



TECHNISCHE
UNIVERSITÄT
WIEN

Geometrische Zufallsgraphen und ihre Anwendungen

Diplomarbeit

Ausgeführt am Institut für
Diskrete Mathematik und Geometrie
der Technischen Universität Wien

unter der Anleitung von
Univ.Prof. Dipl.-Ing. Dr.techn. Michael Drmota

durch

Daniel Rainer, B.Sc.
Matrikelnummer: 01027549
Römergasse 67/5
1160, Wien

Ort, Datum

Unterschrift (Student)

Unterschrift (Betreuer)

Kurzfassung

Die Graphentheorie ist ein Teilgebiet der Mathematik, welches sich aus realen Fragestellungen im 18. Jahrhundert entwickelte und heutzutage Lösungen für viele unterschiedliche Anwendungsgebiete bietet. Etwas jünger ist die Theorie der geometrischen Zufallsgraphen, welche im 20. Jahrhundert von Gilbert ([13]) zur Untersuchung von Kommunikationsnetzwerken eingeführt wurde - hier sind die Graphen nicht mehr ausschließlich rein mathematische Objekte, sondern Objekte in einem geometrischen Raum, in welchen die Lage der Knoten und die Distanz der Knoten eine große Rolle spielen.

In dieser Arbeit werden in Kapitel 2 geometrische Zufallsgraphen eingeführt und grundlegende asymptotische Eigenschaften des Graphen $G(\mathcal{X}_n, r_n)$ für bestimmte Folgen $(r_n)_{(n \geq 1)}$ untersucht. Diese Analysen spielen auch für reale Anwendungen eine große Rolle, da man bei der Verwendung von geometrischen Zufallsgraphen oft bestimmte Eigenschaften, wie beispielsweise den Zusammenhang des Graphen garantieren muss, damit er sinnvoll als Modell für die Anwendungen verwendet werden kann.

In Kapitel 3 werden nun verschiedene, ausgewählte Anwendungen von geometrischen Zufallsgraphen in der Untersuchung von ad-hoc WiFi Netzwerke erläutert und verschiedene bekannte Resultate zusammengefasst:

Betrachtet man Wurm-Epidemien in ad-hoc WiFi Netzwerken, so können diese Epidemien mit einem auf geometrische Zufallsgraphen basierenden Modell analysiert werden, wobei diese Resultate, welche in Abschnitt 3.3 präsentiert werden, helfen sollen, die Verbreitung von Computerwürmern in solch Netzwerken zu verstehen als auch im Weiteren zu verhindern.

Anschließend wird in Abschnitt 3.4 noch ein Routing-Protokoll für die Datenübertragung in ad-hoc WiFi Netzwerken vorgestellt, wobei hier für die Analyse ebenfalls auf geometrische Zufallsgraphen zurückgegriffen wird.

Abstract

Graph theory is a subdomain of mathematics, which arose out of real questions and problems in the 18th century. Today it offers different solutions for various problems and areas. The theory of random geometric graphs, which is younger, was introduced in the 20th century by Gilbert ([13]) to investigate the theory of communication networks. Random geometric graphs aren't anymore only abstract mathematical objects, but objects in a geometric space where the locations of the vertices and the distance of the vertices to each other has an important role.

In this thesis random geometric graphs get introduced and motivated in chapter 2 and some asymptotic properties of the graph $G(\mathcal{X}_n, r_n)$ are given for specific limiting regimes $(r_n)_{(n \geq 1)}$. These properties are important in real applications as well, as for simulations various properties of the random geometric graphs have to be ensured to use them as a useful model for the desired applications.

In chapter 3 some applications of random geometric graphs in the area of ad-hoc WiFi networks are explained and given:

Is there a worm-epidemy in an ad-hoc network, it's possible now to analyse these epidemics by a model based on random geometric graphs. The results, which are given in section 3.3 should help to understand the propagation of worms and prevent propagation of them.

In section 3.4 a routing protocol for communication in an ad-hoc network is presented - the analysis of this protocol is done with a model based on random geometric graphs as well.

Inhaltsverzeichnis

1	Grundlagen und Voraussetzungen	1
1.1	Grundlagen der Graphentheorie	1
1.2	Grundlagen der Wahrscheinlichkeitstheorie	4
1.3	Weitere mathematische Grundlagen	10
2	Geometrische Zufallsgraphen	12
2.1	Definition und Anschauung	12
2.1.1	Kategorien geometrischer Zufallsgraphen	13
2.2	Eigenschaften geometrischer Zufallsgraphen	14
2.2.1	Poissonisierung	15
2.2.2	Minimaler Knotengrad	16
2.2.3	Maximaler Knotengrad	32
2.2.4	Clique-Zahl	36
2.2.5	Chromatische Zahl	40
2.2.6	Zusammenhang	41
2.3	Betweenness-Zentralität in geometrischen Zufallsgraphen	47
2.3.1	Berechnung der Betweenness-Zentralität	48
2.3.2	Berechnung der Betweenness-Zentralität in dichten geometrischen Zufallsgraphen	49
2.3.3	Weitere Zentralitäts-Maße	54
2.4	Ein ähnliches Modell: Erdős-Rényi Zufallsgraphen	57
3	Ad-hoc WiFi Netzwerke	60
3.1	Definition	60
3.2	Geometrische Zufallsgraphen als Modell für ad-hoc WiFi Netzwerke . . .	61
3.3	Ausbreitung von Wurm-Epidemien in ad-hoc WiFi Netzwerken	64
3.3.1	Das Modell	64
3.3.2	Aufbau der Monte-Carlo Simulationen	67
3.3.3	Analyse der Ergebnisse der Monte-Carlo Simulationen	68
3.4	Übermittlungsprotokolle für ad-hoc WiFi Netzwerke	73
3.4.1	Greedy Geographic Routing	73

1 Grundlagen und Voraussetzungen

In diesem Kapitel werden grundlegende Begriffe und Resultate aus der Graphen- und Wahrscheinlichkeitstheorie wiederholt und zusammengefasst, auf welche im weiteren Verlauf der Arbeit zurückgegriffen wird.

1.1 Grundlagen der Graphentheorie

Definition 1.1: Ungerichteter Graph (Diestel [8])

Ein **ungerichteter Graph** $G(V, E)$ ist ein Tupel, bestehend aus einer Knotenmenge V und einer Kantenmenge $E \subseteq V \times V$.

Bemerkung. *Ungerichtet* in diesem Sinne, dass die Menge $e = \{u, v\} \in E$ nicht geordnet ist und somit keine Richtung zwischen den Knoten vorgegeben ist. Handelt es sich bei den Kanten um geordnete Paare (u, v) , so spricht man von einem **gerichteten Graphen**.

Beispiel 1.1. Das Wiener U-Bahn Netz kann durch einen ungerichteten Graphen dargestellt werden: die U-Bahn Stationen sind die Knoten des Graphen, während die Verbindungen zwischen den Stationen als Kanten dargestellt werden. In einer Visualisierung des Graphen könnte man dementsprechend sofort sehen, von welcher Station man zu welcher Station gelangen kann.

Definition 1.2: (Diestel [8])

Sei $G = (V, E)$ ein ungerichteter Graph.

- eine **Schleife** ist eine Kante von einem Knoten zu sich selbst.
- G wird **einfach** genannt, wenn er keine Mehrfachkanten und Schleifen enthält.
- eine Kante $e = \{u, v\}$ heißt **inzident** zu den Knoten u und v .
- zwei Knoten u und v heißen **adjazent**, wenn es eine Kante $e = \{u, v\}$ gibt.

Im weiteren Kontext dieser Arbeit werden nur noch einfache Graphen betrachtet, das heißt es werden nur noch Graphen ohne Schleifen betrachtet.

Für die Analyse der geometrischen Zufallsgraphen in den Abschnitten 2.2.2 und 2.2.3 sind die folgenden Begriffe notwendig:

Definition 1.3: (Diestel [8])

Sei $G = (V, E)$ ein ungerichteter Graph und $v \in V$ ein Knoten von G .

Der **Grad** $d_G(v)$ von v ist die Anzahl der zu v adjazenten Knoten.

Mit $\delta(G) := \min\{d_G(v) \mid v \in V\}$ wird der **Minimalgrad von G** bezeichnet, analog dazu mit $\Delta(G) := \max\{d_G(v) \mid v \in V\}$ der **Maximalgrad von G** .

Mit diesen Definitionen kann bereits der folgende bekannte Satz formuliert werden, welcher auch als Handschlaglemma bekannt ist:

Satz 1.1 (Handschlaglemma - Diestel [8])

Sei $G = (V, E)$ ein einfacher Graph. Dann gilt:

$$\sum_{v \in V} d_G(v) = 2 \cdot |E|$$

Definition 1.4: (Diestel [8])

Ein Graph heißt **vollständig**, wenn es zwischen allen Knoten u, v eine Kante $e = \{u, v\}$ gibt.

Ein **Pfad** P ist eine Folge von Knoten $x_1, x_2, \dots, x_n \in V$, wobei die Knoten x_i paarweise verschieden sind und jeweils durch eine Kante verbunden sind. x_1 und x_n sind dabei die Endpunkte des Pfades, welche von P verbunden werden.

Die Anzahl der Kanten eines Pfades wird als **Länge** des Pfades bezeichnet.

Ein Graph $G = (V, E)$ heißt **zusammenhängend**, wenn je zwei Knoten $u, v \in V$ mit einem Pfad in G verbunden sind.

Durch das Konzept von Pfaden in einem Graphen kann man einen Abstandsbegriff zwischen Knoten $u, v \in V$ eines Graphen $G = (V, E)$ einführen, ohne dabei notwendigerweise Distanzbegriffe eines geometrischen Raums zu verwenden.

Definition 1.5: Abstand zwischen zwei Punkten v und w in einem Graphen (Diestel [8])

Sei $G = (V, E)$ ein ungewichteter Graph, $v, w \in V$ Knoten in G .

Der **Abstand** $d_G(v, w)$ zweier Punkte v und w in einem Graphen $G = (V, E)$ wird als die minimale Länge aller Pfade zwischen v und w definiert.

In Abschnitt 2.2.4 wird die **Clique-Zahl** eines geometrischen Zufallsgraphen analysiert. Hierfür sind folgende Definitionen notwendig:

Definition 1.6: Teilgraph (Diestel [8])

Seien $G = (V, E)$ und $G' = (V', E')$ Graphen. Gilt $V' \subset V$ und $E' \subset E$, so nennt man G' einen **Teilgraphen** von G .

Definition 1.7: Clique und Clique-Zahl (Penrose [25])

Ein vollständiger Teilgraph eines Graphen G wird **Clique** genannt.

Die **Clique-Zahl** $\mathcal{C}$ oder $\mathcal{C}(G)$ eines Graphen G ist die Anzahl der Knoten der größten Clique von G .

Für einen Beweis in Abschnitt 2.2.2 wird der Begriff einer k -separierte Menge benötigt:

Definition 1.8: k -separierte Menge (Penrose [25])

Sei $G = (V, E)$ ein Graph und $U \subseteq V$ eine Teilmenge der Knotenmenge V . Dann heißt U **k -separierte Menge**, wenn $U \neq \emptyset$ und höchstens k Knoten von $V \setminus U$ adjazent zu Knoten in U sind.

Bemerkung. Für eine Menge $U \neq \emptyset \subseteq \mathcal{X} \subset \mathbb{R}^d$ mit $\mathcal{X}$ endlich gilt, dass U k -separiert ist, genau dann wenn $\mathcal{X} [U + B(0, r) \setminus U] < k$ ist.

1.2 Grundlagen der Wahrscheinlichkeitstheorie

Viele Aussagen über Eigenschaften von geometrischen Zufallsgraphen sind Aussagen über das asymptotische Verhalten im Unendlichen - daher werden an dieser Stelle einige wahrscheinlichkeitstheoretische Konvergenzbegriffe zusammengefasst und wiederholt.

Definition 1.9: fast sichere Konvergenz (Kusolitsch [20])

Sind $X, X_1, X_2, \dots$ Zufallsvariablen in einem Wahrscheinlichkeitsraum (Ω, Σ, P) , so spricht man von **fast sicherer Konvergenz** der X_n gegen X dann, wenn die Wahrscheinlichkeit für das Ereignis $\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$ gleich 1 ist.

Man schreibt dafür $X_n \rightarrow X$ fast sicher.

Bemerkung. Bei der fast sicheren Konvergenz gilt die Aussage demnach bis auf eine P -Nullmenge.

Definition 1.10: Konvergenz in Wahrscheinlichkeit (Kusolitsch [20])

Eine Folge von Zufallsvariablen X_n **konvergiert gegen X in Wahrscheinlichkeit**, genau dann wenn $\forall \epsilon > 0 : \mathbb{P}[|X_n - X| \geq \epsilon] \rightarrow 0$ für $n \rightarrow \infty$ gilt.

Satz 1.2 (1. Borel-Cantelli Lemma (Kusolitsch [20]))

Sei $(A_n)_{(n \geq 1)}$ eine Folge von Ereignissen und es gilt $\sum_{i=1}^{\infty} \mathbb{P}[A_i] < \infty$. Dann folgt mit Wahrscheinlichkeit 1:

$$\mathbb{P} \left[\limsup_{i \in \mathbb{N}} A_i \right] = 0$$

Oder alternativ ausgedrückt: das Ereignis A_i tritt nur für endlich viele $i \in \mathbb{N}$ ein.

Viele Beweise in Abschnitt 2.2 können mit Hilfe von Poissonierung (siehe Abschnitt 2.2.1) geführt werden. Hierfür ist der Begriff eines **Poisson-Prozess** notwendig:

Definition 1.11: Poisson-Verteilung (Kusolitsch [20])

Die **Poisson-Verteilung** ist eine diskrete Wahrscheinlichkeitsverteilung, welche durch einen reellwertigen Parameter $\lambda > 0$ definiert wird.

$$P_\lambda(k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \text{ für } k \in \{0, 1, \dots\}$$

Bemerkung. Betrachtet man eine Poisson-verteilte Zufallsvariable X mit Parameter λ , so wird durch λ der Erwartungswert $\mathbb{E}(X)$ und die Varianz $\mathbb{V}(X)$ beschrieben, wie folgende Rechnungen zeigen:

$$\mathbb{E}(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

Durch eine ähnliche Rechnung kann $\mathbb{E}(X^2) = \lambda^2 + \lambda$ gezeigt werden, wodurch mit dem *Steiner'schen Verschiebungssatz* (Kusolitsch [20]) folgt:

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Definition 1.12: Stochastischer Prozess (Kusolitsch [20])

Ein **stochastischer Prozess** auf einem Wahrscheinlichkeitsraum (Ω, Σ, P) ist eine Familie von Zufallsvariablen $\{X_i \mid i \in I\}$.

Definition 1.13: Poisson-Prozess (Penrose [25])

Sei $g: \mathbb{R}^d \rightarrow [0, \infty)$ eine beschränkte, messbare Funktion. Ein **Poisson-Prozess** mit Intensität g ist ein stochastischer Prozess $\mathcal{P}$ in $\mathbb{R}^d$, sodass für eine Borelmenge $A \subset \mathbb{R}^d$ die Zufallsvariable $\mathcal{P}(A)$ Poisson verteilt ist, mit Parameter $\int_A g(x) dx$.

Seien $A_1, A_2, \dots, A_k \subset \mathbb{R}^d$ paarweise disjunkte Borelmengen, so sind ihre Zufallsvariablen $\mathcal{P}(A_i)$ unabhängig für $1 \leq i \leq k$.

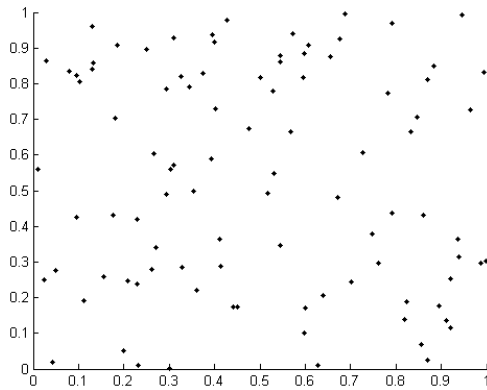
1 Grundlagen und Voraussetzungen

Bemerkung. Definition 1.13 beschreibt den Fall eines *inhomogenen* Poisson-Prozesses. Ein Poisson-Prozess heißt *homogen*, wenn die Zufallsvariable $\mathcal{P}(A)$ Poisson verteilt mit Parameter $g \cdot |A|$ für alle $A \subset \mathbb{R}^d$ ist. Die zweite Bedingung, die Unabhängigkeit der Zufallsvariablen $\mathcal{P}(A_i)$ für paarweise disjunkte Borelmengen $A_i \subset \mathbb{R}^d$, soll auch im Falle eines homogenen Poisson-Prozesses gelten.

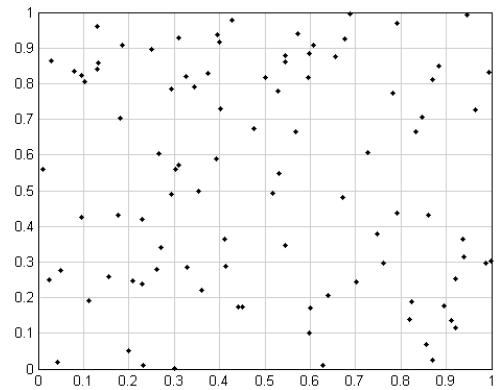
Beispiel 1.2 (Erzeugung eines Poisson-Prozesses von Punkten in $[0, 1]^2$).

Für gegebene, gleichverteilte Punkte $\{X_1, X_2, \dots\}$ im Einheitsquadrat $[0, 1]^2$ wird eine Poisson-Zufallsvariable N_λ mit $\mathbb{E}(N_\lambda) = \lambda([0, 1]^2) \cdot \lambda$ erzeugt. Durch N_λ kann ein Poisson-Prozess $\mathcal{P}_\lambda = \{X_1, \dots, X_{N_\lambda}\}$ definiert werden, wobei N_λ nun die Anzahl der Punkte bestimmt, welche von der Punktmenge in den Poisson-Prozess übernommen werden. Dass die Menge $\mathcal{P}_\lambda$ tatsächlich ein Poisson-Prozess ist, wird in Kapitel 2 in Satz 2.1 gezeigt.

Für Abbildung 1.2 ist für $\lambda = 100$ die Zufallsvariable $N_\lambda = 94$ berechnet worden, anschließend wurden 94 unabhängige, gleichverteilte Punkte im Intervall $[0, 1]^2$ erzeugt und geplottet.



(a) Poisson-Prozess mit Intensität 100



(b) Unterteilung von $[0, 1]^2$

Da es sich bei $\mathcal{P}_\lambda$ um einen Poisson-Prozess handelt, gilt beispielsweise für jedes Quadrat A_{ij} der Unterteilung von $[0, 1]^2$ in Abbildung 1.2 (b), dass die erwartete Anzahl der Punkte in diesem Quadrat gleich $\lambda(A_{ij}) \cdot \lambda = 0,01 \cdot 94 = 0,94$ ist.

Zusätzlich werden noch folgende Aussagen über die Poisson-Verteilung benötigt, welche hier jeweils nur zusammengefasst und ohne Beweis wiedergegeben werden. Beweise für die folgenden Sätze finden sich in *Random Geometric Graphs* von Penrose ([25]).

Die Funktion $H: [0, \infty) \rightarrow [0, \infty)$ wird durch $H(0) = 1$ und $H(a) = 1 - a + a \log(a)$ für $a > 0$ definiert und wird in Kapitel 2 noch öfter benötigt.

Zuerst folgt noch eine Aussage über die Binomial-Verteilung, welche wie die Satz 1.4 eine obere Schranke für die Wahrscheinlichkeit angibt, dass der Wert einer binomialverteilten Zufallsvariable $Bi(n, p)$ größer oder kleiner ist als erwartet.

Definition 1.14: Binomialverteilung $Bi(n, p)$ (Penrose [25])

Die **Binomialverteilung** ist eine diskrete Wahrscheinlichkeitsverteilung welche durch zwei Parameter n und p definiert wird. Dabei beschreibt n die Anzahl der Versuche und p die Erfolgswahrscheinlichkeit $p \in (0, 1)$.

$$B_{n,p}(k) = \binom{n}{k} p^k (1-p)^{(n-k)} \text{ für } k \in \{0, 1, \dots\}$$

Satz 1.3 (Penrose [25])

Seien $n \in \mathbb{N}$, $p \in (0, 1)$, $0 < k < n$ und $\mu = np$. Es gelten folgende Aussagen für die Binomialverteilung $Bi_{n,p}$:

- Wenn $k \geq \mu$ ist, dann gilt:

$$\mathbb{P}[Bi_{n,p} \geq k] \leq e^{(-\mu H(\frac{k}{\mu}))}$$

- Wenn $k \leq \mu$ ist, dann gilt:

$$\mathbb{P}[Bi_{n,p} \leq k] \leq e^{(-\mu H(\frac{k}{\mu}))}$$

- Wenn $k \geq e^2 \mu$ ist, dann gilt:

$$\mathbb{P}[Bi_{n,p} \leq k] \leq e^{-\frac{k}{2} \log(\frac{k}{\mu})}$$

1 Grundlagen und Voraussetzungen

Satz 1.4 (Penrose [25])

Seien $k > 0$, $\lambda > 0$. Es gelten folgende Aussagen für die Poissonverteilung P_λ :

- Wenn $k \geq \lambda$ ist, dann gilt:

$$\mathbb{P}[P_\lambda \geq k] \leq e^{-\lambda H\left(\frac{k}{\lambda}\right)}$$

- Wenn $k \leq \lambda$ ist, dann gilt:

$$\mathbb{P}[P_\lambda \leq k] \leq e^{-\lambda H\left(\frac{k}{\lambda}\right)}$$

- Wenn $k \geq e^2 \lambda$ ist, dann gilt:

$$\mathbb{P}[P_\lambda \geq k] \leq e^{-\left(\frac{k}{2}\right) \log\left(\frac{k}{\lambda}\right)}$$

Satz 1.5 (Penrose [25])

Sei $\lambda \geq 0$ und $k \in \mathbb{N}$, dann gelten folgende Aussagen für die Poissonverteilung P_λ :

- Wenn $k > \lambda$ ist, dann gilt:

$$\mathbb{P}[P_\lambda = k] \geq \left(\frac{e^{-\frac{1}{12k}}}{\sqrt{2\pi k}} \right) e^{-\lambda H\left(\frac{k}{\mu}\right)}$$

- Wenn $k \leq \lambda$ ist, dann gilt:

$$\mathbb{P}[P_\lambda = k] \geq \left(\frac{e^{-\frac{1}{12k}}}{\sqrt{2\pi k}} \right) e^{-k \log\left(\frac{k}{\mu}\right)}$$

Satz 1.6 (Penrose [25])

Sei $\gamma > \frac{1}{2}$, dann existiert eine Konstante $\lambda_1 > 0$ abhängig von γ , sodass für alle $\lambda > \lambda_1$ gilt:

$$\begin{aligned} \mathbb{P}\left[P_\lambda > \lambda + \frac{\lambda^\gamma}{2}\right] &\leq e^{-\frac{\lambda^{2\gamma-1}}{9}} \\ \mathbb{P}\left[P_\lambda < \lambda - \frac{\lambda^\gamma}{2}\right] &\leq e^{-\frac{\lambda^{2\gamma-1}}{9}} \end{aligned}$$

Satz 1.7 (Boolesche Ungleichung (Penrose [25]))

Seien $E_1, \dots, E_n$ Ereignisse eines Wahrscheinlichkeitsraumes. Dann gilt:

$$\mathbb{P}[\cup_{i=1}^n E_i] \leq \sum_{i=1}^n \mathbb{P}[E_i]$$

Diese Ungleichung gilt ebenso für unendlich viele Ereignisse E_i .

1.3 Weitere mathematische Grundlagen

Definition 1.15: Delta-Distribution (Zemanian [32])

Die **Delta-Distribution** ist ein lineares Funktional und erfüllt für eine beliebige Funktion f folgende Gleichung:

$$\langle f, \delta \rangle := f(0)$$

Bemerkung. Durch eine Approximation der Delta-Distribution kann ihr Verhalten aufgezeigt werden und Definition 1.15 sowie die Auswertung der Funktion f an der Stelle 0 motiviert werden.

Hierfür wird für beliebiges $a > 0$ die folgende Funktion δ_a definiert:

$$\delta_a(x) = \frac{1}{\sqrt{\pi a}} \cdot e^{-\frac{x^2}{a}}$$

Die Funktionenfolge δ_a approximiert die Delta-Distribution für $a \rightarrow 0$. Abbildung 1.1 zeigt die Approximationen für $a \in \{0.75, 0.25, 0.15\}$ und veranschaulicht, wie sich die Funktion an der Stelle 0 aufbäumt und an den anderen Stellen den Funktionswert 0 ergibt.

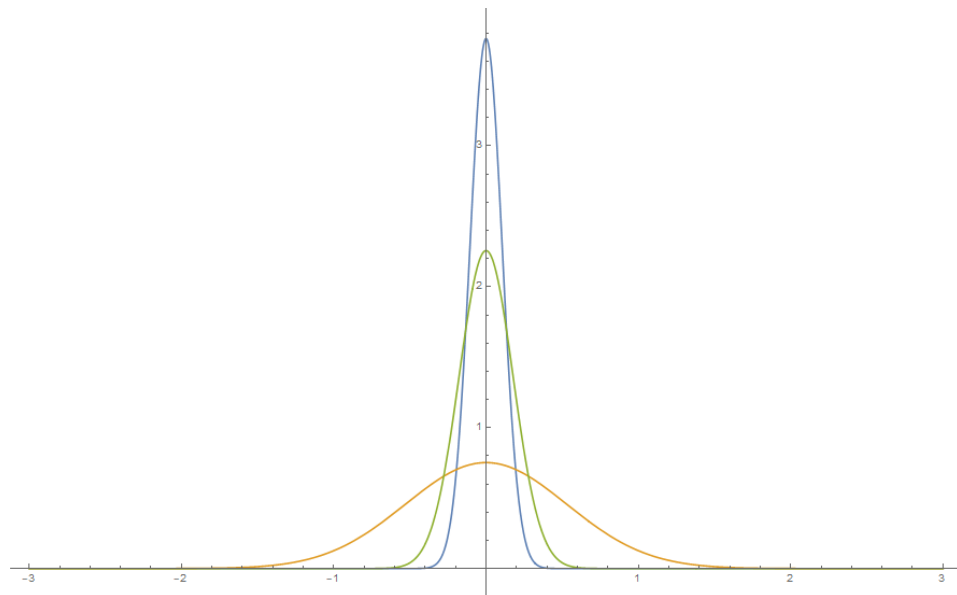


Abbildung 1.1: Approximation der Delta-Distribution $\delta_a(x)$ mit $a \in \{0.75, 0.25, 0.15\}$

Es gilt nun für stetiges und beschränktes f , dass $\lim_{a \rightarrow 0} \int_{-\infty}^{\infty} f(t) \cdot \delta_a(t) dt = \int_{-\infty}^{\infty} f(t) \cdot \delta(t) dt = f(0)$, was die Schreibweise der Integration der Delta-Distribution mittels

$$\langle f, \delta \rangle = \int_{-\infty}^{\infty} f(t) \cdot \delta(t) dt = f(0)$$

motiviert.

Für einen Beweis in Kapitel 2 wird der Begriff der Packing-Nummer benötigt, welcher die maximale Anzahl an disjunkten d -dimensionalen Kugeln angibt, welche in einem Gebiet platziert werden können:

Definition 1.16: Packing-Nummer¹ (Penrose [25])

Sei $U \subseteq \mathbb{R}^d$ und $r > 0$, dann bezeichnet die **Packing-Nummer** $\sigma(U, r)$ das maximale $n \in \mathbb{N}$, so dass n disjunkte Kugeln $B(x, r)$ mit $x \in U$ existieren.

Betrachtet man zwei kompakte Mengen $A, B \subset \mathbb{R}^d$, so kann man mit Hilfe der Brunn-Minkowski Ungleichung das Lebesgue-Maß $\lambda(A + B)$ nach unten abschätzen. Auf diese Aussage wird später in einem Beweis zurückgegriffen.

Satz 1.8 (Brunn-Minkowski Ungleichung (Penrose [25]))

Seien $A, B \subset \mathbb{R}^d$ kompakte und nicht leere Mengen, sowie die Addition der Mengen definiert durch $A + B = \{x + y \mid x \in A, y \in B\}$. Dann gilt für das Lebesgue-Maß λ :

$$\lambda(A + B) \geq \left(\lambda(A)^{\frac{1}{d}} + \lambda(B)^{\frac{1}{d}} \right)^d$$

¹In der englischsprachigen Literatur spricht man auch von *packing number*.

2 Geometrische Zufallsgraphen

2.1 Definition und Anschauung

Um den geometrischen Zufallsgraphen formal zu definieren, muss vorerst noch der geometrische Graph eingeführt werden. Die geometrische Graphentheorie ist ein Teilbereich der Graphentheorie und betrachtet die Knoten $v \in V$ und Kanten $e \in E$ eines Graphen $G = (V, E)$ nicht mehr ausschließlich als abstrakte mathematische Objekte, sondern als Objekte in einem normierten geometrischen Raum.

Definition 2.1: Geometrischer Graph (Penrose [25])

Sei $\mathcal{X} \subset \mathbb{R}^d$ eine endliche Menge an Punkten. $G(\mathcal{X}, r)$ heißt **geometrischer Graph**, wenn er $\mathcal{X}$ als Knotenmenge besitzt, sowie zwei Knoten $u, v \in \mathcal{X}$ genau dann mit einer ungerichteten Kante verbunden sind, wenn $\|u - v\| \leq r$ erfüllt ist.

Verteilt man die Punkte $\mathcal{X} \subset \mathbb{R}^d$ eines geometrischen Graphen nach einer Wahrscheinlichkeitsverteilung zufällig im Raum $\mathbb{R}^d$, so spricht man von einem geometrischen Zufallsgraphen:

Definition 2.2: Geometrischer Zufallsgraph (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte auf $\mathbb{R}^d$, $X_1, X_2, \dots$ unabhängige und identisch verteilte d -dimensionale Zufallsvariablen mit gemeinsamer Dichte f und $\mathcal{X}_n := \{X_1, X_2, \dots, X_n\}$.

Der geometrische Graph $G(\mathcal{X}_n, r)$ wird als **geometrischer Zufallsgraph** bezeichnet.

Bemerkung. Gilbert ([13]) führte die geometrischen Zufallsgraphen 1961 ein und schlug sie als Modell für Kommunikationsnetzwerke vor. Auch heutzutage werden geometrische Zufallsgraphen zur Modellierung und Analyse von ad-hoc WiFi Netzwerken verwendet. Auf die Verwendung von geometrischen Zufallsgraphen als Modell von ad-hoc WiFi Netzwerken wird in dieser Arbeit in Kapitel 3 eingegangen.

Beispiel 2.1. Die Ausbreitung einer Krankheit in einem Wald kann mittels geometrischen Zufallsgraphen modelliert werden, wenn die Krankheit direkt von Baum zu Baum über geringe Distanz übertragen wird. Die Krankheit verbreitet sich demnach im Laufe der Zeit auf alle erreichbaren Bäume, ausgehend vom ersten infizierten Baum.

Hierbei stellt sich die Frage, in wieviele einzelne, nicht verbundene Komponenten der Graph unterteilt ist, um eine Aussage über die Ausbreitung der Krankheit zu treffen. Ist der Graph beispielsweise zusammenhängend, das heißt jeder Knoten ist von jedem erreichbar, dann kann die Krankheit auf alle Individuen übertragen werden.

Beispiel 2.2. Abbildung 2.1 zeigt einen geometrischen Zufallsgraphen, mit 250 gleichverteilten Knoten in $[0, 1]^2$ und $r = 0.1$, das heißt alle Knoten u, v mit $\|u - v\| \leq 0.1$ sind durch eine Kante verbunden. Der so entstandene Graph ist in diesem Falle zusammenhängend und besitzt keine isolierten Knoten.

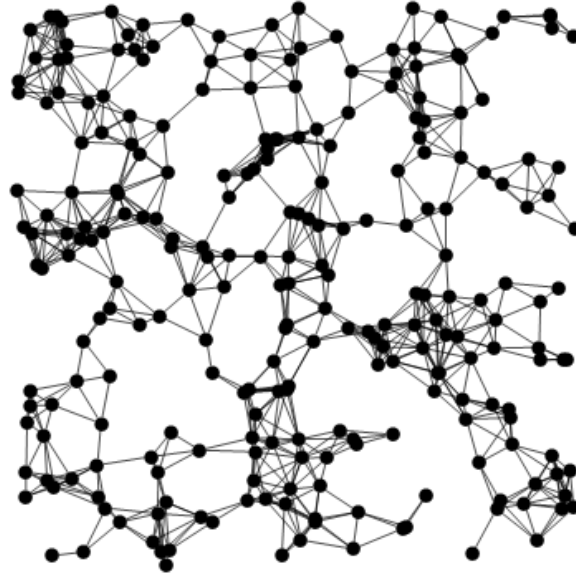


Abbildung 2.1: Geometrischer Zufallsgraph $G(\mathcal{X}_{250}, 0.1)$ mit 250 Knoten

2.1.1 Kategorien geometrischer Zufallsgraphen

In dieser Arbeit werden hauptsächlich asymptotische Ergebnisse des geometrischen Zufallsgraphen $G(\mathcal{X}_n, r_n)$ für $n \rightarrow \infty$ betrachtet, wobei zwischen verschiedenen Folgen von Radien $(r_n)_{(n \geq 1)}$ unterschieden wird.

