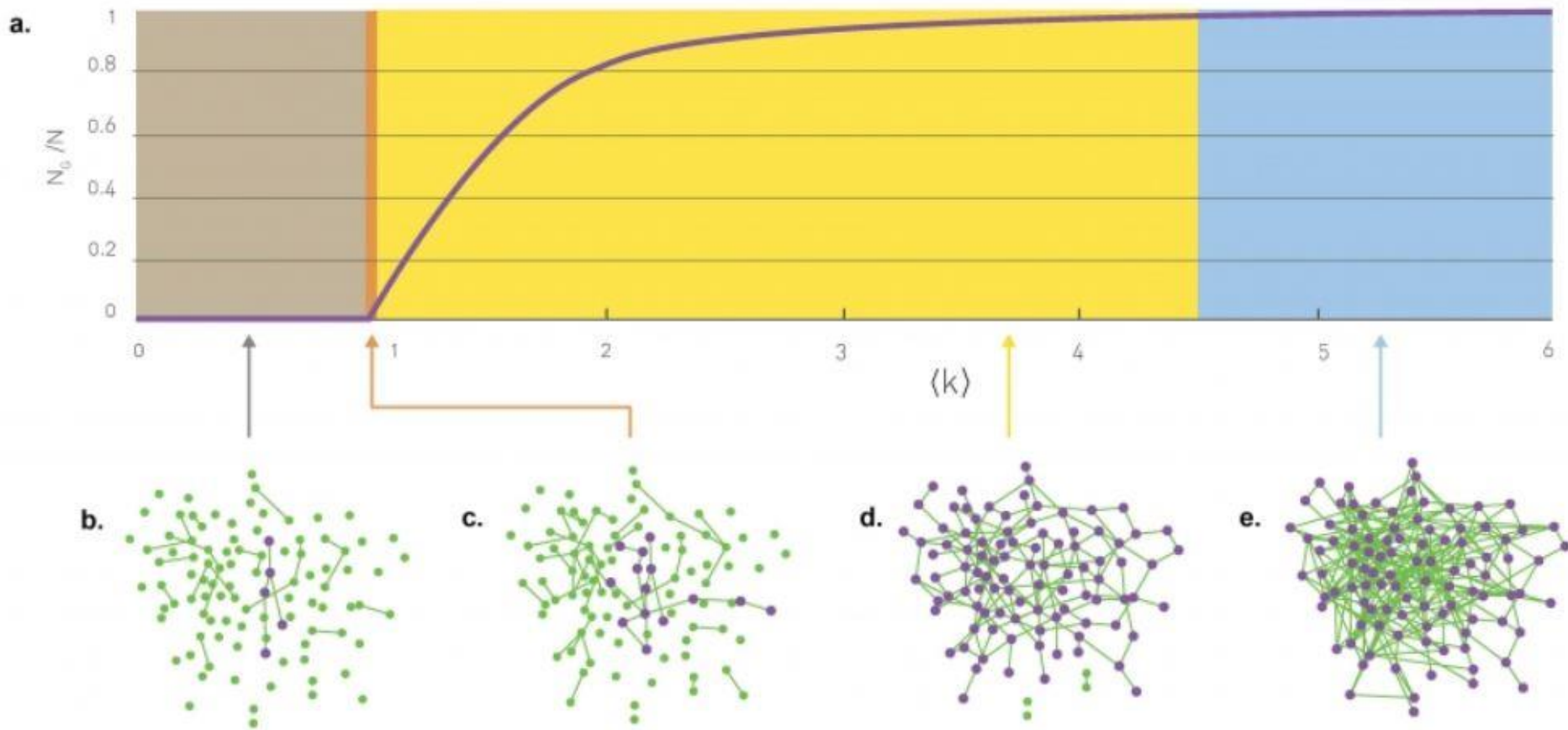


Image 3.7

Evolution of a Random Network

- The relative size of the giant component in function of the average degree $\langle k \rangle$ in the Erdős-Rényi model. The figure illustrates the phase transition at $\langle k \rangle = 1$, responsible for the emergence of a giant component with nonzero N_G
- A sample network and its properties in the four regimes that characterize a random network.

