Seminar "Zufällige Netzwerke"

Konnektivität endlicher Netzwerke

Kathrin Stucke
Henrik Haßfeld

Abbildung: Zufallsgitter

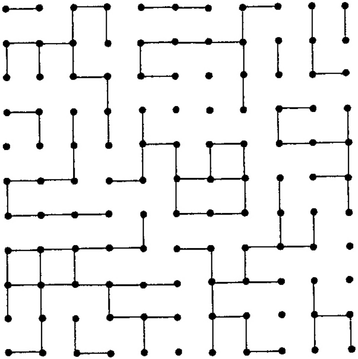
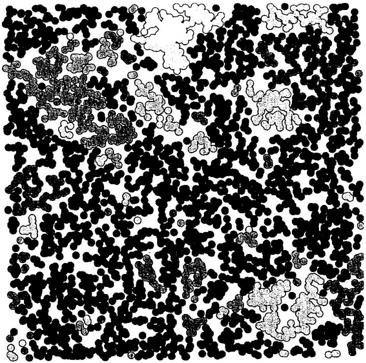
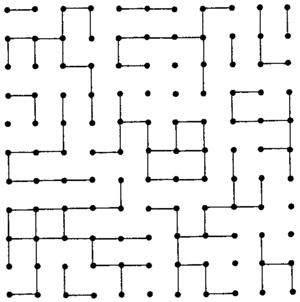


Abbildung: Boolesches Modell

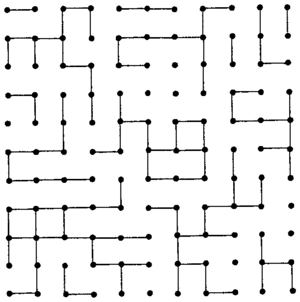


Das Zufallsgitter



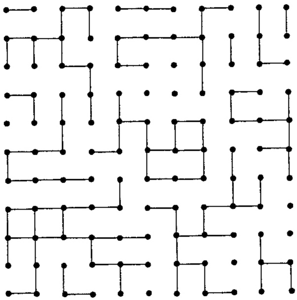
- unendliches, quadratisches Gitter, mit Punkten besetzt

Das Zufallsgitter



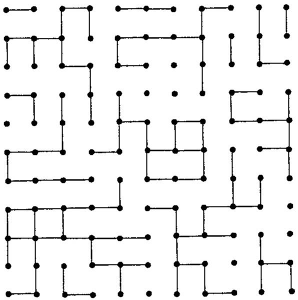
- unendliches, quadratisches Gitter, mit Punkten besetzt
- Modellparameter p (*Kantenwahrscheinlichkeit*): Wahrscheinlichkeit, dass benachbarte Gitterpunkte verbunden sind

Das Zufallsgitter



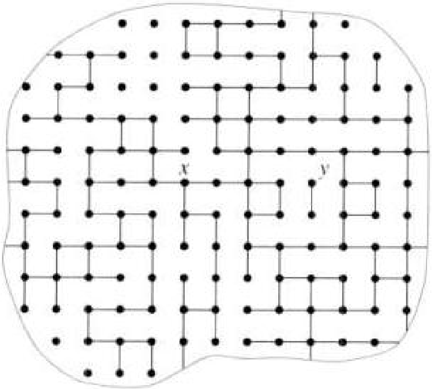
- unendliches, quadratisches Gitter, mit Punkten besetzt
- Modellparameter p (*Kantenwahrscheinlichkeit*): Wahrscheinlichkeit, dass benachbarte Gitterpunkte verbunden sind
- Das Ereignis, dass zwei Punkte miteinander verbunden sind, ist *unabhängig* von allen bereits bestehenden Verbindungen, und tritt immer mit Wahrscheinlichkeit p ein.

Das Zufallsgitter



- unendliches, quadratisches Gitter, mit Punkten besetzt
- Modellparameter p (*Kantenwahrscheinlichkeit*): Wahrscheinlichkeit, dass benachbarte Gitterpunkte verbunden sind
- Das Ereignis, dass zwei Punkte miteinander verbunden sind, ist *unabhängig* von allen bereits bestehenden Verbindungen, und tritt immer mit Wahrscheinlichkeit p ein.
- Wir sprechen dabei auch von offenen und geschlossenen Kanten, wobei geschlossen bedeutet, dass eine Verbindung zwischen den zugehörigen Punkten besteht.