Die Wahl des Radius hat, wie auch die Wahl der Wahrscheinlichkeitsverteilung, große Auswirkungen auf die Struktur und die Eigenschaften des daraus resultierenden geometrischen Zufallsgraphen. Während versucht wird, die Wahrscheinlichkeitsverteilungen

möglichst allgemein anzunehmen, gibt es für die Radien $(r_n)_{(n \geq 1)}$ verschiedene Kategorien, welche unterschieden werden:

- **Thermodynamisches Verhalten¹**: $n \cdot r_n^d \rightarrow c$ mit $c \in \mathbb{R}^+$ und $c > 0$.
In diesem Fall ist der erwartete Knotengrad eines Knoten des geometrischen Zufallsgraphen $G(\mathcal{X}_n, r_n)$ eine Konstante.
- **Dünn besetztes Verhalten¹**: Der Fall $n \cdot r_n^d \rightarrow 0$ wird als dünn besetztes Verhalten bezeichnet.
- **Dicht besetztes Verhalten¹**: Analog wird die Asymptotik $n \cdot r_n^d \rightarrow \infty$ als dicht besetztes Verhalten bezeichnet.
- **Konnektivitätsverhalten¹**: Dieses Verhalten ist ein Spezialfall des dicht besetzten Verhaltens - für den Radius r_n gilt: $r_n^d \cdot \frac{n}{\log(n)} \rightarrow c$ mit $c \in \mathbb{R}^+$.
Der erwartete Knotengrad eines Knoten des geometrischen Zufallsgraphen wächst in diesem Fall logarithmisch in n .

Bemerkung. Betrachtet man einen geometrischen Zufallsgraphen mit gleichverteilten Punkten $\mathcal{X}_n$ in $[0, 1]^2$, so gilt, dass der erwartete Knotengrad eines Knotens v in $G(\mathcal{X}_n, r_n)$ gleich $n\pi r_n^2$ ist.

Im Falle des thermodynamischen Verhaltens folgt $\lim_{n \rightarrow \infty} n\pi r_n^2 = \lim_{n \rightarrow \infty} n\pi \frac{c}{n} = c$, was obige Erklärung über das Verhalten des erwarteten Knotengrades in diesem Falle bestätigt.

Ebenso kann für das Konnektivitätsverhalten der logarithmische Wachstum des erwarteten Knotengrades nachgerechnet werden: $\lim_{n \rightarrow \infty} n\pi r_n^2 = \lim_{n \rightarrow \infty} n\pi \frac{c \cdot \log(n)}{n} = \lim_{n \rightarrow \infty} c \cdot \log(n)$.

2.2 Eigenschaften geometrischer Zufallsgraphen

Wie in den Beispielen zuvor ersichtlich wurde, stellen sich in der Analyse von geometrischen Zufallsgraphen viele Fragen, welche ebenso in der klassischen Graphentheorie gestellt werden. So interessieren wir uns für die typischen Knotengrade, den minimalen und maximalen Grad oder die chromatische Zahl eines geometrischen Zufallsgraphen G , um nur einige wenige Eigenschaften zu nennen.

Da die konkrete Verteilung der Punkte $\mathcal{X}$ im Raum und die Struktur des Graphen a priori nicht bekannt ist, lassen sich nur wenige, sehr allgemeine Resultate der klassischen Graphentheorie direkt auf geometrische Zufallsgraphen anwenden.

¹Englischsprachige Bezeichnungen dieser Verhalten: thermodynamic regime, sparse regime, dense regime, connectivity regime.

In den folgenden Abschnitten werden konkret einige Eigenschaften der geometrischen Zufallsgraphen analysiert und die wichtigsten Aussagen ausgearbeitet. Dies beinhaltet die Analyse des minimalen und maximalen Knotengrades $\delta_n = \delta(G(\mathcal{X}_n, r_n))$ und $\Delta_n = \Delta(G(\mathcal{X}_n, r_n))$ in den Abschnitten 2.2.2 und 2.2.3. In den anschließenden Abschnitten werden die Clique-Zahl $\mathcal{C}_n = \mathcal{C}(G(\mathcal{X}_n, r_n))$, die chromatische Zahl $\chi_n(G) = \chi(G(\mathcal{X}_n, r_n))$ und die Konnektivität $\kappa_n = \kappa(G(\mathcal{X}_n, r_n))$ eines geometrischen Zufallsgraphen $G(\mathcal{X}_n, r_n)$ genauer untersucht.

2.2.1 Poissonisierung

Penrose ([25]) hat 2003 in seinem Werk *Random Geometric Graphs* viele Resultate und Ergebnisse über geometrische Zufallsgraphen erarbeitet und gezeigt. Oftmals wird dabei in Beweisen auf sogenannte *Poissonisierung* zurückgegriffen, um Resultate für geometrische Zufallsgraphen der Form $G(\mathcal{P}_\lambda, r)$ zu zeigen und von diesen auf Resultate für geometrische Zufallsgraphen der Form $G(\mathcal{X}_n, r)$ zu schließen.

Eine wichtige Eigenschaft ist die räumliche Unabhängigkeit, welche die Arbeit mit geometrischen Zufallsgraphen der Form $G(\mathcal{P}_\lambda, r)$ vereinfacht:

Satz 2.1 (Penrose [25])

Sei $\lambda > 0$, N_λ eine Poisson Zufallsvariable, unabhängig von $\mathcal{X}_n = \{X_1, X_2, \dots\}$ dann ist $\mathcal{P}_\lambda := \{X_1, X_2, \dots, X_{N_\lambda}\}$ ein Poisson-Prozess auf $\mathbb{R}^d$ mit Intensität $\lambda f(\cdot)$, wobei f die Dichte der Verteilungen der X_i bezeichnet.

Beweis zu Satz 2.1. Seien die Borelmengen $A_1, \dots, A_k$ so gewählt, dass sie eine Partition von $\mathbb{R}^d$ bilden, das heißt die A_i sind zueinander disjunkt und $\bigcup_{n \in \mathbb{N}} A_i = \mathbb{R}^d$. Für beliebige $n_1, \dots, n_k$ sei $n = \sum_{i=1}^k n_i$. Dann gilt:

$$\begin{aligned} \mathbb{P}\left[\bigcap_{i=1}^k \mathcal{P}_\lambda(A_i) = n_i\right] &= \mathbb{P}[N_\lambda = n] \cdot \mathbb{P}\left[\bigcap_{i=1}^k \{\mathcal{P}_\lambda(A_i) = n_i\} \mid N_\lambda = n\right] \\ &= \left(\frac{e^{-\lambda} \lambda^n}{n!}\right) \cdot \frac{n!}{n_1! \cdots n_k!} \cdot \prod_{i=1}^k \left(\int_{A_i} f(x) \, dx\right)^{n_i} \end{aligned}$$

Wegen $\lambda^n = \lambda^{n_1} \cdots \lambda^{n_k}$ und $\sum_{i=1}^k \int_{A_i} f(x) \, dx = 1$ gilt nun:

$$= \prod_{i=1}^k \left(\frac{e^{-\lambda \int_{A_i} f(x) \, dx} \cdot (\lambda \int_{A_i} f(x) \, dx)^{n_i}}{n_i!} \right) \quad (2.1)$$

Aus Formel (2.1) folgt, dass $\mathcal{P}(A_i)$ für alle $i \in \{1, \dots, k\}$ unabhängige

Poisson-Zufallsvariablen mit dem Erwartungswert $\mathbb{E}[P_\lambda(A_i)] = \int_{A_i} \lambda f(x) dx$ sind, wodurch die Aussage von Satz 2.1 folgt. $\square$

2.2.2 Minimaler Knotengrad

Mit einigen Voraussetzungen an das Gebiet Ω der Wahrscheinlichkeitsdichte f und an die Folge der betrachteten Radien $(r_n)_{(n \geq 1)}$ kann das asymptotische Verhalten des minimalen Knotengrades $\delta_n = \delta(G(\mathcal{X}_n, r_n))$ von $G(\mathcal{X}_n, r_n)$ gezeigt werden. Dafür wird der Radius analysiert, welcher gewählt werden muss um einen bestimmten Knotengrad nicht zu unterschreiten - von dieser Analyse kann auf das Verhalten des minimalen Knotengrades eines geometrischen Zufallsgraphen geschlossen werden.

Die Frage die sich stellt ist, welcher minimaler Knotengrad für gewisse geometrische Zufallsgraphen $G(\mathcal{X}_n, r_n)$ erreicht wird. Damit beantwortet sich beispielsweise auch die Frage, ob es in diesem Graphen *isolierte Punkte* gibt. Isolierte Punkte sind Knoten v mit Knotengrad $d_G(v) = 0$, das heißt, sie sind mit keinem anderen Knoten w des Graphen verbunden.

Für die Aussage von Satz 2.8 wird eine Hilfsfunktion $H: [0, \infty) \rightarrow \mathbb{R}$ benötigt, welche durch $H(0) = 1$ und $H(a) = 1 - a + a \cdot \log(a)$ für $a > 0$ definiert wird. Die eindeutige Umkehrfunktion von $H|_{[0,1]}$ wird mit H_-^{-1} bezeichnet und bildet von $[0, 1]$ in sich selbst ab. Diese Funktion wird später noch einmal in Abschnitt 2.2.3 bei der Analyse des maximalen Knotengrades Δ_n des geometrischen Zufallsgraphen $G(\mathcal{X}_n, r_n)$ benötigt.

Sei $\Omega = \prod_{i=1}^d [a_i, b_i]$ das Produkt von endlich vielen, geschlossenen Intervallen, die im Folgenden verwendete Norm $\|\cdot\|$ eine l_p -Norm und $\mathcal{J} := \{0, 1, \dots, d\}$. f_0 sei das wesentliche Infimum von $f|_\Omega$, für $0 \neq j \in \mathcal{J}$ sei ∂_j die Vereinigung der $(d - j)$ -dimensionalen Ränder von Ω und f_j sei das Infimum von $f|_{\partial_j}$.

Definition 2.3: d -dimensionale Einheitskugel

Die Menge $B(x, r) = \{y \mid \|x - y\| < r, x, y \in \mathbb{R}^d\}$ wird als d -dimensionale Kugel um $x \in \mathbb{R}^d$ mit Radius r bezeichnet.

$B(0, 1)$ ist dementsprechend die **d -dimensionale Einheitskugel**.

Bemerkung. Das Volumen der d -dimensionalen Einheitskugel $\lambda(B(0, 1))$ wird im Folgenden mit $\theta = \lambda(B(0, 1))$ bezeichnet, wobei λ das Lebesgue-Maß ist.

Hilfreich für den Beweis von Satz 2.8 ist die genauere Untersuchung des asymptotischen Verhaltens des größten k -nächsten Nachbarn Radius, von welchem direkt auf das asymptotische Verhalten des minimalen Knotengrades δ_n von $G(\mathcal{X}_n, r_n)$ geschlossen werden kann:

Definition 2.4: größter k -nächsten Nachbarn Radius (Penrose [26])

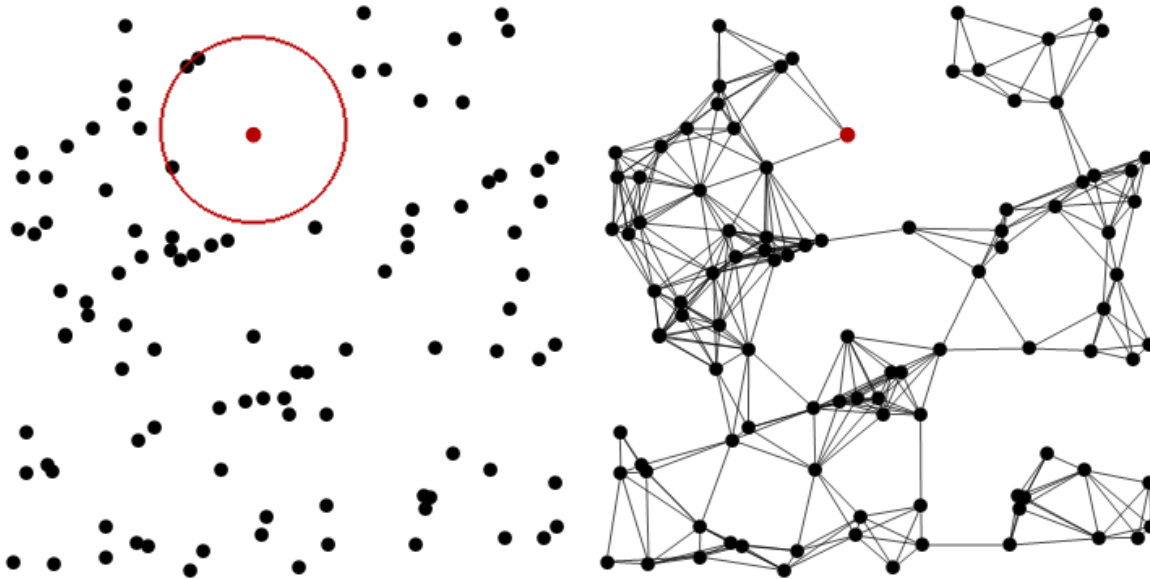
Sei $D_k(X_i)$ für $X_i \in \mathcal{X}$ die Distanz von X_i zu seinen k nächsten Nachbarn in $\mathcal{X}$.

Der **größte k -nächsten Nachbarn Radius**² $M_k(\mathcal{X})$ ist definiert als das Maximum aller $D_k(X_i)$ für $X_i \in \mathcal{X}$, das heißt $M_k(\mathcal{X}) := \max_{X_i \in \mathcal{X}} D_k(X_i)$.

Der größte k -nächsten Nachbarn Radius ist der Radius, welcher sicherstellt, dass jeder Punkt, beziehungsweise Knoten, des geometrischen Graphs $G(\mathcal{X}, M_k(\mathcal{X}))$ mit mindestens k Knoten verbunden ist - damit hat der Graph minimalen Knotengrad k .

Beispiel 2.3 (Erzeugung eines geometrischen Zufallsgraphen mit Minimalgrad 3).

In diesem Beispiel wird ein geometrischer Zufallsgraph mit 100 gleichverteilten Punkten $\mathcal{X}_{100}$ erzeugt und der größte k -nächsten Nachbarn Radius für $k = 3$ berechnet. Anschließend wird der geometrische Zufallsgraph $G(\mathcal{X}_{100}, M_3(\mathcal{X}_{100}))$ mit denselben Punkten $\mathcal{X}_{100}$ neu gezeichnet, wobei $M_3(\mathcal{X}_{100}) = 0.171544$ gilt. Der daraus resultierende Graph hat Minimalgrad 3.



(a) Der rot markierte Punkt nimmt das Maximum über alle $D_3(X_i)$ für $X_i \in \mathcal{X}_{100}$ an

(b) Geometrischer Zufallsgraph $G(\mathcal{X}_{100}, M_3(\mathcal{X}_{100}))$ mit minimalem Knotengrad 3

²In der englischsprachigen Literatur wird der Begriff *largest k -nearest neighbour link* verwendet.

Satz 2.2 zeigt das asymptotische Verhalten von M_k , und dient später als Grundlage im Beweis zur Aussage über das asymptotische Verhalten des minimalen Knotengrades $\delta_n = \delta(G(\mathcal{X}_n, r_n))$ (Satz 2.8).

Satz 2.2 (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte auf Ω , wobei $f|_{\Omega}$ stetig in $x \forall x \in \delta\Omega$ ist. Zusätzlich sei $(k_n)_{(n \geq 1)}$ eine Folge natürlicher Zahlen, wobei $\lim_{n \rightarrow \infty} \frac{k_n}{n} = 0$ und $\lim_{n \rightarrow \infty} \frac{k_n}{\log(n)} = b \in [0, \infty]$ gelte. Wenn $b < \infty$ ist, gelte zusätzlich, dass $(k_n)_{(n \geq 1)}$ nicht fallend ist, sowie $a_0, a_1, \dots, a_{d-1} \in [0, 1)$ definiert werden durch $\frac{a_j}{H(a_j)} = \frac{bd}{d-j}$. Dann gelten folgende Aussagen:

1. Wenn $b = \infty$, dann gilt mit Wahrscheinlichkeit 1:

$$M_{k_n}(\mathcal{X}_n)^d \sim \left(\frac{k_n}{n\theta} \right) \cdot c$$

wobei

$$c = \max \left\{ \frac{2^j}{f_j} \mid 0 \leq j \leq d \right\}$$

gilt.

2. Wenn $b < \infty$, dann gilt mit Wahrscheinlichkeit 1:

$$M_{k_n}(\mathcal{X}_n)^d \sim \left(\frac{\log(n)}{n\theta} \right) \cdot c$$

wobei

$$c = \max \left(\max_{0 \leq j \leq d-1} \left(\frac{2^j(d-j)}{df_j H(a_j)} \right), \frac{2^d \cdot b}{f_d} \right)$$

gilt.

Beweis von Satz 2.2

Um Satz 2.2 zu beweisen, muss etwas weiter ausgeholt werden, da im Beweis die Technik der Poissonisierung verwendet wird. Seien für jedes $n \in \mathbb{N}$ zwei Poisson Zufallsvariablen $N^-(n)$ und $M(n)$ mit den jeweiligen Erwartungswerten $\mathbb{E}[N^-] = n - n^{\frac{3}{4}}$ und $\mathbb{E}[M] = 2n^{\frac{3}{4}}$ definiert, welche voneinander unabhängig sind, sowie unabhängig von $(X_1, X_2, \dots)$ sind.

Mithilfe dieser Poisson Zufallsvariablen können die folgenden zwei Poisson-Prozesse mit den Intensitäten $\left[(n - n^{\frac{3}{4}}) \cdot f(\cdot)\right]$ und $\left[(n + n^{\frac{3}{4}}) \cdot f(\cdot)\right]$ definiert werden:

$$\begin{aligned}\mathcal{P}_n^- &= \{X_1, X_2, \dots, X_{N^-(n)}\} \\ \mathcal{P}_n^+ &= \{X_1, X_2, \dots, X_{N^-(n)+M(n)}\}\end{aligned}$$

Für diese beiden Poisson-Prozesse gilt nun, dass $\mathcal{P}_n^- \subset \mathcal{P}_n^+$ ist. Das Ereignis, dass $\mathcal{X}_n$ von den beiden Poisson-Prozessen $\mathcal{P}_n^-$ und $\mathcal{P}_n^+$ bezüglich der Mengeninklusion eingeschlossen wird, wird als $H_n = \{ \mathcal{P}_n^- \subseteq \mathcal{X}_n \subseteq \mathcal{P}_n^+ \}$ bezeichnet.

Wählt man ein beliebiges $u \geq 0$, so kann für jedes $j \in \{0, 1, \dots, d\}$ eine Funktion $p_n^j(u)$ definiert werden, welche abhängig vom Wert von b wie folgt aussieht:

$$p_n^j(u) = \begin{cases} \left(\frac{2^j u k_n}{n \theta f_j} \right)^{\frac{1}{d}} & \text{für } b = \infty \\ \left(\frac{2^j (d-j) u \log(n)}{n \theta d f_j H(a_j)} \right)^{\frac{1}{d}} & \text{für } b < \infty \text{ und } 0 \leq j \leq d-1 \end{cases} \quad (2.2)$$

Für den Fall $b < \infty$ und $j = d$ wird p_n^d in der folgenden Form definiert:

$$p_n^d(u) = \left(\frac{2^d b u \log(n)}{n \theta f_d} \right)^{\frac{1}{d}} \quad (2.3)$$

Mit Hilfe dieser Konstruktionen kann der folgende Satz gezeigt werden, welcher eine asymptotische untere Schranke für den k -nächsten Nachbarn Radius $M_k(\mathcal{X}_n)$ angibt:

Satz 2.3 (Penrose [25])

Sei $0 < \beta < 1$, $j \in \{0, 1, \dots, d\}$, und f , $(k_n)_{(n \geq 1)}$, $b \in [0, \infty]$ und $a_j \in [0, 1)$ wie in Satz 2.2 definiert. Dann gilt mit Wahrscheinlichkeit 1 für alle hinreichend großen $n \in \mathbb{N}$:

$$M_{k_n}(\mathcal{X}_n) > p_n^j(\beta)$$

Beweis zu Satz 2.3.

1. Fall: $j > 0$

Da $\beta \in (0, 1)$ beliebig und fest gewählt wurde, kann ein ϵ_1 gefunden werden, sodass die folgende Ungleichung gilt:

$$\epsilon_1^{\frac{1}{d}} + \beta^{\frac{1}{d}} < (1 - 4\epsilon_1)^{\frac{1}{d}} \quad (2.4)$$

Für jedes $x \in \mathbb{R}^d$ kann nun das Ereignis $E_n(x)$ definiert werden:

$$E_n(x) = \underbrace{\left\{ \mathcal{P}_n^+ [B(x, p_n^j(1 - 4\epsilon_1))] \right\} \leq k_n}_{E_{n,1}(x)} \cap \underbrace{\left\{ \mathcal{P}_n^- [B(x, p_n^j(\epsilon_1))] \right\} \geq 1}_{E_{n,2}(x)}$$

Gilt $\|y - x\| \leq p_n^j(\epsilon_1)$ und $\|z - x\| \leq p_n^j(\beta)$ für zwei Punkte $y, z \in \mathbb{R}^d$, so folgt mit Hilfe der Dreiecksungleichung $\|z - x\| \leq \|z - y\| + \|y - x\| \leq p_n^j(\epsilon_1) + p_n^j(\beta)$, womit durch Abschätzung (2.4) $\|z - x\| \leq p_n^j(1 - 4\epsilon_1)$ gilt. Betrachtet man nun das oben definierte Ereignis H_n und nimmt an, dass die beiden Ereignisse H_n und $E_n(x)$ gleichzeitig auftreten, das heißt das Ereignis $H_n \cap E_n(x)$ tritt auf, dann gilt folgende Überlegung:

Da $E_{n,2}(x)$ erfüllt wird, existiert ein Punkt $X \in \mathcal{X}_n$ innerhalb der d -dimensionalen Kugel $B(x, p_n^j(\epsilon_1))$. Wegen dem Auftreten von $E_{n,1}(x)$ gibt es höchstens $k_n - 1$ andere Punkte aus $\mathcal{X}_n$ innerhalb von $B(x, p_n^j(\beta))$, woraus $M_{k_n}(\mathcal{X}_n) > p_n^j(\beta)$ folgt. Insgesamt ergibt sich daraus die folgende Abhängigkeit der Ereignisse:

$$\{M_{k_n}(\mathcal{X}_n) \leq p_n^j(\beta)\} \subseteq H_n^C \cup (\cup_{x \in \mathbb{R}^d} E_n(x))^C \quad (2.5)$$

Wenn $b = \infty$ gilt, betrachte ein $x_j \in \partial_j$ mit $f(x_j) < \frac{f_j(1-3\epsilon_1)}{1-4\epsilon_1}$. Für dieses x_j und hinreichend großes $n \in \mathbb{N}$ gilt $F(B(x_j, p_n^j(1-4\epsilon_1))) \leq \frac{(1-3\epsilon_1)k_n}{n}$, womit die erwartete Anzahl an Punkten von $\mathcal{P}_n^+$ innerhalb von $B(x_j, p_n^j(1-4\epsilon_1))$ höchstens $(1-3\epsilon_1)k_n$ ist. Laut Satz 1.4 nimmt die Wahrscheinlichkeit, dass $\mathcal{P}_n^+$ innerhalb von $B(x_j, p_n^j(1-4\epsilon_1))$ mehr als k_n Punkte hat, exponentiell mit k_n ab. Ebenso nimmt die Wahrscheinlichkeit exponentiell in k_n ab, dass $\mathcal{P}_n^-$ keinen Punkt enthält, welcher sich in $B(x_j, p_n^j(\epsilon_1))$ befindet.

Durch Satz 1.6 folgt für das Komplement des Ereignisses H_n^C folgende Ungleichung für hinreichend große $n \in \mathbb{N}$:

$$\mathbb{P}[H_n^C] \leq 2e^{-\frac{n^{\frac{1}{2}}}{9}} \quad (2.6)$$

Nach Voraussetzung gilt $\frac{k_n}{n} \rightarrow \infty$, woraus mit den obigen Überlegungen folgt, dass $\sum_{i=1}^{\infty} 1 - \mathbb{P}[E_i(x_j) \cap H_i] < \infty$ gilt. Daher kann das 1. Lemma von Borel-Cantelli angewandt werden (Satz 1.2), welches besagt, dass das Ereignis $E_n(x_j) \cap H_n$ mit Wahrscheinlichkeit 1 bis auf endlich viele $n \in \mathbb{N}$ eintritt. Durch Formel (2.5) folgt die Aussage für den Fall $b = \infty$.

Wenn $b < \infty$ gilt, kann ϵ_1 in Formel (2.4) so gewählt werden, dass $\epsilon_1 < a_j$ gilt. Wähle $x_1 \in \partial_j$ und $\epsilon_2 > 0$ so, dass $f(x) < \frac{f_j(1-3\epsilon_1)}{1-4\epsilon_1}$ für $x \in (B(x_1, 2\epsilon_2) \cap \Omega)$ gilt und $B(x_1, 2\epsilon_2)$ lediglich j der Hyperebenen schneidet, welche den Rand von Ω bilden.

Im Fall $j < d$ gilt

$$\liminf_{n \rightarrow \infty} \left(\left(\frac{\log(n)}{n} \right)^{(d-j)/d} \sigma(B(x_1, \epsilon_2) \cap \partial_j, p_n^j(1-4\epsilon_1)) \right) > 0 \quad (2.7)$$

da ∂_j $(d-j)$ dimensional ist. Dabei ist $\sigma(U, r)$ die Packing-Nummer, wie in Definition 1.16 definiert.

Im Folgenden wird der Fall $0 < j < d$ mit $b < \infty$ behandelt. Es sei $x \in B(x_1, \epsilon_2) \cap \partial_j$ gewählt, sowie $r < \epsilon_2$. Unter diesen Voraussetzungen gilt, dass $\lambda(B(x, r) \cap \Omega) = 2^{-j} \theta r^d$, wobei λ das Lebesgue-Maß bezeichnet. Für hinreichend große $n \in \mathbb{N}$ gilt damit:

$$\left(n + n^{\frac{3}{4}}\right) F\left(B(x, p_n^j(1 - 4\epsilon_1))\right) \leq \frac{(1 - 2\epsilon_1)(d - j) \log(n)}{dH(a_j)} =: \mu_{n,j}$$

Da nach Voraussetzung $\frac{k_n}{\log(n)} \rightarrow b$ und $\frac{a_j}{H(a_j)} = \frac{bd}{d-j}$ gilt, folgt für das Verhältnis $\frac{k_n}{\mu_{n,j}}$:

$$\lim_{n \rightarrow \infty} \frac{k_n}{\mu_{n,j}} = \frac{bdH(a_j)}{(d - j)(1 - \epsilon_1)} = \frac{a_j}{1 - \epsilon_1} \in (a_j, 1)$$

Mit Hilfe von Satz 1.5 gilt nun für hinreichend große $n \in \mathbb{N}$:

$$\begin{aligned} \mathbb{P}\left[\mathcal{P}_n^+[B(x, p_n^j(1 - 4\epsilon_1))] \leq k_n\right] &\geq \mathbb{P}[P_{\mu_{n,j}} \geq k_n] \\ &\geq \left(\frac{e^{-1} 12k_n}{\sqrt{2\pi k_n}}\right) e^{-\mu_{n,j} H(a_j)} \\ &\geq n^{\frac{(\epsilon_1 - 1)(d - j)}{d}} \end{aligned}$$

Betrachtet man nun die Anzahl der Punkte von $\mathcal{P}_n^+$ innerhalb der Kugel $B(x, p_n^j(1 - 4\epsilon_1))$, so ist die bedingte Verteilung der Anzahl der Punkte von $\mathcal{P}_n^-$ innerhalb von $B(x, p_n^j(\epsilon_1))$ binomial verteilt mit folgendem Parameter:

$$\frac{(n - n^{\frac{3}{4}}) F(B(x, p_n^j(\epsilon_1)))}{(n + n^{\frac{3}{4}}) F(B(x, p_n^j(1 - 4\epsilon_1)))} =: q_{n,j}$$

Es gilt, dass $q_{n,j} > 0$ ist und die erwartete Anzahl an Punkten von $\mathcal{P}_n^+$ innerhalb von $B(x, p_n^j(1 - 4\epsilon_1))$ geht gegen unendlich. Daher existiert ein $\eta > 0$, so dass folgendes gilt:

$$\mathbb{P}[E_n(x)] \geq \eta \cdot n^{\frac{(\epsilon_1 - 1)(d - j)}{d}} \quad (2.8)$$

Wegen Formel (2.8) und der Abschätzung $1 - t \leq e^{-t}$ gilt $\mathbb{P}[E_n(x)^C] \leq e^{-\eta n^{\frac{(\epsilon_1 - 1)(d - j)}{d}}}$. Über letzteren Ausdruck kann wegen Formel (2.7) summiert werden, das heißt es gilt $\sum_{i=1}^{\infty} \mathbb{P}[E_n(x)^C] < \infty$.

Durch die Abschätzung (2.6) und dem ersten Borel-Cantelli Lemma (Satz 1.2) folgt die Aussage von Satz 2.3 im Fall $b < \infty$ und $0 < j < d$.

2 Geometrische Zufallsgraphen

Fehlend ist nun noch der Fall $b < \infty$ und $j = d$. Sollte $b = 0$ sein, so gilt $p_n^d(\beta) = 0$ durch die Definition von p_n^d und es gilt $M_{k_n} > p_n^d(\beta) = 0$. Damit kann $0 < b < \infty$ angenommen werden.

In diesem Fall sei ein Eckpunkt y von B_n gewählt mit $f(y) = f_d$, dann existiert ein $\epsilon_3 > 0$, so dass für hinreichend große $n \in \mathbb{N}$ die Abschätzung $nF[B(y, p_n^d(\epsilon_1))] > \epsilon_3 \cdot \log(n)$ gilt, woraus

$$\mathbb{P}[\mathcal{X}_n[B(y, p_n^d(\epsilon_1))] = 0] \leq n^{-\epsilon_3}$$

gilt.

Wähle $\epsilon_4 = H(\frac{1}{1-\epsilon_1})(1 - 2\epsilon_1)b$, $J \in \mathbb{N}$ mit $J > \max\{\epsilon_4^{-1}, \epsilon_3^{-1}\}$ und definiere $\nu(m) = m^J$, dann gilt für hinreichend große $m \in \mathbb{N}$:

$$\nu(m+1) F(B(y, p_n^d(1 - 4\epsilon_1))) \leq (1 - 2\epsilon_1)b \log(\nu(m))$$

Da sich k_n asymptotisch wie $b \cdot \log(n)$ verhält, folgt durch Satz 1.3 für hinreichend große $m \in \mathbb{N}$ die folgende Abschätzung:

$$\mathbb{P}[\mathcal{X}_{\nu(m+1)}[B(y, p_{\nu(m)}^d(1 - 4\epsilon_1))] \geq k_{\nu(m)}] \leq m^{-\epsilon_4 J}$$

Durch Anwendung des ersten Lemmas von Borel-Cantelli (Satz 1.2) folgt, dass das folgende Ereignis mit Wahrscheinlichkeit 1 für alle bis auf endlich viele $m \in \mathbb{N}$ eintritt:

$$\{\mathcal{X}_{\nu(m+1)}[B(y, p_{\nu(m)}^d(1 - 4\epsilon_1))] \leq k_{\nu(m)}\} \cap \{\mathcal{X}_{\nu(m)}[B(y, p_{\nu(m)}^d(\epsilon_1))] \geq 1\}$$

Durch Anwendung der Dreiecksungleichung und der Beziehung (2.4) impliziert das obige Ereignis, dass für alle $n \in \mathbb{N}$ zwischen $\nu(m)$ und $\nu(m+1)$ die Aussage $M_{k_n}(\mathcal{X}_n) > p_n^d(\beta)$ gilt.