Small Worlds

The *small world phenomenon*, also known as *six degrees of separation*, has long fascinated the general public. It states that if you choose any two individuals anywhere on Earth, you will find a path of at most six acquaintances between them ([Image 3.10](#)). The fact that individuals who live in the same city are only a few handshakes from each other is by no means surprising. The small world concept states, however, that even individuals who are on the opposite side of the globe can be connected to us via a few acquaintances.

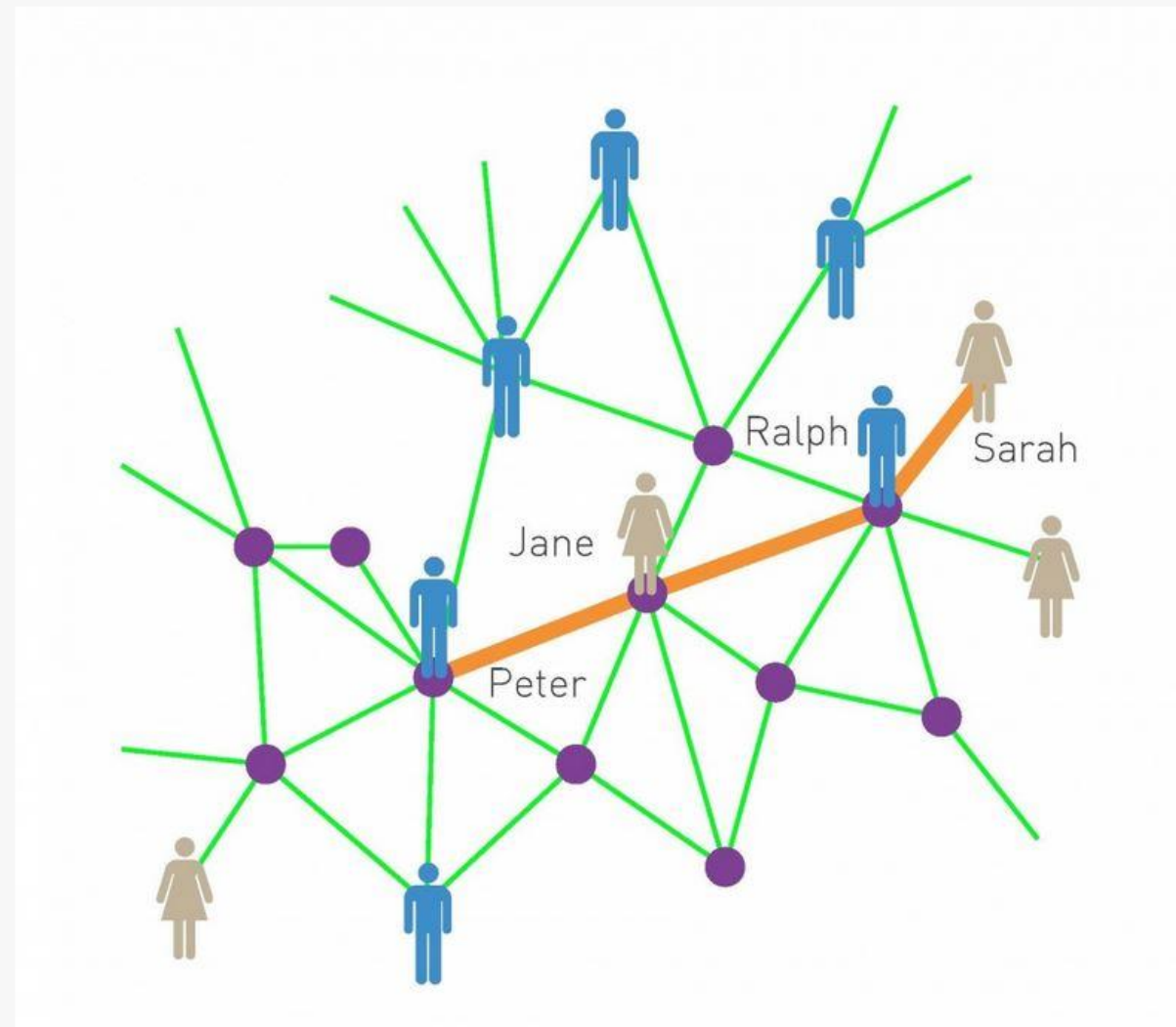
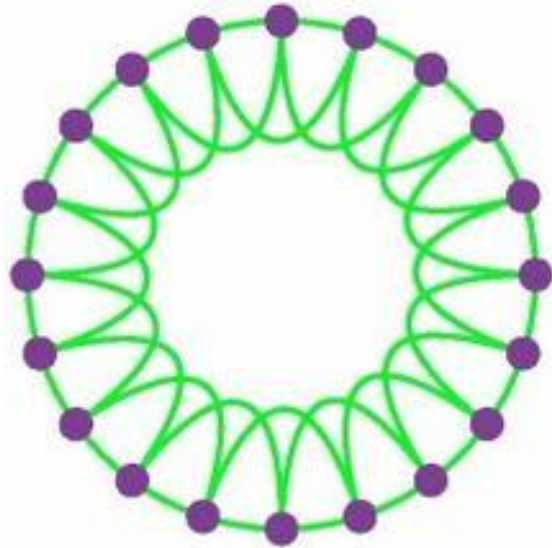