Zufallsgitter — Motivation



Fast-Konnektivität

- **Definition:** *α-fast Konnektivität*
  Sei $\alpha \in (0, 1)$. Eine Konfiguration von offenen und geschlossenen Kanten auf einem quadratischen Gitter der Größe n x n heißt α-fast zusammenhängend, wenn es eine Zusammenhangskomponente von mindestens αn^2 Gitterpunkten gibt.

Fast-Konnektivität

- **Definition:** *α-fast Konnektivität*
  Sei $\alpha \in (0,1)$. Eine Konfiguration von offenen und geschlossenen Kanten auf einem quadratischen Gitter der Größe n x n heißt α-fast zusammenhängend, wenn es eine Zusammenhangskomponente von mindestens αn^2 Gitterpunkten gibt.

- **Definition:** *Perkolationsfunktion*
  Die Perkolationsfunktion $\theta(p)$ ist die Wahrscheinlichkeit, dass sich der Ursprung (oder ein beliebiger Punkt) in einer unendlichen Zusammenhangskomponente von Gitterpunkten befindet.
  Sei C_0 die Menge der Punkte, die mit dem Ursprung verbunden sind.
  Dann gilt:

$$\theta(p) = P_p(|C_0| = \infty).$$

Fast-Konnektivität

- Beobachten Teilgraph G_n der Größe n x n.

- **Theorem 1.1** Sei $\alpha \in (0,1)$ und sei $p_\alpha := \inf\{p, \theta(p) > \alpha\}$. Dann gilt:

  Für $p > p_\alpha$ ist G_n α-fast zusammenhängend a.f.s. (asymptotisch fast-sicher), d.h.

  $$\lim_{n\to\infty} P_p(G_n\, hat\, Zshg's - komponente\, von\, mindestens\, \alpha n^2\, Punkten) = 1.$$

  Für $p < p_\alpha$ ist G_n nicht α-fast zusammenhängend a.f.s.

Volle Konnektivität

- Untersuchen Situation, dass alle Gitterpunkte in G_n miteinander verbunden sind.

Volle Konnektivität

- Untersuchen Situation, dass alle Gitterpunkte in G_n miteinander verbunden sind.
- Um ein komplett verbundenes Cluster zu erhalten, muss die Kantenwahrscheinlichkeit p_n entsprechend mitwachsen.

Volle Konnektivität

- Untersuchen Situation, dass alle Gitterpunkte in G_n miteinander verbunden sind.
- Um ein komplett verbundenes Cluster zu erhalten, muss die Kantenwahrscheinlichkeit p_n entsprechend mitwachsen.
- Zielstellung: Wachstumsrate von p_n.

Volle Konnektivität

- Untersuchen Situation, dass alle Gitterpunkte in G_n miteinander verbunden sind.
- Um ein komplett verbundenes Cluster zu erhalten, muss die Kantenwahrscheinlichkeit p_n entsprechend mitwachsen.
- Zielstellung: Wachstumsrate von p_n.

- **Theorem 1.2** Sei W_n die Anzahl isolierter Gitterpunkte in G_n. Dann gilt: W_n konvergiert in Verteilung gegen eine poisson-verteilte Zufallsvariable mit Parameter λ $(0 < \lambda < \infty)$ genau dann, wenn

$$n^2(1 - p_n)^4 \xrightarrow{n \to \infty} \lambda \tag{1}$$

Volle Konnektivität

- **Definition:** *Totalvariationsabstand*
  Der Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsverteilungen P und Q auf $\mathbb{N}$ ist definiert als

$$d_{TV}(P, Q) = \sup\{|P(A) - Q(A)| : A \subset \mathbb{N}\}$$

Volle Konnektivität

- **Definition:** *Totalvariationsabstand*
  Der Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsverteilungen P und Q auf $\mathbb{N}$ ist definiert als

$$d_{TV}(P,Q) = \sup\{|P(A) - Q(A)| : A \subset \mathbb{N}\}$$

- **Bezeichnungen:**

Volle Konnektivität

- **Definition:** *Totalvariationsabstand*
  Der Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsverteilungen P und Q auf $\mathbb{N}$ ist definiert als