2. Fall: Der Spezialfall $j = 0$ wird in *Random Geometric Graphs* von Penrose ([25]) durch einen anderen Satz mit einem ähnlichen Beweis gezeigt - auf diesen Beweis sei an dieser Stelle verwiesen. $\square$

Nachdem eine untere Schranke für $M_{k_n}(\mathcal{X}_n)$ gefunden wurde, muss nun $M_{k_n}(\mathcal{X}_n)$ noch nach oben beschränkt werden. Dafür wird für alle $n \in \mathbb{N}$ die folgende Identität p_n definiert:

$$p_n = \begin{cases} \frac{k_n^{\frac{1}{d}}}{n} & \text{für } b = \infty \\ \frac{\log(n)^{\frac{1}{d}}}{n} & \text{für } b < \infty \end{cases}$$

Ist eine Folge $(k_n)_{(n \geq 1)}$ gegeben und seien $K > 0, t > 0$ dann kann folgendes Ereignis definiert werden, wobei $\text{diam}(U) := \sup\{\|x - y\| \mid x, y \in U\}$ gilt:

$$E'_n(K, t) = \{\exists \text{ eine } k_n\text{-separierte Menge } U \text{ für den Graphen } G(\mathcal{X}_n, tp_n) \quad (2.9) \\ \text{mit Durchmesser } \text{diam}(U) \leq Kp_n\}$$

Bemerkung. Ist der minimale Knotengrad eines Graphen $G(V, E)$ höchstens k , so existiert ein Knoten $v \in V$ in diesem Graphen, welcher diesen Minimalgrad annimmt, das heißt für diesen Knoten v gilt: $d_G(v) = \delta(G) \leq k$. Das heißt wiederum nach Definition einer k -separierten Menge, dass dieser Knoten v eine k -separierte Menge für diesen Graphen bildet, da er höchstens zu k anderen Knoten in $G(V, E)$ adjazent ist.

Gilt nun $M_{k_n+1}(\mathcal{X}_n) > tp_n$, so folgt, dass das Ereignis $E'_n(K, t)$ eintritt.

Satz 2.4 (Penrose [25])

Seien die Voraussetzungen wie in Satz 2.2 und zusätzlich $K > 0$. t sei so gewählt, dass es die folgenden Ungleichungen erfüllt:

- Ist $b = \infty$:

$$\theta t^d > \max \left\{ \frac{2^j}{f_j} \mid 0 \leq j \leq d \right\} \quad (2.10)$$

- Ist $b < \infty$:

$$\theta t^d > \max \left\{ \max \left\{ \frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1 \right\}, \frac{2^d b}{f_d} \right\} \quad (2.11)$$

Dann gilt mit Wahrscheinlichkeit 1, dass das Ereignis $E'_n(K, t)$ für alle bis auf endlich viele $n \in \mathbb{N}$ eintritt und daher, laut obiger Bemerkung, $M_{k_n} \leq tp_n$ für alle bis auf endlich viele n gilt.

Bemerkung. Der Beweis von Satz 2.4 ist in mehrere kleine Sätze unterteilt, aus welchen gemeinsam sofort die obige Aussage folgt. In den folgenden Sätzen wird bei den Beweisen immer auf die definierten Größen der vorhergehenden Beweise zurückgegriffen und diese weiter verwendet.

Im Folgenden wird mit $c_1 = \text{diam}([0, 1]^d)$ der Durchmesser des d -dimensionalen Einheitswürfels bezeichnet. Je nachdem, ob $b = \infty$ oder $b < \infty$ gilt, werden im Weiteren t und $\epsilon_5 \in (0, 1)$ so gewählt, dass gilt:

- Im Fall $b = \infty$ wird t so gewählt, dass obige Ungleichung (2.10) aus Satz 2.4 erfüllt wird. ϵ_5 erfülle die Ungleichung $\frac{\theta(t-4c_1\epsilon_5)^d}{1+\epsilon_5} > \max \left\{ \frac{2^j}{f_j} \mid 0 \leq j \leq d \right\}$. Zusätzlich sei $J = 1$ gewählt.
- Im Fall $b < \infty$ wird t so gewählt, dass Ungleichung (2.11) erfüllt wird und ϵ_5 so, dass $\frac{\theta(t-4c_1\epsilon_5)^d}{1+\epsilon_5} > \max \left\{ \max \left\{ \frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1 \right\}, \frac{2^d(b+2\epsilon_5)}{f_d} \right\}$ gilt. $J \in \mathbb{N}$ sei so gewählt, dass $\frac{J\epsilon_5}{d} > 1$ und $J\epsilon_5 H(\frac{b+\epsilon_5}{b+2\epsilon_5}) > 2$ gilt.

Wie im Beweis von Satz 2.3 wird auch hier die Abbildung $\nu(m) = m^J$ für $m \in \{1, 2, \dots\}$ betrachtet.

Definiere das Gitter $\epsilon_5 p_n \mathbb{Z}^d = \{\epsilon_5 p_n z \mid z \in \mathbb{Z}^d\}$, sowie für $y \in \epsilon_5 p_n \mathbb{Z}^d$ sei $C_n(y) = \{y\} + [-\frac{\epsilon_5 p_n}{2}, \frac{\epsilon_5 p_n}{2}]^d$ der Würfel mit Seitenlänge $\epsilon_5 p_n$ und Zentrum y . Damit kann das endliche Gitter $\mathcal{L}_n = \{y \in \epsilon_5 p_n \mathbb{Z}^d \mid C_n(y) \cap \Omega \neq \emptyset\}$ definiert werden.

Für beliebiges $K > 0$ definiere $\mathcal{T}_m(K) = \{\tau \subset \mathcal{L}_{\nu(m)} \mid \text{diam}(\tau) \leq 2Kp_{\nu(m)}\}$. Die Menge $\mathcal{T}_m(K)$ beinhaltet dementsprechend nur Teilmengen von $\mathcal{L}_{\nu(m)}$, mit maximalem Durchmesser $2Kp_{\nu(m)}$.

Für $\tau \in \mathcal{T}_m(K)$ kann nun die Menge

$$A_m(\tau) = (\tau + B(0, (t - 2c_1\epsilon_5)p_{\nu(m)})) \setminus (\tau + C_{\nu(m)}(0))$$

definiert werden, wodurch die Ereignisse F_m für die beiden Fälle $b = \infty$ und $b < \infty$ definiert werden können:

- Im Fall $b = \infty$ sei

$$F_m(\tau) = \{\mathcal{X}_m[A_m(\tau)] \leq k_m\} \quad (2.12)$$

.

- Im Fall $b < \infty$ sei

$$F_m(\tau) = \{\mathcal{X}_{\nu(m)}[A_m(\tau)] \leq k_{\nu(m+1)}\} \quad (2.13)$$

.

Mit Hilfe dieser Begrifflichkeiten lässt sich der folgende Satz formulieren:

Satz 2.5 (Penrose [25])

Sei $K > \max(1, t)$. Dann gilt:

Es existiert ein m_0 so, dass für alle $m \geq m_0$ und $\nu(m) \leq n < \nu(m+1)$ das Ereignis $E'_n(K, t)$ wie in Formel (2.9) definiert in der Vereinigung $\cup_{\tau \in \mathcal{T}_m(K)} F_m(\tau)$ der Ereignisse $F_m(\tau)$ enthalten ist.

Beweis zu Satz 2.5.

Es werden die Fälle $b = \infty$ und $b < \infty$ unterschieden.

1. Fall: Wenn $b < \infty$ gilt, dann wird m_0 so gewählt, dass für alle $m \geq m_0$ die Ungleichung

$$c_1 \epsilon_5 p_{\nu(m)} + (t - 2c_1 \epsilon_5) p_{\nu(m)} \leq t p_{\nu(m+1)} \quad (2.14)$$

gilt. Sei $m \geq m_0$ und $\nu(m) \leq n < \nu(m+1)$. Für $U \subset \mathcal{X}_n$ bezeichnet $\tau(U)$ die Diskretisierung von U in $\epsilon_5 p_{\nu(m)} \mathbb{Z}^d$, das ist die Menge aller $z \in \epsilon_5 p_{\nu(m)} \mathbb{Z}^d$ für welche $U \cap C_{\nu(m)} \neq \emptyset$ gilt.

Gilt $\text{diam}(U) \leq K p_n$ so folgt $\text{diam}(\tau(U)) \leq 2K p_{\nu(m)}$ und damit $\tau(U) \in \mathcal{T}_m(K)$ nach Definition von $\mathcal{T}_m(K)$. Nach Formel (2.14) und durch Anwendung der Dreiecksungleichung folgt $A_m(\tau(U)) \subseteq U + B(0, t p_{\nu(m+1)}) \setminus U$.

Ist U eine k_n -separierte Menge für den Graphen $G(\mathcal{X}_n, t p_n)$, dann gilt $\mathcal{X}_n[U + B(0, t p_n) \setminus U] \leq k_{\nu(m+1)}$ und daher wegen oben $\mathcal{X}_{\nu(m)}[A_m(\tau(U))] \leq k_{\nu(m+1)}$, was die Aussage für den Fall $b < \infty$ zeigt.

2. Fall: Wenn $b = \infty$ gilt, kann der obige Beweis analog übernommen werden, wobei Ungleichung (2.14) so geändert wird, sodass gilt:

$$c_1 \epsilon_5 p_{\nu(m)} + (t - 2c_1 \epsilon_5) p_{\nu(m)} \leq t p_m$$

□

Im Weiteren bezeichnet $\mathcal{T}_m^j(K)$ für $j \in \{0, 1, \dots, d\}$ die Menge aller $\tau \in \mathcal{T}_m(K)$ mit der Eigenschaft, dass $\tau + B(0, t p_{\nu(m)})$ genau j der Hyperebenen, welche den Rand von Ω bilden schneiden.

Der zweite Schritt des Beweises für Satz 2.4 ist nun die folgende Aussage:

Satz 2.6 (Penrose [25])

Sei $j \in \{0, 1, \dots, d\}$ und $K > \max(1, t)$, dann gilt:

$$\limsup_{m \rightarrow \infty} |\mathcal{T}_m^j(K)| \cdot p_{\nu(m)}^{d-j} < \infty$$

Beweis zu Satz 2.6.

Betrachtet man ein $\tau \in \mathcal{T}_m(K)$, so bezeichnet $y(\tau)$ das erste Element von τ bezüglich der lexikographischen Ordnung auf $\mathcal{L}_{\nu(m)}$. Zusätzlich sei $\tau' = \{(\epsilon_5 p_{\nu(m)})^{-1}(y - y(\tau)) \mid y \in \tau\}$. Dadurch kann τ durch $y(\tau)$ und τ' eindeutig bestimmt werden.

2 Geometrische Zufallsgraphen

Es gilt ausserdem, dass $\tau' \subset \mathbb{Z}^d \cap B(0, 3\frac{K}{\epsilon_5})$, wobei die Anzahl dieser Teilmengen eine Konstante ist, welche unabhängig von m ist. Daher ist $|T_m^j(K)|$ beschränkt durch eine Konstante mal der Anzahl an Möglichkeiten für $y(\tau)$ mit $\tau \in \mathcal{T}_m^j(K)$.

Wenn $j = 0$ gilt, dann ist die Anzahl der Möglichkeiten für $y(\tau)$ beschränkt durch die Anzahl der Elemente von $\mathcal{L}_n$, daher ist $|\mathcal{T}_m^j(K)|$ beschränkt durch $c \cdot p_{\nu(m)}^{-d}$, womit die Aussage für diesen Fall folgt.

Gilt $j > 0$ und $\tau \in \mathcal{T}_m^j(K)$, dann ist der Abstand von $y(\tau)$ zum Rand ∂_j in der l_∞ -Norm höchstens $3Kp_{\nu(m)}$. Daher kann ein $y_0(\tau) \in \mathcal{L}_n$ gewählt werden, welches $\|y_0(\tau) - y(\tau)\|_\infty \leq 4Kp_{\nu(m)}$ und $C_n(y_0) \cap \partial_j \neq \emptyset$ erfüllt.

Die Anzahl der Möglichkeiten für $y_0(\tau)$, welche die zweite Bedingung erfüllen, ist durch $c \cdot p_n^{-(d-j)}$ beschränkt. Für gegebenes $y_0(\tau)$ ist auch die Anzahl der Möglichkeiten für $y(\tau)$ wegen der ersten Bedingung beschränkt durch eine Konstante. Dadurch folgt die Aussage für den Fall $j > 0$. $\square$

Mit Hilfe dieses Satzes kann nun die wichtige Aussage von Satz 2.7 gezeigt werden, welche den Beweis von Satz 2.2 abschließt.

Satz 2.7

Sei $j \in \{0, 1, \dots, d\}$, dann gilt mit Wahrscheinlichkeit 1, dass das Ereignis $\cup_{\tau \in \mathcal{T}_m^j(K)} F_m(\tau)$ nur für endlich viele m eintritt.

Beweis zu Satz 2.7.

Sei $\tau \in \mathcal{T}_m^j(K)$.

Im Fall $j > 0$ sei angenommen, dass $\Omega = \prod_{i=1}^d [0, \beta_i]$ gilt und die j Hyperebenen, welche den Rand von Ω bilden und von $\tau + B(0, tp_{\nu(m)})$ geschnitten werden, die Hyperebenen $(\pi_1^{-1}(\{0\}), \pi_2^{-1}(\{0\}), \dots, \pi_j^{-1}(\{0\}))$ sind. Dabei ist $\pi_i : \mathbb{R}^d \rightarrow \mathbb{R}$ die Projektion auf die i -te Koordinate.

Definiere für $r > 0$ die Kugel $B^+(0, r)$ durch

$$B^+(0, r) = \{x = (x_1, \dots, x_d) \in B(0, r) \mid x_1 \geq 0, \dots, x_j \geq 0\}$$

, wobei im Fall $j = 0$ gelten soll, dass $B^+(0, r) = B(0, r)$. Ausserdem sei $C'_n(y) = C_n(y) \cap \Omega$. Gilt nun für ein $y \in \tau$, dass $x \in C'_{\nu(m)}(y)$ und $w \in B^+(0, (t - 3c_1\epsilon_5)p_{\nu(m)})$, so folgt $x + w \in \Omega$ und weiters durch die Dreiecksungleichung $x + w \in B(y, t - 2c_1\epsilon_5p_{\nu(m)})$. Daraus folgt:

$$A_m(\tau) \cap \Omega \subseteq \left(\cup_{y \in \tau} C'_{\nu(m)}(y) \right) + B^+(0, (t - 3c_1\epsilon_5p_{\nu(m)})) \setminus \left(\cup_{y \in \tau} C'_{\nu(m)}(y) \right)$$

Die Brunn-Minkowski Ungleichung (Satz 1.8) impliziert die folgende Ungleichung für kompakte, nicht leere Mengen $A, B \in \mathbb{R}^d$:

$$\lambda(A + B) \geq \lambda(A) + \lambda(B),$$

woraus schließlich

$$\lambda(A_m(\tau) \cap \Omega) \geq 2^{-j} \theta \left((t - 3c_1 \epsilon_5) p_{\nu(m)} \right)^d$$

folgt. Für $\tau \in \mathcal{T}_m^j(K)$ und m hinreichend groß gilt deshalb:

$$\mathbb{E}[\mathcal{X}_{\nu(m)}(A_m(\tau))] \geq 2^{-j} f_j \theta(t - 4c_1 \epsilon_5)^d \nu(m) p_{\nu(m)}^d =: \mu_m$$

Im Fall $b = \infty$ gilt $\nu(m) p_{\nu(m)}^d = k_m$ und durch die Bedingung (2.10) folgt für hinreichend große m die Ungleichung $\mu_m \geq (1 + \epsilon_5) k_m$. Betrachtet man nun die Definition des Ereignis $F_m(\tau) = \{\mathcal{X}_m[A_m(\tau)] \leq k_m\}$, so folgt mit Hilfe von Satz 1.3, dass ein $\delta > 0$ existiert, sodass für hinreichend große m und $\tau \in \mathcal{T}_m^j(K)$ die Wahrscheinlichkeit für das Ereignis $F_m(\tau)$ wie folgt abgeschätzt werden kann:

$$\mathbb{P}[F_m(\tau)] \leq e^{-\delta k_m}$$

Damit lässt sich

$$\mathbb{P}\left[\bigcup_{\tau \in \mathcal{T}_m^j(K)} F_m(\tau)\right] \leq |\mathcal{T}_m^j(K)| \cdot e^{-\delta k_m}$$

folgern. Durch die Aussage von Satz 2.6 und der Voraussetzung $\frac{k_m}{\log(m)} \rightarrow \infty$ gilt, dass diese Wahrscheinlichkeiten summierbar sind, das heißt $\sum_{m \in \mathbb{N}} \mathbb{P}\left[\bigcup_{\tau \in \mathcal{T}_m^j(K)} F_m(\tau)\right] < \infty$ ist. Durch das erste Lemma von Borel-Cantelli (Satz 1.2) folgt damit die Aussage des Satzes für den Fall $b = \infty$.

Zu zeigen bleibt noch der Fall $b < \infty$, womit $\nu(m) p_{\nu(m)}^d = \log(\nu(m))$ gilt. Sei $j \in \{0, \dots, d-1\}$. Durch die spezielle Wahl von ϵ_5 so, dass

$$\frac{\theta(t - 4c_1 \epsilon_5)^d}{1 + \epsilon_5} > \max \left\{ \max \left\{ \left(\frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1 \right), \frac{2^d(b + 2\epsilon_5)}{f_d} \right\} \right\}$$

gilt, folgt

$$dH(a_j) \mu_m \geq (d-j)(1 + \epsilon_5) \log(\nu(m)). \quad (2.15)$$

Da nach Voraussetzung im Fall $b < \infty$ auch gilt, dass $\frac{k_{\nu(m+1)}}{\log(m)} \rightarrow b$ folgt durch die Bedingung $\frac{a_j}{H(a_j)} = \frac{bd}{d-j}$ für hinreichend große m :

$$\frac{k_{\nu(m+1)}}{\mu_m} \leq \frac{bdH(a_j)}{d-j} = a_j$$

2 Geometrische Zufallsgraphen

Wenn nun $\tau \in \mathcal{T}_m^j(K)$ ist, gilt wegen der Definition des Ereignis $F_m(\tau) = \{\mathcal{X}_{\nu(m)}[A_m(\tau)] \leq k_{\nu(m+1)}\}$, Satz 1.3 und Ungleichung 2.15 die folgende Abschätzung:

$$\mathbb{P}[F_m(\tau)] \leq e^{-\mu_m H(a_j)} \leq m^{\frac{(1+\epsilon_5)J(j-d)}{d}}$$

Somit gilt nach Satz 2.6 und der Definition von $p_n = \left(\frac{\log(n)}{n}\right)^{\frac{1}{d}}$, dass für hinreichend große m $|\mathcal{T}_m^j(K)| \leq m^{\frac{J(j-d)}{d}}$ und damit:

$$\mathbb{P}\left[\bigcup_{\tau \in \mathcal{T}_m^j(K)} F_m(\tau)\right] \leq |\mathcal{T}_m^j(K)| \cdot m^{\frac{(1+\epsilon_5)J(j-d)}{d}} \leq m^{\frac{J\epsilon_5(j-d)}{d}}$$

Durch die obige Wahl von J ist die Summe über die Wahrscheinlichkeiten wieder summierbar, das heißt $\sum_{m \in \mathbb{N}} \mathbb{P}[\bigcup_{\tau \in \mathcal{T}_m^j(K)} F_m(\tau)] < \infty$ und damit folgt mit dem ersten Lemma von Borel-Cantelli wiederum die Aussage.

Übrig bleibt der Fall $b < \infty$ und $j = d$. In diesem Fall gilt $\mu_m \geq (b + 2\epsilon_5) \cdot \log(\nu(m))$, da ϵ_5 so gewählt wurde, dass $\frac{\theta(t-4c_1\epsilon_5)^d}{1+\epsilon_5} > \max\left\{\max\left\{\frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1\right\}, \frac{2^d(b+2\epsilon_5)}{f_d}\right\}$ gilt.

Da $\frac{k_{\nu(m+1)}}{\log(\nu(m))} \rightarrow b$ gilt, folgt mit Satz 1.3 für hinreichend großes m :

$$\mathbb{P}[F_m(\tau)] \leq e^{-H(\frac{b+\epsilon_5}{b+2\epsilon_5})\epsilon_5 \log(\nu(m))}$$

und damit wegen Satz 2.6:

$$\mathbb{P}\left[\bigcup_{\tau \in \mathcal{T}_m^d(K)} F_m(\tau)\right] \leq cm^{-J\epsilon_5 H(\frac{b+\epsilon_5}{b+2\epsilon_5})}.$$

Über die Wahrscheinlichkeiten kann wieder summiert werden, und durch das erste Lemma von Borel-Cantelli folgt die Aussage für den Fall $b < \infty$ und $j = d$. $\square$

Beweis zu Satz 2.4.

Kombiniert man nun Satz 2.7 mit Satz 2.5 so folgt, dass das Ereignis $E'_n(K, t)$ nur für endlich viele n eintritt - dadurch folgt die Aussage von Satz 2.4, welche eben dies besagt, beziehungsweise anders ausgedrückt, dass $M_{k_n}(\mathcal{X}_n) \leq tp_n$ für alle bis auf endlich viele n gilt. $\square$

Mit diesen Aussagen kann nun das asymptotische Verhalten des größten k -nächsten Nachbarn Radius gezeigt werden:

Beweis zu Satz 2.2.

Im Fall $b = \infty$ ist zu zeigen, dass $\lim_{n \rightarrow \infty} \frac{n\theta M_{k_n}(\mathcal{X}_n)^d}{k_n} = \max\{\frac{2^j}{f_j} \mid 0 \leq j \leq d\}$. Laut Satz 2.3 gilt $M_{k_n}(\mathcal{X}_n) > p_n^j(\beta)$ für alle $\beta \in (0, 1)$, $j \in \{0, 1, \dots, d\}$ und hinreichend große n . Das heißt es gilt $\frac{M_{k_n}(\mathcal{X}_n)^{dn\theta}}{k_n} \geq \frac{2^j}{f_j}$ und mit der Grenzwertbetrachtung $b \rightarrow 1$ folgt schließlich $\frac{M_{k_n}(\mathcal{X}_n)^{dn\theta}}{k_n} \geq \max\{\frac{2^j}{f_j} \mid 0 \leq j \leq d\}$, da die Ungleichung für alle $j \in \{0, 1, \dots, d\}$ gilt.

Laut Satz 2.4 gilt ebenso $M_{k_n}(\mathcal{X}) \leq tp_n$ für alle bis auf endlich viele n , womit $\frac{nM_{k_n}(\mathcal{X}_n)^d}{k_n} \leq t^d$ folgt. Nach der Wahl von t folgt $\frac{n\theta M_{k_n}(\mathcal{X}_n)^d}{k_n} \leq \max\{\frac{2^j}{f_j} \mid 0 \leq j \leq d\}$ und damit die Gleichheit.

Im Fall $b < \infty$ gelten dieselben Aussagen der Sätze 2.3 und 2.4, es ändern sich jedoch die Definitionen von p_n^j , p_n^d und die Wahl von t . Zu zeigen ist $\lim_{n \rightarrow \infty} \frac{n\theta M_{k_n}(\mathcal{X}_n)^d}{k_n} = \max\left\{\max\left\{\frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1\right\}, \frac{2^d b}{f_d}\right\}$.

Durch Satz 2.3 gilt $M_{k_n}(\mathcal{X}_n) > p_n^j(\beta)$ für alle $\beta \in (0, 1)$ und $j \in \{0, \dots, d\}$ - damit folgt mit der Grenzwertbildung $\beta \rightarrow 1$ die folgende Abschätzung nach unten: $\frac{n\theta M_{k_n}(\mathcal{X}_n)^d}{\log(n)} \geq \frac{2^j(d-j)}{df_j H(a_j)}$. Mit Satz 2.4 und der Wahl von t so, dass $\theta t^d > \max\left\{\max\left\{\frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1\right\}, \frac{2^d b}{f_d}\right\}$ folgt die Gleichheit und damit die Aussage von Satz 2.2. $\square$

Asymptotisches Verhalten des minimalen Knotengrades δ_n von $G(\mathcal{X}_n, r_n)$

Mit Hilfe dieser Vorarbeit lässt sich nun die Hauptaussage von Abschnitt 2.2.2 zeigen, welche das asymptotische Verhalten des minimalen Knotengrades $\delta_n = \delta(G(\mathcal{X}_n, r_n))$ von $G(\mathcal{X}_n, r_n)$ beschreibt.

Satz 2.8 (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte auf Ω , wobei $f|_{\Omega}$ stetig in $x \forall x \in \delta\Omega$ ist. Zusätzlich gelte $r_n \rightarrow 0$ und $\frac{n\theta r_n^d}{\log(n)} \rightarrow \alpha \in [0, \infty]$ für $n \rightarrow \infty$ und δ_n bezeichnet den minimalen Knotengrad von $G(\mathcal{X}_n, r_n)$. Dann gelten folgende Aussagen:

1. Wenn $\alpha < \max_{j \in J \setminus \{d\}} \{2^j(d-j)/(df_j)\}$, dann gilt $\delta_n \rightarrow 0$ fast sicher.
2. Wenn $\max_{j \in J \setminus \{d\}} \{2^j(d-j)/(df_j)\} \leq \alpha \leq \infty$, dann gilt mit Wahrscheinlichkeit 1:

$$\delta_n \sim (n \theta r_n^d) \cdot c$$

wobei

$$c = \min \left\{ \min_{0 \leq j \leq d-1} \left\{ \frac{f_j}{2^j} H^{-1} \left(\frac{2^j(d-j)}{df_j a} \right) \right\}, \frac{f_d}{2^d} \right\}$$

gilt.

Bemerkung. Betrachtet man in obigem Satz 2.8 über das asymptotische Verhalten des minimalen Knotengrades δ_n von $G(\mathcal{X}_n, r_n)$ eine beliebige Dimension $d \geq 2$, und setzt vom kompakten Träger Ω von f voraus, dass Ω beschränkt in $\mathbb{R}^d$ und der Rand $\partial\Omega$ eine $(d-1)$ -dimensionale C^2 Untermannigfaltigkeit von $\mathbb{R}^d$ ist, so gilt mit $\mathcal{J} := \{0, 1\}$ für dieses allgemeinere Gebiet im 2. Fall des obigen Satzes dieselbe Aussage, jedoch mit geänderter Konstante c :

$$c = \min_{j \in \mathcal{J}} \left\{ \frac{f_j}{2^j} H^{-1} \left(\frac{2^j(d-j)}{df_j \alpha} \right) \right\}$$

Beweis zu Satz 2.8.

1. Fall: Definiere $k_n = 1 \ \forall n \in \mathbb{N}$, dann gilt $\lim_{n \rightarrow \infty} \frac{k_n}{n} = 0$ und $\lim_{n \rightarrow \infty} \frac{k_n}{\log(n)} = 0$, womit Satz 2.2 angewandt werden kann. Aus $b = 0$ folgt $a_0, \dots, a_{d-1} = 0$ und damit nach Definition von H : $H(a_j) = 1 \ \forall j \in \{1, \dots, d-1\}$. Damit gilt mit Wahrscheinlichkeit 1:

$$\lim_{n \rightarrow \infty} \left(\frac{n\theta(M_1(\mathcal{X}_n))^d}{\log(n)} \right) = \max_{0 \leq j \leq d-1} \left(\frac{2^j(d-j)}{d \cdot f_j H(a_j)} \right) \quad (2.16)$$

Da laut Voraussetzung $\alpha < \max_{j \in J \setminus \{d\}} \{2^j(d-j)/(df_j)\}$ gilt, folgt $M_1(\mathcal{X}_n) > r_n$ für ausreichend großes n , und damit $\delta_n = 0$ für ausreichend großes n .

2. Fall: Es sei $\max_{j \in J \setminus \{d\}} \{2^j(d-j)/(df_j)\} \leq \alpha \leq \infty$. Für $b > 0$ definiere $a_j \in (0, 1)$ mittels $a_j/H(a_j) = bd/(d-j) \ \forall j \in \mathcal{J} \setminus \{d\}$. Zusätzlich definiere $\gamma_j(b) = \frac{2^j(d-j)}{df_j H(a_j)} \ \forall j \in \mathcal{J} \setminus \{d\}$ und $\gamma_d = \frac{2^d b}{f_d}$. $\gamma(b)$ sei das Minimum aller γ_j , das heißt $\gamma(b) = \min_{j \in \mathcal{J}} \gamma_j(b)$.

$(k_n)_{(n \geq 1)}$ sei eine nicht fallende Folge, welche $\frac{k_n}{\log(n)} \rightarrow b$ erfüllt. Damit gilt laut Satz 2.2 mit Wahrscheinlichkeit 1:

$$\liminf_{n \rightarrow \infty} \left(\frac{n\theta M_{k_n}(\mathcal{X}_n)^d}{\log(n)} \right) = \max_{j \in \mathcal{J}} \gamma_j(b) = \gamma(b)$$

$\gamma_j(\cdot)$ ist eine stetige, streng monoton steigende Funktion für alle $j \in \mathcal{J}$. Damit ist auch $\gamma(\cdot)$ eine streng monoton steigende und stetige Funktion. Sei $b < \gamma^{-1}(\alpha) < b'$ und definiere zwei nicht fallende Folgen $(k_n)_{(n \geq 1)}$, sowie $(k'_n)_{(n \geq 1)}$, sodass $\frac{k_n}{\log(n)} \rightarrow b$ und $\frac{k'_n}{\log(n)} \rightarrow b'$ für $n \rightarrow \infty$ gilt. Aus Formel (2.16) folgt, dass $M_{k_n}(\mathcal{X}_n) < r_n < M_{k'_n}(\mathcal{X}_n)$ und damit $k_n < \delta_n < k'_n$. Dividiert man die letzte Gleichung durch $\log(n)$ und bildet den Grenzwert

$n \rightarrow \infty$, so gilt mit Wahrscheinlichkeit 1:

$$b \leq \liminf(\frac{\delta_n}{\log(n)}) \leq \limsup(\frac{\delta_n}{\log(n)}) \leq b'$$

Wählt man nun $b = \lim_{x \rightarrow \alpha, x \leq \alpha} \gamma^{-1}(x)$ und $b' = \lim_{x \rightarrow \alpha, x \geq \alpha} \gamma^{-1}(x)$, dann gilt:

$$\lim_{n \rightarrow \infty} (\frac{\delta_n}{n\theta r_n^d}) = \alpha^{-1} \gamma^{-1}(\alpha) = \alpha^{-1} \min_{j \in \mathcal{J}} \{\gamma_j^{-1}(\alpha)\} \quad (2.17)$$

fast sicher. Wählt man für $j \in \mathcal{J} \setminus \{d\}$ $b = \gamma_j^{-1}(\alpha)$ und $a_j < 1$ so, dass $\frac{a_j}{H(a_j)} = \frac{bd}{(d-j)}$, dann ergeben sich durch die Definition von γ_j folgende Identitäten:

$$a_j = H_-^{-1} \left(\frac{2^j(d-j)}{df_j\alpha} \right) \text{ und } b = \frac{f_j\alpha}{2^j} H_-^{-1} \left(\frac{2^j(d-j)}{df_j\alpha} \right).$$

Damit ergibt sich für die Inversen von γ folgende Darstellung für $j \in \mathcal{J} \setminus \{d\}$:

$$\gamma^{-1}(\alpha) = 2^{-j} f_j \alpha H_-^{-1} \left(\frac{2^j(d-j)}{df_j\alpha} \right), \text{ sowie } \gamma_d^{-1}(\alpha) = 2^{-d} f_d \alpha. \quad (2.18)$$

Im Fall $\alpha < \infty$ ergibt sich die Aussage von Satz 2.8 nun durch einsetzen der Darstellungen 2.18 in 2.16:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\frac{\delta_n}{n\theta r_n^d} \right) &= \alpha^{-1} \min_{j \in \mathcal{J}} \{\gamma_j^{-1}(\alpha)\} = \min_{j \in \mathcal{J}} \{\alpha^{-1} \cdot \gamma_j^{-1}(\alpha)\} \\ &= \min \left\{ \min_{j \in \mathcal{J} \setminus \{d\}} \left\{ \frac{f_j}{2^j} H_-^{-1} \left(\frac{2^j(d-j)}{df_j\alpha} \right) \right\}, \frac{f_d}{2^d} \right\} \end{aligned}$$

Sei $\alpha = \infty$, wenn $\frac{k_n}{\log(n)} \rightarrow \infty$ und $\frac{k_n}{n} \rightarrow 0$, dann gilt laut Satz 2.2 folgende Aussage mit Wahrscheinlichkeit 1, wenn $n \rightarrow \infty$:

$$\frac{n\theta M_{k_n}(\mathcal{X}_n)^d}{k_n} \cdot \min_{j \in \mathcal{J}} \left(\frac{f_j}{2^j} \right) \rightarrow \max_{j \in \mathcal{J}} \left(\frac{2^j}{f_j} \right) \cdot \min_{j \in \mathcal{J}} \left(\frac{f_j}{2^j} \right) = 1 \quad (2.19)$$

Sei $\epsilon > 0$, setze $k_n = \lfloor (1 - \epsilon)n\theta r_n^d \min_{j \in \mathcal{J}} (2^{-j} f_j) \rfloor$ und $k'_n = \lfloor (1 + \epsilon)n\theta r_n^d \min_{j \in \mathcal{J}} (2^{-j} f_j) \rfloor$. Dann gilt $\frac{k_n}{\log(n)} \rightarrow \infty$ und $\frac{k_n}{n} \rightarrow 0$ für $n \rightarrow \infty$. Daraus folgt mit Formel (2.19) mit Wahrscheinlichkeit 1:

$$\frac{n\theta M_{k_n}(\mathcal{X}_n)^d \min_{j \in \mathcal{J}} (2^{-j} f_j)}{(1 - \epsilon)n\theta r_n^d \min_{j \in \mathcal{J}} (2^{-j} f_j)} \rightarrow 1 \quad (2.20)$$

Aus 2.20 folgt $M_{k_n}(\mathcal{X}_n) < r_n$ und $M_{k'_n} > r_n$ für ausreichend großes n . Da $M_{k_n}(\mathcal{X}_n)$ der Radius ist, sodass der Minimalgrad von $G(\mathcal{X}_n, M_{k_n}(\mathcal{X}_n))$ ist, folgt $k_n \leq \delta_n$ und durch dasselbe Argument folgt $\delta_n \leq k'_n$, gesamt ergibt sich mit Wahrscheinlichkeit 1 $k_n \leq \delta_n \leq k'_n$ für ausreichend großes n . Lässt man $\epsilon \downarrow 0$ streben, so erhält man die Aussage von Satz 2.8 für $\alpha = \infty$: $\frac{\delta_n}{n\theta r_n^d} \rightarrow \min_{j \in \mathcal{J}}(2^{-j} f_j)$. $\square$

Bemerkung. Satz 2.8 ist eine Verallgemeinerung einer Aussage, welche bereits 1997 von Appel und Russo für den Spezialfall einer uniformen Verteilung und der l_∞ -Norm in [2] gezeigt wurde.

