Epidemische Informationsausbreitung I

Algorithmen für verteilte Systeme

Sebastian Forster

Universität Salzburg

Informationsausbreitung in Point-to-Point-Netzwerken

Point-to-Point Netzwerke:

- Jeder Knoten kann pro Runde nur mit jeweils einem anderen Knoten eine Verbindung herstellen

Informationsausbreitung in Point-to-Point-Netzwerken

Point-to-Point Netzwerke:

- Jeder Knoten kann pro Runde nur mit jeweils einem anderen Knoten eine Verbindung herstellen
- Bandbreitenbeschränkung: Mit jeder hergestellte Verbindung kann nur eine Nachricht beschränkter Größe gesendet bzw. empfangen werden.

Informationsausbreitung in Point-to-Point-Netzwerken

Point-to-Point Netzwerke:

- Jeder Knoten kann pro Runde nur mit jeweils einem anderen Knoten eine Verbindung herstellen
- Bandbreitenbeschränkung: Mit jeder hergestellte Verbindung kann nur eine Nachricht beschränkter Größe gesendet bzw. empfangen werden.
- Modelliert insbesondere Ende-zu-Ende-Verbindungen über niedere Netzwerkschichten

Informationsausbreitung in Point-to-Point-Netzwerken

Point-to-Point Netzwerke:

- Jeder Knoten kann pro Runde nur mit jeweils einem anderen Knoten eine Verbindung herstellen
- Bandbreitenbeschränkung: Mit jeder hergestellte Verbindung kann nur eine Nachricht beschränkter Größe gesendet bzw. empfangen werden.
- Modelliert insbesondere Ende-zu-Ende-Verbindungen über niedere Netzwerkschichten

Zentrale Fragestellung: Effiziente Verbreitung einer Information an alle Knoten des Netzwerks

Informationsausbreitung in Point-to-Point-Netzwerken

Point-to-Point Netzwerke:

- Jeder Knoten kann pro Runde nur mit jeweils einem anderen Knoten eine Verbindung herstellen
- Bandbreitenbeschränkung: Mit jeder hergestellte Verbindung kann nur eine Nachricht beschränkter Größe gesendet bzw. empfangen werden.
- Modelliert insbesondere Ende-zu-Ende-Verbindungen über niedere Netzwerkschichten

Zentrale Fragestellung: Effiziente Verbreitung einer Information an alle Knoten des Netzwerks

Broadcast durch „Flooding" nicht möglich!

Informationsausbreitung in Point-to-Point-Netzwerken

Point-to-Point Netzwerke:

- Jeder Knoten kann pro Runde nur mit jeweils einem anderen Knoten eine Verbindung herstellen
- Bandbreitenbeschränkung: Mit jeder hergestellte Verbindung kann nur eine Nachricht beschränkter Größe gesendet bzw. empfangen werden.
- Modelliert insbesondere Ende-zu-Ende-Verbindungen über niedere Netzwerkschichten

Zentrale Fragestellung: Effiziente Verbreitung einer Information an alle Knoten des Netzwerks

Broadcast durch „Flooding“ nicht möglich!

Motivation:

- Replizierte Datenbanken in Peer-to-Peer-Netzwerken
- Gleicher Datenbestand an allen Knoten, der konsistent gehalten wird
- Neue Information eines Knotens wird an alle anderen Knoten verteilt

Phone-Call Modell

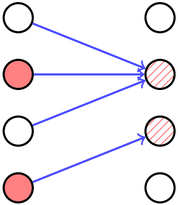
- Netzwerkstruktur: vollständiger Graph mit n Knoten

Phone-Call Modell

- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde kann jeder Knoten einen anderen Knoten anrufen

Phone-Call Modell

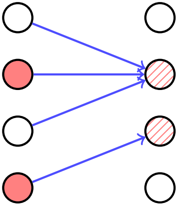
- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde kann jeder Knoten einen anderen Knoten anrufen
- Drei Kommunikationsmodelle:
  1. **Push:** Der Anrufer informiert den Angerufenen



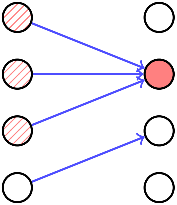
Push

Phone-Call Modell

- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde kann jeder Knoten einen anderen Knoten anrufen
- Drei Kommunikationsmodelle:
  1. **Push:** Der Anrufer informiert den Angerufenen
  2. **Pull:** Der Angerufene informiert den Anrufer



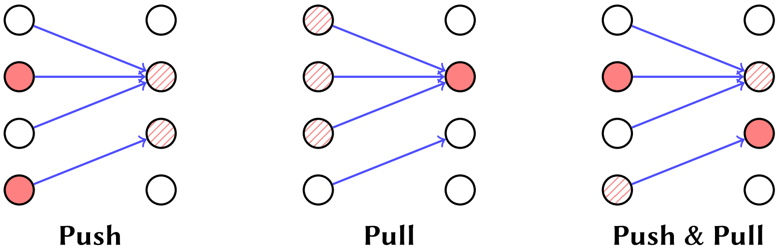
Push



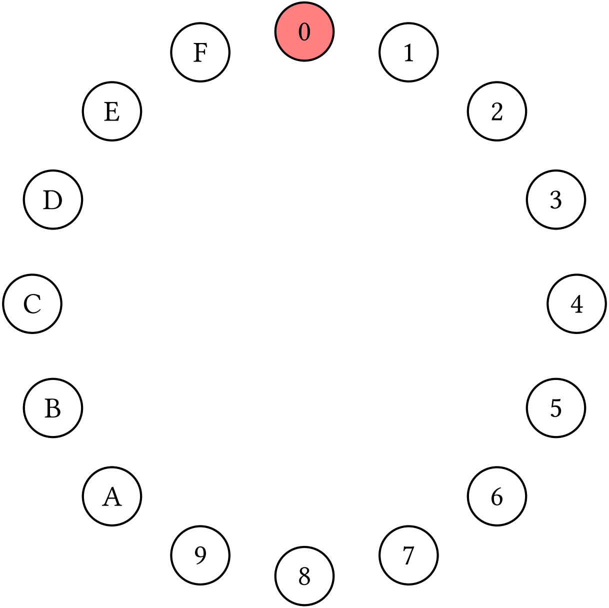
Pull

Phone-Call Modell

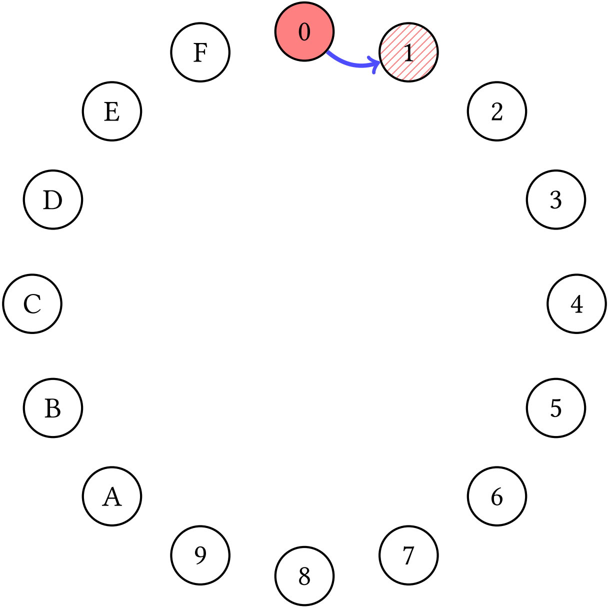
- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde kann jeder Knoten einen anderen Knoten anrufen
- Drei Kommunikationsmodelle:
  1. **Push:** Der Anrufer informiert den Angerufenen
  2. **Pull:** Der Angerufene informiert den Anrufer
  3. **Push & Pull:** Kombination von Push und Pull



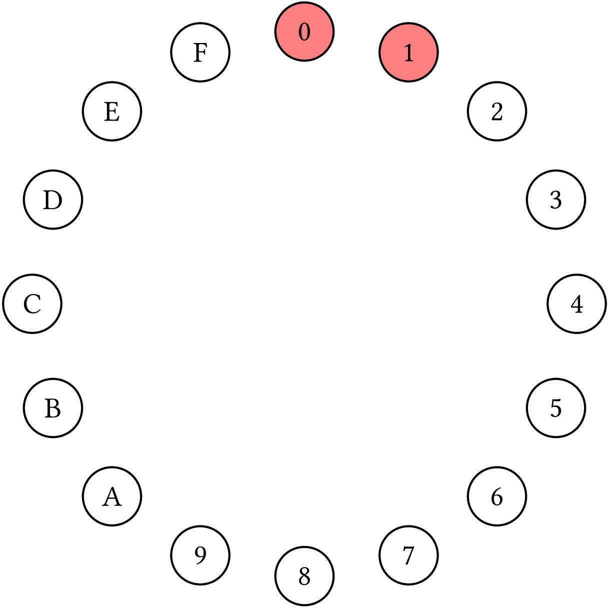
Beispiel: Deterministischer Push-Algorithmus



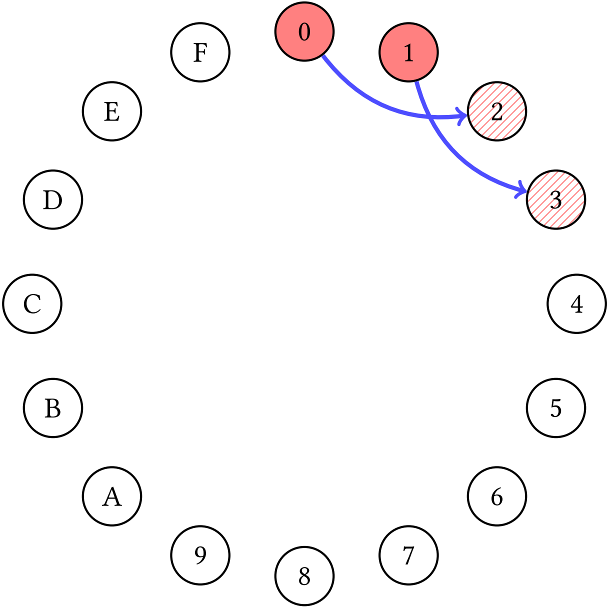
Beispiel: Deterministischer Push-Algorithmus



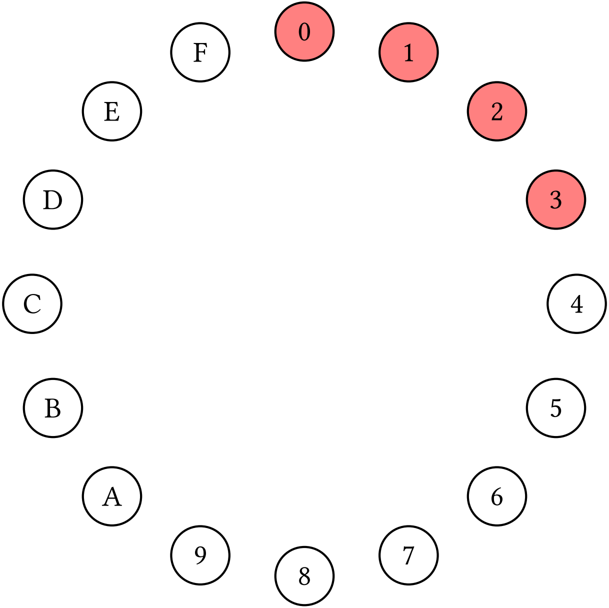
Beispiel: Deterministischer Push-Algorithmus



