Epidemische Informationsausbreitung II

Algorithmen für verteilte Systeme

Sebastian Forster

Universität Salzburg

Random-Phone-Call-Modell

- Netzwerkstruktur: vollständiger Graph mit n Knoten

Random-Phone-Call-Modell

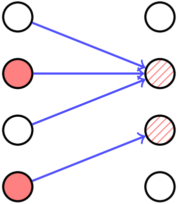
- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an

  Der Einfachheit halber: Knoten können auch sich selbst anrufen

Random-Phone-Call-Modell

- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an

  Der Einfachheit halber: Knoten können auch sich selbst anrufen
- Drei Kommunikationsmodelle:
  1. **Push:** Der Anrufer informiert den Angerufenen

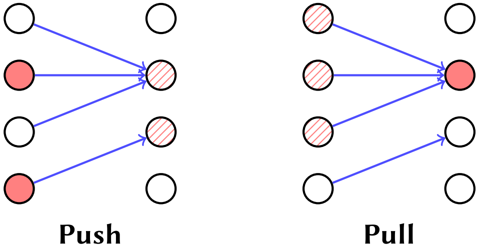


Push

Random-Phone-Call-Modell

- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an

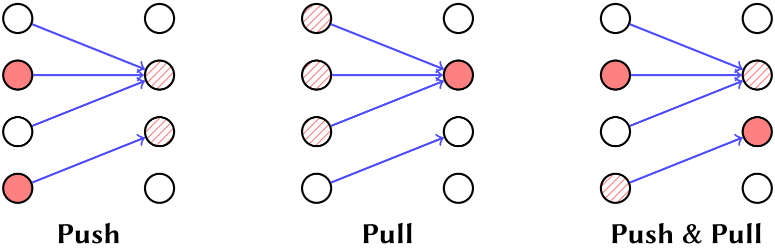
  Der Einfachheit halber: Knoten können auch sich selbst anrufen
- Drei Kommunikationsmodelle:
  1. **Push:** Der Anrufer informiert den Angerufenen
  2. **Pull:** Der Angerufene informiert den Anrufer



Random-Phone-Call-Modell

- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an

  Der Einfachheit halber: Knoten können auch sich selbst 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



Zeit- und Nachrichtenkomplexität

Theorem ([Frieze/Grimmett ’85])

Ausgehend von einem infizierten Knoten, sind um Push-Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ *Runden alle Knoten infiziert.*

Zeit- und Nachrichtenkomplexität

Theorem ([Frieze/Grimmett '85])

Ausgehend von einem infizierten Knoten, sind um Push-Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ Runden alle Knoten infiziert.

- Ähnliche Analyse für $O(\log n)$ Runden im Pull-Modell

Zeit- und Nachrichtenkomplexität

Theorem ([Frieze/Grimmett ’85])

Ausgehend von einem infizierten Knoten, sind um Push-Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ Runden alle Knoten infiziert.

- Ähnliche Analyse für $O(\log n)$ Runden im Pull-Modell
- Pro Anruf eines infizierten Knotens wird eine Nachricht gesendet
  Unabhängig davon ob damit eine neue Infektion stattfindet

Zeit- und Nachrichtenkomplexität

Theorem ([Frieze/Grimmett '85])

Ausgehend von einem infizierten Knoten, sind um Push-Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ Runden alle Knoten infiziert.

- Ähnliche Analyse für $O(\log n)$ Runden im Pull-Modell
- Pro Anruf eines infizierten Knotens wird eine Nachricht gesendet
  Unabhängig davon ob damit eine neue Infektion stattfindet
- Triviale Nachrichtenkomplexität $O(n \log n)$

Zeit- und Nachrichtenkomplexität

Theorem ([Frieze/Grimmett '85])

Ausgehend von einem infizierten Knoten, sind um Push-Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ Runden alle Knoten infiziert.

- Ähnliche Analyse für $O(\log n)$ Runden im Pull-Modell
- Pro Anruf eines infizierten Knotens wird eine Nachricht gesendet
  Unabhängig davon ob damit eine neue Infektion stattfindet
- Triviale Nachrichtenkomplexität $O(n \log n)$

Frage: Wie kann die Nachrichtenkomplexität reduziert werden?