Sie verwendeten in ihrem Werk eine sehr ähnliche Hilfsfunktion, welche auch hier verwendet wurde. Diese wurde durch $F(t) = \frac{1}{t} \log(t) + \frac{1}{t} - 1$ definiert. Die Aussage über das asymptotische Verhalten des minimalen Knotengrades δ_n unter den Einschränkungen lautete in diesem Fall:

Satz 2.9 (Appel und Russo [2])

Sei f_U die Dichte der uniformen Wahrscheinlichkeitsverteilung auf $[0, 1]^d$ und die verwendete Norm die l_∞ -Norm. Sei $(r_n)_{(n \geq 1)}$ eine Folge, welche $\frac{nr_n^d}{\log(n)} \rightarrow \alpha \in (0, \infty]$ für $n \rightarrow \infty$ erfüllt. Dann gelten folgende Aussagen:

1. $\forall \alpha > 0 : \limsup_{n \rightarrow \infty} \frac{\delta_n}{nr_n^d} \leq 1$ gilt fast sicher.
2. Wenn $\alpha < (2d)^{-1}$, dann gilt $\delta_n \rightarrow 0$ fast sicher.
3. Wenn $\alpha > (2d |F(\frac{2}{a})|)^{-1}$ mit $0 < a < 1$ gilt, dann folgt fast sicher:
 $\liminf_{n \rightarrow \infty} \frac{\delta_n}{nr_n^d} \geq a$

Neben dieser Aussage, zeigten sie auch das asymptotische Verhalten des größten k -nächsten Nachbarn Radius $M_k(\mathcal{X})$ in diesem Spezialfall, verwendeten diese Aussage jedoch nicht wie Penrose im Beweis für Satz 2.9:

Satz 2.10 (Appel und Russo [2])

Sei f_U die Dichte der uniformen Wahrscheinlichkeitsverteilung auf $[0, 1]^d$ und die verwendete Norm die l_∞ -Norm. Dann gilt folgende Aussage:

$$M_n(\mathcal{X}_n)^d \sim \frac{\log(n)}{n} \cdot \frac{1}{2^d}$$

2.2.3 Maximaler Knotengrad

Ähnliche Aussagen wie jene für den minimalen Knotengrad können auch für den maximalen Knotengrad $\Delta_n = \Delta(G(\mathcal{X}_n, r_n))$ eines geometrischen Zufallsgraphen $G(\mathcal{X}_n, r_n)$ gezeigt werden. Wie im vorigem Abschnitt, werden auch hier zuvor Eigenschaften des

kleinsten k -nächsten Nachbarn Radius $S_k(\mathcal{X}_n)$ gezeigt, um daraus auf das Verhalten des maximalen Knotengrades zu schließen.

Definition 2.5: kleinster k -nächsten Nachbarn Radius (Penrose [25])

Sei $D_k(X_i)$ für $X_i \in \mathcal{X}$ die Distanz von X_i zu seinen k nächsten Nachbarn in $\mathcal{X}$. Der **kleinste k -nächsten Nachbarn Radius**³ $S_k(\mathcal{X})$ ist definiert als das Minimum aller $D_k(X_i)$ für $X_i \in \mathcal{X}$, das heißt $S_k(\mathcal{X}) := \min_{X_i \in \mathcal{X}} D_k(X_i)$.

Analog zum größten k -nächsten Nachbarn Radius, stellt $S_k(\mathcal{X})$ sicher, dass zumindest ein Punkt des geometrischen Graphen $G(\mathcal{X}, S_k(\mathcal{X}))$ mit k Knoten verbunden ist. Damit hat der resultierende Graphen zumindest maximalen Knotengrad k .

Als erstes Ergebnis ist hier das asymptotische Verhalten des kleinsten k -nächsten Radius $S_k(\mathcal{X}_n)$ aufgeführt, wenn die Folge $(k_n)_{(n \geq 1)}$ zumindest logarithmisch wächst:

Satz 2.11 (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte mit kompaktem Träger, $(k_n)_{(n \geq 1)}$ eine Folge natürlicher Zahlen, welche $\frac{k_n}{\log(n)} \rightarrow b$ und $\frac{k_n}{n} \rightarrow 0$ für ein $b \in (0, \infty]$ erfüllt. Zusätzlich sei $a \geq 1$ durch $\frac{a}{H(a)} = b$ definiert, wobei H die in Abschnitt 2.2.2 definierte Hilfsfunktion bezeichnet. Dann gilt mit Wahrscheinlichkeit 1:

$$S_{k_n}(\mathcal{X}_n)^d \sim \frac{k_n}{n\theta} \cdot \frac{1}{f_{\max} a}$$

Im Falle $b = \infty$ folgt $a = 1$.

Beweisidee zu Satz 2.11

Der Beweis von Satz 2.11 wird recht ähnlich geführt wie der Beweis im vorigen Kapitel 2.2.2 über das asymptotische Verhalten des größten k -nächsten Nachbarn Radius - hier wird der kleinste k -nächsten Nachbarn Radius $S_{k_n}(\mathcal{X}_n)$ geeignet asymptotisch nach oben und unten abgeschätzt, womit man die Aussage von Satz 2.11 schließlich zeigen kann.

Sei für die folgenden Aussagen $p_n(u) = \left(\frac{k_n u}{n\theta f_{\max} a} \right)^{\frac{1}{d}}$ definiert für $u > 0$ und $n \in \mathbb{N}$. Dann gelten folgende zwei Sätze, welche hier lediglich erwähnt werden und nicht mehr bewiesen werden:

Satz 2.12 (Penrose [25])

Sei $\lim_{n \rightarrow \infty} \frac{k_n}{\log(n)} = b \in (0, \infty]$ und $a \geq 1$, wobei a die Gleichung $\frac{a}{H(a)} = b$ erfülle. Zusätzlich sei $\beta > 1$ und $p_n(u)$ wie oben definiert, dann gilt mit Wahrscheinlichkeit 1 für alle hinreichend großen $n \in \mathbb{N}$:

$$S_{k_n}(\mathcal{X}_n) \leq p_n(\beta)$$

³In der englischsprachigen Literatur wird der Begriff *smallest k -nearest neighbour link* verwendet.

Die Formulierung der asymptotischen Abschätzung von $S_{k_n}(\mathcal{X}_n)$ nach unten ist die folgende Aussage:

Satz 2.13 (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte mit kompaktem Träger, $(k_n)_{(n \geq 1)}$ eine Folge, welche $\lim_{n \rightarrow \infty} \frac{k_n}{\log(n)} = b \in (0, \infty]$ erfüllt und $a \geq 1$ so, dass $\frac{a}{H(a)} = b$ gilt. Zusätzlich sei $\beta < 1$ und $p_n(u)$ wie oben definiert. Dann gilt mit Wahrscheinlichkeit 1 für hinreichend große $n \in \mathbb{N}$:

$$S_{k_n}(\mathcal{X}_n) \geq p_n(\beta)$$

Durch Beweis dieser beiden Sätze lässt sich schlussendlich Satz 2.11 direkt und leicht zeigen. Die Beweisdetails dazu finden sich in *Random Geometric Graphs* von Penrose ([25]).

Astymptotisches Verhalten des maximalen Knotengrades Δ_n von $G(\mathcal{X}_n, r_n)$

Für die in Abschnitt 2.2.2 definierte Hilfsfunktion $H: [0, \infty) \rightarrow \mathbb{R}$, mit $H(0) = 1$ und $H(a) = 1 - a + a \cdot \log(a)$ für $a > 0$ wird nun die Umkehrfunktion von $H|_{[1, \infty)}$ benötigt, welche im Folgenden als H_+^{-1} bezeichnet wird, und von $[0, \infty)$ nach $[1, \infty)$ abbildet.

Die folgende Aussage über den maximalen Knotengrad eines geometrischen Zufallsgraphen ist etwas allgemeiner als die analoge Aussage von Satz 2.8 über den minimalen Knotengrad, da weniger starke Voraussetzungen an den Träger Ω der Wahrscheinlichkeitsdichte f gestellt sind. Der Beweis wird unter Verwendung von Satz 2.11 geführt:

Satz 2.14 (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte mit kompaktem Träger Ω , $\alpha \in (0, \infty]$ und $(r_n)_{(n \geq 1)}$ eine Folge, welche $r_n \rightarrow 0$ und $\frac{n\theta r_n^d}{\log(n)} \rightarrow \alpha$ für $n \rightarrow \infty$ erfüllt. Dann gilt folgende Aussage fast sicher:

$$\Delta_n \sim (n\theta r_n^d) \cdot c$$

wobei

$$c = f_{\max} H_+^{-1} \left(\frac{1}{f_{\max} \alpha} \right)$$

gilt.

Beweis zu Satz 2.14.

1. Fall: Sei $\alpha < \infty$. Wähle $b > 0$ und definiere $\alpha > 1$ durch $\frac{\alpha}{H(\alpha)} = b$. Definiere zusätzlich $\gamma(b) = \frac{1}{f_{max}H(a)}$. Ist $(k_n)_{(n \geq 1)}$ eine Folge, welche $\frac{k_n}{\log(n)} \rightarrow b$ erfüllt, dann folgt durch Satz 2.11 mit Wahrscheinlichkeit 1:

$$\lim_{n \rightarrow \infty} \left(\frac{n\theta(S_{k_n}(\mathcal{X}_n))^d}{k_n} \cdot \frac{k_n}{\log(n)} \right) = \frac{b}{f_{max}\alpha} = \gamma(b) \quad (2.21)$$

$\gamma(b)$ ist eine stetige und streng monoton steigende Funktion in b . Wählt man b , sowie b' durch $b < \gamma^{-1}(\alpha) < b'$ und wählt die Folgen $(k_n)_{(n \geq 1)}$, sowie $(k'_n)_{(n \geq 1)}$ durch $\frac{k_n}{\log(n)} \rightarrow b$ und $\frac{k'_n}{\log(n)} \rightarrow b'$, dann folgt mit Formel (2.21) und einer analogen Aussage über $S_{k'_n}$:

$$\lim_{n \rightarrow \infty} \left(\frac{n\theta(S_{k_n}(\mathcal{X}_n))^d}{\log(n)} \right) < \alpha < \lim_{n \rightarrow \infty} \left(\frac{n\theta(S_{k'_n}(\mathcal{X}_n))^d}{\log(n)} \right) \quad (2.22)$$

Aus Formel (2.22) lässt sich für ausreichend großes $n \in \mathbb{N}$ $S_{k_n} < r_n < S_{k'_n}$ folgern, woraus wiederum $k_n \leq \Delta_n \leq k'_n$ folgt. Mit Wahrscheinlichkeit 1 gilt folgende Aussage:

$$b \leq \liminf_{n \rightarrow \infty} \left(\frac{\Delta_n}{\log(n)} \right) \leq \limsup_{n \rightarrow \infty} \left(\frac{\Delta_n}{\log(n)} \right) \leq b'$$

Betrachtet man die Grenzwerte $b = \lim_{x \rightarrow \alpha, x \leq \alpha} \gamma^{-1}(x)$ und $b' = \lim_{x \rightarrow \alpha, x \geq \alpha} \gamma^{-1}(x)$, so folgt $\frac{\Delta_n}{\log(n)} \rightarrow \gamma^{-1}(\alpha)$ fast sicher.

Wähle $b = \gamma^{-1}(\alpha)$. Durch die Definition von γ gilt $H(a) = (f_{max}\alpha)^{-1}$ und folgende Identitäten für α und b :

$$\begin{aligned} a &= H_+^{-1} \left(\frac{1}{f_{max}\alpha} \right) \\ b &= f_{max}\alpha H_+^{-1} \left(\frac{1}{f_{max}\alpha} \right) \end{aligned}$$

Insgesamt folgt daraus die Aussage von Satz 2.14 mit Wahrscheinlichkeit 1:

$$\lim_{n \rightarrow \infty} \left(\frac{\Delta_n}{\log(n)} \cdot \frac{\log(n)}{n\theta r_n^d} \right) = \frac{b}{\alpha} = f_{max} H_+^{-1} \left(\frac{1}{f_{max}\alpha} \right)$$

2. Fall: $\alpha = \infty$. Sei $\epsilon > 0$, setze $k_n = \lfloor (1 + 2\epsilon)n\theta r_n^d f_{max} \rfloor$ und $j_n = \lfloor (1 - 2\epsilon)n\theta r_n^d f_{max} \rfloor$. Dann gilt $\frac{k_n}{\log(n)} \rightarrow \infty$ und $\frac{j_n}{\log(n)} \rightarrow \infty$. Aus Satz 2.11 folgen daraus die Aussagen über S_{k_n} , sowie S_{j_n} : $\lim_{n \rightarrow \infty} \frac{n\theta S_{k_n}(\mathcal{X}_n)^d}{k_n} = \frac{1}{f_{max}}$ und $\lim_{n \rightarrow \infty} \frac{n\theta S_{j_n}(\mathcal{X}_n)^d}{j_n} = \frac{1}{f_{max}}$.

Für ausreichend großes $n \in \mathbb{N}$ gelten folgende Ungleichungen:

$$\begin{aligned} n\theta S_{k_n}(\mathcal{X}_n)^d &\geq \left(\frac{1+\epsilon}{1+2\epsilon}\right) \cdot \left(\frac{k_n}{f_{\max}}\right) > n\theta r_n^d \\ n\theta S_{j_n}(\mathcal{X}_n)^d &\leq \left(\frac{1-\epsilon}{1-2\epsilon}\right) \cdot \left(\frac{k_n}{f_{\max}}\right) < n\theta r_n^d \end{aligned}$$

Das heißt, für ausreichend großes $n \in \mathbb{N}$ gilt: $S_{j_n}(\mathcal{X}_n) < r_n < S_{k_n}(\mathcal{X}_n)$ und damit $j_n \leq \Delta_n \leq k_n$. Mit Wahrscheinlichkeit 1 gilt damit:

$$1 - 2\epsilon \leq \liminf_{n \rightarrow \infty} \left(\frac{\Delta_n}{n\theta f_{\max} r_n^d} \right) \leq \limsup_{n \rightarrow \infty} \left(\frac{\Delta_n}{n\theta f_{\max} r_n^d} \right) \leq 1 + 2\epsilon$$

Da ϵ beliebig klein gewählt ist, folgt die Aussage von Satz 2.14. $\square$

Bemerkung. Die Aussagen über den maximalen Knotengrad und den kleinsten k -nächsten Nachbarn Radius wurden ebenfalls von Appel und Russo 1997 in [1] für den spezielleren Fall einer uniformen Verteilung der Punkte in $[0, 1]^d$ und der l_∞ -Norm gezeigt. Unter der Voraussetzung, dass die Folge der Radien $(r_n)_{(n \geq 1)}$ zusätzlich $\lim_{n \rightarrow \infty} \frac{nr_n^d}{\log(n)} = \infty$ erfüllen, gelten folgende Aussagen:

Satz 2.15 (Appel und Russo [1])

Sei f_U die Dichte der uniformen Wahrscheinlichkeitsverteilung auf $[0, 1]^d$ und die verwendete Norm die l_∞ -Norm. Zusätzlich sei $(r_n)_{(n \geq 1)}$ wie oben beschrieben und es gelte $\limsup_{n \rightarrow \infty} r_n \leq \frac{1}{2}$. Dann gilt fast sicher:

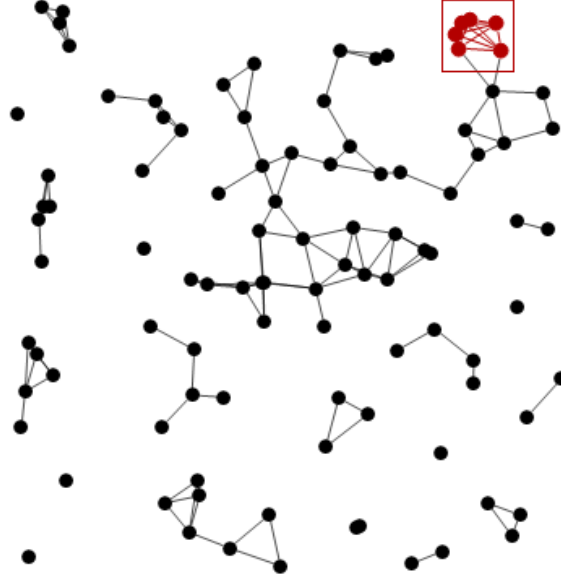
$$\Delta_n \sim 2^d n r_n^d$$

2.2.4 Clique-Zahl

Wählt man für den geometrischen Zufallsgraphen $G(\mathcal{X}, r)$ als Norm die l_∞ -Norm, so ist die Clique-Zahl $\mathcal{C}$ die maximale Nummer an Knoten, welche sich in einem d -dimensionalen Würfel mit Seitenlänge r finden.

Beispiel 2.4. Abbildung 2.2 zeigt einen 2-dimensionalen geometrischen Zufallsgraphen mit 100 Knoten, Radius 0.1 und Clique-Zahl 6. Das rote Quadrat symbolisiert einen 2-dimensionalen Würfel mit Seitenlänge 0.1, welcher die gesamte Clique umfasst.

Für den Fall eines geometrischen Zufallsgraphen, wo $\frac{n}{\log(n)} \cdot r_n^d \rightarrow c \in \mathbb{R}$ gilt, können asymptotische Aussagen über die Clique-Zahl $\mathcal{C}_n = \mathcal{C}(G(\mathcal{X}_n, r_n))$ getätigt werden. Zuvor ist noch eine Aussage über den Grenzradius $r = \rho(\mathcal{X}, \mathcal{C} \geq k)$, welcher sicherstellt, dass der geometrische Zufallsgraph $G(\mathcal{X}, r)$ eine Clique-Zahl von mindestens k hat, notwendig:


 Abbildung 2.2: $G(100, 0.1)$ mit Clique-Zahl 6

Satz 2.16

Sei f eine Wahrscheinlichkeitsdichte mit kompaktem Träger und $(k_n)_{(n \geq 1)}$ eine Folge mit $\frac{k_n}{\log(n)} \rightarrow b \in (0, \infty]$ für $n \rightarrow \infty$. Sei zusätzlich $a \geq 1$ definiert durch $\frac{a}{H(a)} = b$. Dann gilt mit Wahrscheinlichkeit 1:

$$\rho(\mathcal{X}_n, \mathcal{C} \geq k_n)^d \sim \left(\frac{k_n}{n\theta} \right) \cdot \left(\frac{2^d}{f_{\max} \alpha} \right)$$

Beweisidee zu Satz 2.16

Um diese Aussage zu beweisen, wird zuerst wieder eine Hilfsabbildung

$p_n(u) = \left(\frac{uk_n}{n\theta f_{\max} a} \right)^{\frac{1}{d}}$ für $u > 0$ und $n \in \mathbb{N}$ definiert. Anschließend wird gezeigt, dass die folgende Ungleichung für alle $\alpha < 1$ und $\beta > 1$ mit Wahrscheinlichkeit 1 für alle hinreichend große $n \in \mathbb{N}$ gilt:

$$2p_n(\alpha) \leq \rho(\mathcal{X}_n, \mathcal{C} \geq k_n) \leq 2p_n(\beta) \quad (2.23)$$

Um die rechte Seite der Formel (2.23) zu zeigen, betrachtet man einen Graphen

$G(\mathcal{X}_n, r_n)$: Hat dieser Graph einen Knoten $v \in \mathcal{X}_n$ mit Ordnung k_n , so kann man seine Nachbarn mit $w_1, \dots, w_{k_n}$ bezeichnen. Für diese Knoten gilt nun $\|w_i - w_j\| = \|w_i - v + v - w_j\| \leq \|w_i - v\| + \|v - w_j\| \leq 2r_n$ für alle $i, j \leq k_n$. Damit ist $G(\{w_1, \dots, w_{k_n}\}, 2r_n)$ ein vollständiger Graph, woraus $\mathcal{C}(G(\mathcal{X}_n, 2r_n)) \geq k_n$ sowie $\rho(\mathcal{X}_n, \mathcal{C} \geq k_n) \leq 2S_k(\mathcal{X})$ folgen.

Durch Satz 2.12 gilt $S_{k_n}(\mathcal{X}_n) \leq \left(\frac{k_n \beta}{n \theta f_{\max} a}\right)^{\frac{1}{d}} = p_n(\beta)$ für alle $\beta > 1$ und hinreichend große n , woraus schließlich $\rho(\mathcal{X}_n, \mathcal{C} \geq k_n)$ folgt.

Die linke Seite der Ungleichung erfordert etwas mehr Aufwand für den Beweis, dafür sei auf *Random Geometric Graphs* von Penrose ([25]) verwiesen.

Asymptotisches Verhalten der Clique-Zahl $\mathcal{C}_n$ von $G(\mathcal{X}_n, r_n)$

Im Folgenden bezeichne $\mathcal{C}_n$ die Clique-Zahl $\mathcal{C}$ des geometrischen Zufallsgraphen $G(\mathcal{X}_n, r)$. Die Aussage von Satz 2.17 lässt sich ähnlich beweisen, als die Aussage vom letzten Abschnitt über den maximalen Knotengrad von Satz 2.14.

Satz 2.17

Sei f eine Wahrscheinlichkeitsdichte mit kompaktem Träger. Weiters sei $\alpha \in (0, \infty]$ und $(r_n)_{(n \geq 1)}$ eine Folge, welche $\frac{n \theta r_n^d}{\log(n)} \rightarrow \alpha$ für $n \rightarrow \infty$ erfüllt. Dann gilt fast sicher:

$$\mathcal{C}_n \sim (n \theta r_n^d) \cdot c \quad (2.24)$$

wobei

$$c = \frac{f_{\max}}{2^d} \cdot H_+^{-1} \left(\frac{2^d}{f_{\max} \alpha} \right)$$

gilt.

Beweis zu Satz 2.17.

1. Fall: Sei $\alpha < \infty$. Wähle $b > 0$ und definiere $\alpha > 1$ durch $\frac{\alpha}{H(\alpha)} = b$. Definiere zusätzlich $\gamma(b) = \frac{2^d}{f_{\max} H(\alpha)}$. Ist $(k_n)_{(n \geq 1)}$ eine Folge, welche $\frac{k_n}{\log(n)} \rightarrow b$ für $n \rightarrow \infty$ erfüllt, dann folgt durch 2.16 mit Wahrscheinlichkeit 1:

$$\lim_{n \rightarrow \infty} \left(\frac{n \rho(\mathcal{X}_n, \mathcal{C} \geq k_n)^d}{\log(n)} \right) = \frac{2^d b}{f_{\max} a} = \gamma(b) \quad (2.25)$$

$\gamma(b)$ ist eine stetige und streng monoton steigende Funktion in b . Wählt man b , sowie b' durch $b < \gamma^{-1}(\alpha) < b'$ und wählt die Folgen $(k_n) \in \mathbb{N}^{\mathbb{N}}$, sowie $(k'_n) \in \mathbb{N}^{\mathbb{N}}$ durch $\frac{k_n}{\log(n)} \rightarrow b$ und $\frac{k'_n}{\log(n)} \rightarrow b'$, dann folgt mit Formel 2.25 für ausreichend großes $n \in \mathbb{N}$ $\rho(\mathcal{X}_n, \mathcal{C} \geq k_n) < r_n < \rho(\mathcal{X}_n, \mathcal{C} \geq k'_n)$ folgern, woraus wiederum $k_n < \mathcal{C}_n < k'_n$ folgt. Mit Wahrscheinlichkeit 1 gilt folgende Aussage:

$$b \leq \liminf_{n \rightarrow \infty} \left(\frac{\mathcal{C}_n}{\log(n)} \right) \leq \limsup_{n \rightarrow \infty} \left(\frac{\mathcal{C}_n}{\log(n)} \right) \leq b'$$

Betrachtet man die Grenzwerte $b = \lim_{x \rightarrow \alpha, x \leq \alpha} \gamma^{-1}(x)$ und $b' = \lim_{x \rightarrow \alpha, x \geq \alpha} \gamma^{-1}(x)$, so folgt $\frac{\mathcal{C}_n}{\log(n)} \rightarrow \gamma^{-1}(\alpha)$ fast sicher.

Wähle $b = \gamma^{-1}(\alpha)$. Durch die Definition von γ gilt $H(a) = (f_{\max}\alpha)^{-1}$ und folgende Identitäten für α und b :

$$\begin{aligned} a &= H_+^{-1} \left(\frac{2^d}{f_{\max}\alpha} \right) \\ b &= 2^{-d} f_{\max} \alpha H_+^{-1} \left(\frac{2^d}{f_{\max}\alpha} \right) \end{aligned}$$

Insgesamt folgt daraus die Aussage von Satz 2.17 mit Wahrscheinlichkeit 1:

$$\lim_{n \rightarrow \infty} \left(\frac{\mathcal{C}_n}{\log(n)} \cdot \frac{\log(n)}{n\theta r_n^d} \right) = \frac{b}{\alpha} = 2^{-d} f_{\max} H_+^{-1} \left(\frac{2^d}{f_{\max}\alpha} \right)$$

2. Fall: $\alpha = \infty$. Sei $\epsilon > 0$, setze $k_n = \lfloor (1 + 2\epsilon)n\theta(\frac{r_n}{2})^d f_{\max} \rfloor$ und $j_n = \lfloor (1 - 2\epsilon)n\theta(\frac{r_n}{2})^d f_{\max} \rfloor$. Dann gilt $\frac{k_n}{\log(n)} \rightarrow \infty$ und $\frac{j_n}{\log(n)} \rightarrow \infty$. Aus Satz 2.16 folgen daraus die Aussagen über $\rho(\mathcal{X}_n, \mathcal{C}_n \geq k_n)$, sowie $\rho(\mathcal{X}_n, \mathcal{C}_n \geq j_n)$: $\lim_{n \rightarrow \infty} \frac{n\theta\rho(\mathcal{X}_n, \mathcal{C}_n \geq k_n)}{k_n} = \frac{2^d}{f_{\max}}$ und $\lim_{n \rightarrow \infty} \frac{n\theta\rho(\mathcal{X}_n, \mathcal{C}_n \geq j_n)}{k_n} = \frac{2^d}{f_{\max}}$.

Für ausreichend großes $n \in \mathbb{N}$ gelten folgende Ungleichungen:

$$\begin{aligned} n\theta\rho(\mathcal{X}_n, \mathcal{C}_n \geq k_n)^d &\geq \left(\frac{1 + \epsilon}{1 + 2\epsilon} \right) \cdot \left(\frac{2^d k_n}{f_{\max}} \right) > n\theta r_n^d \\ n\theta\rho(\mathcal{X}_n, \mathcal{C}_n \geq j_n)^d &\leq \left(\frac{1 - \epsilon}{1 - 2\epsilon} \right) \cdot \left(\frac{2^d j_n}{f_{\max}} \right) < n\theta r_n^d \end{aligned}$$

Das heißt, für ausreichend großes $n \in \mathbb{N}$ gilt: $\rho(\mathcal{X}_n, \mathcal{C}_n \geq k_n) > r_n > \rho(\mathcal{X}_n, \mathcal{C}_n \geq j_n)$ und damit $j_n \leq \mathcal{C}_n \leq k_n$. Damit gilt folgende Aussage fast sicher:

$$1 - 2\epsilon f_{\max} \leq \liminf_{n \rightarrow \infty} \left(\frac{2^d \mathcal{C}_n}{n\theta r_n^d} \right) \leq \limsup_{n \rightarrow \infty} \left(\frac{2^d \mathcal{C}_n}{n\theta r_n^d} \right) \leq 1 + 2\epsilon f_{\max}$$

Da ϵ beliebig klein gewählt ist, folgt die Aussage von Satz 2.16. $\square$

2.2.5 Chromatische Zahl

Färbt man die Knoten eines Graphen $G(V, E)$ mit verschiedenen Farben $f \in F$, so beschreibt die Abbildung $c: V \rightarrow F$ die Färbung von G , wobei $c(v) \neq c(w)$ für alle $e = (v, w) \in E$ gefordert wird. Das heißt, dass je zwei benachbarte Knoten von G mit unterschiedlichen Farben gefärbt werden müssen.

Existiert eine Färbung von G mit $k = |F|$ Farben, so spricht man von einer k -Färbung von G .

Definition 2.6: Chromatische Zahl (Diestel [8])

Als **chromatische Zahl** $\chi(G)$ wird die minimale Anzahl an Farben bezeichnet, die notwendig ist um den Graphen G korrekt zu färben.

Beispiel 2.5. Eine berühmte Aussage ist der **Vier-Farben-Satz**. Dieser besagt, dass es möglich ist, eine Landkarte mit lediglich 4 Farben so zu färben, dass keine benachbarten Länder mit derselben Farbe gefärbt werden. Dies wurde dadurch bewiesen, dass die chromatische Zahl eines planaren Graphen $\chi(G) = 4$ ist.

Will man das asymptotische Verhalten der chromatischen Zahlen $\chi_n = \chi(G(\mathcal{X}_n, r_n))$ für geometrische Zufallsgraphen $G(\mathcal{X}_n, r_n)$ analysieren, so werden zuvor noch kleine Standardresultate aus der Graphentheorie benötigt, um anschließend die Aussagen der letzten Abschnitte wieder zu verwenden:

Satz 2.18

Für jeden Graphen $G = (V, E)$ gilt $\chi(G) \leq \Delta(G) + 1$.

Beweis zu Satz 2.18.

Nummeriere die Knoten von V mit $v_1, \dots, v_n$ und färbe v_1 mit der Farbe 1. Sind $v_1, \dots, v_i$ bereits gefärbt, färbe v_{i+1} mit der kleinsten Farbe, die für diesen Knoten erlaubt ist, damit kein benachbarter Knoten in derselben Farbe gefärbt ist. Diese Strategie liefert eine Färbung von G .

Da G den maximalen Knotengrad $\Delta(G)$ besitzt, ist jeder Knoten zu maximal $\Delta(G)$ anderen Knoten benachbart, das heißt es werden höchstens $\Delta(G) + 1$ Farben zum Färben des Graphen benötigt. $\square$

Satz 2.19

Für jeden Graphen $G = (V, E)$ gilt $\mathcal{C}(G) \leq \chi(G)$.

Beweis zu Satz 2.19.

Besitzt der Graph G eine Clique der Größe $k = \mathcal{C}(G)$, so heißt dies, dass es einen vollständigen Teilgraphen mit k Knoten in G gibt. Da alle diese k Knoten untereinander benachbart sind, ist es notwendig für eine erlaubte Färbung von G zur Färbung dieser

Knoten k verschiedene Farben zu verwenden.