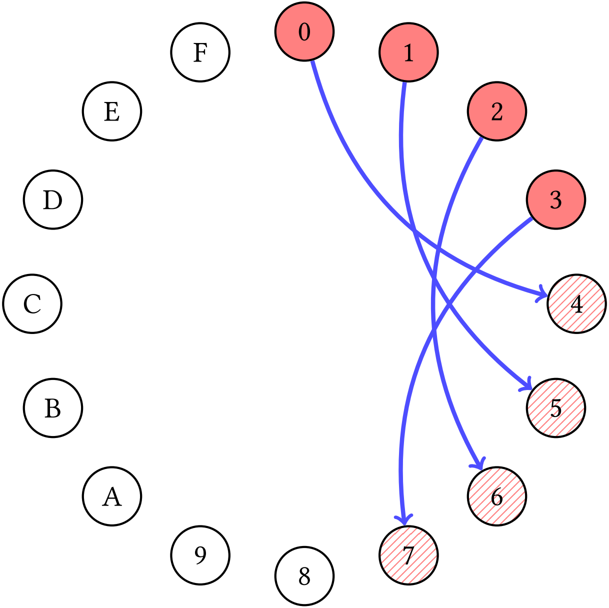
Beispiel: Deterministischer Push-Algorithmus



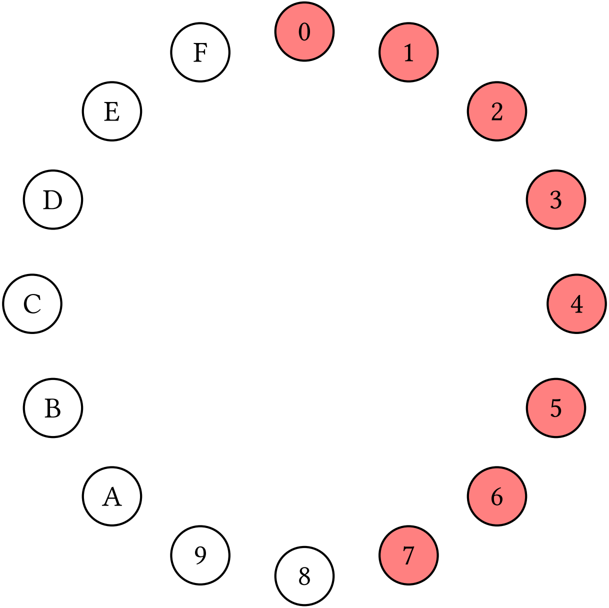
Beispiel: Deterministischer Push-Algorithmus



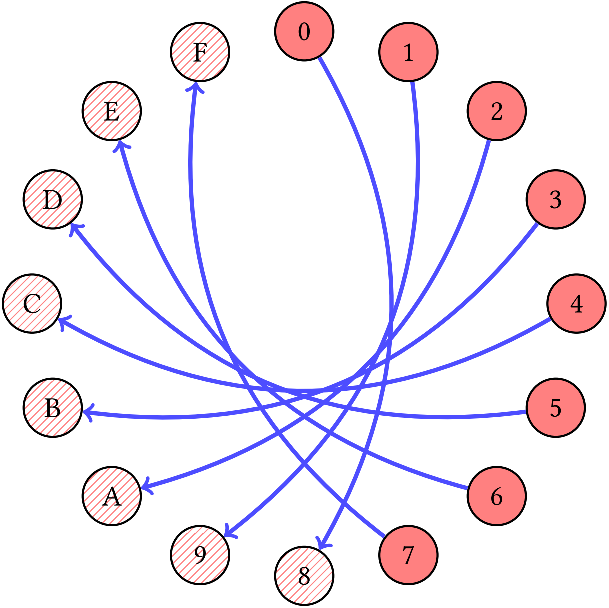
Beispiel: Deterministischer Push-Algorithmus



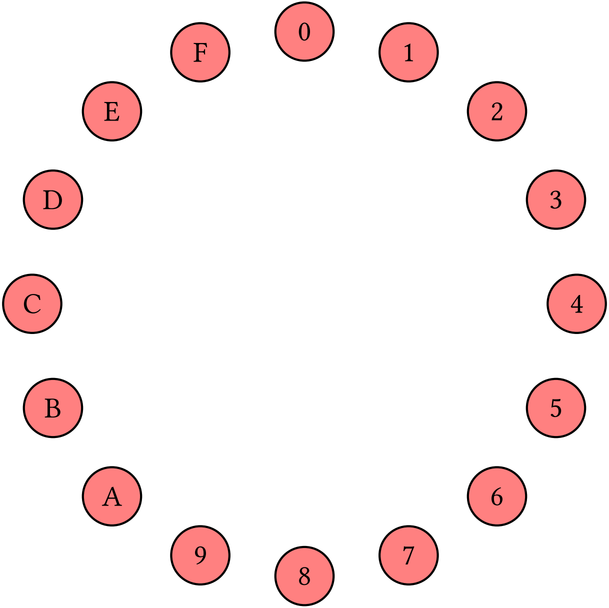
Beispiel: Deterministischer Push-Algorithmus



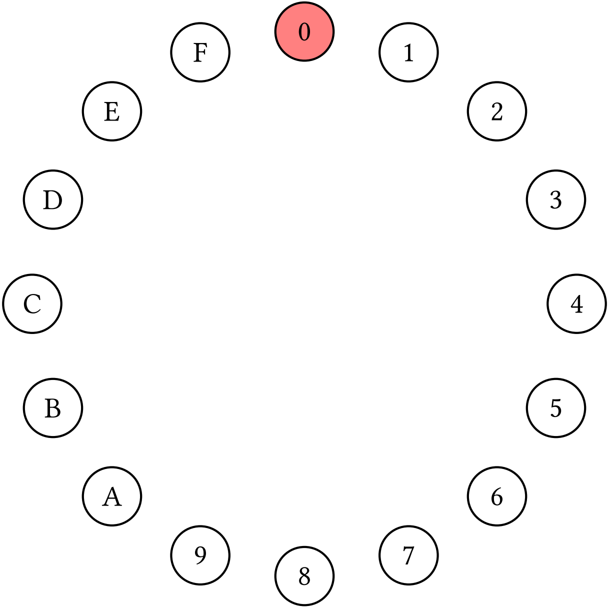
Beispiel: Deterministischer Push-Algorithmus



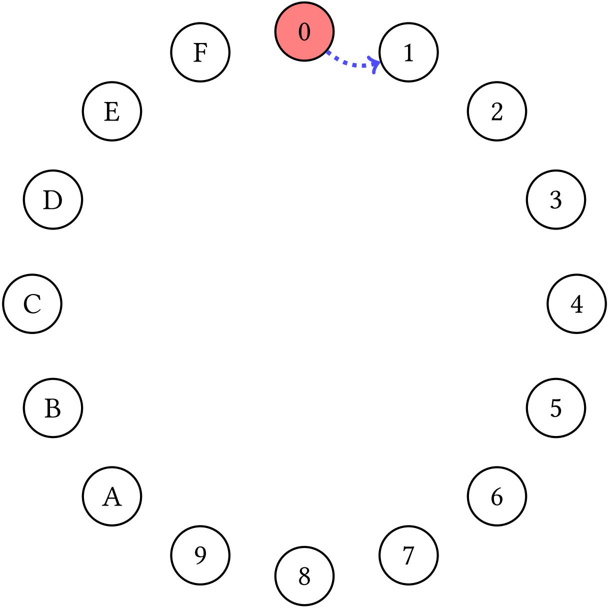
Beispiel: Deterministischer Push-Algorithmus