Image 3.10

Six Degree of Separation

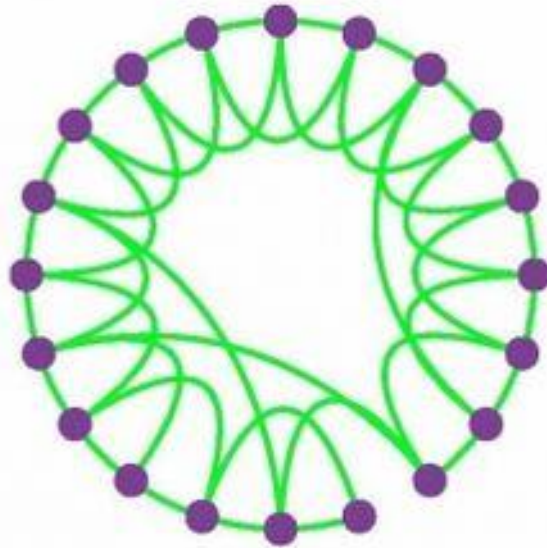
a.

REGULAR



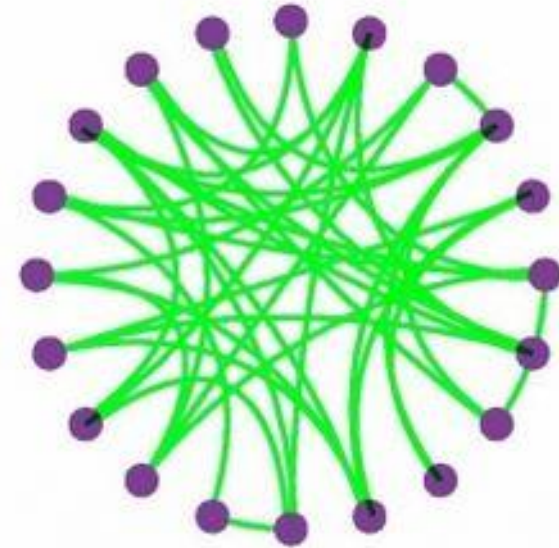
b.

SMALL-WORLD



c.