Damit folgt $k = \mathcal{C}(G) \leq \chi(G)$. □

Bemerkung. Insgesamt folgt durch die beiden Sätze 2.18 und 2.19 die folgende Abschätzung der chromatischen Zahl eines Graphen G durch die Clique-Zahl und den maximalen Grad eines Graphen:

$$\mathcal{C}(G) \leq \chi(G) \leq \Delta(G) + 1 \quad (2.26)$$

Im Falle des Konnektivitätsverhaltens, wenn $\frac{nr_n^d}{\log(n)} \rightarrow c \in \mathbb{R}$ gilt, lassen sich mit Hilfe von Satz 2.14, Satz 2.17 und obiger Ungleichung die asymptotische Beschränktheit von $\chi(G)$ zeigen. Dabei bezeichnet $\chi_n(G)$ die chromatische Zahl des geometrischen Zufallsgraphen $G(\mathcal{X}_n, r_n)$.

Satz 2.20 (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte mit kompaktem Träger. Weiters sei $\alpha \in (0, \infty]$ und $(r_n)_{(n \geq 1)}$ eine Folge, welche $r_n \rightarrow 0$ und $\frac{nr_n^d}{\log(n)} \rightarrow \alpha$ für $n \rightarrow \infty$ erfüllt. Dann gilt:

$$\lim_{n \rightarrow \infty} \left(\frac{\mathcal{C}_n}{n\theta r_n^d} \right) \leq \lim_{n \rightarrow \infty} \left(\frac{\chi_n}{n\theta r_n^d} \right) \leq \lim_{n \rightarrow \infty} \left(\frac{\Delta_n}{n\theta r_n^d} \right) \quad (2.27)$$

beziehungsweise,

$$\frac{f_{\max}}{2^d} H_+^{-1} \left(\frac{2^d}{f_{\max} \alpha} \right) \leq \lim_{n \rightarrow \infty} \left(\frac{\chi_n}{n\theta r_n^d} \right) \leq f_{\max} H_+^{-1} \left(\frac{1}{f_{\max} \alpha} \right). \quad (2.28)$$

Beweis zu Satz 2.20.

Wegen Formel (2.26) gilt: $\mathcal{C}_n \leq \chi_n \leq \Delta_n + 1$, multipliziert man diese Ungleichung mit $\frac{1}{n\theta r_n^d}$, so folgt durch Grenzwertbildung $n \rightarrow \infty$ sofort die Ungleichungskette (2.27).

Wendet man auf (2.27) nun die Ergebnisse der Sätze 2.14 und 2.17 an, so folgt daraus sofort Formel (2.28) und damit die Aussage des Satzes. □

2.2.6 Zusammenhang

Wie in Beispiel 2.1 bereits grob skizziert wurde, spielt es in vielen realen Anwendungen eine große Rolle, ob der geometrische Zufallsgraph zusammenhängend ist, oder nicht. So will man beispielsweise in vielen Anwendungen, welche durch geometrische Zufallsgraphen modelliert werden, zusammenhängende Graphen erhalten, um gewisse notwendige Eigenschaften des Netzwerkes zu erfüllen.

In diesem Abschnitt wird untersucht, unter welchen Bedingungen ein geometrischer Zufallsgraph $G(\mathcal{X}_n, r_n)$ zusammenhängend, beziehungsweise allgemeiner k -zusammenhängend ist.

Definition 2.7: k -zusammenhängend (Diestel [8])

Ein Graph $G = (V, E)$ mit der Ordnung von mindestens $k + 1$ wird **k -zusammenhängend**⁴ genannt, wenn er trotz dem Entfernen von $k - 1$ Knoten $v \in V$ noch zusammenhängend ist.

Die **Zusammenhangszahl**¹ κ von G ist k , wenn G k -zusammenhängend, jedoch nicht mehr $(k + 1)$ -zusammenhängend ist.

Bemerkung. Unter Berücksichtigung von Definition 2.7 heißt ein Graph $G = (V, E)$ zusammenhängend, genau dann, wenn er 1-zusammenhängend ist.

Ähnlich wie bei den Analysen des minimalen und maximalen Knotengrades in Abschnitt 2.2.2 und 2.2.3, wo durch den größten und kleinsten k -Nachbarn Radius ein Grenzradius eingeführt wurde, welcher im Folgenden genauer betrachtet wurde, lässt sich auch hier ein Grenzradius definieren, welcher sicherstellt, dass der resultierende Graph zumindest k -zusammenhängend ist. Dieser Grenzradius wird mit $T_k(\mathcal{X})$ bezeichnet.

In diesem Abschnitt sei $\Omega \subset \mathbb{R}^d$ der Träger der Wahrscheinlichkeitsdichte f und ein Produkt von endlichen Intervallen $\Omega = [0, \omega_1] \times \dots \times [0, \omega_d]$, wobei $d \geq 2$ und $w_i > 0$ gelten solle. Wie auch in den vorigen Abschnitten sei $f_0 = \text{ess inf } f|_{\Omega}$, ∂_j sei die Vereinigung aller $(d - j)$ -dimensionalen Ränder von Ω und f_j sei das Infimum von f auf ∂_j . Zusätzlich sei $f_0 > 0$ und die Menge der Unstetigkeitspunkte von $f|_{\Omega}$ enthält kein Element aus $\delta\Omega$.

Die in Satz 2.21 verwendete Funktion H entspricht der in Abschnitt 2.2.2 definierten Hilfsfunktion $H(a) = 1 - a + a \log(a)$ für $a > 0$ und $H(0) = 1$.

⁴In der englischsprachigen Literatur spricht man auch von k -connected, sowie der connectivity κ .

Satz 2.21 (Penrose ([25]))

Sei $(k_n)_{(n \geq 1)}$ eine Folge von natürlichen Zahlen mit $\lim_{n \rightarrow \infty} \frac{k_n}{n} = 0$, sowie $\lim_{n \rightarrow \infty} \frac{k_n}{\log(n)} = b \in [0, \infty]$. Gilt $b < \infty$, ist die Folge $(k_n)_{(n \geq 1)}$ nicht fallend und $a_0, \dots, a_{d-1} \in [0, 1)$ seien definiert wie folgt:

$$\frac{a_j}{H(a_j)} = \frac{bd}{d-j}$$

Dann gelten folgende Aussagen:

1. Wenn $b = \infty$, dann gilt mit Wahrscheinlichkeit 1:

$$T_{k_n+1}(\mathcal{X}_n)^d \sim \left(\frac{k_n}{n\theta} \right) \cdot \max \left\{ \frac{2^j}{f_j} \mid 0 \leq j \leq d \right\}$$

2. Wenn $b < \infty$, dann gilt mit Wahrscheinlichkeit 1:

$$T_{k_n+1}(\mathcal{X}_n)^d \sim \left(\frac{\log(n)}{n\theta} \right) \cdot c$$

wobei

$$c = \max \left\{ \max \left\{ \frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1 \right\}, \frac{2^d b}{f_d} \right\}$$

gilt.

Beweisidee zu Satz 2.21

Analog zur Definition der Zusammenhangszahl κ lässt sich eine Kanten-Zusammenhangszahl κ^e eines Graphen $G = (V, E)$ definieren:

Definition 2.8: k -Kanten zusammenhängend (Penrose [25])

Ein Graph $G = (V, E)$ ist **k -Kanten zusammenhängend**⁵, wenn er nicht durch das Entfernen von höchstens $(k-1)$ Kanten unzusammenhängend gemacht werden kann.

⁵In der englischsprachigen Literatur wird der Begriff *k-edge-connected* verwendet.

Definition 2.9: Kanten-Zusammenhangszahl κ^e (Penrose [25])

Die **Kanten-Zusammenhangszahl**⁶ κ^e von $G(V, E)$ ist das maximale k , für welches G k -Kanten zusammenhängend ist.

Analog zu obiger Definition von $T_k(\mathcal{X})$ soll nun $T_k^e(\mathcal{X})$ den kleinsten Radius r bezeichnen, welcher sicherstellt, dass $G(\mathcal{X}, r)$ k -Kanten zusammenhängend ist. Es gilt nun folgende Aussage, welche einen Zusammenhang zwischen den Größen $\kappa(G)$, $\kappa^e(G)$ und $\delta(G)$ herstellt:

Satz 2.22 (Havet [16])

Sei $G = (V, E)$ ein Graph mit mindestens zwei Knoten. Dann gilt:

$$\kappa(G) \leq \kappa^e(G) \leq \delta(G)$$

Beweis zu Satz 2.22.

Entfernt man alle Kanten, welche zu einem Knoten $v \in V$ inzident sind, so ist der Graph $G = (V, E)$ nicht mehr zusammenhängend. Daher gilt $\kappa^e(G) \leq \delta(G)$.

Sei $\kappa^e(G) = k$ und $F = \{(v_1, w_1), \dots, (v_k, w_k)\} \subseteq E$ eine Menge an Kanten, wobei die Kanten so gewählt sind, dass $G' = (V, E \setminus F)$ nicht mehr zusammenhängend ist und alle Knoten v_i mit $i \in \{1, \dots, k\}$ im selben Zusammenhangskomponenten C von G' liegen. Ist $G'' = (V \setminus \{v_1, \dots, v_k\}, E)$ nicht zusammenhängend, so folgt $\kappa(G) \leq k$. Andernfalls gilt $C = \{v_1, \dots, v_k\}$ und damit hat der Knoten v_1 höchstens k Nachbarn, nämlich alle v_i mit $v_i \neq v_1$ und den einen Knoten w_i für welchen $v_i = v_1$ gilt. Entfernt man die Nachbarschaft von v_1 vom Graphen $G = (V, E)$, so ist dieser nicht mehr zusammenhängend. Damit folgt $\kappa(G) \leq k$ in beiden Fällen und insgesamt $\kappa(G) \leq \kappa^e(G) \leq \delta(G)$. $\square$

Bemerkung. Aus Satz 2.22 folgt die Aussage über die Grenzwerte der Größen $\kappa(G)$, $\kappa^e(G)$ und $\delta(G)$. Es gilt für eine endliche Knotenmenge $\mathcal{X} \subset \mathbb{R}^d$ mit $|\mathcal{X}| \geq k + 1$ folgende Aussage:

$$T_k(\mathcal{X}) \geq T_k^e(\mathcal{X}) \geq M_k(\mathcal{X}) \quad (2.29)$$

Wegen Satz 2.2 und Formel (2.29) gilt im Fall $b = \infty$:

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{n\theta T_{k_n+1}(\mathcal{X}_n)^d}{k_n} &\geq \lim_{n \rightarrow \infty} \frac{n\theta M_{k_n+1}(\mathcal{X}_n)^d}{k_n} \\ &= \max \left\{ \frac{2^j}{f_j} \mid 0 \leq j \leq d \right\} \end{aligned}$$

⁶In der englischsprachigen Literatur wird der Begriff *edge-connectivity* κ^e verwendet.

und im Fall $b < \infty$:

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{n\theta T_{k_n+1}(\mathcal{X}_n)^d}{\log(n)} &\geq \liminf_{n \rightarrow \infty} \frac{n\theta M_{k_n+1}(\mathcal{X}_n)^d}{\log(n)} \\ &= \max \left\{ \max \left\{ \frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1 \right\}, \frac{2^d b}{f_d} \right\} \end{aligned}$$

Damit bleibt noch die Abschätzung nach oben von $T_{k_n+1}(\mathcal{X}_n)$ zu zeigen. Dafür definiert man für $n > 0$ die Hilfsabbildungen $p_n(u) = \left(\frac{k_n}{n}\right)^{\frac{1}{d}}$ für $b = \infty$ und $p_n(u) = \left(\frac{\log(n)}{n}\right)^{\frac{1}{d}}$ und es sei t so gewählt, dass abhängig von der Größe von b eine der Ungleichungen gilt:

$$\begin{aligned} \theta t^d &> \max \left\{ \frac{2^j}{f_j} \mid 0 \leq j \leq d \right\} \quad \text{für } b = \infty \\ \theta t^d &> \max \left\{ \max \left\{ \frac{2^j(d-j)}{df_j H(a_j)} \mid 0 \leq j \leq d-1 \right\}, \frac{2^d b}{f_d} \right\} \quad \text{für } b < \infty \end{aligned}$$

Mit Hilfe dieser Definitionen kann gezeigt werden, dass mit Wahrscheinlichkeit 1 $T_{k_n+1}(\mathcal{X}_n) \leq tp_n$ für hinreichend große $n \in \mathbb{N}$ gilt. Für die genaue Ausführung des Beweises sei auf *Random Geometric Graphs* von Penrose ([25]) verwiesen.

Asymptotisches Verhalten der Zusammenhangszahl κ_n von $G(\mathcal{X}_n, r_n)$

Dieses Resultat kann dazu verwendet werden um eine Aussage über die Zusammenhangszahl κ_n des geometrischen Zufallsgraphen $G(\mathcal{X}_n, r_n)$ zu treffen. Dabei wird exakt dieselbe Aussage getroffen, welche bereits für den minimalen Knotengrad δ_n in Abschnitt 2.2.2 gezeigt wurde:

Satz 2.23 (Penrose [25])

Sei f eine Wahrscheinlichkeitsdichte auf Ω , wobei $f|_{\Omega}$ stetig in $x \forall x \in \delta\Omega$ ist und Ω zusammenhängend ist. Zusätzlich gelte $r_n \rightarrow 0$, $\frac{n\theta r_n^d}{\log(n)} \rightarrow \alpha \in [0, \infty]$ für $n \rightarrow \infty$ und $d \geq 1$. Dann gelten folgende Aussagen, für $d \geq 2$ und zusammenhängendes Ω :

1. Wenn $\alpha < \max_{j \in J \setminus \{d\}} \{2^j(d-j)/(df_j)\}$, dann gilt $\kappa_n \rightarrow 0$ fast sicher.
2. Wenn $\max_{j \in J \setminus \{d\}} \{2^j(d-j)/(df_j)\} \leq \alpha \leq \infty$, dann gilt mit Wahrscheinlichkeit 1:

$$\kappa_n \sim (n\theta r_n^d) \cdot c$$

wobei

$$c = \min \left\{ \min_{0 \leq j \leq d-1} \left\{ \frac{f_j}{2^j} H^{-1} \left(\frac{2^j(d-j)}{df_j a} \right) \right\}, \frac{f_d}{2^d} \right\}$$

gilt.

Bemerkung. Der Beweis von Satz 2.23 lässt sich durch Satz 2.21 analog zu Satz 2.8 beweisen. Für eine genauere Ausführung wird auf *Random Geometric Graphs* von Penrose ([25]) verwiesen.

2.3 Betweenness-Zentralität in geometrischen Zufallsgraphen

Bereits 1953 beschreibt Shimbel in seinem Werk über strukturelle Parameter von Kommunikationsnetzwerken ([29]) die Betweenness-Zentralität:

Findet zwischen zwei Knoten $u, v \in V$ eines geometrischen Zufallsgraphen $G = (V, E)$ eine Kommunikation statt und gibt es keine direkte Verbindung $(u, v) \in E$ zwischen den beiden Knoten, so gibt es zumindest einen Knoten w , welcher die Kommunikation zwischen den beiden Knoten überträgt. Der Knoten w hat als Zwischenstation in der Kommunikation von u und v nun eine wichtige Rolle und eine gewisse Last zu tragen. Üblicherweise findet solch eine Kommunikation zwischen zwei Knoten entlang des kürzesten Pfades von u und v statt. Zählt man nun alle kürzesten Pfade die durch den Knoten w gehen, so kann man seine Belastung im Netzwerk messen - dieser Wert wird heutzutage als **Betweenness-Zentralität** bezeichnet.

Bezeichnet man mit σ_{uv} die Anzahl der kürzesten Pfade zwischen u und v , sowie mit $\sigma_{uv}(w)$ die Anzahl der kürzesten Pfade zwischen u und v , welche durch den Knoten w führen, so kann man die Betweenness-Zentralität wie folgt berechnen:

Definition 2.10: Betweenness-Zentralität (Giles et al. [14])

Sei $G = (V, E)$ ein Graph, $w \in V$. Die **Betweenness-Zentralität** $C_B(w)$ von w ist gegeben durch:

$$C_B(w) = \frac{1}{2} \sum_{u \in V} \sum_{v \in V} \frac{\sigma_{uv}(w)}{\sigma_{uv}}, \text{ wobei } u \neq v \neq w \text{ gelten soll.} \quad (2.30)$$

Man kann diese Definition der Betweenness-Zentralität noch etwas erweitern und standardisieren, damit die Betweenness-Zentralität $\tilde{C}_B(w)$ nur Werte im Intervall $[0, 1]$ annimmt:

Definition 2.11: Standardisierte Betweenness-Zentralität

Sei $G = (V, E)$ ein Graph, $w \in V$. Weiters sei n die Anzahl der Knoten $|V|$ von G . Die **standardisierte Betweenness-Zentralität** $\tilde{C}_B(w)$ von w ist gegeben durch:

$$\tilde{C}_B(w) = 2 \cdot \frac{C_B(w)}{(n-1) \cdot (n-2)} \in [0, 1]$$

Beispiel 2.6. In einem Kommunikationsnetzwerk mit einem zentralen Router, über welchen sämtliche Nachrichten transportiert werden, hätte dieser beispielsweise Betweenness-Zentralität $\tilde{C}_B(w) = 1$, da sämtliche Kommunikationspfade über diesen verlaufen. Jeder andere Knoten oder Zugangspunkt in diesem Netzwerk hätte eine Betweenness-Zentralität von 0, da keine kürzesten Pfade durch diese Knoten führen.

Abbildung 2.3 verdeutlicht die eben beschriebene Situation. Für jede Kommunikation zwischen den Zugangspunkten 2, 3, 4 oder 5 muss zwingend der Kommunikationsweg über den Router 1 führen.

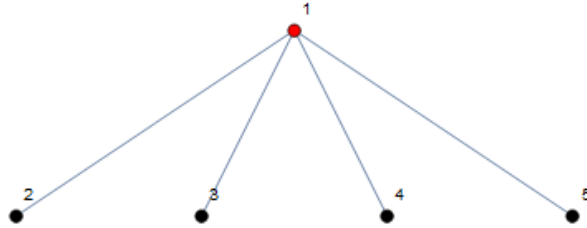


Abbildung 2.3: Kommunikationsnetzwerk mit Router 1 und 4 Zugangspunkten 2, 3, 4 und 5

Die konkrete Berechnung der Betweenness-Zentralität lässt sich in diesem Falle noch einfach per Hand durchführen, da für alle $i, j \in \{2, 3, 4, 5\}$: $\sigma_{ij} = 1$, sowie $\sigma_{ij}(1) = 1$ gilt:

$$\begin{aligned}\tilde{C}_B(1) &= \frac{1}{2} \cdot \left(\frac{\sigma_{23}(1)}{\sigma_{23}} + \frac{\sigma_{24}(1)}{\sigma_{24}} + \frac{\sigma_{25}(1)}{\sigma_{25}} + \frac{\sigma_{32}(1)}{\sigma_{32}} + \frac{\sigma_{34}(1)}{\sigma_{34}} + \frac{\sigma_{35}(1)}{\sigma_{35}} \right. \\ &\quad \left. + \frac{\sigma_{42}(1)}{\sigma_{42}} + \frac{\sigma_{43}(1)}{\sigma_{43}} + \frac{\sigma_{45}(1)}{\sigma_{45}} + \frac{\sigma_{52}(1)}{\sigma_{52}} + \frac{\sigma_{53}(1)}{\sigma_{53}} + \frac{\sigma_{54}(1)}{\sigma_{54}} \right) \cdot \frac{2}{4 \cdot 3} \\ &= \frac{1}{2} \cdot 12 \cdot \frac{2}{12} = 1\end{aligned}$$

Für die Zugangspunkte $i \in \{2, 3, 4, 5\}$ gilt jeweils $\tilde{C}_B(i) = 0$, da keine kürzesten Pfade zwischen zwei beliebigen Knoten in diesem Netzwerk durch diese Knoten führen.

2.3.1 Berechnung der Betweenness-Zentralität

Wie man bereits in Beispiel 2.6 sah, ist die Berechnung der Betweenness-Zentralität bereits für kleine Graphen recht aufwändig, da die Berechnung der kürzesten Pfade innerhalb des Graphen sehr aufwändig ist.

Will man für einen gegebenen Graphen $G = (V, E)$ und jeden seiner Knoten $w \in V$ die Betweenness-Zentralität berechnen, so müssen sämtliche kürzeste Pfade zwischen zwei Knoten $u, v \in V$ gefunden werden, sowie sämtliche kürzeste Pfade welche von $u \in V$ nach $v \in V$ durch den betrachteten Knoten $w \in V$ führen, d.h. die Berechnung von σ_{uv}

und $\sigma_{uv}(w)$ ist für alle $u, v, w \in V$ notwendig.

Verwendet man zur Berechnung der kürzesten Pfade eine modifizierte Version des Floyd-Marshall Algorithmus ([18, 3]), welcher zusätzlich zum kürzesten Pfad zwischen zwei Knoten $u, v \in V$ eines Graphen auch noch die Anzahl der kürzesten Pfade zwischen u und v liefert, so summiert sich der Aufwand der Berechnung auf $\mathcal{O}(|V|^3)$.

Brandes ([4]) veröffentlichte einen effizienteren Algorithmus zur Berechnung der Betweenness-Zentralität aller Knoten $v \in V$ eines ungewichteten Graphen mit Aufwand $\mathcal{O}(|V| \cdot |E|)$. Für einen gewichteten Graphen ist ein Aufwand von $\mathcal{O}(|V| \cdot |E| + |V|^2 \log(|V|))$ notwendig.

2.3.2 Berechnung der Betweenness-Zentralität in dichten geometrischen Zufallsgraphen

In diesem Abschnitt soll eine analoge Aussage zu Formel 2.30 gefunden werden, wobei in diesem Fall die Betweenness-Zentralität von einem dichten geometrischen Zufallsgraphen berechnet wird.

Der geometrische Zufallsgraph soll in diesem Abschnitt dadurch entstehen, dass N Punkte innerhalb eines begrenzten, konvexen Gebiets $\mathcal{V} \subset \mathbb{R}^d$ mit Volumen V zufällig verteilt werden. Die Punkteverteilung innerhalb von $\mathcal{V}$ für die Position $r_i, i \in \{1, \dots, N\}$ soll nach einem uniformen Punktprozess mit Dichte $\rho = \frac{N}{V}$ geschehen. Die Distanz zweier Punkte i und j wird mit $r_{ij} = \|r_i - r_j\|$ bezeichnet.

In den weiteren Berechnungen ist das Gebiet $\mathcal{V}$ ein geschlossener Kreis $\mathcal{D}$ mit Radius R , für welchen $V = \pi R^2$ gilt.

Zwei Punkte i und j sind nun nicht mehr miteinander verbunden, wenn ihre Distanz r_{ij} kleiner einem Grenzwert r ist, hingegen wird ihre Verbindung durch eine Funktion modelliert, welche die Wahrscheinlichkeit einer Kante zwischen i und j angibt und exponentiell mit der Distanz fällt ([6]). Dies ist dadurch motiviert, dass bei realen Kommunikationsnetzwerken die Übertragungsqualität über größere Distanz abnimmt, und die Übertragung nicht mit Überschreiten einer Grenzdistanz abrupt endet.

Hierfür wird eine Ausfallswahrscheinlichkeit P_{out} des Informationsfluss zwischen Knoten i und j definiert, und die Wahrscheinlichkeit einer Kante, beziehungsweise Verbindung zwischen zweier Knoten i und j , mit Distanz r_{ij} ergibt sich als $H(r_{ij}) = 1 - P_{out}$.

Arbeitet man mit einem Single-Input Single-Output (SISO) Modell, so kann die Aus-

fallswahrscheinlichkeit P_{out} als

$$\begin{aligned} P_{out} &= \mathbb{P}(\log_2(1 + SNR \cdot |h|^2) < x) \\ &= \mathbb{P}\left(|h|^2 < \frac{2^x - 1}{SNR}\right) \end{aligned}$$

beschrieben werden.

Dabei ist h der Übertragungskoeffizient, wobei die Zufallsvariable $|h|^2$ als Exponentialverteilt angenommen wird, x ist die Grenzausfallsrate und die Signalstörtrate wird als $SNR \approx r^{-\eta}$ angenommen. Dabei ist r die Distanz zwischen zwei verbundenen Knoten und η die Abfallsrate der Übertragung zwischen den beiden Knoten. Insgesamt kann man nun die Wahrscheinlichkeit einer Kante $H(r_{ij})$ wie folgt modellieren, wobei $\beta \ll 1$ eine Konstante sei:

$$H(r_{ij}) = e^{-\beta r_{ij}^\eta}$$

Für einen beliebigen Punkt i ist die Wahrscheinlichkeit, dass der Punkt auf die Position r_i in $\mathcal{V}$ gesetzt wird, gleich $\frac{1}{V} dr_i$. $\chi_{ij}(\kappa)$ sei die charakteristische Funktion, welche angibt ob κ am Pfad von i nach j liegt, oder nicht. Dementsprechend gilt:

$$\chi_{ij}(\kappa) = \begin{cases} 1 & \kappa \text{ liegt am Pfad zwischen } i \text{ und } j \\ 0 & \text{sonst} \end{cases}$$

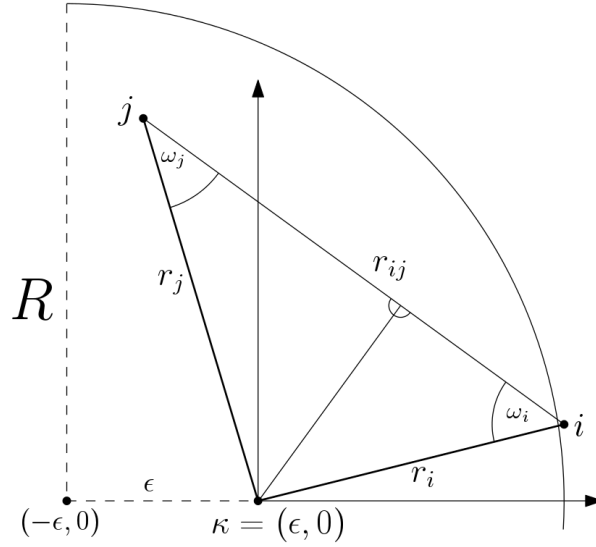
Die Wahrscheinlichkeit, dass ein beliebiges Knotenpaar i und j auf die Positionen r_i und r_j platziert werden, sowie ein kürzester Pfad zwischen i und j durch κ führt ergibt sich dementsprechend als $\frac{1}{V^2} dr_i dr_j \chi_{ij}(\kappa)$.

Betrachtet man nun alle möglichen Positionen, welche r_i und r_j in $\mathcal{V}$ einnehmen können, und summiert die Ergebnisse auf, so ergibt sich für die **erwartete Betweenness-Zentralität von κ** unter einer zufälligen Positionierung der Knoten in $\mathcal{V}$ für $\rho \rightarrow \infty$:

$$g(\kappa) = \frac{1}{V^2} \int_{\mathcal{V}} \int_{\mathcal{V}} \chi_{ij}(\kappa) dr_i dr_j \quad (2.31)$$

Betrachtet man 3 verschiedene Knoten i, j, κ , welche an 3 beliebige Positionen r_i, r_j, r_κ in $\mathcal{D}$ positioniert werden, so lässt sich $\kappa_\perp = \kappa_\perp(r_i, r_j)$ als die Länge der Normale von κ zur Verbindungslinie r_{ij} zwischen i und j definieren. Dabei soll das Koordinatensystem auf κ zentriert sein, während der Kreis $\mathcal{D}$ mit Radius R sein Zentrum in $(-\epsilon, 0)$ hat. Abbildung 2.4 veranschaulicht die soeben beschriebene Situation.

Weiters betrachtet man die Delta-Funktion $\delta(\kappa_\perp(r_i, r_j))$ und definiert $\kappa_\perp$ so, dass genau dann eine eindeutige Nullstelle vorliegt, wenn κ direkt am Pfad von i nach j liegt. Unter


 Abbildung 2.4: Zusammenhang zwischen κ , r_i , r_j und $\kappa_{\perp}$

dieser Voraussetzung gilt:

$$\int_{\mathcal{V}} \int_{\mathcal{V}} \chi_{ij} dr_i dr_j = \int_{\mathcal{V}} \int_{\mathcal{V}} \delta(\kappa_{\perp}) dr_i dr_j \quad (2.32)$$

Um $\delta(\kappa_{\perp}(r_i, r_j))$ auszudrücken und eine eindeutige Nullstelle zu garantieren, wenn κ am Pfad von i nach j liegt, sind Approximationen von $\sin(x)$ notwendig. Für kleine $x \ll 1$ gilt: $\sin(x) \approx x$. Generell kann die Distanz κ zur Verbindungslinie zwischen i und j durch folgende Gleichungen ausgedrückt werden:

$$\frac{\kappa_{\perp}}{r_i} = \sin(\omega_i), \text{ sowie } \frac{\kappa_{\perp}}{r_j} = \sin(\omega_j) \quad (2.33)$$

Während in Abbildung 2.4 die Innenwinkel am Knoten i und j jeweils mit ω_i und ω_j bezeichnet wurden, soll im Folgenden mit $(\theta_j - \theta_i)$ die Summe zu π bezeichnet werden. Wendet man nun die Approximation des Sinus auf die Gleichungen 2.33 an und addiert diese miteinander, so ergibt sich folgende Darstellung, wenn $\kappa_{\perp} \ll 1$:

$$\omega_i + \omega_j = \pi + \theta_i - \theta_j = \kappa_{\perp} \cdot \left(\frac{1}{r_i} + \frac{1}{r_j} \right) \quad (2.34)$$

In Formel 2.34 wird $\kappa_{\perp} = 0$, genau dann wenn $\pi + \theta_i - \theta_j = 0$ ist, wodurch man das Argument $\kappa_{\perp}$ innerhalb der Delta-Funktion durch $\frac{\pi + \theta_i - \theta_j}{\frac{1}{r_i} + \frac{1}{r_j}}$ ersetzen kann.

Dadurch kann nun also der Integrand des Integrals 2.32 wie folgt umgeschrieben werden:

$$\delta(\kappa_{\perp}) = \delta\left(\frac{\pi + \theta_i - \theta_j}{\frac{1}{r_i} + \frac{1}{r_j}}\right) = \delta(\pi + \theta_i - \theta_j) \left(\frac{1}{r_i} + \frac{1}{r_j}\right)$$

Schreibt man nun das doppelte Integral von der Ausgangsformel 2.31 um, und setzt die soeben erhaltene Identität für $\delta(\kappa_{\perp})$ ein, so ergibt sich, wenn zusätzlich $\mathcal{V} = \mathcal{D}$ beachtet wird:

$$\begin{aligned} g(\epsilon) &= \frac{1}{2V^2} \int_{\mathcal{D}} \int_{\mathcal{D}} \chi_{ij}(\kappa) \, dr_j \, dr_i \\ &= \frac{1}{2V^2} \int_0^{2\pi} \int_0^{2\pi} \int_0^{r(\theta_i)} \int_0^{r(\theta_j)} r_i \cdot r_j \cdot \delta(\kappa_{\perp}) \, dr_j \, dr_i \, d\theta_j \, d\theta_i \\ &= \frac{1}{2V^2} \int_0^{2\pi} \int_0^{2\pi} \int_0^{r(\theta_i)} \int_0^{r(\theta_j)} r_i \cdot r_j \cdot \left(\frac{1}{r_i} + \frac{1}{r_j}\right) \cdot \delta(\pi + \theta_i - \theta_j) \, dr_j \, dr_i \, d\theta_j \, d\theta_i \end{aligned}$$

Durch Integration der Delta-Distribution erhält man laut Definition 1.15 nach Integration der beiden äußeren Integrale anschließend:

$$\begin{aligned} g(\epsilon) &= \frac{1}{2V^2} \int_0^{2\pi} \int_0^{2\pi} \int_0^{r(\theta_i)} \int_0^{r(\theta_j)} r_i \cdot r_j \cdot \left(\frac{1}{r_i} + \frac{1}{r_j}\right) \cdot \delta(\pi + \theta_i - \theta_j) \, dr_j \, dr_i \, d\theta_j \, d\theta_i \\ &= \frac{1}{2V^2} \int_0^{2\pi} \int_0^{2\pi} \int_0^{r(\theta_i)} \int_0^{r(\theta_j)} r_i \cdot r_j \cdot \delta(\pi + \theta_i - \theta_j) \, dr_j \, dr_i \, d\theta_j \, d\theta_i \\ &= \frac{1}{2V^2} \int_0^{2\pi} \int_0^{2\pi} \int_0^{r(\theta_i)} \left(\frac{r^2(\theta_j)}{2} + r_i \cdot r(\theta_j)\right) \cdot \delta(\pi + \theta_i - \theta_j) \, dr_i \, d\theta_j \, d\theta_i \\ &= \frac{1}{2V^2} \int_0^{2\pi} \int_0^{2\pi} \left(r(\theta_i) \frac{r^2(\theta_j)}{2} + r(\theta_j) \frac{r^2(\theta_i)}{2}\right) \cdot \delta(\pi + \theta_i - \theta_j) \, d\theta_j \, d\theta_i \\ &= \frac{1}{2V^2} \int_0^{2\pi} \frac{r(\theta_i) \cdot r^2(\theta_i + \pi)}{2} + \frac{r(\theta_i + \pi) \cdot r^2(\theta_i)}{2} \, d\theta_i \\ &= \frac{1}{4V^2} \int_0^{2\pi} r(\theta_i) \cdot r(\theta_i + \pi) \cdot (r(\theta_i) + r(\theta_i + \pi)) \, d\theta_i \end{aligned}$$

Für die nächsten Integrationsschritte ist die Parametrisierung des Kreises $\mathcal{D}$ mit Mittelpunkt $(-\epsilon, 0)$ notwendig, welche mittels $r(\theta) = \sqrt{R^2 - \epsilon^2} \cdot \sin(\theta)^2 - \epsilon \cdot \cos(\theta)$ beschrieben werden kann. Zusätzlich werden noch einige Eigenschaften der trigonometrischen Funktionen $\sin$ und $\cos$ ausgenutzt: $\cos(x) = -\cos(x + \pi)$, $\sin^2(x) = \sin^2(x + \pi)$, sowie $\sin^2(x) + \cos^2(x) = 1$.