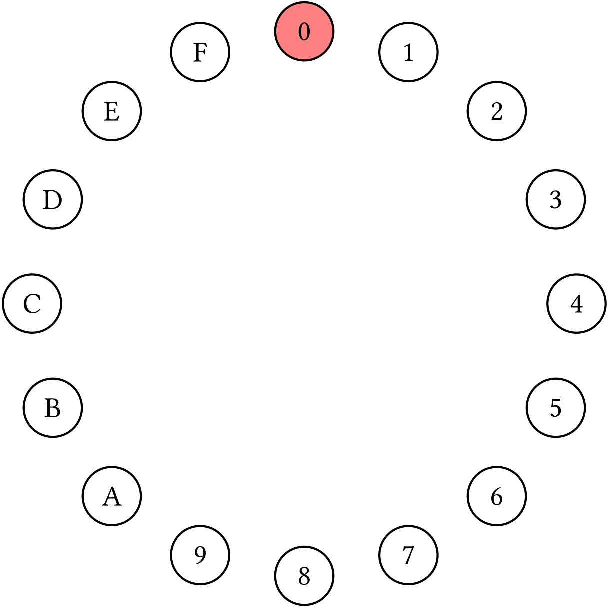
Mangelnde Fehlertoleranz



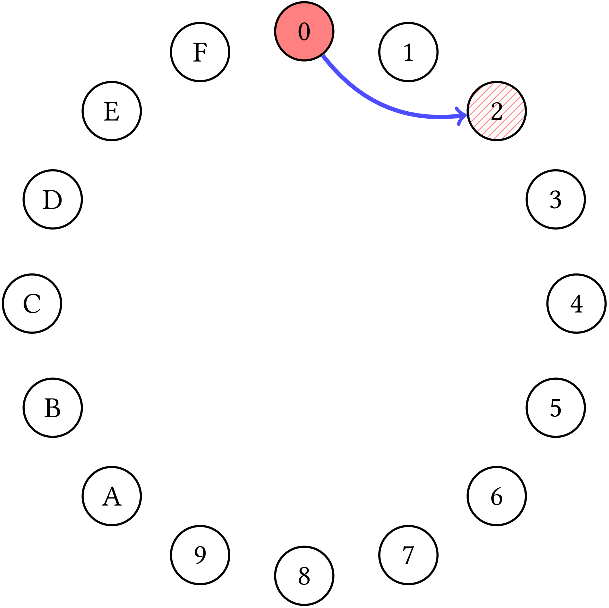
Mangelnde Fehlertoleranz



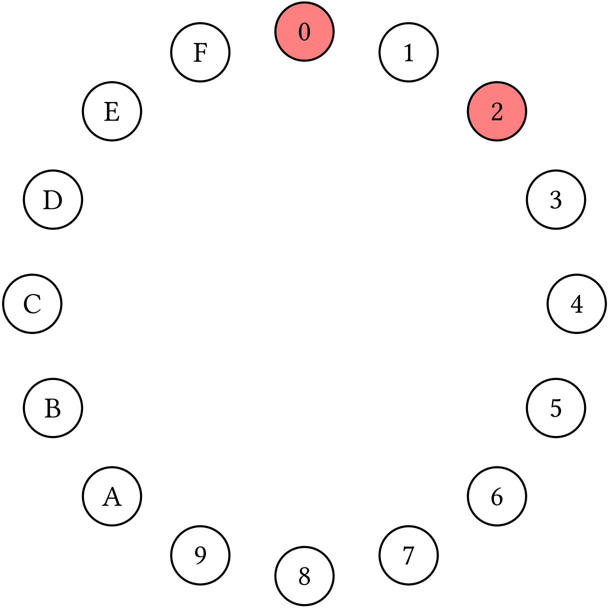
Mangelnde Fehlertoleranz



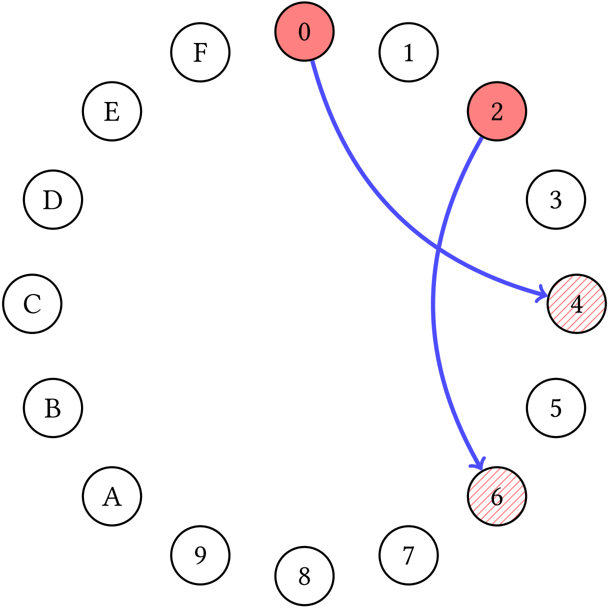
Mangelnde Fehlertoleranz



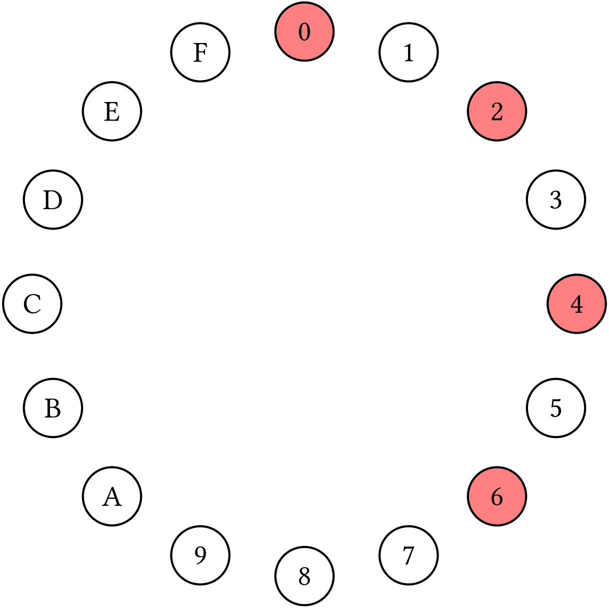
Mangelnde Fehlertoleranz



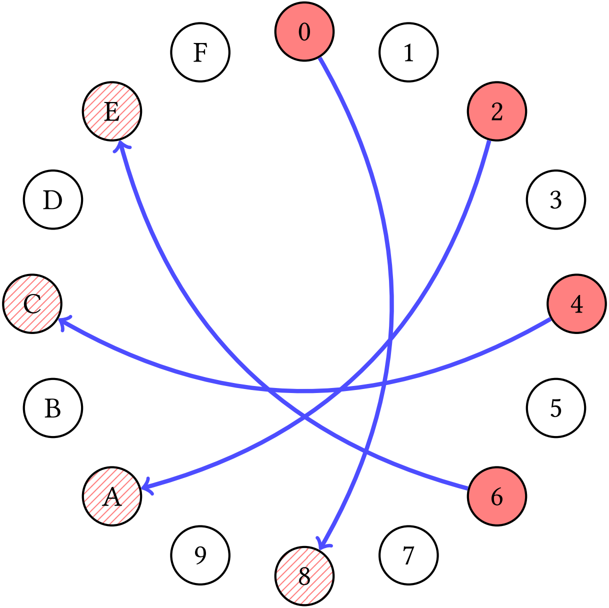
Mangelnde Fehlertoleranz



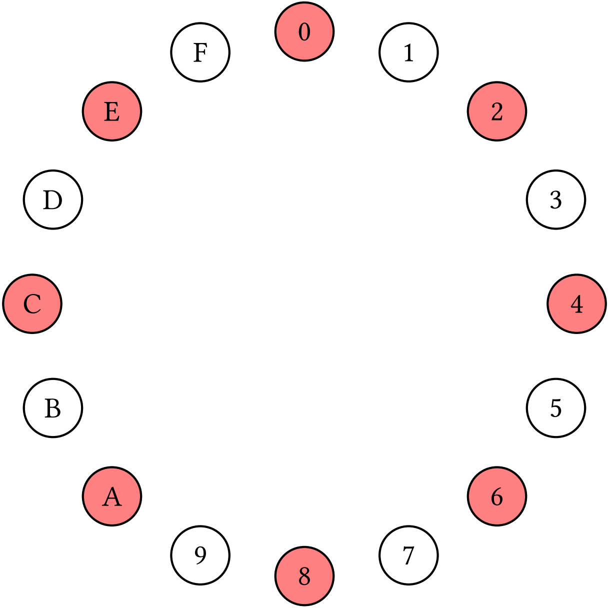
Mangelnde Fehlertoleranz



Mangelnde Fehlertoleranz



Mangelnde Fehlertoleranz



Beobachtungen

Deterministischer Algorithmus:

- Insgesamt $O(n)$ Nachrichten in $O(\log n)$ Runden

Beobachtungen

Deterministischer Algorithmus:

- Insgesamt $O(n)$ Nachrichten in $O(\log n)$ Runden
  Optimal, da jeder Knoten mit einer Nachricht informiert werden muss und sich Anzahl informierter Knoten in jeder Runde bestenfalls verdoppeln kann

Beobachtungen

Deterministischer Algorithmus:

- Insgesamt $O(n)$ Nachrichten in $O(\log n)$ Runden
  Optimal, da jeder Knoten mit einer Nachricht informiert werden muss und sich Anzahl informierter Knoten in jeder Runde bestenfalls verdoppeln kann
- Nachteile:
  1. Keine Fehlertoleranz bei unzuverlässigen Knoten oder Verbindungen

Beobachtungen

Deterministischer Algorithmus:

- Insgesamt $O(n)$ Nachrichten in $O(\log n)$ Runden
  Optimal, da jeder Knoten mit einer Nachricht informiert werden muss und sich Anzahl informierter Knoten in jeder Runde bestenfalls verdoppeln kann
- Nachteile:
  1. Keine Fehlertoleranz bei unzuverlässigen Knoten oder Verbindungen
  2. Jedem Knoten müssen alle anderen Knoten bekannt sein

Beobachtungen

Deterministischer Algorithmus:

- Insgesamt $O(n)$ Nachrichten in $O(\log n)$ Runden

  Optimal, da jeder Knoten mit einer Nachricht informiert werden muss und sich Anzahl informierter Knoten in jeder Runde bestenfalls verdoppeln kann
- Nachteile:
  1. Keine Fehlertoleranz bei unzuverlässigen Knoten oder Verbindungen
  2. Jedem Knoten müssen alle anderen Knoten bekannt sein
  3. Übertragung auf Pull Modell unklar

Epicast als Ersatz für Broadcast [Demers et al. 87]

Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht

Epicast als Ersatz für Broadcast [Demers et al. 87]

Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht
- Jeder Knoten gibt die Information durch Zufallsauswahl weiter, bis sie allen Knoten bekannt ist

Epicast als Ersatz für Broadcast [Demers et al. 87]

Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht
- Jeder Knoten gibt die Information durch Zufallsauswahl weiter, bis sie allen Knoten bekannt ist
- Vorteile: schnell, robust, einfach
- Nachteil: Nachrichtenoverhead

Epicast als Ersatz für Broadcast [Demers et al. 87]

Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht
- Jeder Knoten gibt die Information durch Zufallsauswahl weiter, bis sie allen Knoten bekannt ist
- Vorteile: schnell, robust, einfach
- Nachteil: Nachrichtenoverhead

Random-Phone-Call Modell:

- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an

Epicast als Ersatz für Broadcast [Demers et al. 87]

Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht
- Jeder Knoten gibt die Information durch Zufallsauswahl weiter, bis sie allen Knoten bekannt ist
- Vorteile: schnell, robust, einfach
- Nachteil: Nachrichtenoverhead

Random-Phone-Call Modell:

- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an
- Der Einfachheit halber: Knoten können auch sich selbst anrufen

Epicast als Ersatz für Broadcast [Demers et al. 87]

Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht
- Jeder Knoten gibt die Information durch Zufallsauswahl weiter, bis sie allen Knoten bekannt ist
- Vorteile: schnell, robust, einfach
- Nachteil: Nachrichtenoverhead

Random-Phone-Call Modell:

- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an
- Der Einfachheit halber: Knoten können auch sich selbst anrufen
- Somit: Knoten u ruft Knoten v mit Wahrscheinlichkeit $\frac{1}{n}$ an

Epicast als Ersatz für Broadcast [Demers et al. 87]

Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht
- Jeder Knoten gibt die Information durch Zufallsauswahl weiter, bis sie allen Knoten bekannt ist
- Vorteile: schnell, robust, einfach
- Nachteil: Nachrichtenoverhead

Random-Phone-Call Modell:

- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an
- Der Einfachheit halber: Knoten können auch sich selbst anrufen
- Somit: Knoten u ruft Knoten v mit Wahrscheinlichkeit $\frac{1}{n}$ an
- Zufallsauswahl ohne explizite Kenntnis der anderen Knoten

Epicast als Ersatz für Broadcast [Demers et al. 87]

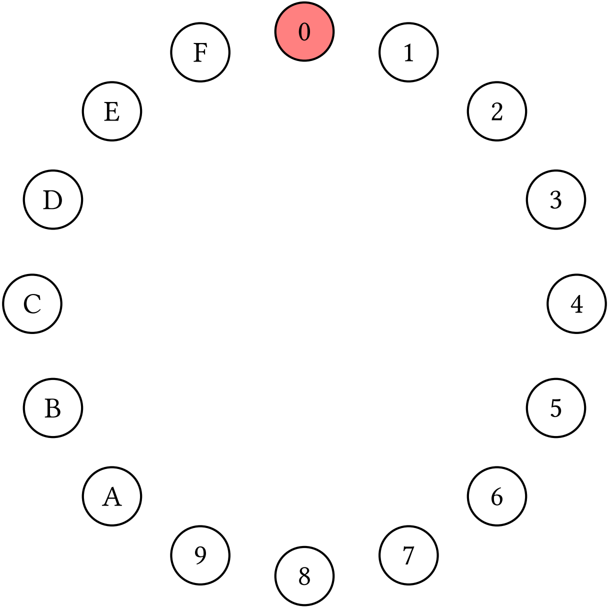
Epicast:

- Ausbreitung der Information wie ein Virus oder ein Gerücht
- Jeder Knoten gibt die Information durch Zufallsauswahl weiter, bis sie allen Knoten bekannt ist
- Vorteile: schnell, robust, einfach
- Nachteil: Nachrichtenoverhead

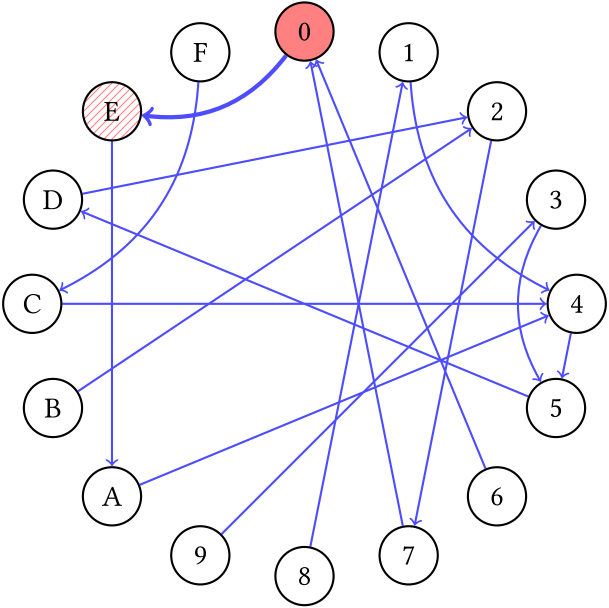
Random-Phone-Call Modell:

- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an
- Der Einfachheit halber: Knoten können auch sich selbst anrufen
- Somit: Knoten u ruft Knoten v mit Wahrscheinlichkeit $\frac{1}{n}$ an
- Zufallsauswahl ohne explizite Kenntnis der anderen Knoten
  Zufällige Adressierung wird in vielen Peer-to-Peer-Netzwerken unterstützt