Zeit- und Nachrichtenkomplexität

Theorem ([Frieze/Grimmett ’85])

Ausgehend von einem infizierten Knoten, sind um Push-Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ *Runden alle Knoten infiziert.*

- Ähnliche Analyse für $O(\log n)$ Runden im Pull-Modell
- Pro Anruf eines infizierten Knotens wird eine Nachricht gesendet
  Unabhängig davon ob damit eine neue Infektion stattfindet
- Triviale Nachrichtenkomplexität $O(n \log n)$

Frage: Wie kann die Nachrichtenkomplexität reduziert werden?

Theorem ([Karp et al. ’00])

Ausgehend von einem infizierten Knoten, sind um Push&Pull-Modell mit hoher Wahrscheinlichkeit nach $O(\log n)$ *Runden mit insgesamt* $O(n \log \log n)$ *gesendeten Nachrichten alle Knoten infiziert.*

Beweisstrategie

Einteilung in drei Phasen:

1. **Wachstum:** $I(t) \leq \frac{n}{\ln n}$
   Dauer: $O(\log n)$ Runden
   #Nachrichten: $O(n)$

Beweisstrategie

Einteilung in drei Phasen:

1. **Wachstum:** $I(t) \leq \frac{n}{\ln n}$
   Dauer: $O(\log n)$ Runden
   #Nachrichten: $O(n)$
2. **Schrumpfung:** $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$

Beweisstrategie

Einteilung in drei Phasen:

1. **Wachstum:** $I(t) \leq \frac{n}{\ln n}$
   Dauer: $O(\log n)$ Runden
   #Nachrichten: $O(n)$
2. **Schrumpfung:** $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $G(t) < \sqrt{3cn(\ln n)^5}$
   Dauer: $O(1)$ Runden
   #Nachrichten: $O(n)$

Beweisstrategie

Einteilung in drei Phasen:

1. **Wachstum:** $I(t) \leq \frac{n}{\ln n}$
   Dauer: $O(\log n)$ Runden
   #Nachrichten: $O(n)$
2. **Schrumpfung:** $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $G(t) < \sqrt{3cn(\ln n)^5}$
   Dauer: $O(1)$ Runden
   #Nachrichten: $O(n)$

In jeder Phase: Garantie mit hoher Wahrscheinlichkeit

Gesamt: $O(\log n)$ Runden und $O(n \log \log n)$ Nachrichten mit hoher Wahrscheinlichkeit

Beweisstrategie

Einteilung in drei Phasen:

1. **Wachstum:** $I(t) \leq \frac{n}{\ln n}$
   Dauer: $O(\log n)$ Runden
   #Nachrichten: $O(n)$
2. **Schrumpfung:** $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $G(t) < \sqrt{3cn(\ln n)^5}$
   Dauer: $O(1)$ Runden
   #Nachrichten: $O(n)$

In jeder Phase: Garantie mit hoher Wahrscheinlichkeit

Gesamt: $O(\log n)$ Runden und $O(n \log \log n)$ Nachrichten mit hoher Wahrscheinlichkeit

Annahme: n ist „groß genug“, d.h., größer als eine in der Analyse festgelegte Konstante n_0

Notation für Analyse

- n: Anzahl der Knoten
- $I(t)$: Anzahl der infizierten Knoten am Beginn von Runde t
- $G(t)$: Anzahl der gesunden Knoten
- $i(t) = \frac{I(t)}{n}$: relativer Anteil der infizierten Knoten
- $g(t) = \frac{G(t)}{n}$: relativer Anteil der gesunden Knoten

Achtung: Zufallsvariablen

Notation für Analyse

- n: Anzahl der Knoten
- $I(t)$: Anzahl der infizierten Knoten am Beginn von Runde t
- $G(t)$: Anzahl der gesunden Knoten
- $i(t) = \frac{I(t)}{n}$: relativer Anteil der infizierten Knoten
- $g(t) = \frac{G(t)}{n}$: relativer Anteil der gesunden Knoten