Mit der Parametrisierung von $r(\theta)$ kann nun der Integrand $r(\theta_i) \cdot r(\theta_i + \pi) \cdot (r(\theta_i) + r(\theta_i + \pi))$ vereinfacht werden. Es gelten folgende Identitäten:

$$\begin{aligned}
 r(\theta_i) \cdot r(\theta_i + \pi) &= \left(\sqrt{R^2 - \epsilon^2 \sin^2(\theta_i)} - \epsilon \cos(\theta_i) \right) \quad (2.35) \\
 &\quad \cdot \left(\sqrt{R^2 - \epsilon^2 \sin^2(\theta_i + \pi)} - \epsilon \cos(\theta_i + \pi) \right) \\
 &= \left(\sqrt{R^2 - \epsilon^2 \sin^2(\theta_i)} - \epsilon \cos(\theta_i) \right) \cdot \left(\sqrt{R^2 - \epsilon^2 \sin^2(\theta_i)} + \epsilon \cos(\theta_i) \right) \\
 &= R^2 - \epsilon^2 \sin^2(\theta_i) + \underbrace{\sqrt{\dots} \cdot \epsilon \cos(\theta_i) - \sqrt{\dots} \cdot \epsilon \cos(\theta_i)}_{=0} - \epsilon^2 \cos^2(\theta_i) \\
 &= R^2 - \epsilon^2 \cdot \underbrace{(\sin^2(\theta_i) + \cos^2(\theta_i))}_{=1} \\
 &= R^2 - \epsilon^2 \quad (2.36)
 \end{aligned}$$

sowie:

$$\begin{aligned}
 r(\theta_i) + r(\theta_i + \pi) &= \sqrt{R^2 - \epsilon \sin^2(\theta_i)} - \epsilon \cos(\theta_i) + \sqrt{R^2 - \epsilon \sin^2(\theta_i + \pi)} - \epsilon \cos(\theta_i + \pi) \\
 &= \sqrt{R^2 - \epsilon \sin^2(\theta_i)} - \epsilon \cos(\theta_i) + \sqrt{R^2 - \epsilon \sin^2(\theta_i)} + \epsilon \cos(\theta_i) \\
 &= 2 \cdot \sqrt{R^2 - \epsilon \sin^2(\theta_i)} \quad (2.37)
 \end{aligned}$$

Mithilfe von Formel 2.35 und Formel 2.37 kann das Integral nun weiter berechnet werden:

$$\begin{aligned}
 g(\epsilon) &= \frac{1}{4V^2} \int_0^{2\pi} r(\theta_i) \cdot r(\theta_i + \pi) \cdot (r(\theta_i) + r(\theta_i + \pi)) \, d\theta_i \\
 &= \frac{1}{2V^2} \int_0^{2\pi} (R^2 - \epsilon^2) \cdot \sqrt{R^2 - \epsilon^2 \sin^2(\theta_i)} \, d\theta_i \\
 &= \frac{1}{2V^2} (R^2 - \epsilon^2) \cdot 4 \int_0^{\frac{\pi}{2}} \sqrt{R^2 - \epsilon^2 \sin^2(\theta_i)} \, d\theta_i \\
 &= \frac{4 \cdot (R^2 - \epsilon^2)}{2V^2} \int_0^{\frac{\pi}{2}} R \cdot \sqrt{1 - \frac{\epsilon^2}{R^2} \sin^2(\theta_i)} \, d\theta_i \\
 &= \frac{2 \cdot (R^2 - \epsilon^2)}{\pi^2 R^3} \cdot E\left(\frac{\epsilon}{R}\right)
 \end{aligned}$$

Für die letzte Identität wird ein vollständiges elliptisches Integral 2. Art verwendet, für welches gilt:

$$E(k) = \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \sin^2(\theta)} \, d\theta$$

Durch Division von $g(\epsilon)$ durch den maximalen Wert, welchen g annehmen kann, kann $g(\epsilon)$ normalisiert werden. Damit gilt $g^*(\epsilon) \cdot g(0) = g(\epsilon)$:

Satz 2.24 (Giles et al. [14])

Unter den bisher getroffenen Voraussetzungen, welche für die Herleitung von $g(\epsilon)$ getroffen wurden, kann $g^(\epsilon)$ geschrieben werden als:*

$$g^*(\epsilon) = \frac{2}{\pi} \cdot (1 - \epsilon^2) \cdot E(\epsilon) \in [0, 1]$$

Dabei bezeichnet $E(\epsilon)$ das elliptische Integral 2. Art.

Entwickelt man die Darstellung von $g^*(\epsilon)$ in einer Umgebung nahe des Zentrums, das heißt für $\epsilon \ll 1$, sowie nahe des Randes für $\epsilon \approx R$, erhält man die folgenden Approximationen für g^* :

Satz 2.25 (Giles et al. [14])

$$\text{Reihenentwicklung um } 0: g^*(\epsilon) = 1 - \frac{5\epsilon^2}{R^2} + \frac{13\epsilon^4}{64R^4} + \mathcal{O}(\epsilon^6)$$

$$\text{Reihenentwicklung um } R: g^*(\epsilon) = \frac{4(R - \epsilon)}{\pi R} + \mathcal{O}((R - \epsilon)^2)$$

Die Approximationen zeigen eine quadratische Veränderung der Betweenness-Zentralität nahe des Zentrums, sowie eine lineare Veränderung in der Nähe des Randes des Gebietes $\mathcal{D}$.

2.3.3 Weitere Zentralitäts-Maße

Neben der Betweenness-Zentralität, welche in diesem Kapitel bereits recht detailliert beschrieben wurde, gibt es noch einige weitere Maße, welche die Zentralität und Bedeutung von Knoten $v \in V$ in einem Graphen $G = (V, E)$ messen.

Sabidussi ([28]) führte 1966 die **Punkt-Zentralität** ein, welche die Entfernungen des ausgewiesenen Knotens $v \in V$ zu allen anderen Knoten $w \neq v \in V$ miteinbezieht. Ein Punkt ist in diesem Sinne zentral, wenn er den kleinst möglichen Abstand zu allen anderen Punkten besitzt.

$$C_P(v) = \frac{1}{\sum_{w \neq v \in V} d_G(v, w)}$$

Die Punkt-Zentralität versagt leider bei nicht zusammenhängenden Graphen, d.h. wenn es Knoten $v, w \in V$ gibt, welche nicht mit einem Pfad miteinander verbunden sind: dann gilt nämlich $d_G(v, w) = \infty$ und dementsprechend $C_P(v) = 0$. Rochat ([27]) nahm sich 2009 diesem Problem an, und entwickelte die **harmonische Zentralität**, welche ähnlich der Punkt-Zentralität ist und aber auch für nicht zusammenhängende Graphen plausible Ergebnisse liefert.

$$C_H(v) = \sum_{w \neq v \in V} \frac{1}{d_G(v, w)}$$

Legt man besonders Wert auf den größten Abstand, welcher ein Knoten $v \in V$ zu allen anderen Knoten w hat, kann man die **Graph-Zentralität** $C_G(v)$ eines Knoten betrachten, welche 1995 von Hage und Harary in [15] beschrieben wurde:

$$C_G(v) = \frac{1}{\max_{w \in V} d_G(v, w)}$$

Unter der Annahme, dass Knoten $v \in V$ eines Graphen viel Einfluss auf deren benachbarten Knoten $w \in V$ haben, kann man einen Knoten v zentral nennen, wenn er viele Nachbarn besitzt. Dies hat Nieminen in [24] vorgeschlagen und daraus resultierte die **Grad-Zentralität**. Die Grad-Zentralität ordnet jedem Knoten $(v_1, \dots, v_n)$ eines Graphen seinen Knotengrad $d_G(v_i)$ zu: $(d_G(v_1), \dots, d_G(v_n))$. Ein Knoten mit hohem Knotengrad wird dabei als zentral betrachtet.

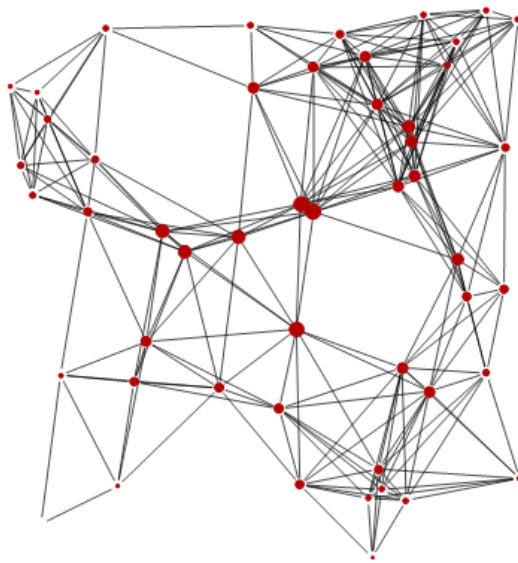
$$C_d(v) = d_G(v)$$

Beispiel 2.7. In Abbildung 2.5 wurden die vier oben beschriebenen Arten der Zentralität an einem geometrischen Zufallsgraphen mit 50 Knoten und einem Radius von 0.3 für jeden Knoten in $G(\mathcal{X}_{50}, 0.3)$ berechnet und visualisiert. Die Werte wurden dabei normalisiert, sodass der Wertebereich der Zentralitäten im Intervall $[0, 1]$ liegt.

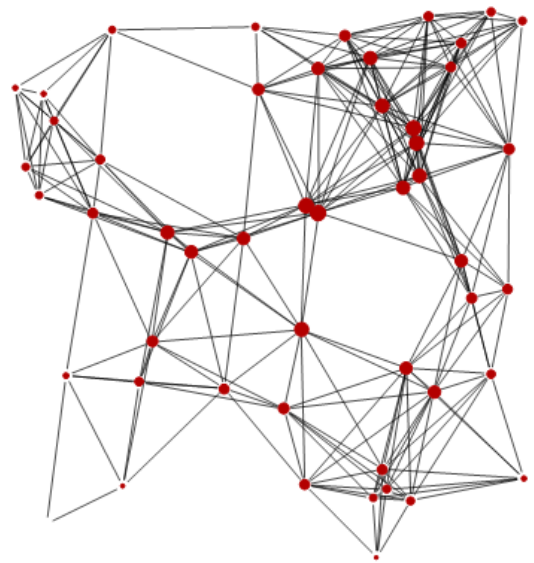
Hat ein Knoten $v \in \mathcal{X}_{50}$ eine große Zentralität, das heißt ist $C(v)$ nahe bei 1, so wird der Knoten in der Darstellung des Graphen größer gedruckt, als ein Knoten w mit kleinerer Zentralität.

In diesen Visualisierungen sieht man die Unterschiede der Zentralitäts-Maße, und wie sich die Zentralität unterschiedlich auf die Knoten aufteilt. Die Punkt-Zentralität C_P entspricht in diesem Falle im Wesentlichen der harmonischen Zentralität C_H , da es sich bei dem Graphen um einen zusammenhängenden Graphen handelt, und der gravierende Unterschied der harmonischen Zentralität erst bei einem nicht zusammenhängenden Graphen klar ersichtlich wird.

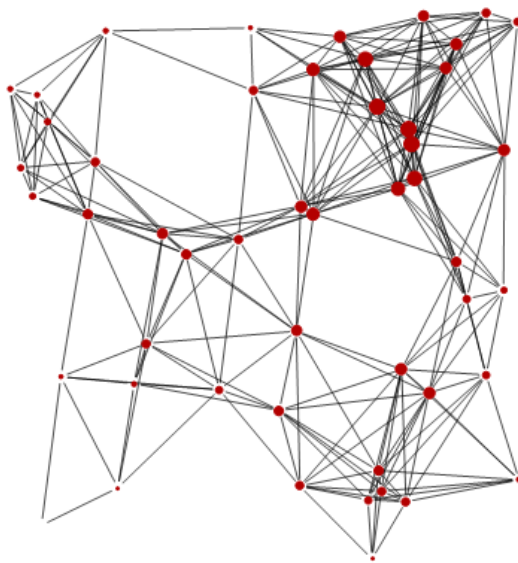
Insbesondere, wenn man die Grad-Zentralität und die Graph-Zentralität mit den anderen Zentralitätsmaßen vergleicht, fallen große Unterschiede in der Verteilung der zentralen



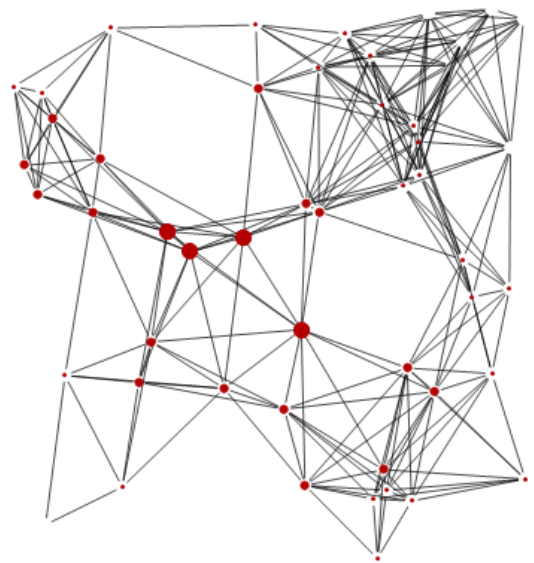
(a) Punkt-Zentralität C_P



(b) Harmonische Zentralität C_H



(c) Grad-Zentralität C_d



(d) Graph-Zentralität C_G

Abbildung 2.5: Verschiedene Zentralitäts-Maße visualisiert an einem geometrischen Zufallsgraphen $G(\mathcal{X}_{50}, 0.3)$

Punkte auf. So weisen die Knoten, welche man selbst als geographisch zentral im Graph erachten würde, in Abbildung 2.5 (d) eine hohe Graph-Zentralität auf.

Die Grad-Zentralität ist wie zu erwarten in Bereichen, wo sich viele Knoten in naher Distanz häufen in einem geometrischen Zufallsgraphen höher als in Gebieten des Graphen, wo kaum Verbindungen zu anderen Knoten herrschen.

Allgemein kann man sagen, dass die Wahl des Zentralitäts-Maßes immer an die konkrete Anwendung und ihr Ziel angepasst werden muss, da die Maße selbst unterschiedliche Ergebnisse und Lösungsmengen an Knoten erzielen, welche je nach Problemstellung mehr oder weniger hilfreich sein können.

Bemerkung. Die in diesem Abschnitt beschriebenen und verglichenen Zentralitäts-Maße sollen zeigen, dass es in diesem Bereich viele weitere Maße gibt, für welche sich eine genauere Betrachtung lohnt. Aktuell sind lediglich die oben ausgeführten Resultate von Giles et al. ([14]) über die Betweenness-Zentralität im Zusammenhang mit geometrischen Zufallsgraphen bekannt. Für die weiteren Maße wurden noch keinerlei asymptotische Berechnungsmöglichkeiten für geometrischen Zufallsgraphen veröffentlicht.

2.4 Ein ähnliches Modell: Erdős-Rényi Zufallsgraphen

Erdős und R nyi f hrten 1959 in ihrem Werk *On random graphs* ([9]) **Zufallsgraphen** $G(n, N)$ und $G(n, p)$ ein, welche in manchen Anwendungsf llen eine Alternative zu geometrischen Zufallsgraphen sein k nnen, beziehungsweise die Problemstellung besser darstellen. Die Struktur der Zufallsgraphen von Erd s und R nyi unterscheidet sich hierbei wesentlich von der Struktur der bisher behandelten geometrischen Zufallsgraphen.

Fixiert man n als die Anzahl der Knoten $\{v_1, \dots, v_n\}$ und M als die Anzahl der Kanten $\{e_1, \dots, e_M\}$ eines Graphen, so gibt es $g := \binom{n}{M}$ verschiedene Graphen mit exakt dieser Anzahl von Knoten und Kanten. Sei $\Gamma(n, M) := \{G(V, E) : |V| = n, |E| = M\}$ die Menge dieser Graphen mit g Elementen.

Definition 2.12: Erd s-R nyi Zufallsgraph ([9])

Ein **Erd s-R nyi Zufallsgraph** $G(n, M)$ ist ein zuf llig ausgew hlter Graph aus der Menge $\Gamma(n, M)$ der Graphen mit n Knoten und M Kanten. Die Wahrscheinlichkeit einen gewissen Graphen zu erhalten ist dabei gleichverteilt.

W hrend man bei dieser Methode einen Zufallsgraphen zu erzeugen genau  ber die Kardinalit ten der Knoten- und Kantenmenge informiert ist, gibt es eine alternative Methode andere Zufallsgraphen zu erzeugen:

F r eine gegebene Wahrscheinlichkeit $p \in (0, 1]$ betrachtet man wiederum Graphen auf

einer n -elementigen Knotenmenge $\{v_1, \dots, v_n\}$. Je zwei Knoten v_i und v_j werden nun mit Wahrscheinlichkeit p durch eine Kante $e_{ij} = \{v_i, v_j\}$ miteinander verbunden.

Definition 2.13: Erdős-Rényi Zufallsgraph ([10])

Ein **Erdős-Rényi Zufallsgraph** $G(n, p)$ ist ein Graph mit n Knoten $\{v_1, \dots, v_n\}$, wobei je zwei Knoten mit Wahrscheinlichkeit p durch eine Kante verbunden sind.

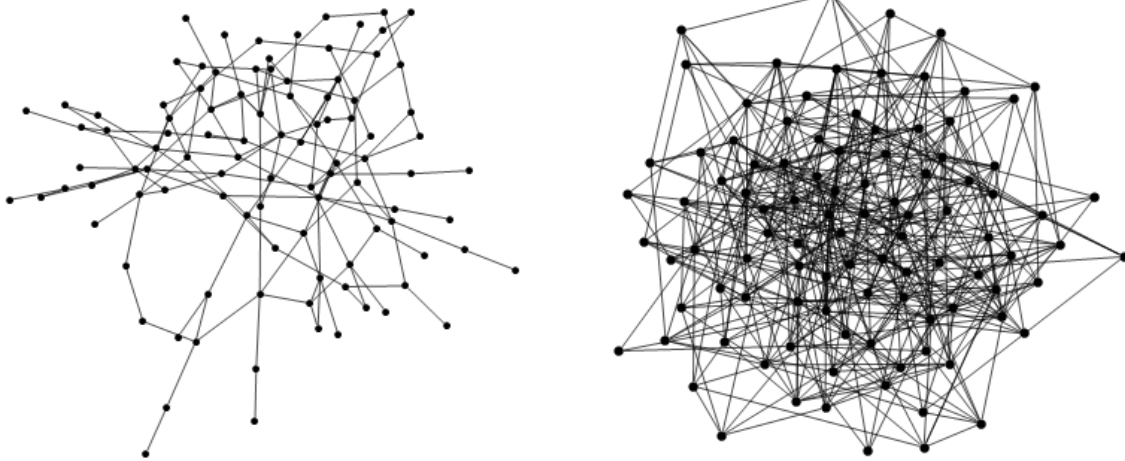
Bemerkung. Der Graph $G(n, p)$ hat nach Definition genau n Knoten und die erwartete Anzahl der Kanten ist $|E| = \binom{n}{2}p$.

Verteilt man die Knoten $v_i \in \{v_1, \dots, v_n\}$ nach einer gegebenen Wahrscheinlichkeitsverteilung in $\mathbb{R}^d$, so könnte man die Graphen $G(n, M)$ und $G(n, p)$ ebenso als ein Modell für geometrische Zufallsgraphen betrachten. Die Analyse dieser Graphen unterscheidet sich jedoch grundlegend von der Analyse der geometrischen Zufallsgraphen $G(\mathcal{X}, r)$, welche in den letzten Abschnitten besprochen wurde.

Beispiel 2.8. In diesem Beispiel wurden zwei Erdős-Rényi Zufallsgraphen generiert, wobei die Knotenanzahl auf 100 festgelegt wurde - die Darstellung dieser Graphen wird in Abbildung 2.6 gezeigt.

Im ersten Fall, dem Zufallsgraphen $G(100, 150)$ wurde zufällig ein Graph aus der Menge $\Gamma(100, 150)$ ausgewählt und der erhaltene Graph ist in diesem Fall zusammenhängend.

Für den zweiten Zufallsgraphen $G(100, 0.1)$ wurde ebenso eine 100-elementige Knotenmenge gewählt, und zwei Knoten v_i, v_j sollen mit der relativ geringen Wahrscheinlichkeit von $p = 0.1$ durch eine Kante miteinander verbunden werden. Der resultierende Graph ist in diesem Fall ebenfalls zusammenhängend, und die erwartete Anzahl der Knoten beträgt $|E| \approx \binom{100}{2} p = 4950 \cdot 0.1 = 495$. Die tatsächliche Anzahl der Kanten weicht in diesem Fall leicht ab, und beträgt 467.



(a) Zufallsgraph $G(100, 150)$ mit 100 Knoten und 150 Kanten

(b) Zufallsgraph $G(100, 0.1)$ mit 100 Knoten und Kantenwahrscheinlichkeit $p = 0.1$

Abbildung 2.6: Veranschaulichung der 2 Typen des Erdős-Rényi Zufallsgraphen $G(n, M)$ und $G(n, p)$

3 Ad-hoc WiFi Netzwerke

Ein potenzielles Anwendungsgebiet von geometrischen Zufallsgraphen ist die Verwendung dieser Graphen als Modell für ad-hoc WiFi Netzwerke. Damit kann dieser Typ von Netzwerk unter theoretischen Gesichtspunkten genauer analysiert werden, beziehungsweise können damit neue technologische Entwicklungen entstehen oder erforscht werden.

Nach einer genauen Definition der im Folgenden verwendeten Begrifflichkeiten um ad-hoc WiFi Netzwerke wird in Abschnitt 3.3 die Verbreitung von Würmern in solch Netzwerken genauer untersucht. Durch die stetig zunehmende Verbreitung von mobilen Kommunikationsgeräten wie Notebooks, Tablets und Smartphones steigt auch die Anzahl von Computervürmern, welche speziell für diese Art von Geräten und Netzwerktopologien entwickelt werden. Dadurch ist eine intensive Auseinandersetzung mit den potenziellen Gefahren für diese Geräte und Netzwerke sehr wichtig.

Anschließend wird in Abschnitt 3.4 ein Kommunikationsprotokoll zur effektiven Nachrichtenübertragung zwischen den Geräten in ad-hoc WiFi Netzwerken vorgestellt. Greedy geographic Routing ist eine einfache Möglichkeit, um Kommunikation zwischen Geräten in einem Netzwerk zu ermöglichen. Dabei wird auch analysiert, wie zuverlässig dieses Routing-Verfahren in beliebigen ad-hoc WiFi Netzwerken funktioniert und welche technischen Anforderungen dabei an die Geräte gestellt werden müssen.

3.1 Definition

Ein **ad-hoc WiFi Netzwerk** ist ein autonomes, dezentrales Netzwerk zur direkten Kommunikation von mobilen und stationären Geräten, ohne die Verwendung von spezieller Infrastruktur (wie beispielsweise einem Router, Switch oder einem kabellosen Zugriffspunkt) oder anderen Netzwerken, welche die Kommunikation ermöglichen. Die Geräte kommunizieren hierbei kabellos über Funkverbindungen.

Anders als in herkömmlichen sternförmigen Netzwerken, wo die Datenübertragung ständig über die bereitgestellte Infrastruktur läuft, werden in ad-hoc WiFi Netzwerken die Daten von Gerät zu Gerät übertragen. Einerseits findet die Übertragung entweder direkt von Gerät zu Gerät statt, wenn diese zueinander in entsprechender Distanz sind, sodass eine Funkverbindung zwischen ihnen besteht. Andererseits, wenn die Geräte ausserhalb der Funkreichweite für eine direkte Datenübertragung sind, werden die Daten über andere Geräte weitergeleitet, welche dabei die Rolle eines Router übernehmen. Das heißt, dass in solch Netzwerken jedes Gerät auch eine aktive Rolle bei der Datenübertragung

spielt und nicht nur aktiv ist, wenn es selbst Daten empfängt oder versendet - für mobile und batteriebetriebene Geräte wirkt sich dies mitunter in einer hohen Belastung des Akkus aus. Auf welchem Weg die Daten zwischen den Geräten transportiert werden, legen dabei Routing-Algorithmen fest. [17, 12]

Es finden sich aktuell bereits zahlreiche reale Anwendungsfälle in welchen im täglichen Leben auf ad-hoc Netzwerke zurückgegriffen wird:

- Automatisierung, Sensorik und Monitoring (intelligente Elektronik, Datenanalyse)
- Information: touristische Informationen, welche eine lokale oder zeitliche Abhängigkeit aufweisen
- Mobile Zugriffspunkte für Internetzugriff
- Kommunikation in privatem Kreise ohne Internetzugriff bei beispielsweise auf Kongressen, Messen oder Meetings

Die großen **Vorteile** bei dieser Type von Netzwerk sind, dass keinerlei Infrastruktur oder spezielle Geräte für den Aufbau des Netzwerks notwendig sind und somit Kosten reduziert werden können. Ebenso wird die Last bei der Datenübertragung sofern möglich durch die Routing-Algorithmen fair aufgeteilt, sodass kein Gerät unter voller Auslastung stehen muss und ein Ausfall eines Gerätes die Stabilität des Netzwerks im Wesentlichen nicht negativ beeinflusst, da die Datenübertragung noch weiterhin über andere Geräte stattfinden kann.

Im Gegensatz dazu ist es von Nachteil, dass jedes Gerät in diesem Netzwerk eine aktive Rolle als Router spielt und demnach einen höheren Stromverbrauch aufweist. Ebenso ist es notwendig, dass die Geräte möglichst immer aktiv sind um eine zuverlässige Datenübertragung zu gewährleisten und einen kompletten Zusammenbruch oder Ausfall des Netzwerkes zu verhindern.

3.2 Geometrische Zufallsgraphen als Modell für ad-hoc WiFi Netzwerke

Für die Entwicklung neuer Protokolle und Algorithmen für ad-hoc WiFi Netzwerke ist es notwendig, diese Netzwerke durch ein geeignetes mathematisches Modell zu beschreiben, um die Korrektheit dieser Neuerungen direkt beweisen zu können. Ebenfalls ist es sehr hilfreich, wenn man die Performance und Zuverlässigkeit von beispielsweise Routing-Algorithmen in diesem Modell analysieren kann, sodass eine Vergleichbarkeit zu anderen Algorithmen gegeben ist. Dabei ist es wichtig, das Modell relativ einfach zu wählen, sodass die gewünschten Untersuchungen mit vertretbarem Aufwand durchgeführt werden können, aber die modelltheoretischen Annahmen trotzdem noch die reale Situation gut

beschreiben und keine essentiellen Parameter im Modell fehlen.

Man kann die Geräte, welche in einem ad-hoc WiFi Netzwerk miteinander kommunizieren als Knoten v_i eines Graphen interpretieren und zwei Knoten v_i und v_j mit einer (gerichteten) Kante verbinden, falls eine direkte Kommunikation zwischen den Geräten möglich ist, beziehungsweise falls eine direkte, einseitige Kommunikation zwischen Gerät i und j möglich ist. Damit eignen sich prinzipiell Graphen als Modell für solch Netzwerke.

Da die Kommunikation in solch ad-hoc WiFi Netzwerken recht spezifisch ist, da die Geräte untereinander mit durch Funk erreichbare Geräte kommunizieren, eignet sich jedoch nicht jeder Typ von Graphen als Modell für diese Netzwerke. So widerspiegeln die Erdős-Rényi Zufallsgraphen $G(n, M)$, sowie $G(n, p)$, die reale Situation von ad-hoc WiFi Netzwerken relativ schlecht. In beiden dieser Graphen sind Kanten zwischen den Knoten zufällig gewählt, was bedeuten würde, dass ein Gerät im ad-hoc WiFi Netzwerk zufällig mit einem anderen Gerät direkt kommunizieren könnte, unabhängig von geometrischer Lage der beiden Geräte zueinander. Dies ist aber in der Realität nicht der Fall, im Gegenteil ist die direkte Kommunikation zweier Geräte üblicherweise abhängig von der Entfernung dieser Geräte zueinander.

Dieser geometrische Zusammenhang, der eine direkte Übertragung von Information zwischen zwei Geräten möglich macht, wird jedoch von geometrischen Zufallsgraphen $G(\mathcal{X}, r)$ sehr gut modelliert. Als Vereinfachung ist hierbei notwendig, dass für jedes Gerät die maximale Übertragungsdistanz r gleich sein muss. Dies ist zwar nicht immer der Fall, kann aber im Grunde vernachlässigt werden, da man beispielsweise in realen Anwendungen eine hohe Qualität der Sender voraussetzen kann und somit eine minimale Übertragungsdistanz fordern kann, welche von jedem Gerät erreicht werden muss. Etwas problematisch ist es in diesem Modell noch, dass die Übertragungsdistanz üblicherweise nicht bis zu einer fixen Grenzdistanz gewährleistet ist, sondern mit der Distanz abnimmt. Für die Untersuchung von Netzwerken ist es daher oft notwendig, durch die physikalischen Bedingungen welche für Funkübertragung gelten Gleichungen herzuleiten, welche eine fixe Übertragungsdistanz rechtfertigen, sodass die Übertragung von Informationen noch fehlerfrei und zuverlässig durchgeführt werden kann. In Abschnitt 2.3.2 und Abschnitt 3.3.1 sind solch Herleitungen genauer ausgeführt und begründet. [19]

Das Modell mit geometrischen Zufallsgraphen erfasst ad-hoc WiFi Netzwerke zu einem bestimmten Zeitpunkt sehr gut. Zu einem bestimmten Zeitpunkt bedeutet, dass die Kommunikationsgeräte in diesem Modell zufällig in einem 2-dimensionalen Raum angeordnet sind, eine Bewegung der Geräte - die beispielsweise für mobile Geräte durchaus in der Realität stattfindet - aber nicht möglich ist. Hierbei gibt es aber ebenso Modelle, welche die Bewegungen in solch Netzwerken simulieren, welche man auf den geometrischen Zufallsgraphen bei Bedarf anwenden kann und dementsprechend Beobachtungen über Veränderungen in diesen Netzwerken machen kann.

Beispiel 3.1. Abbildung 3.1 (a) zeigt ein ad-hoc WiFi Netzwerk mit 13 Geräten, wobei die Geräte in diesem Fall Mobiltelefone, Laptops und stationäre Router darstellen. Man sieht anhand der Pfeile, dass jedes Gerät zu jedem Gerät in geringer Distanz eine Verbindung aufbaut und dadurch die abgebildete Netzwerkstruktur entsteht. Dieses Netzwerk ist zusammenhängend, d.h. Informationen können über dieses Netzwerk von jedem Gerät zu jedem anderen gesendet werden. Es gibt kein Gerät, welches zu nicht zu mindestens einem anderen Gerät verbunden ist.

Die rechte Abbildung 3.1 (b) zeigt den geometrischen Zufallsgraphen $G(\mathcal{X}_{13}, 3)$ im Gebiet $[0, 10]^2$ mit einer uniformen Verteilung der Knoten. Das heißt, in diesem Modell wurde ein Raum von $10m^2 \times 10m^2$ simuliert, in welchem sich 13 Geräte befinden, welche eine Funkreichweite von 3 Metern besitzen. Dieser Graph könnte als Modell für das reale ad-hoc WiFi Netzwerk dienen und zeigt ähnliche Charakteristiken.