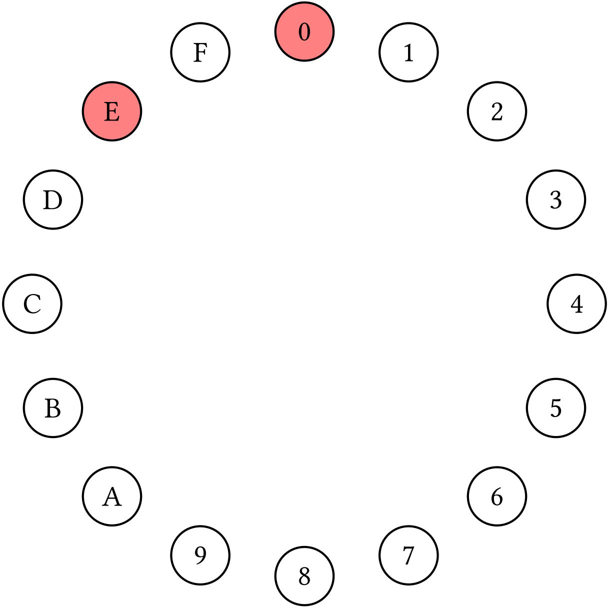
Beispiel Push Modell



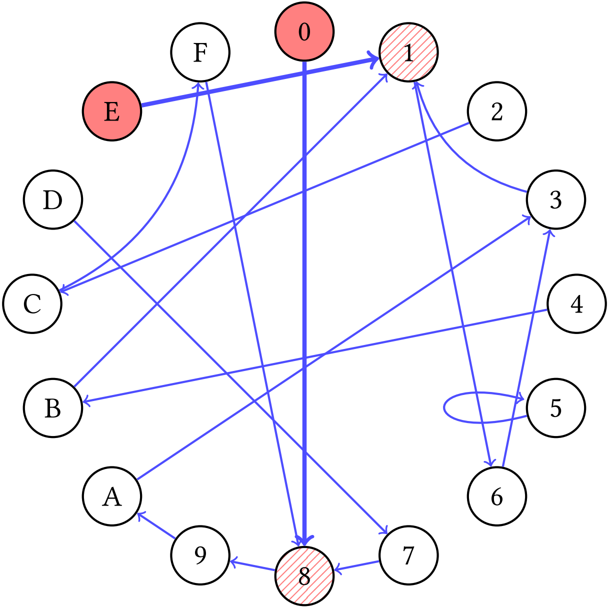
Beispiel Push Modell



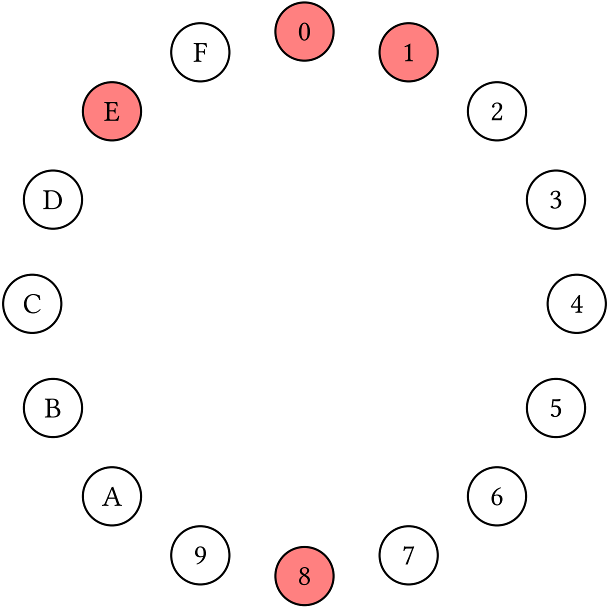
Beispiel Push Modell



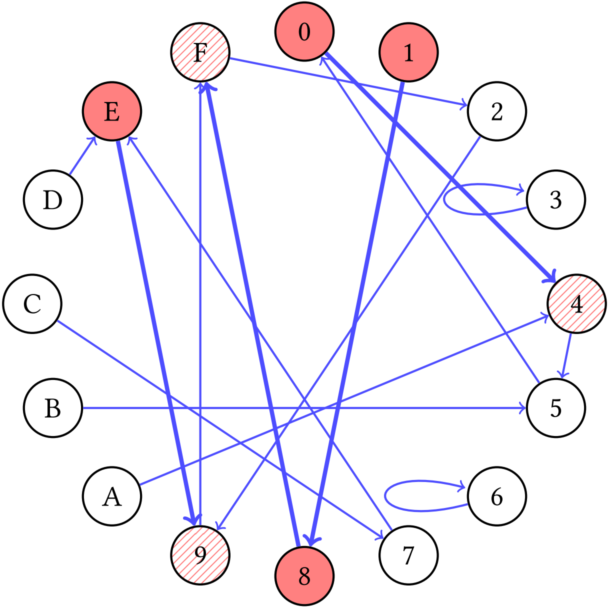
Beispiel Push Modell



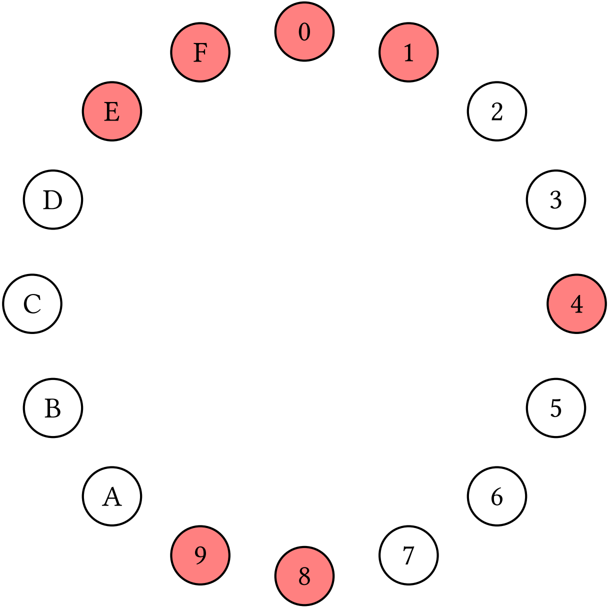
Beispiel Push Modell



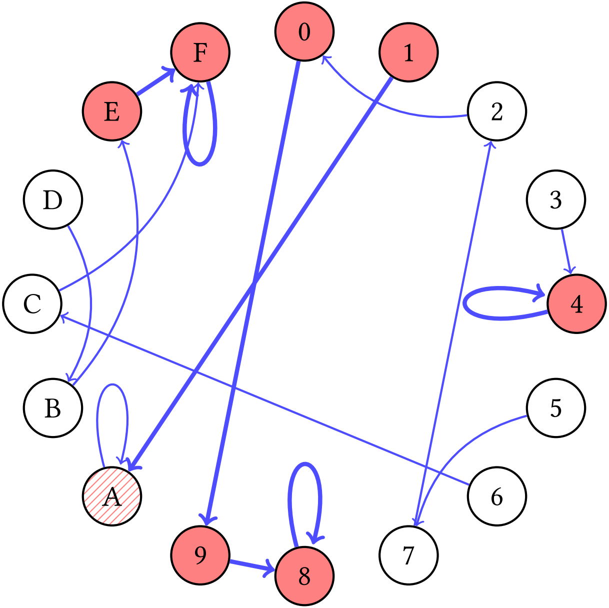
Beispiel Push Modell



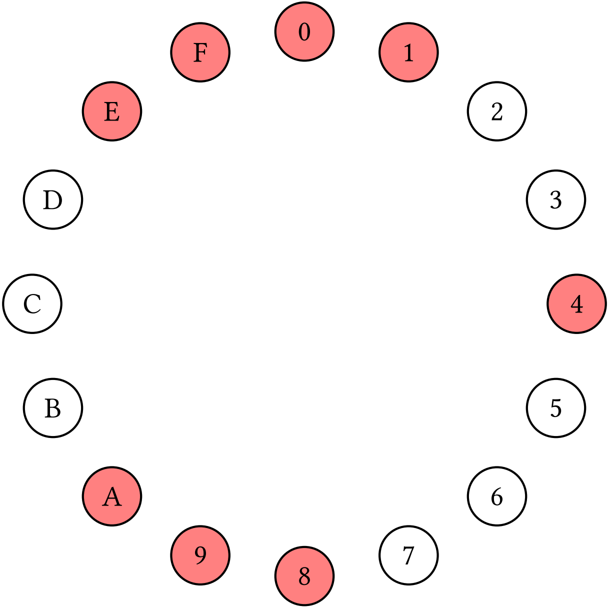
Beispiel Push Modell



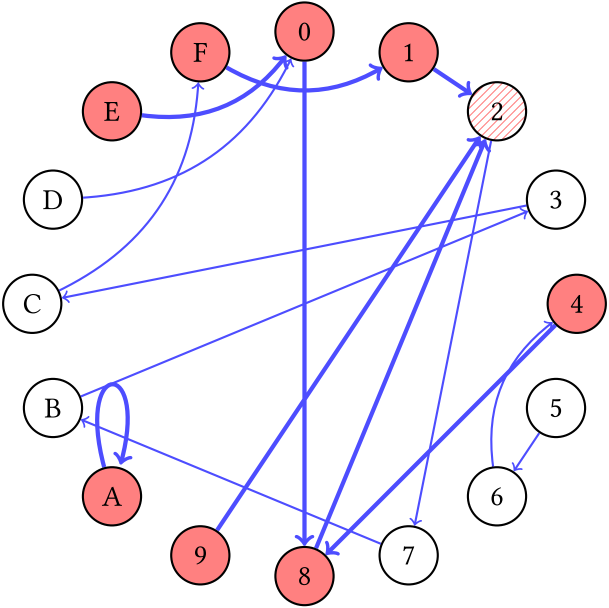
Beispiel Push Modell



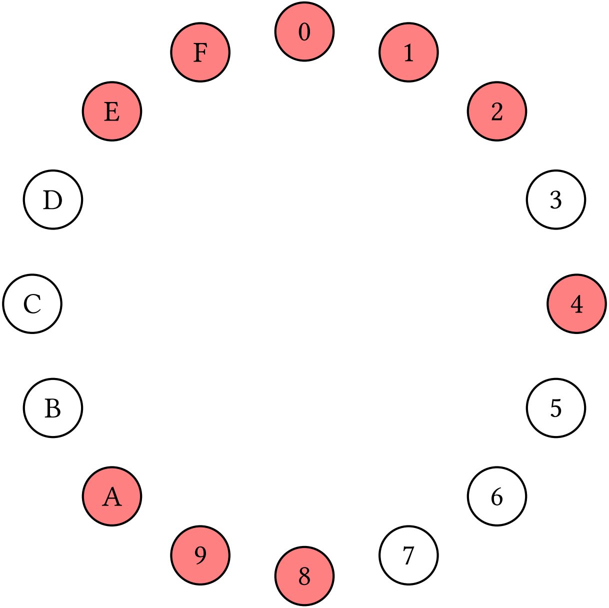
Beispiel Push Modell



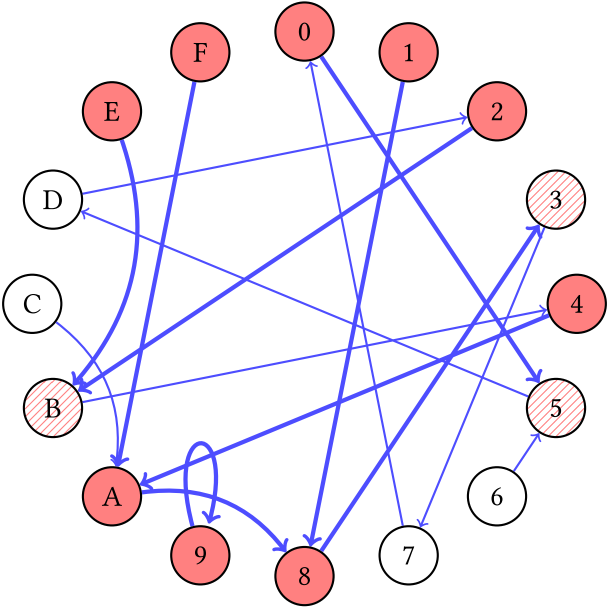
Beispiel Push Modell



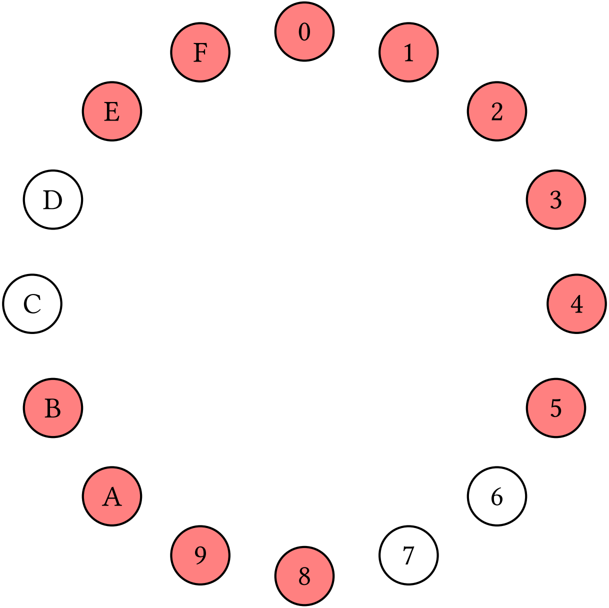
Beispiel Push Modell



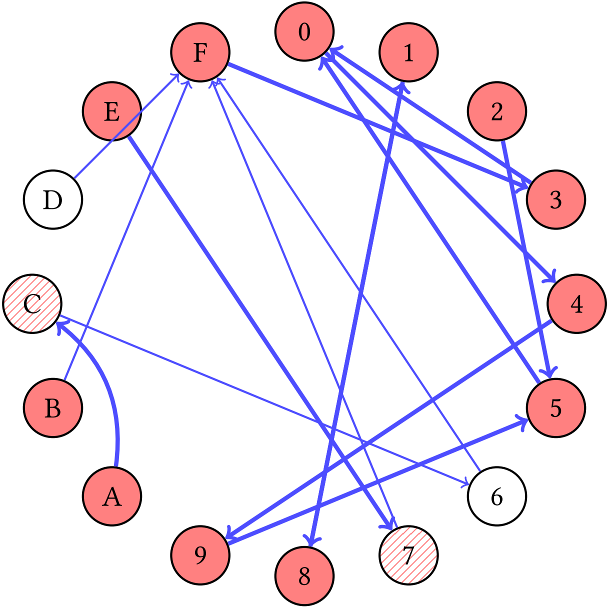
Beispiel Push Modell



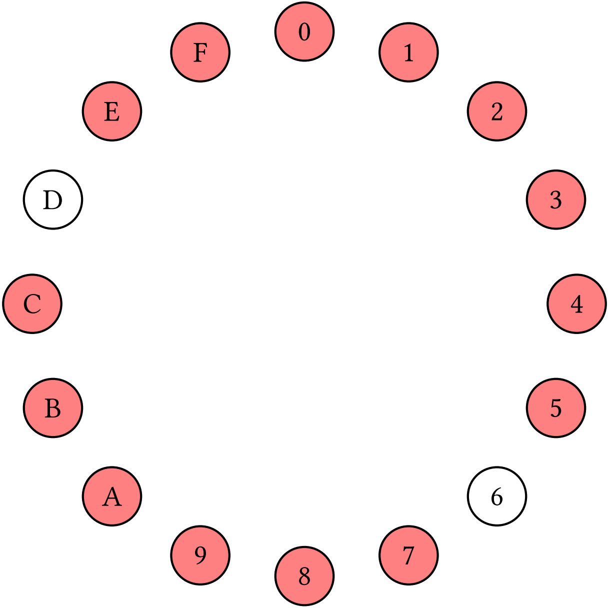
Beispiel Push Modell



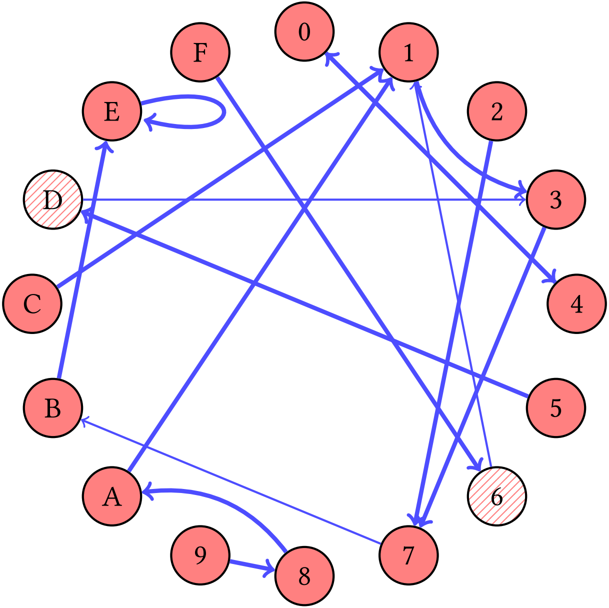
Beispiel Push Modell



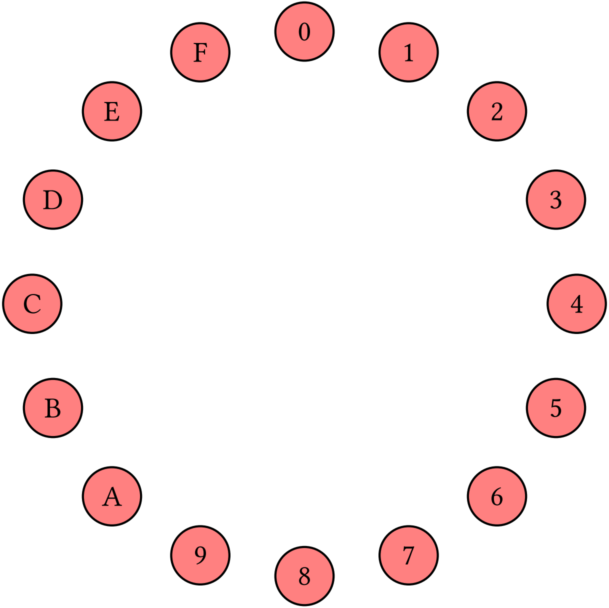
Beispiel Push Modell



Beispiel Push Modell



Beispiel Push Modell



Theoretische Analyse Push Modell

Frage: Wie lange es dauert bis alle Knoten informiert sind, wenn anfangs ein Knoten informiert ist?

Theoretische Analyse Push Modell

Frage: Wie lange es dauert bis alle Knoten informiert sind, wenn anfangs ein Knoten informiert ist?

Theorem ([Frieze/Grimmett '85])

Ausgehend von einem informierten Knoten, sind im Push Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ Runden alle n Knoten informiert.

Theoretische Analyse Push Modell

Frage: Wie lange es dauert bis alle Knoten informiert sind, wenn anfangs ein Knoten informiert ist?

Theorem ([Frieze/Grimmett '85])

Ausgehend von einem informierten Knoten, sind im Push Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ *Runden alle* n *Knoten informiert.*

Anmerkungen:

- Mit zusätzlichem Zeitstempel kann überprüft werden, ob mit hoher Wahrscheinlichkeit bereits alle Knoten informiert sind und Ausbreitung der Information gestoppt werden kann

Theoretische Analyse Push Modell

Frage: Wie lange es dauert bis alle Knoten informiert sind, wenn anfangs ein Knoten informiert ist?

Theorem ([Frieze/Grimmett '85])

Ausgehend von einem informierten Knoten, sind im Push Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ Runden alle n Knoten informiert.

Anmerkungen:

- Mit zusätzlichem Zeitstempel kann überprüft werden, ob mit hoher Wahrscheinlichkeit bereits alle Knoten informiert sind und Ausbreitung der Information gestoppt werden kann
- Robustheit:
  - Pro fehlerbehafteter Runde erhöht sich Gesamtzahl benötigter Runden um eins
  - Analyse kann auf Modell mit Weitergabewahrscheinlichkeit erweitert werden

Theoretische Analyse Push Modell

Frage: Wie lange es dauert bis alle Knoten informiert sind, wenn anfangs ein Knoten informiert ist?

Theorem ([Frieze/Grimmett '85])

Ausgehend von einem informierten Knoten, sind im Push Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ Runden alle n Knoten informiert.

