



# Komplexe Netzwerke

## Erdős–Rényi Zufallsgraph-Modell

Dr. Matthias Scholz

[www.network-science.org/SS2009.html](http://www.network-science.org/SS2009.html)

### Modell

Mit Computersimulationen lassen sich Modelle empirisch überprüfen.

Idee  $\Rightarrow$  Algorithmus und Computersimulation  $\Rightarrow$  Modell (Graph)

Hier: Modelle, die das Entstehen von Netzwerken erklären sollen.

Idee A: Kanten sind zufällig  $\rightarrow$  Zufallsgraph-Modell

Idee B: neue Knoten verbinden sich bevorzugt mit stark vernetzten Knoten  $\rightarrow$  Attraktivitätsmodell

Überprüfung: Stimmen die Eigenschaften (Skalenfreiheit, Kleine-Welt-Phänomen) des generierten Graphen mit denen realer Netzwerke überein?

Der Nutzen eines Modells ist zum einen ein besseres Verständnis realer Vorgänge und zum anderen die Möglichkeit der Vorhersage und Analyse: Wie verhält sich das Netzwerk oder ein System unter bestimmten Bedingungen (unter verschiedenen Modell-Parametern)?

### 3 Erdős–Rényi Zufallsgraph-Modell (ER-Modell)

(*Erdős–Rényi random graph model*)

- Idee: Verbindungen zwischen Knoten sind zufällig.
- **Graph-Generierungsmodell**
- Der Zufallsgraph ist ein klassisches Modell in der Graphentheorie. Als erstes von Anatol Rapoport beschrieben, dann ausführlich von Paul Erdős und Alfréd Rényi analysiert (Erdős and Rényi, 1960).
- **Zufallsgraphen sind nicht skalenfrei.** Das Zufallsgraph-Modell kann zwar die in realen Netzwerken vereinzelt vorkommende exponentielle Gradverteilung reproduzieren, nicht aber die häufig zu beobachtende polynomielle Gradverteilung (Potenzgesetz) bei skalenfreien Netzwerken.

#### 3.1 Algorithmus

1. Initialisierung mit  $N$  isolierten Knoten

2. Jedes Knotenpaar wird mit der Wahrscheinlichkeit  $p$  durch eine Kante verbunden.

Alle Kanten werden gleich wahrscheinlich und unabhängig voneinander erzeugt.

### 3.2 Binomialverteilung

Die Grade der Knoten im Zufallsgraphen sind binomialverteilt.

Die Wahrscheinlichkeit  $P(k)$ , dass ein Knoten den Grad  $k$  hat (d.h.  $k$  Kanten besitzt) ist:

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

$p$  ist die Wahrscheinlichkeit, dass eine bestimmte Kante existiert.  
 $(1-p)$  ist die Wahrscheinlichkeit, dass eine bestimmte Kante *nicht* existiert.  
 $p^k$  ist die Wahrscheinlichkeit, dass  $k$  bestimmte (ausgewählte) Kanten existieren.

$p^k(1-p)^{n-k}$  ist die Wahrscheinlichkeit, dass *exakt*  $k$  ausgewählte Kanten existieren und die anderen  $n-k$  Kanten nicht existieren

$P(k)$  ist die Wahrscheinlichkeit, dass von allen möglichen Kanten genau  $k$  Kanten existieren, egal in welcher Anordnung, dazu brauchen wir die Anzahl der möglichen Varianten, die durch den Binomialkoeffizienten gegeben ist:

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$  Binomialkoeffizient: Wieviel verschiedene Möglichkeiten gibt es,  $k$  Kanten aus einer Menge von  $n$  Kanten auszuwählen.  
( $n = N - 1$ ,  $N$  ist die Anzahl aller Knoten)

Die Form der Binomialverteilung hängt sowohl von der Wahrscheinlichkeit  $p$  der Kanten als auch von der Anzahl  $N$  der Knoten ab. Bei großer Wahrscheinlichkeit  $p$  oder großer Anzahl  $N$  nähert sich die Binomialverteilung einer Gaußverteilung an und bei kleiner Wahrscheinlichkeit  $p$  oder kleiner Anzahl  $N$  einer Exponentialverteilung, siehe Abbildung 1.

Die Form der Binomialverteilung bleibt gleich, wenn  $pn$  konstant bleibt,  $\lambda = pn$  ist der Mittelwert der Binomialverteilung.

### 3.3 Poissonverteilung

Die Binomialverteilung kann durch eine Poissonverteilung approximiert werden.

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} ; \quad \lambda = pn$$

Für große  $\lambda$  hat die Poissonverteilung die Form einer Normal- bzw. Gauß-Verteilung. Die Poissonverteilung ist die Grenzverteilung der Binomialverteilung für kleine  $p$  und große  $n$  ( $p \rightarrow 0$ ,  $n \rightarrow \infty$ ).