Achtung: Zufallsvariablen

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

Notation für Analyse

- n: Anzahl der Knoten
- $I(t)$: Anzahl der infizierten Knoten am Beginn von Runde t
- $G(t)$: Anzahl der gesunden Knoten
- $i(t) = \frac{I(t)}{n}$: relativer Anteil der infizierten Knoten
- $g(t) = \frac{G(t)}{n}$: relativer Anteil der gesunden Knoten

Achtung: Zufallsvariablen

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

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

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl infizierter Knoten, die gesunde Knoten anrufen (Push)
  - $B(t)$: Anzahl gesunder Knoten, die infizierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl infizierter Knoten, die infizierte Knoten anrufen (Push&Pull)

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl infizierter Knoten, die gesunde Knoten anrufen (Push)
  - $B(t)$: Anzahl gesunder Knoten, die infizierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl infizierter Knoten, die infizierte Knoten anrufen (Push&Pull)
- Es gilt: $M(t) = A(t) + B(t) + 2C(t) \leq 3I(t) + B(t)$

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl infizierter Knoten, die gesunde Knoten anrufen (Push)
  - $B(t)$: Anzahl gesunder Knoten, die infizierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl infizierter Knoten, die infizierte Knoten anrufen (Push&Pull)
- Es gilt: $M(t) = A(t) + B(t) + 2C(t) \leq 3I(t) + B(t)$
- Jeder Knoten kann nur ein Mal infiziert werden: $\sum_{t=1}^{t^*} B(t) \leq n$

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl infizierter Knoten, die gesunde Knoten anrufen (Push)
  - $B(t)$: Anzahl gesunder Knoten, die infizierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl infizierter Knoten, die infizierte Knoten anrufen (Push&Pull)
- Es gilt: $M(t) = A(t) + B(t) + 2C(t) \leq 3I(t) + B(t)$
- Jeder Knoten kann nur ein Mal infiziert werden: $\sum_{t=1}^{t^*} B(t) \leq n$
- Anzahl der in der Wachstumsphase gesendeten Nachrichten:

$$\sum_{t=1}^{t^*} M(t) \leq \sum_{t=1}^{t^*} (3I(t) + B(t))$$

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl infizierter Knoten, die gesunde Knoten anrufen (Push)
  - $B(t)$: Anzahl gesunder Knoten, die infizierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl infizierter Knoten, die infizierte Knoten anrufen (Push&Pull)
- Es gilt: $M(t) = A(t) + B(t) + 2C(t) \leq 3I(t) + B(t)$
- Jeder Knoten kann nur ein Mal infiziert werden: $\sum_{t=1}^{t^*} B(t) \leq n$
- Anzahl der in der Wachstumsphase gesendeten Nachrichten:

$$\sum_{t=1}^{t^*} M(t) \leq \sum_{t=1}^{t^*} (3I(t) + B(t)) = \sum_{t=1}^{t^*} 3I(t) + \sum_{t=1}^{t^*} B(t)$$

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl infizierter Knoten, die gesunde Knoten anrufen (Push)
  - $B(t)$: Anzahl gesunder Knoten, die infizierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl infizierter Knoten, die infizierte Knoten anrufen (Push&Pull)
- Es gilt: $M(t) = A(t) + B(t) + 2C(t) \leq 3I(t) + B(t)$
- Jeder Knoten kann nur ein Mal infiziert werden: $\sum_{t=1}^{t^*} B(t) \leq n$
- Anzahl der in der Wachstumsphase gesendeten Nachrichten:

$$\sum_{t=1}^{t^*} M(t) \leq \sum_{t=1}^{t^*} (3I(t) + B(t)) = \sum_{t=1}^{t^*} 3I(t) + \sum_{t=1}^{t^*} B(t) \leq \frac{n}{\ln n} \cdot t^* + n$$