Anmerkungen:

- Mit zusätzlichem Zeitstempel kann überprüft werden, ob mit hoher Wahrscheinlichkeit bereits alle Knoten informiert sind und Ausbreitung der Information gestoppt werden kann
- Robustheit:
  - Pro fehlerbehafteter Runde erhöht sich Gesamtzahl benötigter Runden um eins
  - Analyse kann auf Modell mit Weitergabewahrscheinlichkeit erweitert werden
- Analyse des Zufallsprozesses auch für andere Graphstrukturen möglich

Notation für Analyse

- n: Anzahl der Knoten
- $I(t)$: Anzahl der informierten Knoten am Beginn von Runde t
- $U(t)$: Anzahl der uninformierten Knoten
- $i(t) = \frac{I(t)}{n}$: relativer Anteil der informierten Knoten
- $u(t) = \frac{U(t)}{n}$: relativer Anteil der uninformierten Knoten

Achtung: Zufallsvariablen

Notation für Analyse

- n: Anzahl der Knoten
- $I(t)$: Anzahl der informierten Knoten am Beginn von Runde t
- $U(t)$: Anzahl der uninformierten Knoten
- $i(t) = \frac{I(t)}{n}$: relativer Anteil der informierten Knoten
- $u(t) = \frac{U(t)}{n}$: relativer Anteil der uninformierten Knoten

Achtung: Zufallsvariablen

Für jede Runde t gilt: $i(t) + u(t) = 1$

Notation für Analyse

- n: Anzahl der Knoten
- $I(t)$: Anzahl der informierten Knoten am Beginn von Runde t
- $U(t)$: Anzahl der uninformierten Knoten
- $i(t) = \frac{I(t)}{n}$: relativer Anteil der informierten Knoten
- $u(t) = \frac{U(t)}{n}$: relativer Anteil der uninformierten Knoten

Achtung: Zufallsvariablen

Für jede Runde t gilt: $i(t) + u(t) = 1$

Ordnung der Knoten im Beweis: Bei der Analyse von Runde t nehmen wir eine beliebige Ordnung der informierten Knoten an und bezeichnen die Knoten mit $1, \ldots, I(t)$

Allgemeine Analyse I

Frage

Wie hoch ist $I(t+1)$ in Abhängigkeit von $I(t)$?

Allgemeine Analyse I

Frage

Wie hoch ist $I(t+1)$ in Abhängigkeit von $I(t)$?

Analysetrick:

Jeder Knoten der erstmalig informiert wird, wird seinem kleinsten informierten Anrufer zugeschrieben.

Allgemeine Analyse I

Frage

Wie hoch ist $I(t+1)$ in Abhängigkeit von $I(t)$?

Analysetrick:

Jeder Knoten der erstmalig informiert wird, wird seinem kleinsten informierten Anrufer zugeschrieben.

Definition

Ein Knoten j führt in Runde t eine **erstmalige Informationsweitergabe** durch, wenn

- j am Beginn von Runde t informiert ist,
- der von j in Runde t angerufene Knoten am Beginn von Runde t uninformiert ist und
- der von j in Runde t angerufene Knoten in Runde t von keinem informierten Knoten, der kleiner ist als j, angerufen wird

Allgemeine Analyse II

Notation:

$\mathcal{I}(t)$: Menge der informierten Knoten am Beginn von Runde t

$c_j(t)$: Knoten, der von Knoten j in Runde t angerufen wird

Allgemeine Analyse II

Notation:

$\mathcal{I}(t)$: Menge der informierten Knoten am Beginn von Runde t

$c_j(t)$: Knoten, der von Knoten j in Runde t angerufen wird

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } j \text{ eine erstmalige Informationsweitergabe durchführt} \\ 0 & \text{andernfalls} \end{cases}$$

Allgemeine Analyse II

Notation:

$\mathcal{I}(t)$: Menge der informierten Knoten am Beginn von Runde t

$c_j(t)$: Knoten, der von Knoten j in Runde t angerufen wird

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } j \text{ eine erstmalige Informationsweitergabe durchführt} \\ 0 & \text{andernfalls} \end{cases}$$

Äquivalent:

$$X_j(t) = \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \dots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Allgemeine Analyse II

Notation:

$\mathcal{I}(t)$: Menge der informierten Knoten am Beginn von Runde t

$c_j(t)$: Knoten, der von Knoten j in Runde t angerufen wird

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } j \text{ eine erstmalige Informationsweitergabe durchführt} \\ 0 & \text{andernfalls} \end{cases}$$

Äquivalent:

$$X_j(t) = \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \dots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Analyse des Gegenereignisses:

$$\Pr[X_j(t) = 0] = \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \dots, c_{j-1}(t)\}]$$

Allgemeine Analyse II

Notation:

$\mathcal{I}(t)$: Menge der informierten Knoten am Beginn von Runde t

$c_j(t)$: Knoten, der von Knoten j in Runde t angerufen wird

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } j \text{ eine erstmalige Informationsweitergabe durchführt} \\ 0 & \text{andernfalls} \end{cases}$$

Äquivalent:

$$X_j(t) = \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Analyse des Gegenereignisses:

$$\begin{aligned} \Pr[X_j(t) = 0] &= \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\}] \\ &= \frac{|\mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\}|}{n} \end{aligned}$$

Allgemeine Analyse II

Notation:

$\mathcal{I}(t)$: Menge der informierten Knoten am Beginn von Runde t

$c_j(t)$: Knoten, der von Knoten j in Runde t angerufen wird

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } j \text{ eine erstmalige Informationsweitergabe durchführt} \\ 0 & \text{andernfalls} \end{cases}$$

Äquivalent:

$$X_j(t) = \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Analyse des Gegenereignisses:

$$\begin{aligned} \Pr[X_j(t) = 0] &= \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\}] \\ &= \frac{|\mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\}|}{n} \leq \frac{2I(t)}{n} = 2i(t) \end{aligned}$$

Allgemeine Analyse II

Notation:

$\mathcal{I}(t)$: Menge der informierten Knoten am Beginn von Runde t

$c_j(t)$: Knoten, der von Knoten j in Runde t angerufen wird

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } j \text{ eine erstmalige Informationsweitergabe durchführt} \\ 0 & \text{andernfalls} \end{cases}$$

Äquivalent:

$$X_j(t) = \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \dots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Analyse des Gegenereignisses:

$$\begin{aligned} \Pr[X_j(t) = 0] &= \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \dots, c_{j-1}(t)\}] \\ &= \frac{|\mathcal{I}(t) \cup \{c_1(t), \dots, c_{j-1}(t)\}|}{n} \leq \frac{2I(t)}{n} = 2i(t) \end{aligned}$$

$\Rightarrow \Pr[X_j(t) = 1] = 1 - \Pr[X_j(t) = 0] \geq 1 - 2i(t)$

Analyse des Erwartungswerts

Analysiere erwarteten Anstieg von $i(t)$

Analyse des Erwartungswerts

Analysiere erwarteten Anstieg von $i(t)$

$F(t) := I(t+1) - I(t)$ Anzahl der in Runde t erstmalig informierten Knoten

$$\begin{aligned} \mathrm{Ex}[F(t)] = \mathrm{Ex}\left[\sum_{j=1}^{I(t)} X_j(t)\right] &= \sum_{j=1}^{I(t)} \mathrm{Ex}[X_j(t)] \quad \text{(Linearität des Erwartungswerts)} \\ &= \sum_{j=1}^{I(t)} \Pr[X_j(t) = 1] \geq \sum_{j=1}^{I(t)} (1 - 2i(t)) = I(t) - 2i(t)I(t) \end{aligned}$$

Analyse des Erwartungswerts

Analysiere erwarteten Anstieg von $i(t)$

$F(t) := I(t+1) - I(t)$ Anzahl der in Runde t erstmalig informierten Knoten

$$\begin{aligned}\mathrm{Ex}[F(t)] = \mathrm{Ex}\left[\sum_{j=1}^{I(t)} X_j(t)\right] &= \sum_{j=1}^{I(t)} \mathrm{Ex}[X_j(t)] \qquad \text{(Linearität des Erwartungswerts)} \\ &= \sum_{j=1}^{I(t)} \Pr[X_j(t) = 1] \geq \sum_{j=1}^{I(t)} (1 - 2i(t)) = I(t) - 2i(t)I(t)\end{aligned}$$

Erwartungswert des relativen Anteils informierter Knoten nach Runde t:

$$\begin{aligned}\mathrm{Ex}[i(t+1)] = \mathrm{Ex}\left[i(t) + \frac{F(t)}{n}\right] &= i(t) + \frac{\mathrm{Ex}[F(t)]}{n} \geq i(t) + \frac{I(t) - 2i(t)I(t)}{n} \\ &= i(t) + i(t) - 2i(t)^2 = 2i(t) - 2i(t)^2\end{aligned}$$

Analyse des Erwartungswerts

Analysiere erwarteten Anstieg von $i(t)$

$F(t) := I(t+1) - I(t)$ Anzahl der in Runde t erstmalig informierten Knoten

$$\mathrm{Ex}[F(t)] = \mathrm{Ex}\left[\sum_{j=1}^{I(t)} X_j(t)\right] = \sum_{j=1}^{I(t)} \mathrm{Ex}[X_j(t)] \qquad \text{(Linearität des Erwartungswerts)}$$

$$= \sum_{j=1}^{I(t)} \Pr[X_j(t) = 1] \geq \sum_{j=1}^{I(t)} (1 - 2i(t)) = I(t) - 2i(t)I(t)$$

Erwartungswert des relativen Anteils informierter Knoten nach Runde t:

$$\mathrm{Ex}[i(t+1)] = \mathrm{Ex}\left[i(t) + \frac{F(t)}{n}\right] = i(t) + \frac{\mathrm{Ex}[F(t)]}{n} \geq i(t) + \frac{I(t) - 2i(t)I(t)}{n}$$

$$= i(t) + i(t) - 2i(t)^2 = 2i(t) - 2i(t)^2$$

Intuition für Analyse:

- $2i(t)^2$ anfangs vernachlässigbar
- Annähernd Verdopplung in jeder Runde

Beweisstrategie

Einteilung in zwei Phasen:

1. **Wachstum:** $I(t) \leq \frac{1}{3}n$
   Dauer: $O(\log n)$ Runden
2. **Schrumpfung:** $I(t) > \frac{1}{3}n$
   Dauer: $O(\log n)$ Runden

Gesamtzahl an Runden: $O(\log n)$

In jeder Phase: Garantie mit hoher Wahrscheinlichkeit

Wachstum: $I(t) \leq \frac{n}{3}$

$I(t)$ wächst um Faktor $> \frac{7}{6}$ in jeder Runde mit konstanter Wahrscheinlichkeit

Wachstum: $I(t) \le \frac{n}{3}$

$I(t)$ wächst um Faktor $> \frac{7}{6}$ in jeder Runde mit konstanter Wahrscheinlichkeit

Chernoff Bound (Lower Tail)