$$d_{TV}(P,Q) = \sup\{|P(A) - Q(A)| : A \subset \mathbb{N}\}$$

- **Bezeichnungen:**
  - $\mathcal{I}$ beliebige Indexmenge, $i \in \mathcal{I}$

Volle Konnektivität

- **Definition:** *Totalvariationsabstand*
  Der Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsverteilungen P und Q auf $\mathbb{N}$ ist definiert als

$$d_{TV}(P,Q) = \sup\{|P(A) - Q(A)| : A \subset \mathbb{N}\}$$

- **Bezeichnungen:**
  - $\mathcal{I}$ beliebige Indexmenge, $i \in \mathcal{I}$
  - 1_i Indikatorvariable mit $\mathbb{E}(1_i) = q_i$

Volle Konnektivität

- **Definition:** *Totalvariationsabstand*
  Der Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsverteilungen P und Q auf $\mathbb{N}$ ist definiert als

$$d_{TV}(P,Q) = \sup\{|P(A) - Q(A)| : A \subset \mathbb{N}\}$$

- **Bezeichnungen:**
  - $\mathcal{I}$ beliebige Indexmenge, $i \in \mathcal{I}$
  - 1_i Indikatorvariable mit $\mathbb{E}(1_i) = q_i$
  - $\lambda = \sum_{i \in \mathcal{I}} q_i < \infty$

Volle Konnektivität

- **Definition:** *Totalvariationsabstand*
  Der Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsverteilungen P und Q auf $\mathbb{N}$ ist definiert als

$$d_{TV}(P, Q) = \sup\{|P(A) - Q(A)| : A \subset \mathbb{N}\}$$

- **Bezeichnungen:**
  - $\mathcal{I}$ beliebige Indexmenge, $i \in \mathcal{I}$
  - 1_i Indikatorvariable mit $\mathbb{E}(1_i) = q_i$
  - $\lambda = \sum_{i \in \mathcal{I}} q_i < \infty$
  - $W = \sum_{i \in \mathcal{I}} 1_i$
    $\Rightarrow \mathbb{E}(W) = \lambda$

Volle Konnektivität

- **Definition:** *Totalvariationsabstand*
  Der Totalvariationsabstand zwischen zwei Wahrscheinlichkeitsverteilungen P und Q auf $\mathbb{N}$ ist definiert als

$$d_{TV}(P, Q) = \sup\{|P(A) - Q(A)| : A \subset \mathbb{N}\}$$

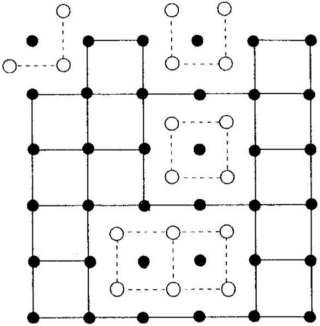
- **Bezeichnungen:**
  - $\mathcal{I}$ beliebige Indexmenge, $i \in \mathcal{I}$
  - 1_i Indikatorvariable mit $\mathbb{E}(1_i) = q_i$
  - $\lambda = \sum_{i \in \mathcal{I}} q_i < \infty$
  - $W = \sum_{i \in \mathcal{I}} 1_i$
    $\Rightarrow \mathbb{E}(W) = \lambda$
  - $Po(\lambda)$ poisson-verteilte ZV mit Parameter λ

Volle Konnektivität

- **Theorem 1.3** *Chen-Stein-Schranke*
  Seien $X_1, \ldots, X_k$ unabhängige Zufallsvariablen über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ und für jedes i aus einer beliebigen Indexmenge $\mathcal{I}$ sei $1_i(\omega) = 1_{A_i}((X_1, \ldots, X_k)(\omega))$ für ein $A_i \in \mathfrak{B}(\mathbb{R}^k)$. Dann gilt:

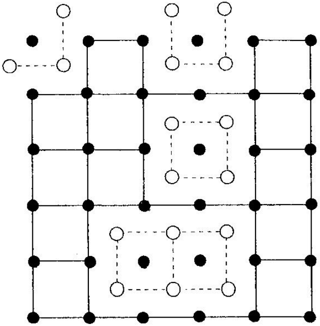
$$d_{TV}(W, Po(\lambda)) \leq \frac{1 - e^{-\lambda}}{\lambda}(\mathrm{Var}\, W - \lambda + 2 \sum_{i \in \mathcal{I}} q_i^2) \tag{2}$$