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

Analyse Push-Modell: Wachstumsphase dauert $t^* = O(\log n)$ Runden

(Wir dürfen annehmen, dass $\frac{n}{\ln n} \leq \frac{n}{3}$)

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl infizierter Knoten, die gesunde Knoten anrufen (Push)
  - $B(t)$: Anzahl gesunder Knoten, die infizierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl infizierter Knoten, die infizierte Knoten anrufen (Push&Pull)
- Es gilt: $M(t) = A(t) + B(t) + 2C(t) \leq 3I(t) + B(t)$
- Jeder Knoten kann nur ein Mal infiziert werden: $\sum_{t=1}^{t^*} B(t) \leq n$
- Anzahl der in der Wachstumsphase gesendeten Nachrichten:

$$\begin{aligned}\sum_{t=1}^{t^*} M(t) \leq \sum_{t=1}^{t^*} (3I(t) + B(t)) &= \sum_{t=1}^{t^*} 3I(t) + \sum_{t=1}^{t^*} B(t) \\ &\leq \frac{n}{\ln n} \cdot t^* + n = O(n)\end{aligned}$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft: $g(t)$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft: $g(t)$

Somit:

$$\mathrm{Ex}[G(t+1)] = G(t) \cdot g(t) = \frac{(G(t))^2}{n}$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft: $g(t)$

Somit:

$$\mathrm{Ex}[G(t+1)] = G(t) \cdot g(t) = \frac{(G(t))^2}{n}$$

$$\mathrm{Ex}[g(t+1)] = \frac{G(t) \cdot g(t)}{n} = (g(t))^2$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft: $g(t)$

Somit:

$$\mathrm{Ex}[G(t+1)] = G(t) \cdot g(t) = \frac{(G(t))^2}{n}$$

$$\mathrm{Ex}[g(t+1)] = \frac{G(t) \cdot g(t)}{n} = (g(t))^2$$

Allgemein:

$$\mathrm{Ex}[g(t+k)] = (g(t))^{(2^k)}$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft: $g(t)$

Somit:

$$\mathrm{Ex}[G(t+1)] = G(t) \cdot g(t) = \frac{(G(t))^2}{n}$$

$$\mathrm{Ex}[g(t+1)] = \frac{G(t) \cdot g(t)}{n} = (g(t))^2$$

Allgemein:

$$\mathrm{Ex}[g(t+k)] = (g(t))^{(2^k)}$$

Beobachtung

Für $G(t) < n - \frac{n}{\ln n}$ (und somit $g(t) < 1 - \frac{1}{\ln n}$) und $k \geq 2 \log_2(\ln n)$ gilt:

$$(g(t))^{(2^k)} < \left(1 - \frac{1}{\ln n}\right)^{(\ln n)^2} \leq \frac{1}{e^{\ln n}} = \frac{1}{n}$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft: $g(t)$

Somit:

$$\mathrm{Ex}[G(t+1)] = G(t) \cdot g(t) = \frac{(G(t))^2}{n}$$

$$\mathrm{Ex}[g(t+1)] = \frac{G(t) \cdot g(t)}{n} = (g(t))^2$$

Allgemein:

$$\mathrm{Ex}[g(t+k)] = (g(t))^{(2^k)}$$

Beobachtung

Für $G(t) < n - \frac{n}{\ln n}$ (und somit $g(t) < 1 - \frac{1}{\ln n}$) und $k \geq 2 \log_2(\ln n)$ gilt:

$$(g(t))^{(2^k)} < \left(1 - \frac{1}{\ln n}\right)^{(\ln n)^2} \leq \frac{1}{e^{\ln n}} = \frac{1}{n}$$

Idee: Zeige mit Chernoff Bound, dass quadratisches Schrumpfen mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase auftritt