Seien $X_1, \ldots, X_k \in \{0, 1\}$ *unabhängige* binäre Zufallsvariablen mit $\Pr[X_j = 1] \ge p$ und sei $\mu := pn$. Dann gilt für jedes $\delta \in [0, 1]$:

$$\Pr\left[\sum_{j=1}^{k} X_j \le (1-\delta) \cdot \mu\right] \le \frac{1}{e^{\frac{\delta^2}{2} \cdot \mu}}$$

Setze $p = \frac{1}{3}$, $\mu = \frac{1}{3} I(t)$, $\delta = \frac{1}{2}$

Wachstum: $I(t) \leq \frac{n}{3}$

$I(t)$ wächst um Faktor $> \frac{7}{6}$ in jeder Runde mit konstanter Wahrscheinlichkeit

Chernoff Bound (Lower Tail)

Seien $X_1, \ldots, X_k \in \{0, 1\}$ *unabhängige* binäre Zufallsvariablen mit $\Pr[X_j = 1] \geq p$ und sei $\mu := pn$. Dann gilt für jedes $\delta \in [0, 1]$:

$$\Pr\left[\sum_{j=1}^{k} X_j \leq (1 - \delta) \cdot \mu\right] \leq \frac{1}{e^{\frac{\delta^2}{2} \cdot \mu}}$$

Setze $p = \frac{1}{3}$, $\mu = \frac{1}{3} I(t)$, $\delta = \frac{1}{2}$

$\Pr[X_j(t) = 1] \geq 1 - 2i(t) \geq 1 - 2 \cdot \frac{1}{3} = \frac{1}{3}$

Wachstum: $I(t) \leq \frac{n}{3}$

$I(t)$ wächst um Faktor $> \frac{7}{6}$ in jeder Runde mit konstanter Wahrscheinlichkeit

Chernoff Bound (Lower Tail)

Seien $X_1, \ldots, X_k \in \{0, 1\}$ *unabhängige* binäre Zufallsvariablen mit $\Pr[X_j = 1] \geq p$ und sei $\mu := pn$. Dann gilt für jedes $\delta \in [0, 1]$:

$$\Pr\left[\sum_{j=1}^{k} X_j \leq (1-\delta)\cdot\mu\right] \leq \frac{1}{e^{\frac{\delta^2}{2}\cdot\mu}}$$

Setze $p = \frac{1}{3}$, $\mu = \frac{1}{3}I(t)$, $\delta = \frac{1}{2}$

$\Pr[X_j(t) = 1] \geq 1 - 2i(t) \geq 1 - 2\cdot\frac{1}{3} = \frac{1}{3}$

Wahrscheinlichkeit, dass $F(t) \leq \frac{1}{6}I(t)$ (Wachstum nicht stark genug):

$$\Pr\left[\sum_{j=1}^{I(t)} X_j(t) \leq \frac{1}{6}\cdot I(t)\right] = \Pr\left[\sum_{j=1}^{I(t)} X_j(t) \leq (1-\frac{1}{2})\cdot\frac{1}{3}I(t)\right]$$

$$\leq \frac{1}{e^{\frac{(\frac{1}{2})^2}{2}\cdot\frac{1}{3}I(t)}} = \frac{1}{e^{\frac{1}{24}I(t)}} \leq \frac{1}{e^{\frac{1}{24}}}$$

Phase 1: Wachstum (Fortsetzung)

Einteilung in **gute** und **schlechte** Runden:

Phase 1: Wachstum (Fortsetzung)

Einteilung in **gute** und **schlechte** Runden:

- Runde t ist **schlecht** falls $I(t+1) \leq \frac{7}{6} I(t)$

- Runde t ist **gut** falls $I(t+1) > \frac{7}{6} I(t)$

Phase 1: Wachstum (Fortsetzung)

Einteilung in **gute** und **schlechte** Runden:

- Runde t ist **schlecht** falls $I(t+1) \leq \frac{7}{6} I(t)$
  $\Pr[I(t+1) \leq \frac{7}{6} I(t)] \leq \Pr[F(t) \leq \frac{1}{6} I(t)] \leq \frac{1}{e^{\frac{1}{24}}}$
- Runde t ist **gut** falls $I(t+1) > \frac{7}{6} I(t)$

Phase 1: Wachstum (Fortsetzung)

Einteilung in **gute** und **schlechte** Runden:

- Runde t ist **schlecht** falls $I(t+1) \leq \frac{7}{6}I(t)$
  $\Pr[I(t+1) \leq \frac{7}{6}I(t)] \leq \Pr[F(t) \leq \frac{1}{6}I(t)] \leq \frac{1}{e^{\frac{1}{24}}}$
- Runde t ist **gut** falls $I(t+1) > \frac{7}{6}I(t)$
  $\Pr[I(t+1) > \frac{7}{6}I(t)] = 1 - \Pr[I(t+1) \leq \frac{7}{6}I(t)] \geq 1 - \frac{1}{e^{\frac{1}{24}}} \geq \frac{4}{100}$

Phase 1: Wachstum (Fortsetzung)

Einteilung in **gute** und **schlechte** Runden:

- Runde t ist **schlecht** falls $I(t+1) \leq \frac{7}{6} I(t)$
  $\Pr[I(t+1) \leq \frac{7}{6} I(t)] \leq \Pr[F(t) \leq \frac{1}{6} I(t)] \leq \frac{1}{e^{\frac{1}{24}}}$
- Runde t ist **gut** falls $I(t+1) > \frac{7}{6} I(t)$
  $\Pr[I(t+1) > \frac{7}{6} I(t)] = 1 - \Pr[I(t+1) \leq \frac{7}{6} I(t)] \geq 1 - \frac{1}{e^{\frac{1}{24}}} \geq \frac{4}{100}$

Analyse:

- Nach spätestens $\log_{7/6} \frac{n}{3}$ guten Runden gilt $I(t) > \frac{1}{3} n$

Phase 1: Wachstum (Fortsetzung)

Einteilung in **gute** und **schlechte** Runden:

- Runde t ist **schlecht** falls $I(t+1) \leq \frac{7}{6} I(t)$
  $\Pr[I(t+1) \leq \frac{7}{6} I(t)] \leq \Pr[F(t) \leq \frac{1}{6} I(t)] \leq \frac{1}{e^{\frac{1}{24}}}$
- Runde t ist **gut** falls $I(t+1) > \frac{7}{6} I(t)$
  $\Pr[I(t+1) > \frac{7}{6} I(t)] = 1 - \Pr[I(t+1) \leq \frac{7}{6} I(t)] \geq 1 - \frac{1}{e^{\frac{1}{24}}} \geq \frac{4}{100}$

Analyse:

- Nach spätestens $\log_{7/6} \frac{n}{3}$ guten Runden gilt $I(t) > \frac{1}{3} n$
- Erneute Anwendung der Chernoff-Bound: Bei $O(\log n)$ Runden treten mit hoher Wahrscheinlichkeit $\log_{7/6} \frac{n}{3}$ gute Runden auf

Phase 1: Wachstum (Fortsetzung)

Einteilung in **gute** und **schlechte** Runden:

- Runde t ist **schlecht** falls $I(t+1) \leq \frac{7}{6} I(t)$
  $\Pr[I(t+1) \leq \frac{7}{6} I(t)] \leq \Pr[F(t) \leq \frac{1}{6} I(t)] \leq \frac{1}{e^{\frac{1}{24}}}$
- Runde t ist **gut** falls $I(t+1) > \frac{7}{6} I(t)$
  $\Pr[I(t+1) > \frac{7}{6} I(t)] = 1 - \Pr[I(t+1) \leq \frac{7}{6} I(t)] \geq 1 - \frac{1}{e^{\frac{1}{24}}} \geq \frac{4}{100}$

Analyse:

- Nach spätestens $\log_{7/6} \frac{n}{3}$ guten Runden gilt $I(t) > \frac{1}{3} n$
- Erneute Anwendung der Chernoff-Bound: Bei $O(\log n)$ Runden treten mit hoher Wahrscheinlichkeit $\log_{7/6} \frac{n}{3}$ gute Runden auf

Somit:

Phase 1 (Wachstum) benötigt mit hoher Wahrscheinlichkeit $O(\log n)$ Runden.

Unabhängigkeit?

Achtung: Chernoff Bound gilt nur für unabhängige Variablen!

Unabhängigkeit?

Achtung: Chernoff Bound gilt nur für unabhängige Variablen!

Aber: $X_j(t)$'s sind nicht unabhängig!

Unabhängigkeit?

Achtung: Chernoff Bound gilt nur für unabhängige Variablen!

Aber: $X_j(t)$'s sind nicht unabhängig!

Beispiel:

- $X_j(t) = 1$ impliziert, dass die Möglichkeit Knoten $c_j(t)$ zu informieren für andere Knoten ausgeschlossen wird

Unabhängigkeit?

Achtung: Chernoff Bound gilt nur für unabhängige Variablen!

Aber: $X_j(t)$'s sind nicht unabhängig!

Beispiel:

- $X_j(t) = 1$ impliziert, dass die Möglichkeit Knoten $c_j(t)$ zu informieren für andere Knoten ausgeschlossen wird
- Dies verringert die Wahrscheinlichkeit für $X_{j'}(t) = 1$ für $j' \neq j$

Unabhängigkeit?

Achtung: Chernoff Bound gilt nur für unabhängige Variablen!

Aber: $X_j(t)$'s sind nicht unabhängig!

Beispiel:

- $X_j(t) = 1$ impliziert, dass die Möglichkeit Knoten $c_j(t)$ zu informieren für andere Knoten ausgeschlossen wird
- Dies verringert die Wahrscheinlichkeit für $X_{j'}(t) = 1$ für $j' \neq j$
- Somit: Keine unabhängigen Ereignisse

Unabhängigkeit?

Achtung: Chernoff Bound gilt nur für unabhängige Variablen!

Aber: $X_j(t)$'s sind nicht unabhängig!