Dualer Graph



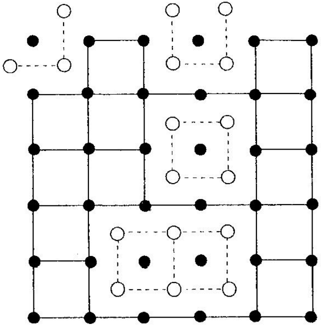
- dualer Punkt in jedes Gitterquadrat (am Rand fortsetzen)

Dualer Graph



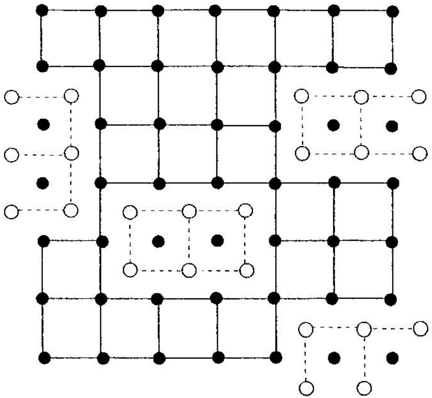
- dualer Punkt in jedes Gitterquadrat (am Rand fortsetzen)
- Verbindung zwischen zwei dualen Punkten wird gesetzt, wenn diese Verbindung keine Verbindung des ursprünglichen Graphen schneidet.

Dualer Graph



- dualer Punkt in jedes Gitterquadrat (am Rand fortsetzen)
- Verbindung zwischen zwei dualen Punkten wird gesetzt, wenn diese Verbindung keine Verbindung des ursprünglichen Graphen schneidet.
- ein Punkt (des ursprünglichen Graphen) ist *isoliert*, wenn er vom dualen Graphen umschlossen ist.

Dualer Graph



Volle Konnektivität

- **Theorem 1.4** Sei $p_n = 1 - \frac{c_n}{\sqrt{n}}$ und sei A_n das Ereignis, dass G_n keine isolierten Gitterpunkte enthält. Dann gilt:

$$\lim_{n\to\infty} P(A_n) = e^{-c^4} \text{ genau dann, wenn } \; c_n \longrightarrow c \; \text{ für } \; n \to \infty$$

Volle Konnektivität

- **Theorem 1.4** Sei $p_n = 1 - \frac{c_n}{\sqrt{n}}$ und sei A_n das Ereignis, dass G_n keine isolierten Gitterpunkte enthält. Dann gilt:

$$\lim_{n\to\infty} P(A_n) = e^{-c^4} \text{ genau dann, wenn } \; c_n \longrightarrow c \; \text{ für } \; n \to \infty$$

- **Proposition 1.5** Sei $p_n = 1 - \frac{c_n}{\sqrt{n}}$, wobei $c_n \to c \in (0, \infty)$. Dann gibt es asymptotisch fast-sicher (genau) eine Zusammenhangskomponente, und alle Punkte, die nicht Teil dieser Zusammenhangskopmponente sind, sind isoliert.

Volle Konnektivität

- **Theorem 1.4** Sei $p_n = 1 - \frac{c_n}{\sqrt{n}}$ und sei A_n das Ereignis, dass G_n keine isolierten Gitterpunkte enthält. Dann gilt:

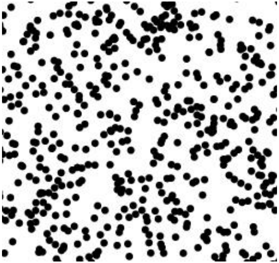
$$\lim_{n\to\infty} P(A_n) = e^{-c^4} \text{ genau dann, wenn } c_n \longrightarrow c \text{ für } n \to \infty$$

- **Proposition 1.5** Sei $p_n = 1 - \frac{c_n}{\sqrt{n}}$, wobei $c_n \to c \in (0, \infty)$. Dann gibt es asymptotisch fast-sicher (genau) eine Zusammenhangskomponente, und alle Punkte, die nicht Teil dieser Zusammenhangskopmponente sind, sind isoliert.

- **Korollar 1.6** Sei $p_n = 1 - \frac{c_n}{\sqrt{n}}$.
Dann ist G_n voll zusammenhängend a.f.s., genau dann, wenn $c_n \to 0$.