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$
- $\mu := \mathrm{Ex}[G(t+1)] = \frac{(G(t))^2}{n}$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$
- $\mu := \mathrm{Ex}[G(t+1)] = \frac{(G(t))^2}{n}$
- Setze $\delta = \frac{1}{(\ln n)^2}$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$
- $\mu := \mathrm{Ex}[G(t+1)] = \frac{(G(t))^2}{n}$
- Setze $\delta = \frac{1}{(\ln n)^2}$
- In Schrumpfungsphase: $G(t) \geq \sqrt{3cn(\ln n)^5}$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$
- $\mu := \mathrm{Ex}[G(t+1)] = \frac{(G(t))^2}{n}$
- Setze $\delta = \frac{1}{(\ln n)^2}$
- In Schrumpfungsphase: $G(t) \geq \sqrt{3cn(\ln n)^5}$

$$\Pr\left[\sum_{j=1}^{G(t)} X_j(t) > (1+\delta) \cdot \mu\right] \leq \frac{1}{e^{\frac{\delta^2}{3} \cdot \mu}}$$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$
- $\mu := \operatorname{Ex}[G(t+1)] = \frac{(G(t))^2}{n}$
- Setze $\delta = \frac{1}{(\ln n)^2}$
- In Schrumpfungsphase: $G(t) \geq \sqrt{3cn(\ln n)^5}$

$$\Pr\left[\sum_{j=1}^{G(t)} X_j(t) > (1+\delta) \cdot \mu\right] \leq \frac{1}{e^{\frac{\delta^2}{3} \cdot \mu}} = \frac{1}{e^{\frac{\delta^2}{3} \cdot \frac{(G(t))^2}{n}}}$$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$
- $\mu := \mathrm{Ex}[G(t+1)] = \frac{(G(t))^2}{n}$
- Setze $\delta = \frac{1}{(\ln n)^2}$
- In Schrumpfungsphase: $G(t) \geq \sqrt{3cn(\ln n)^5}$

$$\Pr\left[\sum_{j=1}^{G(t)} X_j(t) > (1+\delta)\cdot\mu\right] \leq \frac{1}{e^{\frac{\delta^2}{3}\cdot\mu}} = \frac{1}{e^{\frac{\delta^2}{3}\cdot\frac{(G(t))^2}{n}}}$$

$$\leq \frac{1}{e^{\frac{\delta^2}{3}\cdot\frac{3cn(\ln n)^5}{n}}}$$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden gesunden Knoten

$$X_j(t) := \begin{cases} 1 & \text{falls gesunder Knoten } j \text{ in Runde } t \text{ gesunden Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $G(t+1) = \sum_{j=1}^{G(t)} X_j(t)$
- $\mu := \mathrm{Ex}[G(t+1)] = \frac{(G(t))^2}{n}$
- Setze $\delta = \frac{1}{(\ln n)^2}$
- In Schrumpfungsphase: $G(t) \geq \sqrt{3cn(\ln n)^5}$

$$\Pr\left[\sum_{j=1}^{G(t)} X_j(t) > (1+\delta)\cdot\mu\right] \leq \frac{1}{e^{\frac{\delta^2}{3}\cdot\mu}} = \frac{1}{e^{\frac{\delta^2}{3}\cdot\frac{(G(t))^2}{n}}}$$

$$\leq \frac{1}{e^{\frac{\delta^2}{3}\cdot\frac{3cn(\ln n)^5}{n}}} = \frac{1}{e^{c\ln n}} = \frac{1}{n^c}$$

Schrumpfung: $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$

Quadratisches Schrumpfen:

$$G(t+1) \leq (1+\delta) \cdot (G(t))^2 \quad \text{bzw.} \quad g(t+1) \leq (1+\delta) \cdot (g(t))^2$$

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Schrumpfung: $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$

Quadratisches Schrumpfen:

$$G(t+1) \leq (1+\delta) \cdot (G(t))^2 \quad \text{bzw.} \quad g(t+1) \leq (1+\delta) \cdot (g(t))^2$$

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Allgemein:

$$g(t+k) \leq (1+\delta)^{(2^k)-1} \cdot (g(t))^{(2^k)} \leq (1+\delta)^{(2^k)} \cdot (g(t))^{(2^k)}$$