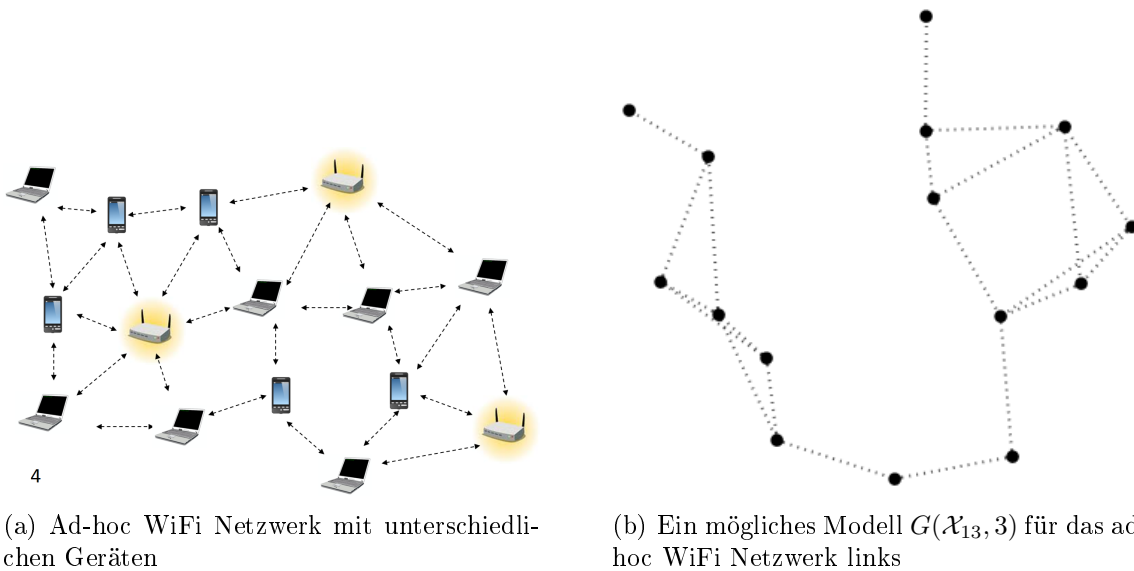


Abbildung 3.1: Ein ad-hoc WiFi Netzwerk mit 13 Geräten und die Gegenüberstellung eines möglichen Modells für dieses Netzwerk als geometrischer Zufallsgraph

3.3 Ausbreitung von Wurm-Epidemien in ad-hoc WiFi Netzwerken

Mit Hilfe von geometrischen Zufallsgraphen kann die Ausbreitung von Wurm-Epidemien in ad-hoc WiFi Netzwerken genauer analysiert werden. Würmer sind Schadprogramme, welche sich komplett eigenständig und ohne zusätzliches menschliches Handeln auf andere Geräte verbreiten. Die Verbreitung geschieht dabei direkt von Gerät zu Gerät. Die eigenständige Analyse der Ausbreitung von Computerwürmern in ad-hoc WiFi Netzwerken ist sehr wichtig, um Gegenmaßnahmen gegen die Ausbreitung sowie Verteidigungsstrategien gegen potenzielle Angriffe zu entwickeln.

Nekovee ([23]) untersuchte das Verhalten von Computerwürmern in ad-hoc WiFi Netzwerken anhand umfangreicher Monte-Carlo Simulationen. Dafür verwendete er als Modell für das ad-hoc WiFi Netzwerk geometrische Zufallsgraphen.

3.3.1 Das Modell

Modell für das Kommunikationsnetzwerk als geometrischer Zufallsgraph

Über Funk kommunizierende Geräte des ad-hoc WiFi Netzwerks werden als Knoten in einem 2-dimensionalen Raum interpretiert. Wie bereits in Abschnitt 3.2 angedeutet, ist nun eine Schwierigkeit, die Qualität der Funkverbindung abzuschätzen um die Distanz anzugeben, für welche eine zuverlässige Funkverbindung zwischen zwei Geräten aufgebaut werden kann. Die Qualität der Funkverbindung zwischen zwei Geräten i und j hängt dabei von der Distanz der Geräte ab, wobei durch Signalausfälle und Freiraumdämpfung die Qualität der Funkverbindung zusätzlich reduziert wird.

Bezeichnet man mit r_{ij} die Distanz zwischen den Knoten i und j im geometrischen Zufallsgraphen, mit P^i die Signalstärke bei der Übermittlung der Information von Knoten i und mit P^{ij} die Signalstärke bei Empfang des Signals, so kann die durchschnittliche empfangene Signalstärke P^{ij} durch das Pathloss-Model wie folgt beschrieben werden:

$$P^{ij} = \frac{P^i}{cr_{ij}^\alpha} \quad (3.1)$$

Die Konstante c hängt unter anderem von der Frequenz ab, über welche das Signal versandt wird, während α von den räumlichen Umständen abhängt. Je nachdem, ob die Funkverbindung in Innenräumen oder im Freien aufgebaut wird, kann $2 \leq \alpha \leq 5$ gewählt werden. Für den Fall einer Ausbreitung der Funksignale im freien Raum wird $\alpha = 2$ gewählt.

Bezeichnet man mit β den Grenzwert der Freiraumdämpfung und mit ν den Rauschpegel des Gerätes j , so kann nur dann eine zuverlässige Datenübertragung zwischen den Geräten i und j stattfinden, wenn folgende Gleichung gilt:

$$\frac{P^{ij}}{\nu} \geq \beta \quad (3.2)$$

Für den Einsatz von geometrischen Zufallsgraphen als Modell für diese Situation ist es aber notwendig, die Distanz, beziehungsweise den Radius anzugeben, unter welchen eine zuverlässige Funkverbindung zwischen den Geräten aufgebaut werden kann. Durch die Formeln 3.1 und 3.2 kann die maximale Reichweite des Funksignals ausgehend von Gerät i beschrieben werden:

$$\begin{aligned} \frac{P^{ij}}{\nu} &\geq \beta \\ \Leftrightarrow P^{ij} &\geq \nu\beta \\ \Leftrightarrow \frac{P^i}{cr_{ij}^\alpha} &\geq \nu\beta \\ \Leftrightarrow r_{ij} &\leq \left(\frac{P^i}{c\nu\beta} \right)^{\frac{1}{\alpha}} \end{aligned} \quad (3.3)$$

Formel 3.3 beschreibt nun die maximale Distanz $r_{ij}^{max} = \left(\frac{P^i}{c\nu\beta} \right)^{\frac{1}{\alpha}}$ welche zwischen zwei Geräten i und j angenommen werden darf, um eine zuverlässige Funkverbindung aufzubauen. Nimmt man nun zusätzlich an, dass für jedes Gerät die Signalstärke der Übermittlung der Information P^i gleich ist, so ist r_{ij}^{max} für jedes beliebige Paar an Geräten i und j in diesem Netzwerk ebenfalls gleich. Damit kann diese Situation der Funkkommunikation zwischen mobilen Geräten passend durch den geometrischen Zufallsgraphen $G(\mathcal{X}, r_{ij}^{max})$ beschrieben werden.

Modell für die Ausbreitung der Würmer

Es wird davon ausgegangen, dass in diesem ad-hoc WiFi Netzwerk kein Internetzugriff besteht, und sich die Würmer direkt von Gerät zu Gerät übertragen. Dabei findet eine direkte Übertragung des Wurms statt und es werden nur Geräte attackiert, zu welchen eine direkte Funkverbindung vom Angreifer hergestellt werden kann. Es werden keine Geräte eingesetzt, welche diese schadhaften Pakete lediglich weiterleiten, wodurch auch keinerlei Routing-Algorithmen für die Ausbreitung des Wurms in diesem Netzwerk notwendig sind - das heißt, dass diese Angriffsmethode im Endeffekt alle ad-hoc WiFi Netzwerke betreffen kann, wo eine direkte Kommunikation zwischen den Geräten in Funkreichweite möglich ist.

Üblicherweise können infizierte Geräte, welche bereits mit dem Wurm verseucht sind, durch ein Update der Virenschutz-Software (Patch) geschützt werden.

Dadurch ergeben sich für ein Gerät drei verschiedene Zustände, welche im Laufe einer Wurm-Epidemie angenommen werden können:

- **ungeschützte Geräte:** diese Geräte sind nicht vom Wurm infiziert, sind jedoch durch den Mangel eines Patches nicht gegen den Wurm geschützt.
Die Anzahl der ungeschützten Geräte zum Zeitpunkt t im ad-hoc WiFi Netzwerk wird im Folgenden mit $U(t)$ bezeichnet.
- **infizierte Geräte:** Geräte, welche bereits vom Wurm verseucht sind, und noch keinen Patch erhalten haben.
Die Anzahl der infizierten Geräte zum Zeitpunkt t wird mit $I(t)$ bezeichnet.
- **geschützte Geräte:** Geräte, welche durch den Patch geschützt wurden und immun gegen den Wurm sind.
Die Anzahl der geschützten Geräte zum Zeitpunkt t wird mit $G(t)$ bezeichnet.

Ist die Anzahl der Geräte des ad-hoc WiFi Netzwerks N , so gilt für jeden beliebigen Zeitpunkt t : $N = U(t) + I(t) + G(t)$, das heißt, jedes Gerät hat zu jedem Zeitpunkt exakt einen dieser 3 Zustände.

Definition 3.1: SIR-Modell (Moreno et al. [21])

Ein Modell für eine Population, welche von einer Epidemie erfasst wird und die gesamte Population in 3 disjunkte Gruppen einteilt (anfällig - infiziert - entfernt $\leftrightarrow$ susceptible - infected - removed) wird in der Epidemiologie auch als **SIR-Modell** bezeichnet.

Dieses Modell wird häufig bei der Untersuchung von Epidemien verwendet.

Bemerkung. Bei dem oben eingeführten Modell für die Geräte des ad-hoc WiFi Netzwerks, auf welchen sich die Wurm-Epidemie in den späteren Simulationen ausbreitet, handelt es sich um ein SIR-Modell der Population der Geräte.

Erhalten ungeschützte Geräte von einem infizierten Gerät den Wurm, so werden sie mit einer Wahrscheinlichkeit von $\lambda \in (0, 1]$ infiziert und sind damit selbst Überträger des Wurms. Infizierte Geräte versuchen zu jedem Zeitpunkt t andere benachbarte Geräte mit dem Wurm zu infizieren. Bereits infizierte Geräte können anschließend mit einer Wahrscheinlichkeit von $\delta \in (0, 1]$ einen Patch erhalten und sind somit gegen den Wurm geschützt. Dieses Ausbreitungsverhalten wird für jeden Zeitschritt t der einzelnen Monte-Carlo Simulationen wiederholt.

MAC - Medium Access Control

Zusätzlich zu den Analysen des Verhaltens der Wurm-Epidemie in einem ad-hoc Wi-Fi Netzwerk im Modell eines geometrischen Zufallsgraphen wurden die Simulationen auch für ein solches Netzwerk mit Medium Access Control (MAC) durchgeführt. Die Medienzugriffssteuerung, beziehungsweise Medium-Access Control, ist ein Standard zur Datenübertragung in Funknetzwerken und gehört zur 2. Schicht des OSI - Modells, welches einen Standard für Netzwerkprotokolle darstellt. Die Medienzugriffssteuerung regelt den Zugriff auf Daten in einem Netzwerk, sodass es nicht durch gleichzeitige Zugriffe auf ein bestimmtes Gerät zu Kommunikationsstörungen oder Datenverlusten führt.

Nevokee übernahm eine Regelung der Medienzugriffssteuerung für seine Monte-Carlo Simulationen, welche mit LBT (listen - before - talk) bezeichnet wird: Bevor ein Gerät eine Datenübertragung an ein anderes Gerät startet, muss es überprüfen, ob das Zielgerät im Moment selbst keine Übertragungen durchführt. Nur in diesem Fall darf eine Datenübertragung an das Zielgerät gestartet werden.

Konkret bedeutet dies im Falle der Monte-Carlo Simulationen, dass wenn ein infiziertes Gerät i einen Wurm versendet, alle anderen infizierten Geräte j , welche sich in der Reichweite zur direkten Kommunikation mit i befinden blockiert werden und in diesem Zeitschritt selbst keinen Wurm verbreiten dürfen, da es zu Zugriffskollisionen führen könnte. Wenn im Folgenden bei der Analyse der Monte-Carlo Simulationen von dem Modell des geometrischen Zufallsgraphen mit MAC-Protokoll auftritt, dann wird bei der Ausbreitung der Würmer diese Regel noch zusätzlich betrachtet. Dies ermöglicht eine Analyse des Einflusses eines MAC-Protokolls auf das Verhalten einer Wurm-Epidemie in einem ad-hoc WiFi Netzwerk.

3.3.2 Aufbau der Monte-Carlo Simulationen

Anhand der Modelle für das ad-hoc WiFi Netzwerke und des Ausbreitungsmodell des Wurms führte Nevokee ([23]) Monte-Carlo Simulationen durch, um das Verhalten der Ausbreitung des Wurms im Netzwerk zu analysieren. Dafür wurden für jeden Zeitschritt t jeder einzelnen Simulation die Anzahl der ungeschützten, geschützten und infizierten Geräte $U(t), G(t), I(t)$ betrachtet.

Das zugrunde liegende ad-hoc WiFi Netzwerk besteht aus N Geräten, welche in einem Bereich von $A^2 = 1000m \times 1000m$ gleichverteilt platziert werden, wobei die Reichweite der Funksignale der Geräte für jedes Gerät $50m$ beträgt. Es wurden verschiedene Simulationen mit unterschiedlicher Anzahl an Geräten und somit unterschiedlicher Dichte an Geräten $\rho = \frac{N}{A^2}$ durchgeführt, jeweils für $N \in \{4000, 6000, 8000, 10000, 20000\}$. Zusätzlich wurden nur zusammenhängende geometrische Zufallsgraphen betrachtet, wobei vor jedem Run der Monte-Carlo Simulation überprüft wurde, ob der erhaltene geometrische Zufallsgraph zusammenhängend ist. Wurde ein geometrischer Zufallsgraph generiert, welcher nicht zusammenhängend war, so wurde dieser Run abgebrochen und sofort mit

einer neuen Simulationsrun begonnen.

Bei den Berechnungen wurde die Wahrscheinlichkeit δ , dass ein Gerät einen Patch erhält als $\delta = 1$ fixiert, da der zeitliche Verlauf der Ausbreitung des Wurms lediglich vom Verhältnis zwischen der Ausbreitungswahrscheinlichkeit des Wurms λ und der Wahrscheinlichkeit eines Patches δ abhängt, das heißt es hängt lediglich von $\frac{\lambda}{\delta}$ ab.

Für eine Monte-Carlo Simulation wurde dabei die Ausbreitung der Wurm - Epidemie in 500 geometrischen Zufallsgraphen nach obigem Modell berechnet, und über die verschiedenen Ergebnisse dabei der Mittelwert gebildet. Jeder Simulationsschritt startet dabei damit, dass ein zufällig ausgewählter Knoten des Graphen mit dem Wurm infiziert wird, und die Ausbreitung findet so lange wie in Abschnitt 3.3.1 definiert statt, bis die Epidemie ausstirbt.

Zusätzlich zum Modell mit geometrischen Zufallsgraphen, wurden die Monte-Carlo Simulationen auch für ein Modell mit geometrischen Zufallsgraphen und dem Protokoll MAC durchgeführt, als auch für die in Abschnitt 2.4 definierten Erdős-Rényi Zufallsgraphen. Die Bezeichnungen in den Plots des Originalpapers von Nekovee sind dabei in Englisch, wobei folgende Beziehungen gelten:

- RG (random graph) entspricht dem Erdős-Rényi Zufallsgraph
- RGG (random geometric graph) entspricht dem geometrischen Zufallsgraphen
- RGG + MAC entspricht dem geometrischen Zufallsgraphen mit dem MAC-Protokoll

Die Ergebnisse der Simulationen mit dem Modell des Erdős-Rényi Zufallsgraphen (RG) lassen sich oftmals mit bekannten Ergebnissen aus der Verbreitung von Würmern im Internet vergleichen - da dort die Ausbreitung von Würmern ähnlich zufällig stattfindet, wie zwei Geräte in einem Erdős-Rényi Zufallsgraphen miteinander durch eine Kante verbunden sind. Konkret findet die Wurmverteilung im Internet oft so statt, dass die Würmer ganze IP (Internet Protocol) - Adressbereiche nach potenziell gefährdeten Geräten absuchen und anschließend die gefundenen Geräte attackieren ([5]). Die Auswahl der Geräte findet dabei zufällig statt, und nicht wie im Falle eines ad-hoc WiFi Netzwerks über einen geometrischen Zusammenhang bezüglich der Lage der Geräte zueinander.

3.3.3 Analyse der Ergebnisse der Monte-Carlo Simulationen

In den Monte-Carlo Simulationen wurden sämtliche Werte von $U(t)$, $I(t)$, $G(t)$ für jeden Zeitschritt t gespeichert, sodass anhand von Plots die Resultate der Simulationen sehr schön sichtbar werden und analysiert werden können. Die hier präsentierten Ergebnisse entsprechen den originalen Plots von Nekovee in *Worm epidemics in wireless ad hoc networks* ([23]).

Ausbreitung der Epidemie

In der folgenden Abbildung sind die berechneten Daten zu einem Netzwerk mit einer Größe von $N = 10000$ gehörig, wobei die Wahrscheinlichkeit λ , dass ein Gerät infiziert wird mit $\lambda = 0.1$ gewählt wurde. Wie bereits vorhin erläutert, ist die Wahrscheinlichkeit δ eines infizierten Gerätes einen Patch zu erhalten $\delta = 1$.

Abbildung 3.2 zeigt den Anteil an infizierten Geräten $\frac{I(t)}{N}$ in diesem Netzwerk über den gesamten zeitlichen Verlauf der Epidemie. Man sieht sofort, dass für jedes Modell (RG, RGG, RGG + MAC) die Epidemie vollständig gestoppt werden kann, und der Anteil der infizierten Geräte sich mit unterschiedlicher Geschwindigkeit wieder gegen 0 bewegt. Auffallend ist jedoch, dass sich die Epidemie im Modell des Erdős-Rényi Zufallsgraphen (RG) exponentiell ausbreitet, während in den anderen beiden Modellen (RGG, RGG + MAC) die Ausbreitung weit langsamer steigt.

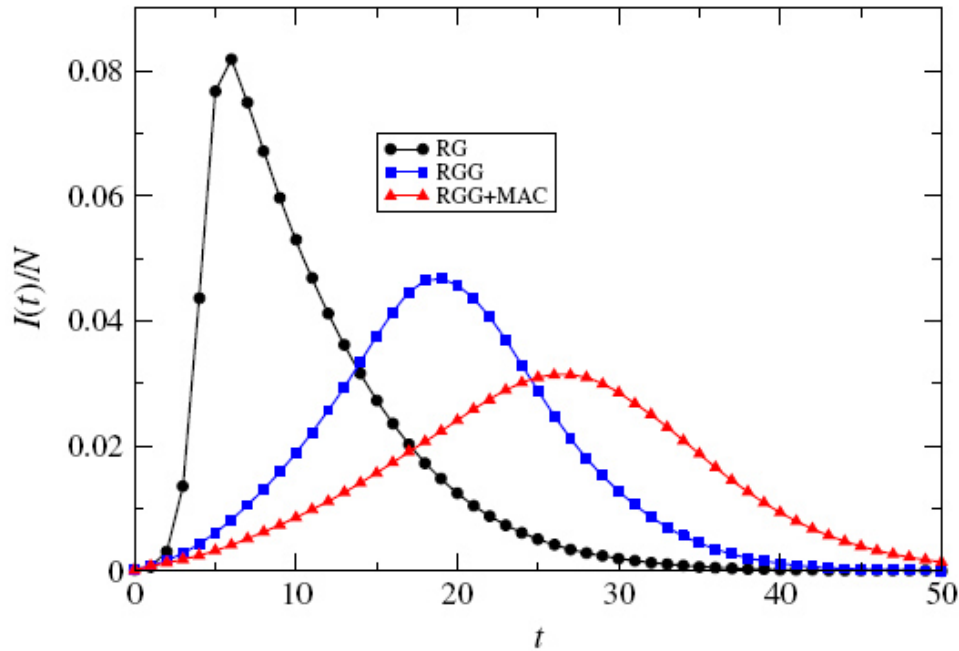


Abbildung 3.2: Anteil an infizierten Geräten $\frac{I(t)}{N}$ im zeitlichen Verlauf der Epidemie - Abbildung aus [23]

Die langsamere Ausbreitung in den beiden Modellen, welche geometrische Zufallsgraphen verwenden, wird durch die topologische Struktur der geometrischen Zufallsgraphen erklärt: durch die geometrische Lage der Geräte zueinander und durch eine große Gruppenbildung der Geräte innerhalb des Netzwerkes kann eine langsamere Ausbreitung des Anteils an infizierten Geräten begründet werden.

Ist zusätzlich Medium-Access Control aktiviert, so wächst der Anteil an infizierten Geräten noch langsamer, ebenso ist der maximale Anteil an infizierten Geräten geringer als

bei den anderen beiden Modellen. Dies hängt damit zusammen, dass viele benachbarte infizierte Geräte gleichzeitig den Wurm verbreiten möchten, dies aber durch MAC verhindert wird, wodurch viele Geräte für das Versenden von Datenpaketen im Netzwerk blockiert werden.

Der exponentielle Wachstum des Anteils an infizierten Geräten im Netzwerkmodell des Erdős-Rényi Zufallsgraphen deckt sich mit den bekannten Resultaten, welche ein exponentielles Wachstum für Würmer im Internet vorhersagen, wie beispielsweise Robert und Chen in *Worm Epidemics in High-Speed Networks* ([5]) begründen: Dort wird die Verbreitung der Epidemie mit logistischem Wachstum modelliert und durch diese Formel ausgedrückt:

$$I(t) = \frac{I(0) \cdot N}{I(0) + (N - I(0)) \cdot e^{-\lambda N t}}$$

Dabei ist $I(0)$ die anfangs infizierte Menge an Geräten, N die Anzahl aller Geräte in diesem Netzwerk und λ die Wahrscheinlichkeit ein Gerät zu infizieren. In der Anfangsphase, in der noch wenige Geräte der gesamten Menge an Geräten infiziert ist, das heißt wenn $I(0) \ll N$ gilt, folgt das exponentielle Wachstum welches auch in Abbildung 3.2 sichtbar ist:

$$I(t) = I(0) \cdot e^{\lambda N t}$$

Neben diesen Resultaten, wurde auch analysiert, wie die Ausbreitung der Epidemie mit der Anzahl der Geräte, beziehungsweise der Dichte der Geräte in diesem Netzwerks zusammenhängt. Dafür wurde untersucht, wie schnell die maximale Ausbreitung der Epidemie $\max_{t \in T} \frac{I(t)}{N}$ erreicht wird - der Zeitschritt, für welchen das Maximum angenommen wird, wird dabei als T_{max} bezeichnet. Je kleiner T_{max} ist, desto schneller schreitet die Ausbreitung der Epidemie in dem Netzwerk voran. Abbildung 3.3 zeigt $T_{max}(N)$ für die beiden Modelle mit geometrischen Zufallsgraphen, einmal mit Medium-Access Control und einmal ohne. Die Gesamtanzahl N der Geräte nimmt dabei die folgenden Werte an: $N = \{4000, 6000, 10000, 20000\}$, womit sich schrittweise die Dichte an Geräten in dem definierten Gebiet von $1000 \times 1000m^2$ erhöht.

Es zeigt sich, dass die Ausbreitung der Epidemie bei einer hohen Gerätedichte im Netzwerk sehr schnell stattfindet, während bei einer geringeren Dichte an Geräten der maximale Anteil der infizierten Geräte erst langsamer erreicht wird. Diese Ergebnisse bestätigen sich, da T_{max} für eine steigende Geräteanzahl N monoton fallend ist.

Die Auswirkungen des MAC-Protokolls wirken sich am Besten bei einer geringen Gerätedichte aus, wo die Ausbreitung der Epidemie bis hin zum maximalen Anteil an infizierten Geräten um circa 40% länger dauert als im selben Modell ohne aktivierte Medienzugriffsbeschränkung. Bei einer höheren Gerätedichte zeigt das Medium-Access Control immer noch Wirkung in der Verlangsamung der Ausbreitung der Epidemie, jedoch eine weit geringere.

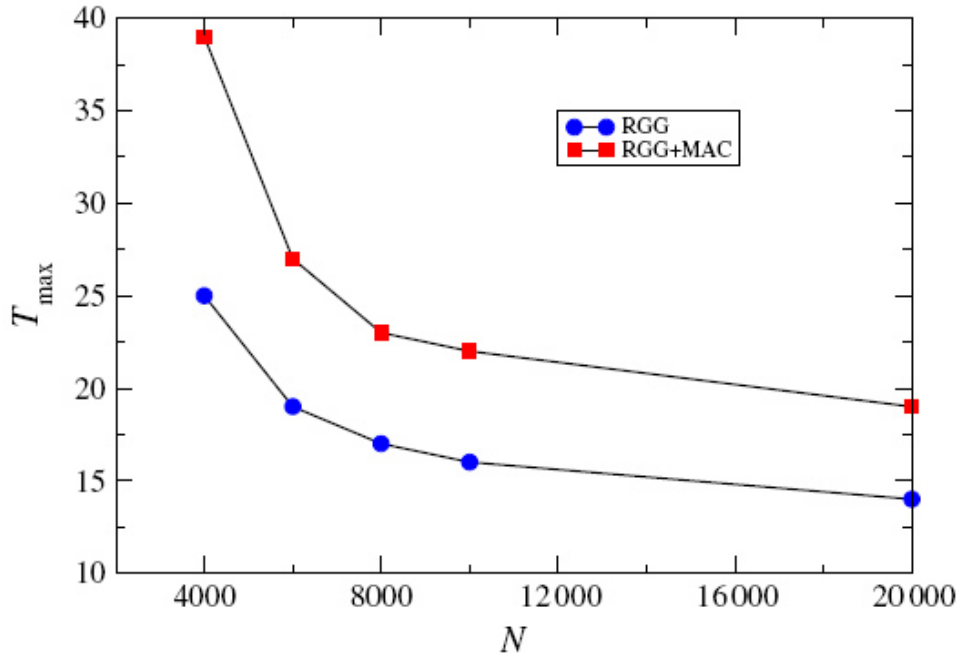


Abbildung 3.3: Zeit bis zum Erreichen des maximalen Anteils an infizierten Geräten in Abhängigkeit zur Anzahl der Geräte N - Abbildung aus [23]

Epidemischer Grenzwert λ_C

Murray führte in seinem Werk *Mathematical biology. I* ([22]) einige wichtige Konzepte für die Untersuchung von biologischen Epidemien ein - dabei konzentrierte er sich auf Epidemien, welche auf das in Abschnitt 3.3.1 erläuterte SIR-Modell aufbauen. Daher sind diese Konzepte auch für die von Nekovee durchgeführten Simulationen anwendbar.

Grundlegend bei der Untersuchung von Epidemien ist die Suche nach einem epidemischen Grenzwert λ_C - das ist die minimale Wahrscheinlichkeit der Verbreitung eines Wurms, welche notwendig ist, dass sich die Epidemie in dem Netzwerk ausbreiten kann. Das heißt, dass für Werte $\lambda < \lambda_C$ keine Ausbreitung der Epidemie stattfindet und die Ausbreitung der Epidemie nur für Werte $\lambda > \lambda_C$ möglich ist.

Moreno, Pastor-Satorras und Vespignani zeigten in *Epidemic outbreaks in complex heterogeneous networks* ([21]), dass ein starker Zusammenhang zwischen dem epidemischen Grenzwert λ_C , dass sich eine Epidemie ausbreitet und der epidemischen Auftretungshäufigkeit $\tilde{G}_\infty = \lim_{t \rightarrow \infty} G(t)$ besteht:

Sie sagen, dass wenn $\lambda < \lambda_C$ gilt, ist $\tilde{G}_\infty$ sehr nahe bei 0 - und umgekehrt, wenn $\lambda > \lambda_C$ gilt, und sich somit eine Epidemie ausbreiten kann, gilt $\tilde{G}_\infty > 0$ und ein endlicher Anteil aller Geräte wird infiziert. Da nur bereits infizierte Geräte einen Patch gegen den Wurm erhalten können, kann $\tilde{G}_\infty$ auch als die gesamte Anzahl an infizierten Geräten im Laufe der Epidemie angesehen werden.

Nekovee änderte die Formel für die epidemische Auftretungshäufigkeit $\tilde{G}_\infty$ leicht ab, indem er diese durch die gesamte Anzahl an Geräten dividierte, und erhielt damit:

$$G_\infty = \frac{\lim_{t \rightarrow \infty} G(t)}{N} \quad (3.4)$$

Formel 3.4 half ihm anschließend bei der Untersuchung des epidemischen Grenzwertes λ_C , indem er G_∞ als Funktion der Verbreitungswahrscheinlichkeit λ des Wurms betrachtete. Abbildung 3.4 zeigt die Darstellung der Funktion $G_\infty(\lambda)$ und die Ergebnisse der Simulationen für alle drei betrachteten Modelle (RG, RGG, RGG + MAC).

Man erkennt deutlich den oben erklärten Zusammenhang zwischen dem epidemischen Grenzwert λ_C und der epidemischen Auftretungshäufigkeit G_∞ , da sich für sehr kleine Werte für λ in keinem der drei analysierten Modelle eine Epidemie entwickelt. Auffallend ist jedoch, dass der Grenzwert λ_C sich für alle Modelle unterscheidet, und die Epidemie sich beim Modell mit dem Erdős-Rényi Zufallsgraphen bei geringeren Werten für λ ausbreiten kann, als im Modell des geometrischen Zufallsgraphen mit und ohne Medium-Access Control.

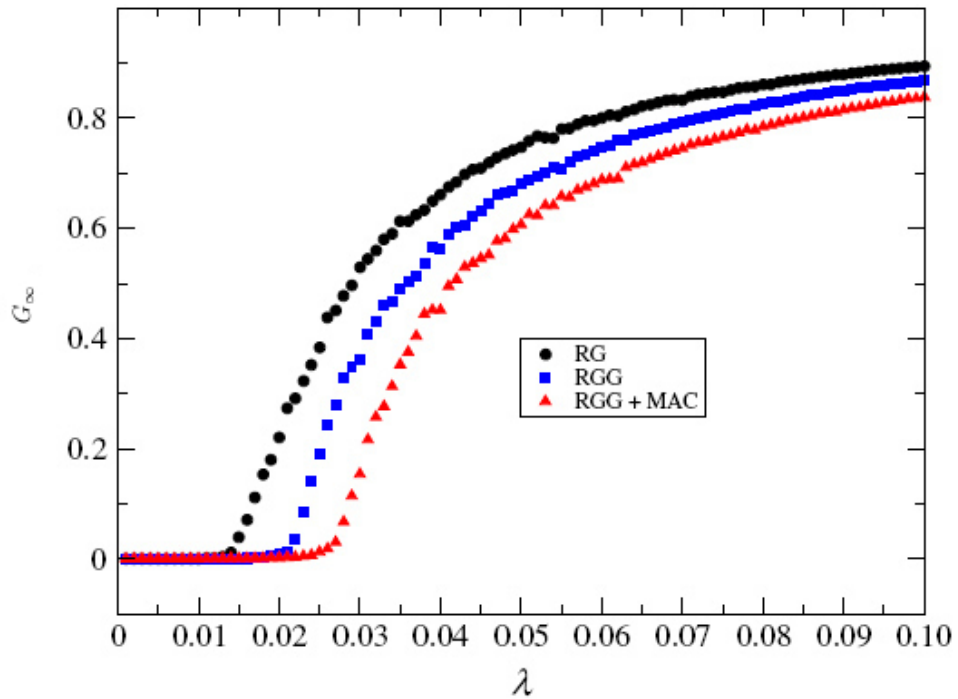


Abbildung 3.4: G_∞ als Funktion von λ - Abbildung aus [23]

Dass die Entwicklung einer Epidemie in den Modellen mit geometrischen Zufallsgraphen erst mit einer höheren Verbreitungswahrscheinlichkeit λ des Wurms stattfinden kann,

kann durch die topologische Struktur des geometrischen Zufallsgraphen begründet werden. Die Infektion eines Gerätes in einem Modell mit geometrischen Zufallsgraphen hängt stark mit der Infektion seiner direkten Nachbarn zusammen, womit sich die Reproduktionsrate der Epidemie verringert und die durchschnittliche Anzahl an neuen Infektionen deutlich geringer ist, als in Netzwerken die auf einem Modell eines Erdős-Rényi Zufallsgraphen basieren.

Medium-Access Control verzögert die Ausbreitung einer Epidemie, beziehungsweise kann diese sich erst für noch höhere Werte von λ ausbreiten, das heißt, dass der epidemische Grenzwert λ_C für die Modelle mit geometrischen Zufallsgraphen oberhalb des epidemischen Grenzwertes für Netzwerke mit Zufallsgraphen befindet. Ebenso wird durch das MAC-Protokoll die gesamte Ausbreitung der Epidemie eingedämmt, im Vergleich zu den anderen beiden Modellen.