Poisson-Prozess

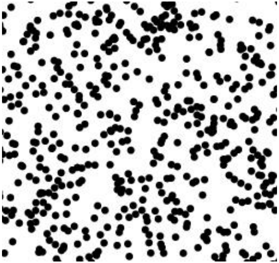
Definition: *Homogener Poisson-Prozess in* $\mathbb{R}^2$
Eine zufällige Menge von Punkten $X \subset \mathbb{R}^2$ heißt *homogener Poisson-Prozess mit Intensität* λ, falls



Poisson-Prozess

Definition: *Homogener Poisson-Prozess in* $\mathbb{R}^2$
Eine zufällige Menge von Punkten $X \subset \mathbb{R}^2$ heißt *homogener Poisson-Prozess mit Intensität* λ, falls

(i) Für paarweise disjunkte (Borel-)Mengen $D_1, \ldots, D_k \subset \mathbb{R}^2$ sind die Zufallsvariablen $X(D_1), \ldots, X(D_k)$ unabhängig, wobei X(D) die Anzahl der Punkte von X in D bezeichnet.



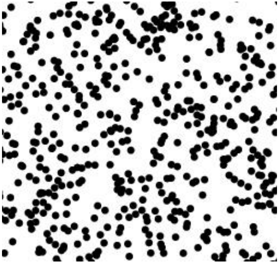
Poisson-Prozess

Definition: *Homogener Poisson-Prozess in* $\mathbb{R}^2$
Eine zufällige Menge von Punkten $X \subset \mathbb{R}^2$ heißt *homogener Poisson-Prozess mit Intensität* λ, falls

(i) Für paarweise disjunkte (Borel-)Mengen $D_1, \ldots, D_k \subset \mathbb{R}^2$ sind die Zufallsvariablen $X(D_1), \ldots, X(D_k)$ unabhängig, wobei X(D) die Anzahl der Punkte von X in D bezeichnet.

(ii) Für jede beschränkte (Borel-)Menge $D \subset \mathbb{R}^2$ gilt für jedes k=0,1,2,... :

$$P(X(D) = k) = e^{-\lambda |D|} \frac{(\lambda |D|)^k}{k!}$$



Poisson-Prozess

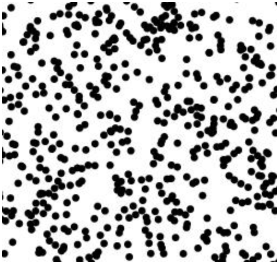
Definition: *Homogener Poisson-Prozess in* $\mathbb{R}^2$
Eine zufällige Menge von Punkten $X \subset \mathbb{R}^2$ heißt *homogener Poisson-Prozess mit Intensität* λ, falls

(i) Für paarweise disjunkte (Borel-)Mengen $D_1, \ldots, D_k \subset \mathbb{R}^2$ sind die Zufallsvariablen $X(D_1), \ldots, X(D_k)$ unabhängig, wobei X(D) die Anzahl der Punkte von X in D bezeichnet.

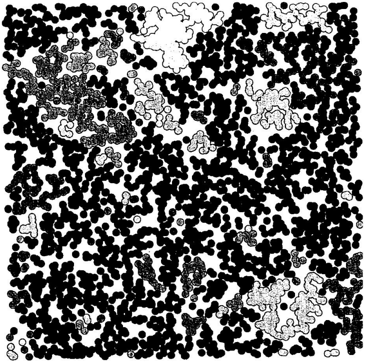
(ii) Für jede beschränkte (Borel-)Menge $D \subset \mathbb{R}^2$ gilt für jedes k=0,1,2,... :

$$P(X(D) = k) = e^{-\lambda |D|} \frac{(\lambda |D|)^k}{k!}$$

Beachte Es gilt: $\mathbb{E}(X(D)) = \lambda |D|$.