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

Schrumpfung: $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$

Quadratisches Schrumpfen:

$$G(t+1) \leq (1+\delta) \cdot (G(t))^2 \quad \text{bzw.} \quad g(t+1) \leq (1+\delta) \cdot (g(t))^2$$

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Allgemein:

$$g(t+k) \leq (1+\delta)^{(2^k)-1} \cdot (g(t))^{(2^k)} \leq (1+\delta)^{(2^k)} \cdot (g(t))^{(2^k)}$$

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

Mit $G(t) < n - \frac{n}{\ln n}$ (also $g(t) < 1 - \frac{1}{\ln n}$) und $k = \lceil 2 \log_2(\ln n) \rceil = \lceil \log_2 \frac{1}{\delta} \rceil$:

Schrumpfung: $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$

Quadratisches Schrumpfen:

$$G(t+1) \leq (1+\delta) \cdot (G(t))^2 \quad \text{bzw.} \quad g(t+1) \leq (1+\delta) \cdot (g(t))^2$$

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Allgemein:

$$g(t+k) \leq (1+\delta)^{(2^k)-1} \cdot (g(t))^{(2^k)} \leq (1+\delta)^{(2^k)} \cdot (g(t))^{(2^k)}$$

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

Mit $G(t) < n - \frac{n}{\ln n}$ (also $g(t) < 1 - \frac{1}{\ln n}$) und $k = \lceil 2 \log_2(\ln n) \rceil = \lceil \log_2 \frac{1}{\delta} \rceil$:

$$(1+\delta)^{(2^k)}(g(t))^{(2^k)} \leq (1+\delta)^{\frac{1}{\delta} \cdot 2} (g(t))^{(2^k)}$$

Schrumpfung: $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$

Quadratisches Schrumpfen:

$$G(t+1) \leq (1+\delta) \cdot (G(t))^2 \quad \text{bzw.} \quad g(t+1) \leq (1+\delta) \cdot (g(t))^2$$

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Allgemein:

$$g(t+k) \leq (1+\delta)^{(2^k)-1} \cdot (g(t))^{(2^k)} \leq (1+\delta)^{(2^k)} \cdot (g(t))^{(2^k)}$$

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

Mit $G(t) < n - \frac{n}{\ln n}$ (also $g(t) < 1 - \frac{1}{\ln n}$) und $k = \lceil 2 \log_2(\ln n) \rceil = \lceil \log_2 \frac{1}{\delta} \rceil$:

$$(1+\delta)^{(2^k)}(g(t))^{(2^k)} \leq (1+\delta)^{\frac{1}{\delta} \cdot 2} (g(t))^{(2^k)} < e^2 \cdot \frac{1}{n}$$

Mit Grenzwertdefinition der Eulerschen Zahl: $\left(1 + \frac{1}{x}\right)^x \leq e$ für $x > 0$

Schrumpfung: $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$

Quadratisches Schrumpfen:

$$G(t+1) \leq (1+\delta) \cdot (G(t))^2 \quad \text{bzw.} \quad g(t+1) \leq (1+\delta) \cdot (g(t))^2$$

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Allgemein:

$$g(t+k) \leq (1+\delta)^{(2^k)-1} \cdot (g(t))^{(2^k)} \leq (1+\delta)^{(2^k)} \cdot (g(t))^{(2^k)}$$

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

Mit $G(t) < n - \frac{n}{\ln n}$ (also $g(t) < 1 - \frac{1}{\ln n}$) und $k = \lceil 2 \log_2(\ln n) \rceil = \lceil \log_2 \frac{1}{\delta} \rceil$:

$$(1+\delta)^{(2^k)}(g(t))^{(2^k)} \leq (1+\delta)^{\frac{1}{\delta} \cdot 2} (g(t))^{(2^k)} < e^2 \cdot \frac{1}{n} \leq \frac{\sqrt{3cn(\ln n)^5}}{n}$$

(Wir dürfen annehmen, dass $\sqrt{3cn(\ln n)^5} \geq e$)

Mit Grenzwertdefinition der Eulerschen Zahl: $\left(1 + \frac{1}{x}\right)^x \leq e$ für $x > 0$