RANDOM

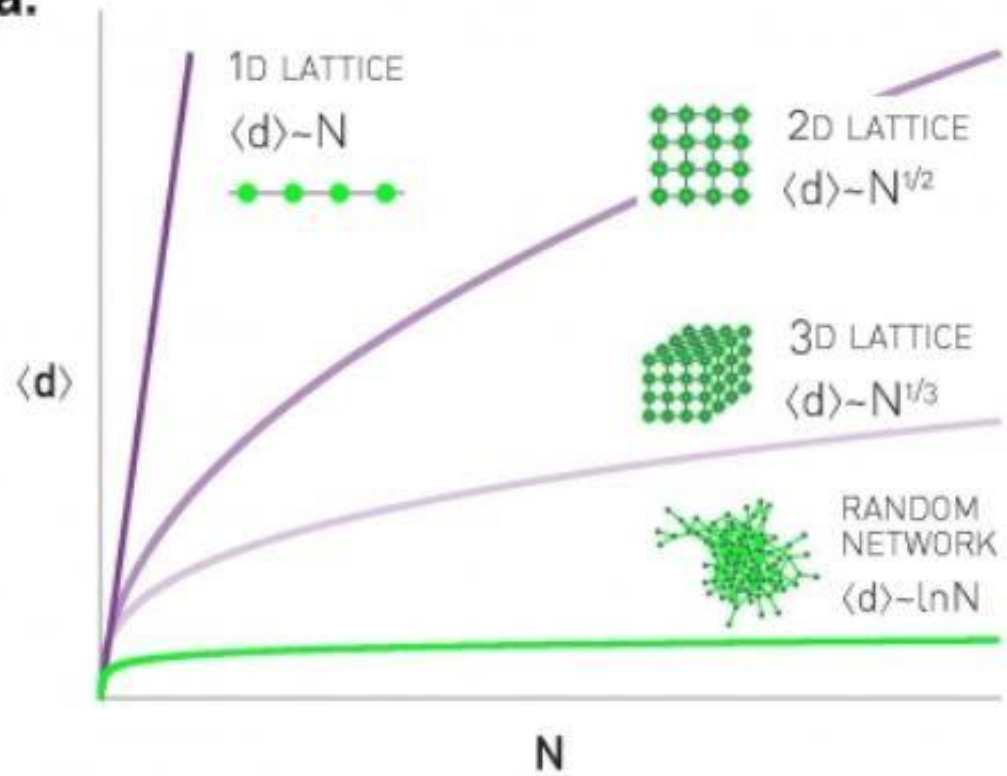
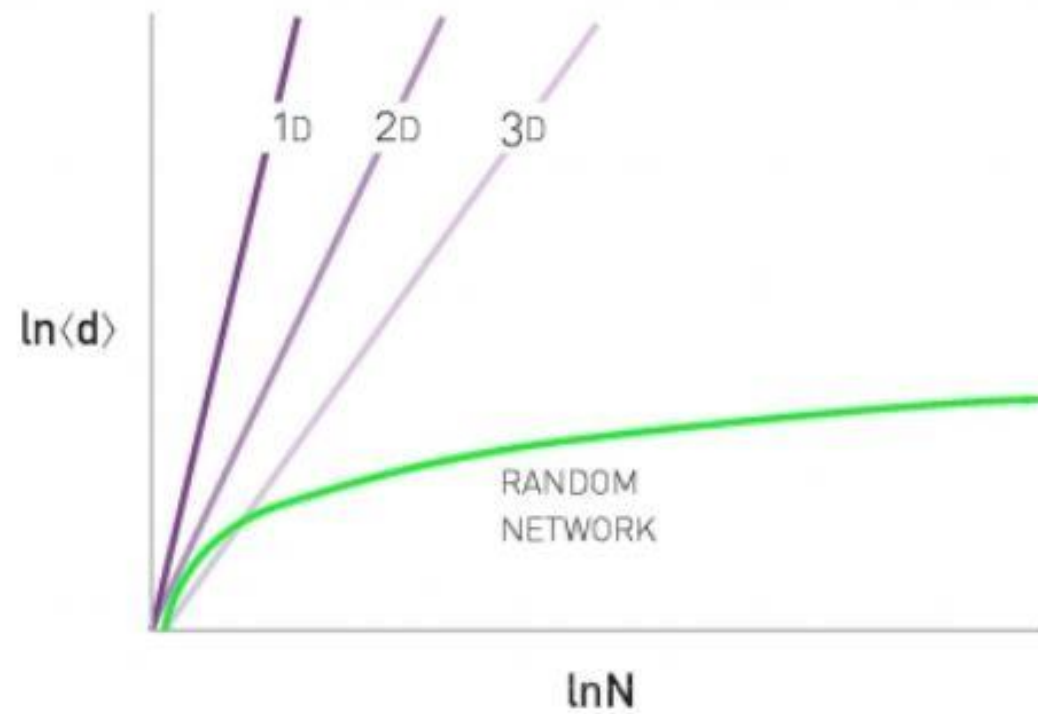