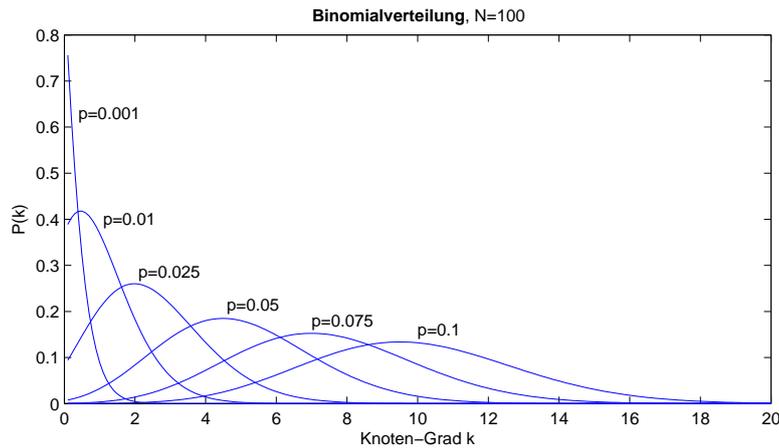


Abbildung 1: **Binomialverteilung.** Die Binomialverteilung hat bei großen Wahrscheinlichkeiten  $p$  die Form einer Gaußverteilung. Mit  $p \rightarrow 0$  (bei konstanter Knotenzahl  $N$ ) wird die Verteilung zunehmend schief bis hin zu einer Exponentialverteilung. Die Exponentialverteilung zeigt sich im Zufallsgraphen aber erst, wenn der Graph in kleine Teilgraphen zerfällt ( $Np < 1$ , siehe  $\rightarrow$  Phasenübergang).

### 3.4 Adaptiertes Modell (beliebige Gradverteilung)

Mit einem modifizierten Algorithmus (Newman et al., 2001) ist es möglich, Zufallsgraphen mit jeder beliebigen Gradverteilung zu generieren, also auch mit einer Gradverteilung, die dem Potenzgesetz folgt. Dazu gibt man die Verteilung explizit vor.

Aus praktischen Gründen wählt man statt einer Gradverteilung eine Grad-Sequenz  $k_1, k_2, k_3, k_4, \dots, k_N$  für die einzelnen Knoten  $i = 1, \dots, N$ . Jedem Knoten  $i$  wird also ein individueller Grad  $k_i$  zugeordnet. Entsprechend den vorgegebenen Graden werden die Knoten dann zufällig verbunden.

Mit solchem universellen Modell ist es möglich, Zufallsgraphen mit beliebiger Gradverteilung zu generieren mit denen man dann verschiedene Analysen durchführen kann.

Allerdings, ist ein extrem universelles Modell nicht sehr hilfreich, um zu verstehen, wie reale Netzwerke wirklich entstehen oder warum reale Netzwerke bestimmte Gradverteilungen haben.

### 3.5 Phasenübergang

Mit zunehmender Wahrscheinlichkeit  $p$  geht der Graph von *isolierten* Knoten über zu einem *zusammenhängenden* Graphen bis hin zu einem *vollständig* vernetzten Graphen. Der Übergang von isolierten Knoten oder kleinen Teilgraphen zu einem zusam-

menhängenden Graphen findet relativ plötzlich bei  $Np = 1$  statt. Es formt sich ein großer zusammenhängender Teilgraph (*giant component*), der den Hauptteil der Knoten enthält. Da sich dieser Vorgang ähnlich verhält wie beispielsweise das plötzliche Erstarren von Wasser zu Eis bei sinkender Temperatur wird er häufig als *Phasenübergang* bezeichnet.

Der Phasenübergang hat eine wichtige Bedeutung für den Zusammenhalt eines Netzwerkes. Wenn zu viele Knoten entfernt werden oder ausfallen, dann wird bei  $Np = 1$  der Phasenübergang erreicht und das Netz zerfällt relativ plötzlich in viele kleine Teilnetze oder isolierte Knoten. Netzweite Kommunikation oder Interaktion ist dann nicht mehr möglich ( $\rightarrow$  Robustheit).

Ein **isolierter Knoten** ist ein Knoten, an dem keine Kante anliegt.

Ein **zusammenhängender Graph** ist ein Graph, in dem es von jedem Knoten zu jedem anderen einen Weg (Kantenzug) gibt, anderenfalls besteht er aus zwei oder mehr Teilgraphen (Komponenten).

Ein **vollständiger Graph** (*complete graph*) ist ein Graph, in dem jeder Knoten mit jedem anderen Knoten direkt verbunden ist. Er besitzt alle  $\frac{N(N-1)}{2}$  möglichen Kanten. Ein vollständiger Teilgraph wird auch als Clique bezeichnet.

## Literatur

Erdős, P., Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61, 1960.

Newman, M.E.J., Strogatz, S.H., Watts, D.J. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64(2):026118, 2001. doi: 10.1103/PhysRevE.64.026118. URL <http://link.aps.org/doi/10.1103/PhysRevE.64.026118>.