Schrumpfung: $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$

Quadratisches Schrumpfen:

$$G(t+1) \leq (1+\delta) \cdot (G(t))^2 \quad \text{bzw.} \quad g(t+1) \leq (1+\delta) \cdot (g(t))^2$$

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Allgemein:

$$g(t+k) \leq (1+\delta)^{(2^k)-1} \cdot (g(t))^{(2^k)} \leq (1+\delta)^{(2^k)} \cdot (g(t))^{(2^k)}$$

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

Mit $G(t) < n - \frac{n}{\ln n}$ (also $g(t) < 1 - \frac{1}{\ln n}$) und $k = \lceil 2 \log_2(\ln n) \rceil = \lceil \log_2 \frac{1}{\delta} \rceil$:

$$(1+\delta)^{(2^k)}(g(t))^{(2^k)} \leq (1+\delta)^{\frac{1}{\delta} \cdot 2} (g(t))^{(2^k)} < e^2 \cdot \frac{1}{n} \leq \frac{\sqrt{3cn(\ln n)^5}}{n}$$

(Wir dürfen annehmen, dass $\sqrt{3cn(\ln n)^5} \geq e$)

Mit Grenzwertdefinition der Eulerschen Zahl: $\left(1 + \frac{1}{x}\right)^x \leq e$ für $x > 0$

Somit: Mit hoher Wahrscheinlichkeit weniger als $k = O(\log \log n)$ Runden für Schrumpfungsphase

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Wahrscheinlichkeit, dass gesunder Knoten in $2c + 3$ Runden nur gesunde Knoten anruft:

$$\leq p^{2c+3}$$

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Wahrscheinlichkeit, dass gesunder Knoten in $2c + 3$ Runden nur gesunde Knoten anruft:

$$\leq p^{2c+3}$$

Wahrscheinlichkeit, dass in $2c + 3$ Runden mindestens einer der gesunden Knoten nur gesunde Knoten anruft (und damit gesund bleibt):

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Wahrscheinlichkeit, dass gesunder Knoten in $2c + 3$ Runden nur gesunde Knoten anruft:

$$\leq p^{2c+3}$$

Wahrscheinlichkeit, dass in $2c + 3$ Runden mindestens einer der gesunden Knoten nur gesunde Knoten anruft (und damit gesund bleibt):

(Union Bound)

$$\leq \sqrt{3cn(\ln n)^5} \cdot p^{2c+3}$$

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Wahrscheinlichkeit, dass gesunder Knoten in $2c + 3$ Runden nur gesunde Knoten anruft:

$$\leq p^{2c+3}$$

Wahrscheinlichkeit, dass in $2c + 3$ Runden mindestens einer der gesunden Knoten nur gesunde Knoten anruft (und damit gesund bleibt):

(Union Bound)

$$\leq \sqrt{3cn(\ln n)^5} \cdot p^{2c+3} = \frac{\left(\sqrt{3cn(\ln n)^5}\right)^{2c+4}}{n^{2c+3}}$$

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Wahrscheinlichkeit, dass gesunder Knoten in $2c+3$ Runden nur gesunde Knoten anruft:

$$\leq p^{2c+3}$$

Wahrscheinlichkeit, dass in $2c+3$ Runden mindestens einer der gesunden Knoten nur gesunde Knoten anruft (und damit gesund bleibt):

(Union Bound)

$$\leq \sqrt{3cn(\ln n)^5} \cdot p^{2c+3} = \frac{\left(\sqrt{3cn(\ln n)^5}\right)^{2c+4}}{n^{2c+3}}$$

$$= \frac{(3c(\ln n)^5)^{c+2} \cdot n^{c+2}}{n^{2c+3}}$$

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Wahrscheinlichkeit, dass gesunder Knoten in $2c + 3$ Runden nur gesunde Knoten anruft:

$$\leq p^{2c+3}$$

Wahrscheinlichkeit, dass in $2c + 3$ Runden mindestens einer der gesunden Knoten nur gesunde Knoten anruft (und damit gesund bleibt):