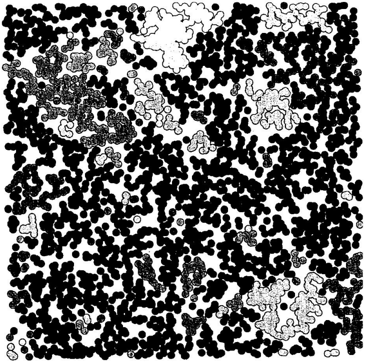


Das Boolesche Modell (X, λ, r)



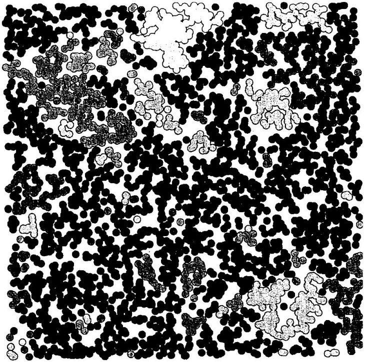
- X homogener Poisson-Prozess mit Intensität λ

Das Boolesche Modell (X, λ, r)



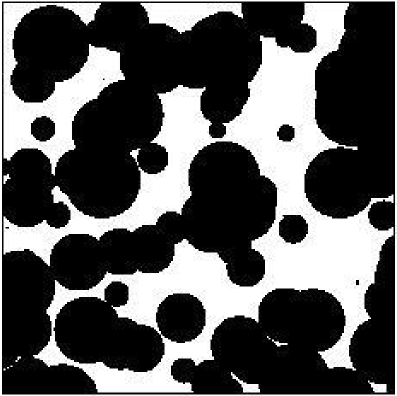
- X homogener Poisson-Prozess mit Intensität λ
- r ist der Radius der Kreisscheiben, deren Zentren die Punkte des Poisson-Prozesses sind

Das Boolesche Modell (X, λ, r)



- X homogener Poisson-Prozess mit Intensität λ
- r ist der Radius der Kreisscheiben, deren Zentren die Punkte des Poisson-Prozesses sind
- zwei Punkte sind verbunden, wenn sich ihre Kreisscheiben überlappen.

Boolesches Modell – Motivation



Konnektivität im Booleschen Modell

- Betrachten Boolesches Modell $(X, \lambda = 1, r > 0)$

Konnektivität im Booleschen Modell

- Betrachten Boolesches Modell $(X, \lambda = 1, r > 0)$
- Die Perkolationsfunktion θ(r) ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Punkt Teil einer unendlichen Zusammenhangskomponente ist,

Konnektivität im Booleschen Modell

- Betrachten Boolesches Modell $(X, \lambda = 1, r > 0)$
- Die Perkolationsfunktion θ(r) ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Punkt Teil einer unendlichen Zusammenhangskomponente ist,
- präziser: Wie hoch ist die Wahrscheinlichkeit, dass
    - unter der Bedingung $0 \in X$
    - 0 in einer unendlichen Zusammenhangskomponente liegt?

Konnektivität im Booleschen Modell

- Betrachten Boolesches Modell $(X, \lambda = 1, r > 0)$
- Die Perkolationsfunktion θ(r) ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Punkt Teil einer unendlichen Zusammenhangskomponente ist,
- präziser: Wie hoch ist die Wahrscheinlichkeit, dass
  - unter der Bedingung $0 \in X$
  - 0 in einer unendlichen Zusammenhangskomponente liegt?
- Dies führt auf den Begriff der *Palm-Verteilung*

$$P_N^0(A) = \lim_{\epsilon \to 0} P(A \,|\, X \cap B(o, \epsilon) \neq \emptyset)$$

Konnektivität im Booleschen Modell

- Beobachten Teilgraph $G_n(r)$ in einem $\sqrt{n} \times \sqrt{n}$ Quadrat $B_n \subset \mathbb{R}^2$.

Konnektivität im Booleschen Modell

- Beobachten Teilgraph $G_n(r)$ in einem $\sqrt{n} \times \sqrt{n}$ Quadrat $B_n \subset \mathbb{R}^2$.
- Außerdem sei $N_\infty(B_n)$ die Anzahl der Punkte des Poisson-Prozesses X in B_n, die Teil einer unendlichen Zusammenhangskomponente des Booleschen Modells im $\mathbb{R}^2$ sind.

Konnektivität im Booleschen Modell

- Beobachten Teilgraph $G_n(r)$ in einem $\sqrt{n} \times \sqrt{n}$ Quadrat $B_n \subset \mathbb{R}^2$.
- Außerdem sei $N_\infty(B_n)$ die Anzahl der Punkte des Poisson-Prozesses X in B_n, die Teil einer unendlichen Zusammenhangskomponente des Booleschen Modells im $\mathbb{R}^2$ sind.