Beispiel:

- $X_j(t) = 1$ impliziert, dass die Möglichkeit Knoten $c_j(t)$ zu informieren für andere Knoten ausgeschlossen wird
- Dies verringert die Wahrscheinlichkeit für $X_{j'}(t) = 1$ für $j' \neq j$
- Somit: Keine unabhängigen Ereignisse

Idee: Es genügt, den „Grad an Unabhängigkeit" auf folgende Aussage einzuschränken:

$$\Pr[X_j(t) = 1 \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \geq 1 - 2i(t)$$

Stochastische Dominanz

Lemma (Doerr 2011)

Seien $X_1, \ldots, X_k$ ***beliebige*** *binäre Zufallsvariablen und seien* $Y_1, \ldots, Y_k$ ***unabhängige*** *binäre Zufallsvariablen. Wenn für alle* j *und alle* $x_1, \ldots, x_{j-1} \in \{0, 1\}$ *gilt, dass*

$$\Pr[X_j = 1 \mid X_1 = x_1, \ldots, X_{j-1} = x_{j-1}] \geq \Pr[Y_j = 1]$$

dann gilt für alle $T \geq 0$

$$\Pr\left[\sum_{j=1}^{k} X_j \leq T\right] \leq \Pr\left[\sum_{j=1}^{k} Y_j \leq T\right]$$

Stochastische Dominanz

Lemma (Doerr 2011)

Seien $X_1, \ldots, X_k$ ***beliebige*** *binäre Zufallsvariablen und seien* $Y_1, \ldots, Y_k$ ***unabhängige*** *binäre Zufallsvariablen. Wenn für alle* j *und alle* $x_1, \ldots, x_{j-1} \in \{0, 1\}$ *gilt, dass*

$$\Pr[X_j = 1 \mid X_1 = x_1, \ldots, X_{j-1} = x_{j-1}] \geq \Pr[Y_j = 1]$$

dann gilt für alle $T \geq 0$

$$\Pr\left[\sum_{j=1}^{k} X_j \leq T\right] \leq \Pr\left[\sum_{j=1}^{k} Y_j \leq T\right]$$

Abschätzung von $\Pr\left[\sum_{j=1}^{k} Y_j \leq T\right]$ durch Chernoff Bound

Stochastische Dominanz

Lemma (Doerr 2011)

Seien $X_1, \ldots, X_k$ ***beliebige*** *binäre Zufallsvariablen und seien* $Y_1, \ldots, Y_k$ ***unabhängige*** *binäre Zufallsvariablen. Wenn für alle* j *und alle* $x_1, \ldots, x_{j-1} \in \{0, 1\}$ *gilt, dass*

$$\Pr[X_j = 1 \mid X_1 = x_1, \ldots, X_{j-1} = x_{j-1}] \geq \Pr[Y_j = 1]$$

dann gilt für alle $T \geq 0$

$$\Pr\left[\sum_{j=1}^{k} X_j \leq T\right] \leq \Pr\left[\sum_{j=1}^{k} Y_j \leq T\right]$$

Abschätzung von $\Pr\left[\sum_{j=1}^{k} Y_j \leq T\right]$ durch Chernoff Bound

Setze $Y_j = 1$ mit Wahrscheinlichkeit $1 - 2i(t)$ und 0 andernfalls

$$\Pr[X_j = 1 \mid X_1 = x_1, \ldots, X_{j-1} = x_{j-1}] \geq 1 - 2i(t) = \Pr[Y_j = 1]$$

Erweiterte Analyse der Zufallsvariablen

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Erweiterte Analyse der Zufallsvariablen

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Für jede Belegung $x_1, \ldots, x_{j-1}$ gilt:

$$\begin{aligned} &\Pr[X_j(t) = 0 \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \\ &\qquad = \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \end{aligned}$$

Erweiterte Analyse der Zufallsvariablen

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Für jede Belegung $x_1, \ldots, x_{j-1}$ gilt:

$$\begin{aligned} \Pr[X_j(t) = 0 \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \\ = \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \\ \leq \frac{I(t) + j - 1}{n} \leq \frac{2I(t)}{n} = 2i(t) \end{aligned}$$

Erweiterte Analyse der Zufallsvariablen

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Für jede Belegung $x_1, \ldots, x_{j-1}$ gilt:

$$\begin{aligned} &\Pr[X_j(t) = 0 \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \\ &\qquad = \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \\ &\qquad \leq \frac{I(t) + j - 1}{n} \leq \frac{2I(t)}{n} = 2i(t) \end{aligned}$$

$$\Rightarrow \Pr[X_j(t) = 1 \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \geq 1 - 2i(t)$$

Erweiterte Analyse der Zufallsvariablen

Definiere Zufallsvariable für jeden Knoten $1 \leq j \leq I(t)$:

$$X_j(t) := \begin{cases} 1 & \text{falls } c_j(t) \notin \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \\ 0 & \text{andernfalls} \end{cases}$$

Für jede Belegung $x_1, \ldots, x_{j-1}$ gilt:

$$\begin{aligned} &\Pr[X_j(t) = 0 \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \\ &\qquad = \Pr[c_j(t) \in \mathcal{I}(t) \cup \{c_1(t), \ldots, c_{j-1}(t)\} \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \\ &\qquad \leq \frac{I(t) + j - 1}{n} \leq \frac{2I(t)}{n} = 2i(t) \end{aligned}$$

$\Rightarrow \Pr[X_j(t) = 1 \mid X_1(t) = x_1, \ldots, X_{j-1}(t) = x_{j-1}] \geq 1 - 2i(t)$

Weitere Anwendung der Chernoff Bound wie vorher

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n}$$

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n} = \left(\left(1 - \frac{1}{n}\right)^{n}\right)^{\frac{1}{3}}$$

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n} = \left(\left(1 - \frac{1}{n}\right)^{n}\right)^{\frac{1}{3}} \leq \left(\frac{1}{e}\right)^{\frac{1}{3}} = \frac{1}{e^{1/3}}$$

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n} = \left(\left(1 - \frac{1}{n}\right)^{n}\right)^{\frac{1}{3}} \leq \left(\frac{1}{e}\right)^{\frac{1}{3}} = \frac{1}{e^{1/3}}$$

Ungleichung: $(1 - \frac{1}{x})^x \leq \frac{1}{e}$ für $x \geq 1$

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n} = \left(\left(1 - \frac{1}{n}\right)^{n}\right)^{\frac{1}{3}} \leq \left(\frac{1}{e}\right)^{\frac{1}{3}} = \frac{1}{e^{1/3}}$$

Ungleichung: $(1 - \frac{1}{x})^x \leq \frac{1}{e}$ für $x \geq 1$

Wahrscheinlichkeit, dass (fixierter) uninformierter Knoten in $3(c+1)\ln n$ aufeinanderfolgenden Runden nicht informiert wird (Bernoulli-Experiment):

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n} = \left(\left(1 - \frac{1}{n}\right)^{n}\right)^{\frac{1}{3}} \leq \left(\frac{1}{e}\right)^{\frac{1}{3}} = \frac{1}{e^{1/3}}$$

Ungleichung: $(1 - \frac{1}{x})^x \leq \frac{1}{e}$ für $x \geq 1$

Wahrscheinlichkeit, dass (fixierter) uninformierter Knoten in $3(c+1)\ln n$ aufeinanderfolgenden Runden nicht informiert wird (Bernoulli-Experiment):

$$= \left(\left(1 - \frac{1}{n}\right)^{I(t)}\right)^{3(c+1)\ln n}$$

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n} = \left(\left(1 - \frac{1}{n}\right)^{n}\right)^{\frac{1}{3}} \leq \left(\frac{1}{e}\right)^{\frac{1}{3}} = \frac{1}{e^{1/3}}$$

Ungleichung: $(1 - \frac{1}{x})^x \leq \frac{1}{e}$ für $x \geq 1$

Wahrscheinlichkeit, dass (fixierter) uninformierter Knoten in $3(c+1)\ln n$ aufeinanderfolgenden Runden nicht informiert wird (Bernoulli-Experiment):

$$= \left(\left(1 - \frac{1}{n}\right)^{I(t)}\right)^{3(c+1)\ln n} \leq \left(\frac{1}{e^{1/3}}\right)^{3(c+1)\ln n}$$

Schrumpfung: $I(t) > \frac{1}{3}n$

Wechsel der Analyse: **Schrumpfung** uninformierter Knoten statt Wachstum informierter Knoten

Wahrscheinlichkeit, dass uninformierter Knoten nicht informiert wird: $\left(1 - \frac{1}{n}\right)^{I(t)}$

Obere Schranke:

$$\left(1 - \frac{1}{n}\right)^{I(t)} \leq \left(1 - \frac{1}{n}\right)^{\frac{1}{3}n} = \left(\left(1 - \frac{1}{n}\right)^{n}\right)^{\frac{1}{3}} \leq \left(\frac{1}{e}\right)^{\frac{1}{3}} = \frac{1}{e^{1/3}}$$

Ungleichung: $(1 - \frac{1}{x})^x \leq \frac{1}{e}$ für $x \geq 1$

Wahrscheinlichkeit, dass (fixierter) uninformierter Knoten in $3(c+1)\ln n$ aufeinanderfolgenden Runden nicht informiert wird (Bernoulli-Experiment):

$$= \left(\left(1 - \frac{1}{n}\right)^{I(t)}\right)^{3(c+1)\ln n} \leq \left(\frac{1}{e^{1/3}}\right)^{3(c+1)\ln n} = \frac{1}{e^{(c+1)\ln n}} = \frac{1}{n^{c+1}}$$

Schrumpfung: $I(t) \geq \frac{1}{3}n$ (Fortsetzung)

Mit **Union Bound**: Wahrscheinlichkeit, dass mindestens einer der uninformierten Knoten in $3(c+1)\ln n$ aufeinanderfolgenden Runden nicht informiert wird:

$$\leq n \cdot \frac{1}{n^{c+1}} = \frac{1}{n^c}$$

Schrumpfung: $I(t) \geq \frac{1}{3}n$ (Fortsetzung)

Mit **Union Bound**: Wahrscheinlichkeit, dass mindestens einer der uninformierten Knoten in $3(c+1)\ln n$ aufeinanderfolgenden Runden nicht informiert wird:

$$\leq n \cdot \frac{1}{n^{c+1}} = \frac{1}{n^c}$$

Gegenereignis: Jeder uninformierte Knoten wird in $3(c+1)\ln n$ aufeinanderfolgenden Runden informiert

Wahrscheinlichkeit $\geq 1 - \frac{1}{n^c}$

Schrumpfung: $I(t) \geq \frac{1}{3}n$ (Fortsetzung)

Mit **Union Bound**: Wahrscheinlichkeit, dass mindestens einer der uninformierten Knoten in $3(c+1)\ln n$ aufeinanderfolgenden Runden nicht informiert wird:

$$\leq n \cdot \frac{1}{n^{c+1}} = \frac{1}{n^c}$$

Gegenereignis: Jeder uninformierte Knoten wird in $3(c+1)\ln n$ aufeinanderfolgenden Runden informiert

Wahrscheinlichkeit $\geq 1 - \frac{1}{n^c}$

Somit:

Phase 2 (Schrumpfung) benötigt mit hoher Wahrscheinlichkeit $O(\log n)$ Runden.

Zusammenfassung

Informationsausbreitung im Push Modell:

1. **Wachstum:** $I(t) \leq \frac{1}{3}n$
   Dauer: $O(\log n)$ Runden mit hoher Wahrscheinlichkeit
2. **Schrumpfung:** $I(t) > \frac{1}{3}n$
   Dauer: $O(\log n)$ Runden mit hoher Wahrscheinlichkeit

Informationsausbreitung im Pull Modell:

- $O(\log n)$ Runden mit hoher Wahrscheinlichkeit
- Analyse benötigt ähnliche Beweistechniken wie im Push Modell

Quellen

Der Inhalt dieser Vorlesungseinheit basiert zum Teil auf Vorlesungseinheiten von Robert Elsässer und Christian Schindelhauer.

Literatur:

- Alan J. Demers et al. „Epidemic Algorithms for Replicated Database Maintenance“. *Operating Systems Review* 22(1): 8–32 (1988)
- Benjamin Doerr, Marvin Künnemann. „Tight Analysis of Randomized Rumor Spreading in Complete Graphs“. In: *Proc. of the Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*. 2014, S. 82–91
- Alan M. Frieze, Geoffrey R. Grimmett. „The shortest-path problem for graphs with random arc-lengths“. *Discrete Applied Mathematics* 10(1): 57–77 (1985)