(Union Bound)

$$\leq \sqrt{3cn(\ln n)^5} \cdot p^{2c+3} = \frac{\left(\sqrt{3cn(\ln n)^5}\right)^{2c+4}}{n^{2c+3}}$$

$$= \frac{(3c(\ln n)^5)^{c+2} \cdot n^{c+2}}{n^{2c+3}} \qquad = \frac{(3c(\ln n)^5)^{c+2}}{n^c}$$

Schluss: $G(t) < \sqrt{3cn(\ln n)^5}$

Wahrscheinlichkeit, dass gesunder Knoten gesunden Knoten anruft:

$$\leq p := \frac{\sqrt{3cn(\ln n)^5}}{n}$$

Wahrscheinlichkeit, dass gesunder Knoten in $2c + 3$ Runden nur gesunde Knoten anruft:

$$\leq p^{2c+3}$$

Wahrscheinlichkeit, dass in $2c + 3$ Runden mindestens einer der gesunden Knoten nur gesunde Knoten anruft (und damit gesund bleibt):

(Union Bound)

$$\leq \sqrt{3cn(\ln n)^5} \cdot p^{2c+3} = \frac{\left(\sqrt{3cn(\ln n)^5}\right)^{2c+4}}{n^{2c+3}}$$

$$= \frac{(3c(\ln n)^5)^{c+2} \cdot n^{c+2}}{n^{2c+3}} \qquad = \frac{(3c(\ln n)^5)^{c+2}}{n^c} \leq \frac{1}{n^c}$$

(Wir dürfen annehmen, dass $n \geq (3c(\ln n)^5)^{c+2}$)

Zusammenfassung: Push & Pull

Einteilung in drei Phasen:

1. **Wachstum:** $I(t) \leq \frac{n}{\ln n}$
   Dauer: $O(\log n)$ Runden
   #Nachrichten: $O(n)$
2. **Schrumpfung:** $n - \frac{n}{\ln n} > G(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $G(t) < \sqrt{3cn(\ln n)^5}$
   Dauer: $O(1)$ Runden
   #Nachrichten: $O(n)$

In jeder Phase: Garantie mit hoher Wahrscheinlichkeit

Gesamt: $O(\log n)$ Runden und $O(n \log \log n)$ Nachrichten mit hoher Wahrscheinlichkeit

Techniken

- Standard-Tools:
  - Union Bound
  - Chernoff Bound
  - Binomialverteilung
  - Abschätzung durch Gegenereignis
  - $(1 - \frac{1}{x})^x \leq \frac{1}{e}$
  - $(1 + \frac{1}{x})^x \leq e$

Techniken

- Standard-Tools:
  - Union Bound
  - Chernoff Bound
  - Binomialverteilung
  - Abschätzung durch Gegenereignis
  - $(1 - \frac{1}{x})^x \leq \frac{1}{e}$
  - $(1 + \frac{1}{x})^x \leq e$
- Chernoff-Bound benötigt untere Schranke an Erwartungswert
  $\rightarrow$ Spezialfall zur Einhaltung dieser Schranke muss abgedeckt werden

Techniken

- Standard-Tools:
  - Union Bound
  - Chernoff Bound
  - Binomialverteilung
  - Abschätzung durch Gegenereignis
  - $(1 - \frac{1}{x})^x \leq \frac{1}{e}$
  - $(1 + \frac{1}{x})^x \leq e$
- Chernoff-Bound benötigt untere Schranke an Erwartungswert
  $\longrightarrow$ Spezialfall zur Einhaltung dieser Schranke muss abgedeckt werden
- Analyse „kippt“ sobald gewisser Anteil an Knoten infiziert wurde, sowohl bei Push also auch bei Pull

Quellen

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

Literatur:

- Richard M. Karp, Christian Schindelhauer, Scott Shenker, Berthold Vöcking. „Randomized Rumor Spreading“. In: *Proc. of the Symposium on Foundations of Computer Science (FOCS)*. 2000, S. 565–574