- **Proposition 2.1** Es gilt $\theta(r) = \mathbb{E}(N_\infty(B_1))$.

Fast-Konnektivität

- **Definition:** α-*fast Konnektivität*
  Sei $\alpha \in (0,1)$. $G_n(r)$ heißt α-*fast zusammenhängend* $:\Leftrightarrow$ $G_n(r)$ enthält eine Zusammenhangskomponente von mindestens αn Punkten.

Fast-Konnektivität

- **Definition:** α-*fast Konnektivität*
  Sei $\alpha \in (0,1)$. $G_n(r)$ heißt α-*fast zusammenhängend* $:\Leftrightarrow$ $G_n(r)$ enthält eine Zusammenhangskomponente von mindestens αn Punkten.

- **Beachte:** Zwei Unterschiede zum diskreten Fall:
  1. αn Punkte statt αn^2. Begründung: B_n ist $\sqrt{n} \times \sqrt{n}$ Quadrat
  2. die *durchschnittliche Anzahl* an Punkten in B_n ist n (im diskreten Fall konnte die Anzahl der Punkte exakt angegeben werden).

Fast-Konnektivität

- **Definition:** *α-fast Konnektivität*
  Sei $\alpha \in (0,1)$. $G_n(r)$ heißt *α-fast zusammenhängend* $:\Leftrightarrow$ $G_n(r)$ enthält eine Zusammenhangskomponente von mindestens αn Punkten.

- **Beachte:** Zwei Unterschiede zum diskreten Fall:
  1. αn Punkte statt αn^2. Begründung: B_n ist $\sqrt{n} \times \sqrt{n}$ Quadrat
  2. die *durchschnittliche Anzahl* an Punkten in B_n ist n (im diskreten Fall konnte die Anzahl der Punkte exakt angegeben werden).

- **Theorem 2.2** Sei $r_\alpha = \inf\{r : \theta(r) > \alpha\}$. Für $\alpha \in (0,1)$ gilt:
  - Für $r > r_\alpha$ ist $G_n(r)$ α-fast zusammenhängend a.f.s.
  - Für $r < r_\alpha$ ist $G_n(r)$ nicht α-fast zusammenhängend a.f.s.

Volle Konnektivität

- Untersuchen Situation, dass <u>alle</u> Punkte in B_n verbunden sind.

Volle Konnektivität

- Untersuchen Situation, dass alle Punkte in B_n verbunden sind.
- Wenn wir aber ein komplett verbundenes Cluster innerhalb einer wachsenden Box B_n betrachten, muss der Radius r mitwachsen.

Volle Konnektivität

- Untersuchen Situation, dass <u>alle</u> Punkte in B_n verbunden sind.
- Wenn wir aber ein komplett verbundenes Cluster innerhalb einer wachsenden Box B_n betrachten, muss der Radius r mitwachsen.
- Zielstellung: exakte Wachstumsrate für r_n.

Dazu betrachten wir zunächst das folgende vorläufige Ergebnis:

Volle Konnektivität

- Untersuchen Situation, dass alle Punkte in B_n verbunden sind.
- Wenn wir aber ein komplett verbundenes Cluster innerhalb einer wachsenden Box B_n betrachten, muss der Radius r mitwachsen.
- Zielstellung: exakte Wachstumsrate für r_n.

Dazu betrachten wir zunächst das folgende vorläufige Ergebnis:

- **Theorem 2.3** Sei $\pi r_n^2 = \beta \log n$.
  - Für $\beta > \frac{5\pi}{4}$ ist $G_n(r)$ a.f.s. voll zusammenhängend, aber für $\beta < \frac{1}{8}$ nicht.

Volle Konnektivität

- Ein stärkeres Resultat, das die exakte Wachstumsrate für r_n betrifft, ist das folgende

  Theorem 2.4 Sei $\pi(2r_n)^2 = \log n + \beta_n$.
  $\Rightarrow G_n(r)$ ist voll zusammenhängend a.f.s. $\Leftrightarrow \beta_n \stackrel{n\to\infty}{\longrightarrow} \infty$

Literatur

Massimo Franceschetti, Ronald Meester (2007):
"Random Networks for Communication", Cambridge University Press (§3.1-§3.3)

Fragen?