$p = 0$



$p = 1$

Increasing randomness

a.**b.**

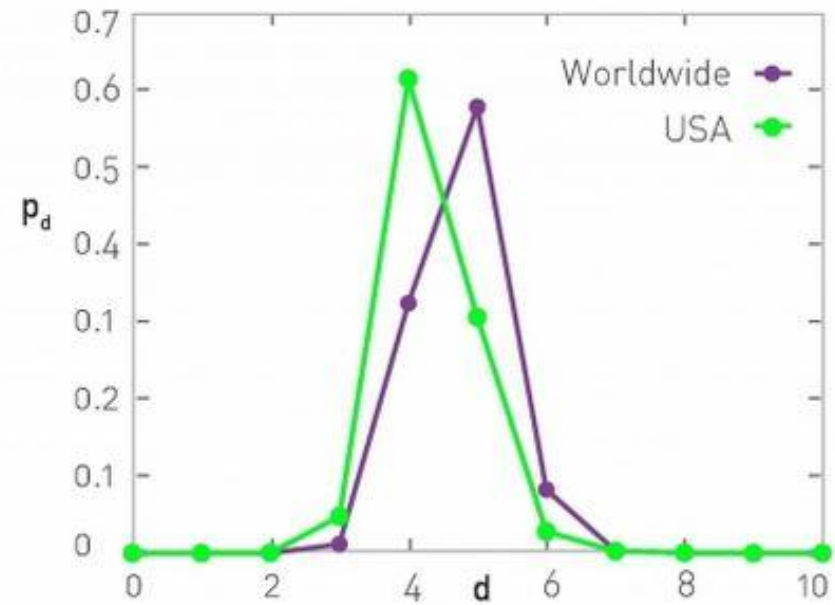
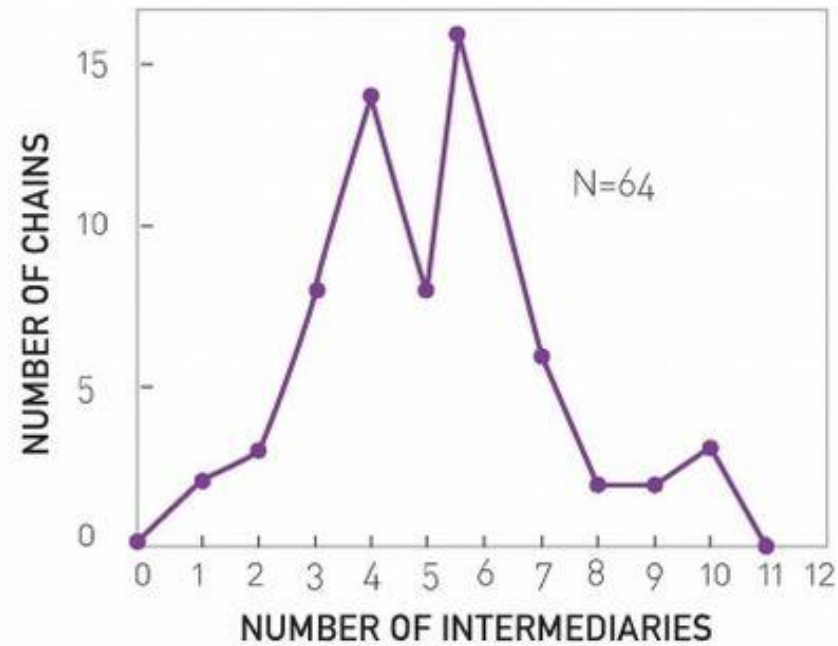


