Ernst Moritz Arndt Universität Greifswald 24. 4. 2009

Komplexe Netzwerke

Gradverteilung der Knoten,
Skalenfreiheit, Potenzgesetz

Dr. Matthias Scholz

www.network-science.org/SS2009.html

1 Begriffe und Definitionen

1.1 Graph

Graph G=(V,E)
V ist eine endliche Menge von Knoten
E ist eine endliche Menge von Kanten

Knoten $v_i \in V; \quad i = 1, ..., N$
verbunden durch Kanten $e \in E$
N ist die Anzahl der Knoten

1.2 Adjazenzmatrix (Nachbarschaftsmatrix)

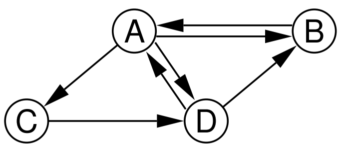
$A = N$ x N quadratische Matrix
$a_{ij} = 1$; falls Kante von Knoten i zum Knoten j existiert
$a_{ij} = 0$; sonst

für einen ungerichteten Graphen ist A symmetrisch: $a_{ij} = a_{ji}, \quad A^T = A$

Ein **gerichteter Graph** oder **Digraph** (***directet graph***) ist ein Graph, in dem alle Kanten eine Richtung haben.

Beispiel (gerichteter Graph):

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$



1.2.1 Gewichtete Graphen

Ein **gewichteter Graph** oder **bewerteter Graph** ist ein Graph, bei dem jeder Kante eine Zahl zugeordnet ist.

- **Abstandsmatrix (Distanzmatrix)** D
  d_{ij} ist der Abstand zwischen Knoten i und j
  Beispiel: Straßennetz (Entfernung zwischen zwei Städten)
- **Ähnlichkeitsmatrix** (z.B.: Korrelationsmatrix) C
  Beispiel: Gen-Gen-Netzwerk (Korrelation der Genexpression)

2 Gradverteilung der Knoten

(*node-degree distribution*)

Grad eines Knoten (*node-degree*): Der Grad k_i eines Knoten i ist die Anzahl seiner Kanten.

Der **Eingangsgrad** (*in-degree*) ist die Anzahl der Kanten, die zu einem Knoten hinführen (gerichteter Graph).
Der Eingangsgrad k^{in} eines Knotens i lässt sich über die Summe der Spalte i in der Adjazenzmatrix berechnen:

$$k_i^{in} = \sum_{j=1}^{N} a_{ji}$$

Der **Ausgangsgrad** (*out-degree*) ist die Anzahl der Kanten, die von einem Knoten wegführen (gerichteter Graph). Der Ausgangsgrad k^{out} eines Knotens i lässt sich über die Summe der Zeile i in der Adjazenzmatrix berechnen:

$$k_i^{out} = \sum_{j=1}^{N} a_{ij}$$

Für ungerichtete Graphen ist der Eingangsgrad eines Knoten gleich dem Ausgangsgrad.

2.1 Gradverteilung

Die Verteilung der Knotengrade ist eines der wichtigsten Merkmale zur Charakterisierung von Netzwerken. Sie zeigt, wie unterschiedlich die Kanten auf die einzelnen Knoten im Netzwerk verteilt sein können.

Reale Netzwerke zeigen in vielen Fälle eine Verteilung, die dem Potenzgesetz folgt (Barabási and Albert, 1999), siehe Abbildung 1. Bei einigen Netzwerken ist aber eine leichte oder auch stärkere Abweichung vom Potenzgesetz zu beobachten, bis hin zu einer Exponentialverteilung, siehe beispielsweise Amaral et al. (2000).

In gerichteten realen Netzwerken sind der Eingangsgrad und der Ausgangsgrad häufig sehr ähnlich verteilt.