3.4 Übermittlungsprotokolle für ad-hoc WiFi Netzwerke

Findet zwischen zwei Geräten i und j in einem ad-hoc WiFi Netzwerk eine Kommunikation statt, so kann die Nachricht entweder direkt zwischen den Geräten ausgetauscht werden, sofern sie zueinander in Funkreichweite sind (single-hop Übertragung), oder es sind weitere Geräte notwendig, welche die Nachricht vom Absender zum Empfänger weiterleiten (multi-hop Übertragung). Wie bereits in Abschnitt 3.1 kurz erläutert, sind für die Auswahl der Geräte, welche die Nachricht lediglich zum Empfänger übertragen, Routing-Algorithmen zuständig. ([31])

Solch Routing-Algorithmen steuern die Kommunikation zwischen den Geräten in WiFi Netzwerken und sorgen dafür, dass die Informationen welche gesendet werden, zuverlässig beim Empfänger ankommen.

3.4.1 Greedy Geographic Routing

Eine relativ einfache Art von Algorithmus, welcher ein Routing von Daten in WiFi Netzwerken steuern kann, ist das sogenannte Greedy geographic Routing.

Definition 3.2: Greedy Algorithmen (Cormen [7])

Greedy-Algorithmen sind Algorithmen, welche Schritt für Schritt das lokale Optimum für eine Teillösung auswählen und dadurch versuchen das globale Optimum als Lösung für die Problemstellung zu erzielen.

Bemerkung. Dabei ist es nicht notwendigerweise so, dass Greedy Algorithmen durch die Wahl des lokalen Optimums immer das globale Optimum als Lösung finden, abhängig von der Implementierung und der konkreten Problemstellung gelingt es jedoch oftmals sehr gute Näherungen zum globalen Optimum zu finden.

Finn führte in *Routing and Addressing Problems in Large Metropolitan-Scale Internet-works* ([11]) das **Greedy geographic Routing** ein, welches das Austauschen von Nachrichten in einem Netzwerk über Greedy-Heuristiken ermöglichen soll:

Definition 3.3: Greedy geographic Routing (Finn [11], Stojmenovic [30])

Eine Datenpaket soll von Gerät s zu Gerät d gesendet werden, wobei jedes Gerät eine maximale Übertragungsdistanz von r besitzt. **Greedy geographic Routing** regelt den Transport des Datenpakets über folgende Regeln:

1. Gilt $\|s - d\| \leq r$, so kann das Paket direkt von Gerät zu Gerät übertragen werden.
2. Gilt $\|s - d\| > r$, so wird das Paket zu jenem benachbarten Gerät v gesendet werden, welches von s aus erreichbar ist und für welches $\|v - d\|$ minimiert wird.

Tritt Fall 2 ein, so übernimmt das Gerät v , welches das Paket erhalten hat die Rolle vom Absender s und das Verfahren wird so lange wiederholt, bis Fall 1 eintritt und das Datenpaket erfolgreich an das Empfangsgerät d übermittelt wurde.

Bemerkung. Hierbei handelt es sich um ein **lokalisierendes Routing-Verfahren**: die Geräte im Netzwerk müssen dabei immer wissen, wo sie selbst und ihre Nachbarn innerhalb der maximalen Übertragungsdistanz sich räumlich befinden. Ebenfalls muss die geographische Lage des Empfangsgerätes d bekannt sein. ([30])

Beispiel 3.2. In einem ad-hoc WiFi Netzwerk mit 14 Geräten soll eine Nachricht von Gerät s nach Gerät d gesendet werden - dabei wird Greedy geographic Routing verwendet, um den Transport der Nachricht zu ermöglichen.

In diesem Beispiel besitzen alle Geräte dieselbe maximale Übertragungsdistanz r , womit alle Geräte Nachrichten an andere Geräte mit maximaler Distanz r weiterleiten können. Abbildung 3.5 zeigt das ad-hoc WiFi Netzwerk und den Weg über welchen die Nachricht transportiert wird, um schließlich beim Empfänger d anzukommen. Dabei wurden für alle Geräte, welche an der Übermittlung der Nachricht beteiligt sind, die maximale Übertragungsdistanz r als gepunkteter Kreis eingezeichnet, um die Möglichkeiten an Geräten sichtbar zu machen, an welche die Nachricht von diesem Gerät aus weitergereicht werden könnte.

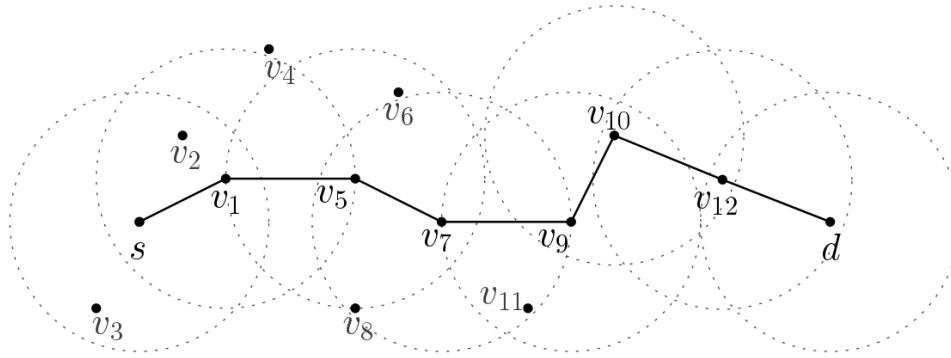


Abbildung 3.5: Skizze der Funktionsweise von Greedy geographic Routing anhand eines Gerätes s , welches ein Datenpaket an ein Gerät d sendet

Wie leicht ersichtlich ist, gilt beim Versenden der Nachricht, dass der Empfänger d vom Absender s nicht direkt erreichbar ist, das heißt es gilt $\|s - d\| > r$ und damit muss ein anderer Knoten gefunden werden, welcher die Nachricht näher zum Ziel transportiert. Das Gerät s hat 3 verschiedene Geräte in Reichweite (v_1, v_2, v_3), wobei für v_1 gilt, dass die Distanz $\|v_1 - d\|$ zum Empfangsgerät d minimiert wird. Daher wird die Nachricht an v_1 gesandt, welches die Rolle des Absenders übernimmt und mit derselben Regelung versucht, die Nachricht an d zuzustellen.

Schlussendlich findet sich durch Greedy geographic Routing eine Route von s nach d , welche über die Geräte $v_1, v_5, v_7, v_9, v_{10}, v_{12}$ verläuft.

Bemerkung. Durch das Auftreten von lokalen Minima, das heißt mit der Existenz von Geräten, welche keine näheren Geräte zum Empfänger d in ihrer Reichweite finden, kann das Greedy geographic Routing versagen und die Nachricht dadurch nicht zum Empfänger zugestellt werden.

Vermeidung von lokalen Minima um eine zuverlässige Datenübertragung durch Greedy geographic Routing zu erhalten

In dem Werk *Asymptotic Critical Transmission Radii for Greedy Forward Routing in Wireless Ad Hoc Networks* ([31]) von Wan et al. wird eine minimale Übertragungreichweite für die Geräte in einem ad-hoc WiFi Netzwerk erarbeitet, sodass mit Greedy geographic Routing eine zuverlässige Zustellung von Nachrichten in diesem Netzwerk garantiert werden kann, ohne dass der in der obigen Bemerkung beschriebene Fall eintreten kann, dass sich das Verfahren in einer Sackgasse befindet und eine korrekte Zustellung der Nachricht nicht mehr möglich ist.

Betrachtet man ein beliebiges Gerät s in einem ad-hoc WiFi Netzwerk, welches eine Nachricht an das Empfangsgerät d senden oder weiterleiten soll und bezeichnet die restli-

chen Geräte in diesem Netzwerk mit v_i für $i \in I \subset \mathbb{N}$, dann gelten folgende Überlegungen:

Ein Gerät, welches ein Datenpaket von s nach d weiterleiten kann, muss im Schnitt der Kreise um s und d mit Radius $\|s - d\|$ liegen, damit die Weiterleitung erfolgreich durchgeführt werden kann, das heißt in diesem Gebiet (siehe Abbildung 3.6):

$$B(s, \|s - d\|) \cap B(d, \|s - d\|)$$

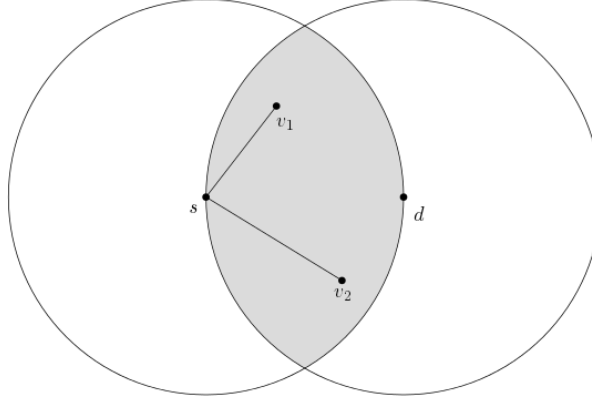


Abbildung 3.6: Veranschaulichung der Fläche $B(s, \|s - d\|) \cap B(d, \|s - d\|)$ (hellgrau) in welchen sich Geräte befinden müssen, um eine Nachricht von s nach d weiterleiten zu können - Abbildung aus [31]

Denn kann ein Gerät v_i das Datenpaket für s nach d weiterleiten, dann muss sich dieses nach den Regeln des Greedy geographic Routing näher an d befinden, als das Gerät s , von welchem die Nachricht ursprünglich kommt. Damit gilt $\|d - v_i\| < \|d - s\|$ und damit $v_i \in B(d, \|s - d\|)$.

Findet sich jedoch kein Gerät, welches die Nachricht weiterleiten kann, so muss für eine zuverlässige Übertragung der Nachricht die Übertragungsreichweite so gesetzt werden, dass eine direkte Übertragung von s nach d möglich ist, das heißt die Übertragungsreichweite vom Sender s müsste auf $\|s - d\|$ gesetzt werden.

Dadurch folgt, dass Geräte, welche eine Nachricht für das Gerät s weiterleiten können, in $B(s, \|s - d\|)$ liegen müssen.

Man kann nun aber die Übertragungsdistanz vom Sendegerät s so wählen, dass man genau ein Gerät findet, an welches s die Nachricht weiterleiten kann, ohne die Übertragungsdistanz so groß zu wählen, dass s die Nachricht direkt an den Empfänger d übertragen kann. Wählt man als Übertragungsdistanz r_s für s nun:

$$r_s = \min(\|v_1 - s\|, \|v_2 - s\|) \quad (3.5)$$

, so gilt im Beispiel von Abbildung 3.6 $r_s = \|v_1 - s\|$ und damit kann die Nachricht von s nach v_1 weitergeleitet werden.

Die Überlegung, die Übertragungsdistanz eines Gerätes so wie in Formel 3.5 zu wählen, dass sich genau ein Gerät findet an welches die Nachricht zum Ziel weitergeleitet werden kann, kann für alle Geräte durchgeführt werden um eine einheitliche Übertragungsdistanz zu finden, sodass diese Situation für jedes Gerät in diesem Netzwerk erfüllt wird. Bezeichnet man mit $\mathcal{X} \subset \mathbb{R}^2$ eine Menge an Punkten, beziehungsweise identifiziert man mit $\mathcal{X}$ die Standorte von verschiedenen Geräte in einem Netzwerk in der Ebene, so kann durch:

$$\rho(\mathcal{X}) = \max_{\substack{(s,d) \in \mathcal{X}^2 \\ s \neq d}} \left(\min_{v \in B(d, \|s-d\|) \cap \mathcal{X}} \|v - s\| \right) \quad (3.6)$$

die minimale Übertragungsdistanz angegeben werden, sodass für jedes Paar an Sender s und Empfänger d sich Geräte v finden, welche die Nachricht im Netzwerk weiterleiten können. Dadurch lässt sich der folgende Satz definieren:

Satz 3.1 (Wan et al. [31])

Sei $\mathcal{X} \subset \mathbb{R}^2$ eine Menge von Standorten von Geräten in einem ad-hoc WiFi Netzwerk und $\rho(\mathcal{X})$ wie in Formel 3.6 definiert und gleichzeitig die Übertragungsdistanz der Geräte in diesem Netzwerk.

Dann kann Greedy geographic Routing zuverlässig durchgeführt werden und ist frei von lokalen Minima. Dabei ist $\rho(\mathcal{X})$ der minimale Wert für den dies gilt.

Beweis zu Satz 3.1.

Nach der Definition von $\rho(\mathcal{X})$ und den obigen Überlegungen folgt sofort, dass eine Kommunikation zwischen zwei beliebigen Geräten s und d in dem Netzwerk zuverlässig durchgeführt werden kann.

Betrachtet man die Punkte $(s, d) \in \mathcal{X}^2$, für welche das Maximum in $\rho(\mathcal{X})$ angenommen wird und setzt die Übertragungsdistanz r_s von Gerät s kleiner als $\rho(\mathcal{X})$, das heißt $r_s < \rho(\mathcal{X})$. Dann kann weder s selbst die Nachricht direkt zum Gerät d übertragen, noch gibt es andere Geräte $v \in \mathcal{X}$, welche die Nachrichten von s nach d weiterleiten könnten. Dementsprechend wäre s ein lokales Minimum und die Nachrichtenübertragung würde nicht mehr zuverlässig funktionieren.

Daher ist $\rho(\mathcal{X})$ der minimale und optimale Wert um die Zuverlässigkeit von Greedy geographic Routing zu garantieren. $\square$

Zusätzlich zu diesem Ergebnis, zeigten Wan et al. in [31] ausserdem noch das asymptotische Verhalten von $\rho(\mathcal{X}_n)$, wenn $\mathcal{X}_n$ einen Poisson-Prozess mit Dichte n über einem

konvexen, kompakten Gebiet $\mathbb{D}$ beschreibt. Dies entspricht der Antwort auf die Frage, wie groß der Radius r eines geometrischen Zufallsgraphen mit Punktmenge $\mathcal{X}_n$ gewählt werden muss, damit Greedy geographic Routing in diesem Modell für ein ad-hoc Wi-Fi Netzwerk funktioniert. Zusätzlich liefert diese Analyse noch interessante Ansätze für reale Anwendungen.

Es gilt folgendes Verhältnis zwischen $\rho(\mathcal{X}_n)$ und $\sqrt{\frac{\ln(n)}{\pi n}}$ fast sicher:

$$\lim_{n \rightarrow \infty} \frac{\rho(\mathcal{X}_n)}{\sqrt{\frac{\ln(n)}{\pi n}}} = \sqrt{\frac{1}{\left(\frac{2}{3} - \frac{\sqrt{3}}{2\pi}\right)}} \approx 1.6 \quad (3.7)$$

Durch Formel 3.7 sei im Folgenden $\beta_0 := \frac{1}{\left(\frac{2}{3} - \frac{\sqrt{3}}{2\pi}\right)} \approx 1.6^2$ gewählt.

Satz 3.2 (Wan et al. [31])

Seien $\mathbb{D}, \mathcal{X}_n, \beta_0$ wie oben beschrieben. Dann gilt für alle $\epsilon > 0$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[(1 - \epsilon) \sqrt{\frac{\beta_0 \ln(n)}{n\pi}} \leq \rho(\mathcal{X}_n) \leq (1 + \epsilon) \sqrt{\frac{\beta_0 \ln(n)}{n\pi}} \right] = 1 \quad (3.8)$$

Aus Satz 3.2, wessen Beweis den Hauptteil des Papers von Wan et al. in [31] einnimmt, folgt der folgende Satz:

Satz 3.3 (Wan et al. [31])

Sei $\sqrt{\frac{\beta \ln(n)}{n\pi}}$ die Übertragungsdistanz von allen Geräten in einem ad-hoc WiFi Netzwerk und β_0 wie oben definiert. Dann gilt:

- wenn $\beta > \beta_0$ folgt, dass asymptotisch fast sicher die Kommunikation zwischen zwei beliebigen Geräten in diesen Gerät zuverlässig durchgeführt werden kann.
- wenn $\beta < \beta_0$ folgt, dass asymptotisch fast sicher die Kommunikation nicht zwischen jeden beliebigen Geräten zuverlässig durchgeführt werden kann.

Sätze 3.3 und 3.2 geben also Aufschluss darüber, welche Übertragungsdistanz die Geräte eines ad-hoc WiFi Netzwerkes haben sollten, um zuverlässiges Routing zu ermöglichen. Der verlässliche Aufbau von Kommunikation zwischen Geräten in einem ad-hoc Wi-Fi Netzwerken gehört mitunter zu den wichtigsten Punkten bei der Verwendung dieser

Netzwerke.

Basierend auf Satz 3.2 kann bei gegebener und fixierter Übertragungsdistanz von Signalen der Geräte in einem ad-hoc WiFi Netzwerk abgeschätzt werden, wieviele Geräte in einer bestimmten Region notwendig sind um das Funktionieren des Greedy geographic Routing zu garantieren und ein zusammenhängendes Netzwerk zu erhalten. Dies ist eine wichtige Information, da oftmals die gegebene Hardware nicht nach Wunsch modifiziert werden kann und dementsprechend die notwendige Anzahl an Geräten zum Aufbau eines zuverlässigen Netzwerkes bedacht werden kann.

Kann in einem Netzwerk die zuverlässige Übertragung von Nachrichten via Greedy geographic Routing garantiert werden, so kann auf komplexe Routing-Algorithmen verzichtet werden und sämtliche Entscheidungen bezüglich Routing von Datenpaketen können durch das Greedy-Verfahren getroffen werden, welches relativ einfach zu implementieren ist.

Simulationen zu Greedy geographic Routing

Für eine verschiedene Anzahl an Geräten $N \in \{200, 400, 800\}$, aufgeteilt in einem Einheitskreis $\mathbb{D} = B(0, 1)$ wurden für jede Netzwerkgröße 400 ad-hoc WiFi Netzwerke anhand von uniform verteilten Poisson Prozessen generiert. Anschließend wurde die minimale Übertragungsdistanz $\rho(\mathcal{X}_N)$ laut Formel 3.6 berechnet, als auch der approximative Wert ρ_N laut Satz 3.2, welche miteinander verglichen wurden. Die 400 verschiedenen Werte von $\rho(\mathcal{X}_N)$ für eine bestimmte Anzahl an Geräten N im Netzwerk wurde dabei gemittelt.

In Tabelle 3.1 finden sich die berechneten Werte, welche aus den Simulationen von Wan et al. folgten. Es zeigt sich, dass die konkreten Werte $\rho(\mathcal{X}_N)$ in allen Simulationen etwas höher sind, als die von Satz 3.2 asymptotische Näherung für ρ_N - das heißt, die Konvergenz wird erst für eine größere Anzahl an Geräten im Netzwerk wirklich sichtbar. Dennoch sind die Werte ρ_N nicht all zu schlecht für kleine Netzwerkgrößen, der Fehler $\frac{|\rho(\mathcal{X}_N) - \rho_N|}{\rho_N}$ beläuft sich auf etwa 20%.

N	$\rho(\mathcal{X}_N)$	ρ_N
200	0.1808	0.1469
400	0.1332	0.1104
800	0.1000	0.0825

Tabelle 3.1: Vergleich von $\rho(\mathcal{X}_N)$ und ρ_N für verschiedene Netzwerkgrößen - Daten der Simulationen aus [31]

Da $\rho_N < \rho(\mathcal{X}_N)$ für alle simulierten Netzwerke gilt, folgt, dass der Näherung ρ_N für $\rho(\mathcal{X}_N)$ aus Satz 3.2 den wahren Wert leicht unterschätzt. Würde man die maximale

Übertragungsdistanz r der Geräte mit ρ_N wählen, könnte in diesen Netzwerken Greedy geographic Routing nicht zuverlässig funktionieren, da die Übertragungreichweite zu gering wäre.

Weiters wurden in den Simulationen die Auswirkungen untersucht, wenn die maximale Übertragungsdistanz der Geräte in diesen Netzwerken einen Wert kleiner als der optimalen Übertragungsdistanz $\rho(\mathcal{X}_N)$ annimmt. Für kleinere Übertragungsdistanzen $r < \rho(\mathcal{X}_N)$ der Geräte in dem Netzwerk wurde die Anzahl der Paare $(u, v) \in \mathcal{X}_N$ berechnet, für welche eine zuverlässige Kommunikation zwischen u und v durch Greedy geographic Routing möglich ist. Ist hingegen die Übertragungsdistanz groß genug gewählt, funktioniert Greedy geographic Routing zuverlässig und dementsprechend gilt für alle Paare $(u, v) \in \mathcal{X}_N$, dass eine zuverlässige Kommunikation möglich ist.

Ausgehend von der optimalen Übertragungsdistanz $\rho(\mathcal{X}_N)$ wurde bei den verschiedenen Netzwerken als maximale Übertragungsdistanz r der Geräte folgende Werte gewählt:

$$\begin{aligned} r &= \rho(\mathcal{X}_N) \cdot s \\ s &\in \{0.80, 0.85, 0.90, 0.95, 0.99\} \end{aligned}$$

Dabei ist s ein Faktor, welcher die maximale Übertragungsdistanz verringert. Anschließend wurde das Verhältnis zwischen der Anzahl der Geräte, zwischen welchen eine zuverlässige Kommunikation noch möglich ist, und der Gesamtanzahl $\binom{N}{2}$ aller Paare $(u, v) \in \mathcal{X}_N$ berechnet.

Abbildung 3.7 zeigt die Ergebnisse der oben beschriebenen Analyse von der Auswirkung einer Änderung der Übertragungsdistanz der Geräte auf die Zuverlässigkeit von Greedy geographic Routing. Es zeigt sich, dass in schwach besetzten Netzwerken mit einer geringen Gerätedichte die Zuverlässigkeit des Routing signifikant stärker abnimmt als in dicht besetzten Netzwerken bei derselben Reduktion von $\rho(\mathcal{X}_N)$ durch s .

Trotz Verringerung der maximalen Übertragungsdistanz der Geräte um 20% des optimalen Wertes zeigt das Greedy geographic Routing im schlechtesten Fall des schwach besetzten Netzwerks mit $N = 200$ Geräten trotzdem noch eine beachtliche Zuverlässigkeit, indem 92% aller Gerätepaare in diesem Netzwerk noch miteinander kommunizieren können.

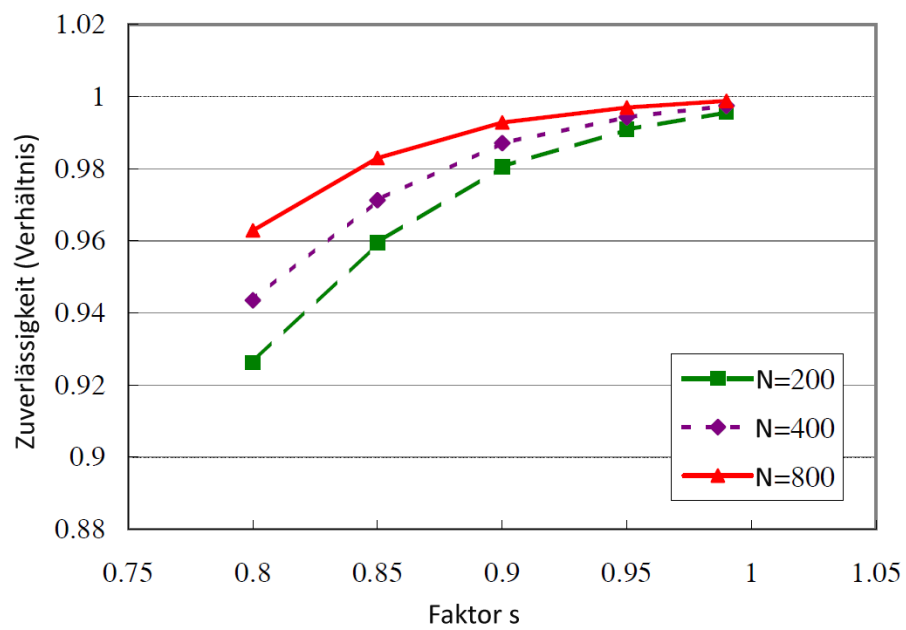


Abbildung 3.7: Zuverlässigkeit von Greedy geographic Routing - Abbildung aus [31]

Literaturverzeichnis

- [1] Martin J. B. Appel and Ralph P. Russo. The maximum vertex degree of a graph on uniform points in $[0, 1]^d$. *Advances in Applied Probability*, 29(3):567–581, 1997.
- [2] Martin J. B. Appel and Ralph P. Russo. The minimum vertex degree of a graph on uniform points in $[0, 1]^d$. *Advances in Applied Probability*, 29(3):582–594, 1997.
- [3] Vladimir Batagelj. Semirings for social networks analysis. *The Journal of Mathematical Sociology*, 19(1):53–68, 1994.
- [4] Ulrik Brandes. A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology*, 25(2):163–177, 2001.
- [5] Thomas M Chen and J-M Robert. Worm epidemics in high-speed networks. *Computer*, 37(6):48–53, 2004.
- [6] Justin Coon, Carl P. Dettmann, and Orestis Georgiou. Full connectivity: Corners, edges and faces. *Journal of Statistical Physics*, 147(4):758–778, Jun 2012.
- [7] Thomas H Cormen. *Introduction to algorithms*. MIT press, 2009.
- [8] R. Diestel. *Graphentheorie*. Springer-Lehrbuch Masterclass. Springer, 2010.
- [9] P Erdős and A Rényi. On random graphs i, publi. math.
- [10] Paul Erdos and Alfréd Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist*, 38(4):343–347, 1961.
- [11] Gregory G Finn. Routing and addressing problems in large metropolitan-scale internetworks. isi research report. 1987.
- [12] Orestis Georgiou. Random graphs and wireless communication networks. 2016.
- [13] E. N. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):533–543, 1961.
- [14] A. P. Giles, O. Georgiou, and C. P. Dettmann. Betweenness centrality in dense random geometric networks. In *2015 IEEE International Conference on Communications (ICC)*, pages 6450–6455, June 2015.
- [15] Per Hage and Frank Harary. Eccentricity and centrality in networks. *Social Networks*, 17(1):57 – 63, 1995.

- [16] Frédéric Havet. Vorlesungsaufzeichnungen aus der vorlesung combinatorial optimization.
- [17] Jeroen Hoebeke, Ingrid Moerman, Bart Dhoedt, and Piet Demeester. An overview of mobile ad hoc networks: applications and challenges. *Journal-Communications Network*, 3(3):60–66, 2004.
- [18] John E Hopcroft and Jeffrey D Ullman. *Data structures and algorithms*. Addison-Wesley, 1987.
- [19] Hichem Kenniche and Vlady Ravelomananana. Random geometric graphs as model of wireless sensor networks. In *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on*, volume 4, pages 103–107. IEEE, 2010.
- [20] Norbert Kusolitsch. *Maß-und Wahrscheinlichkeitstheorie: Eine Einführung*. Springer-Verlag, 2014.
- [21] Yamir Moreno, Romualdo Pastor-Satorras, and Alessandro Vespignani. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 26(4):521–529, 2002.
- [22] James D Murray. Mathematical biology. i, volume 17 of interdisciplinary applied mathematics, 2002.
- [23] Maziar Nekovee. Worm epidemics in wireless ad hoc networks. *New Journal of Physics*, 9(6):189, 2007.
- [24] Juhani Nieminen. On the centrality in a graph. *Scandinavian journal of psychology*, 15(1):332–336, 1974.
- [25] M. Penrose. *Random Geometric Graphs*. Oxford studies in probability. Oxford University Press, 2003.
- [26] Mathew D. Penrose. A strong law for the largest nearest-neighbour link between random points. *Journal of the London Mathematical Society*, 60(3):951, 1999.
- [27] Yannick Rochat. Closeness centrality extended to unconnected graphs: The harmonic centrality index. In *ASNA*, number EPFL-CONF-200525, 2009.
- [28] Gert Sabidussi. The centrality index of a graph. *Psychometrika*, 31(4):581–603, Dec 1966.
- [29] Alfonso Shimbel. Structural parameters of communication networks. *The bulletin of mathematical biophysics*, 15(4):501–507, 1953.
- [30] Ivan Stojmenovic. Position-based routing in ad hoc networks. *IEEE communications magazine*, 40(7):128–134, 2002.

- [31] P-J Wan, C-W Yi, Lixin Wang, Frances Yao, and Xiaohua Jia. Asymptotic critical transmission radii for greedy forward routing in wireless ad hoc networks. *IEEE Transactions on communications*, 57(5), 2009.
- [32] Armen H Zemanian. *Distribution theory and transform analysis: an introduction to generalized functions, with applications*. Courier Corporation, 1965.

Abbildungsverzeichnis

1.1	Approximation der Delta-Distribution $\delta_a(x)$ mit $a \in \{0.75, 0.25, 0.15\}$. . .	10
2.1	Geometrischer Zufallsgraph $G(\mathcal{X}_{250}, 0.1)$ mit 250 Knoten	13
2.2	$G(100, 0.1)$ mit Clique-Zahl 6	37
2.3	Kommunikationsnetzwerk mit Router 1 und 4 Zugangspunkten 2, 3, 4 und 5	48
2.4	Zusammenhang zwischen κ , r_i , r_j und $\kappa_{\perp}$	51
2.5	Verschiedene Zentralitäts-Maße visualisiert an einem geometrischen Zufallsgraphen $G(\mathcal{X}_{50}, 0.3)$	56
2.6	Veranschaulichung der 2 Typen des Erdős-Rényi Zufallsgraphen $G(n, M)$ und $G(n, p)$	59
3.1	Ein ad-hoc WiFi Netzwerk mit 13 Geräten und die Gegenüberstellung eines möglichen Modells für dieses Netzwerk als geometrischer Zufallsgraph	63
3.2	Anteil an infizierten Geräten $\frac{I(t)}{N}$ im zeitlichen Verlauf der Epidemie - Abbildung aus [23]	69
3.3	Zeit bis zum Erreichen des maximalen Anteils an infizierten Geräten in Abhängigkeit zur Anzahl der Geräte N - Abbildung aus [23]	71
3.4	G_{∞} als Funktion von λ - Abbildung aus [23]	72
3.5	Skizze der Funktionsweise von Greedy geographic Routing anhand eines Gerätes s , welches ein Datenpaket an ein Gerät d sendet	75
3.6	Veranschaulichung der Fläche $B(s, s - d) \cap B(d, s - d)$ (hellgrau) in welchen sich Geräte befinden müssen, um eine Nachricht von s nach d weiterleiten zu können - Abbildung aus [31]	76
3.7	Zuverlässigkeit von Greedy geographic Routing - Abbildung aus [31] . . .	81

Symbolverzeichnis

$B(0, 1)$	d -dimensionale Einheitskugel	16
$B_{n,p}(k)$	Binomial-Verteilung mit Parameter n und p	7
$C_B(w)$	Betweenness-Zentralität eines Knotens w	46
$\mathcal{C}$	Clique-Zahl des Graphen G	4
$\Delta(G)$	Maximalgrad von G	3
$\delta(G)$	Minimalgrad von G	3
$d_G(v)$	Knotengrad von v im Graphen G	3
$d_G(v, w)$	Distanz zweier Knoten v und w im Graphen G	4
$f _A$	Einschränkung einer Funktion $f : D \rightarrow B$ auf den Definitionsbereich $A \subset D$	1
f_U	Dichte der uniformen Wahrscheinlichkeitsverteilung	32
$G(\mathcal{X}_n, r)$	Geometrischer Zufallsgraph mit Knotenmenge $\mathcal{X}_n$ und Radius r	12
$G(n, M)$	Erdős-Rényi Zufallsgraph mit n Knoten und M Kanten	56
$G(n, p)$	Erdős-Rényi Zufallsgraph mit n Knoten und Kantenwahrscheinlichkeit p	57
$G(V, E)$	Ungerichteter Graph mit Knotenmenge V und Kantenmenge E	2
κ	Zusammenhangszahl κ eines Graphen G	41
λ	Lebesgue-Maß	16
l_∞	l_∞ -Norm	31
$M_k(\mathcal{X})$	größter k -nächsten Nachbarn Radius der Punkte $\mathcal{X}$	17
$\sigma(U, r)$	Packing-Nummer	11
$P_\lambda(k)$	Poisson-Verteilung mit Parameter λ	6

$S_k(\mathcal{X})$	kleinster k -nächsten Nachbarn Radius der Punkte $\mathcal{X}$	32
σ_{uv}	Anzahl der kürzesten Pfade zwischen den Knoten u und v	46
$\sigma_{uv}(w)$	Anzahl der kürzesten Pfade zwischen den Knoten u und v , welche durch den Knoten w führen	46
θ	Volumen der d-dimensionalen Einheitskugel $B(0, 1)$	16
$\chi(G)$	Chromatische Zahl des Graphen G	39