Image 3.12

Six Degrees? From Milgram to Facebook

- In Milgram's experiment 64 of the 296 letters made it to the recipient. The figure shows the length distribution of the completed chains, indicating that some letters required only one intermediary, while others required as many as ten. The mean of the distribution was 5.2, indicating that on average six 'handshakes' were required to get a letter to its recipient. The playwright John Guare renamed this 'six degrees of separation' two decades later. After [25].
- The distance distribution, p_d , for all pairs of Facebook users worldwide and within the US only. Using Facebook's N and L (3.19) predicts the average degree to be approximately 3.90, not far from the reported four degrees. After [18].

19 Degrees of Separation

How many clicks do we need to reach a randomly chosen document on the Web? The difficulty in addressing this question is rooted in the fact that we lack a complete map of the WWW—we only have access to small samples of the full map. We can start, however, by measuring the WWW's average path length in samples of increasing sizes, a procedure called *finite size scaling*. The measurements indicate that the average path length of the WWW increases with the size of the network as [21]

$$\langle d \rangle \approx 0.35 + 0.89 \ln N$$

In 1999 the WWW was estimated to have about 800 million documents [22], in which case the above equation predicts $\langle d \rangle \approx 18.69$. In other words in 1999 two randomly chosen documents were on average 19 clicks from each other, a result that became known as *19 degrees of separation*. Subsequent measurements on a sample of 200 million documents found $\langle d \rangle \approx 16$ [23], in good agreement with the $\langle d \rangle \approx 17$ prediction. Currently the WWW is estimated to have about trillion nodes ($N \sim 10^{12}$), in which case the formula predicts $\langle d \rangle \approx 25$. Hence $\langle d \rangle$ is not fixed but as the network grows, so does the distance between two documents.

The average path length of 25 is much larger than the proverbial six degrees (BOX 3.7). The difference is easy to understand: The WWW has smaller average degree and larger size than the social network. According to (3.19) both of these differences increase the Web's diameter.

Network	N	L	$\langle k \rangle$	$\langle d \rangle$	d_{max}	$\ln N / \ln \langle k \rangle$
Internet	192,244	609,066	6.34	6.98	26	6.58
WWW	325,729	1,497,134	4.60	11.27	93	8.31
Power Grid	4,941	6,594	2.67	18.99	46	8.66
Mobile-Phone Calls	36,595	91,826	2.51	11.72	39	11.42
Email	57,194	103,731	1.81	5.88	18	18.4
Science Collaboration	23,133	93,437	8.08	5.35	15	4.81
Actor Network	702,388	29,397,908	83.71	3.91	14	3.04
Citation Network	449,673	4,707,958	10.43	11.21	42	5.55
E. Coli Metabolism	1,039	5,802	5.58	2.98	8	4.04
Protein Interactions	2,018	2,930	2.90	5.61	14	7.14

Table 3.2

Six Degrees of Separation

The average distance $\langle d \rangle$ and the maximum distance d_{max} for the ten reference networks. The last column provides $\langle d \rangle$ predicted by (3.19), indicating that it offers a reasonable approximation to the measured $\langle d \rangle$. Yet, the agreement is not perfect - we will see in the next chapter that for many real networks (3.19) needs to be adjusted. For directed networks the average degree and the path lengths are measured along the direction of the links.