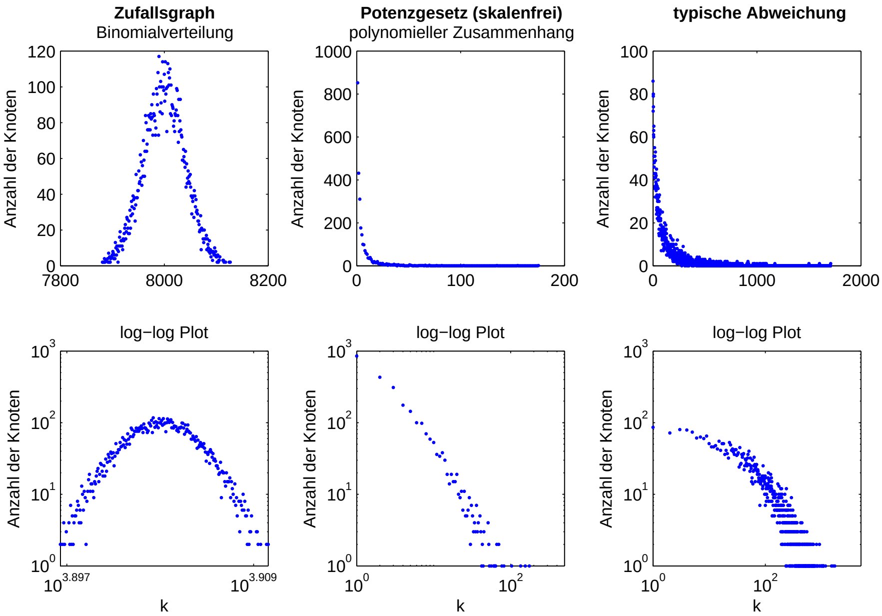


Abbildung 1: **Gradverteilungen.** Links: klassische Annahme, dass die Grade aller Knoten um einen Durchschnittswert schwanken (→ **Zufallsgraphmodell**).
Mitte: typische 'ungleiche' Gradverteilung realer Netzwerke (→ **Potenzgesetz**; → **Attraktivitätsmodell**).
Rechts: öfters zu beobachtende Abweichung vom Potenzgesetz in Form einer leichten oder stärkeren Abschwächung bei kleinen Graden k. Die Abweichung kann so stark sein, dass die Gradverteilung eher der schnell abfallenden Kurve einer Exponentialverteilung als dem Potenzgesetz entspricht.
Die Verteilungen sind jeweils mit linearen Skalen (oben) und mit logarithmischen Skalen (unten) dargestellt.

Graphgenerierungsmodelle: Das Zufallsgraphmodell erzeugt eine Binomialverteilung (bzw. Poissonverteilung) die sich bei großen Wahrscheinlichkeiten p der Glockenkurve einer Gaußverteilung (Normalverteilung) annähert und bei sehr kleinen Wahrscheinlichkeiten p einer Exponentialverteilung, aber nie einer polynomiellen Verteilung (Potenzgesetz). Das Attraktivitätsmodell erzeugt eine Gradverteilung die (nur) dem Potenzgesetz folgt.

Potenzgesetz (*power-law*): Das Potenzgesetz bezeichnet einen polynomiellen Zusammenhang $f(x) = ax^b$, der bei komplexen Netzwerken häufig zwischen den Graden k der Knoten und ihrer Wahrscheinlichkeit $P(k)$ zu beobachten ist:

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

Ein Potenzgesetz ergibt sich aus exponentiellem Wachstum, wenn sowohl die Anzahl

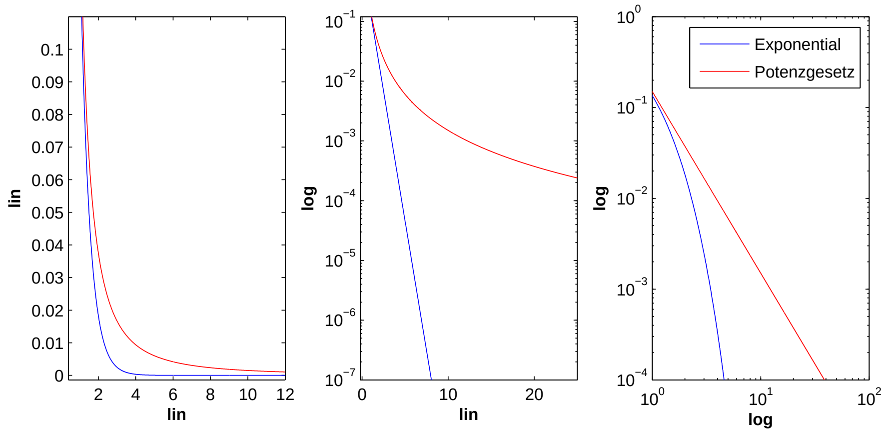


Abbildung 2: **Exponentialverteilung versus Potenzgesetz.** Eine Exponentialverteilung $P(k) \sim e^{-\lambda k}$ fällt schneller ab als eine Verteilung nach dem Potenzgesetz $P(k) \sim k^{-\gamma}$. Die Wahrscheinlichkeit stark vernetzter Knoten ist daher bei einer Exponentialverteilung viel geringer als beim Potenzgesetz.

(Knoten) als auch die Größe (Grad der Knoten) sich exponentiell ändert, siehe auch Abbildung 2.

Bei einer Gradverteilung nach dem Potenzgesetz sind auch extrem stark vernetzte Knoten nicht unwahrscheinlich.

Skalenfreie Netzwerke (*scale-free networks*): Für Netzwerke mit der Eigenschaft des polynomiellen Zusammenhangs (Potenzgesetz) zwischen Wahrscheinlichkeit $P(k)$ und Grad k wurde von Albert-László Barabási auch die Bezeichnung 'skalenfrei' eingeführt.

Netzwerke mit einer Gradverteilung, die dem Potenzgesetz folgt, werden als skalenfreie Netzwerke bezeichnet.

Skalenfreiheit (Skaleninvarianz) spielt eine große Rolle in der Chaostheorie und bedeutet, dass auf unterschiedlichen Skalen immer wieder die gleichen Charakteristiken (Muster) auftreten (Selbstähnlichkeit).

Skalenfreiheit, Skaleninvarianz oder auch Skalenunabhängigkeit beschreibt die Eigenschaft eines Zustands, Vorgangs oder Verhältnisses, bei dem auch bei Veränderung der Betrachtungsgrößen (Skalierung) die Charakteristik weitestgehend erhalten bleibt: $f(x) \sim f(cx)$.

Literatur

Amaral, L.A.N., Scala, A., Barthélémy, M., Stanley, H.E. Classes of small-world networks. *PNAS*, 97(21):11149–11152, 2000.

Barabási, A., Albert, R. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.