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

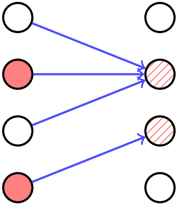
- Netzwerkstruktur: vollständiger Graph mit n Knoten
- Kommunikation findet in synchronen Runden statt
- In jeder Runde ruft jeder Knoten einen uniform zufällig gewählten Knoten an

  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

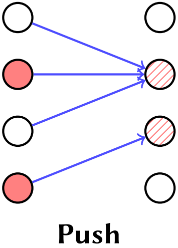


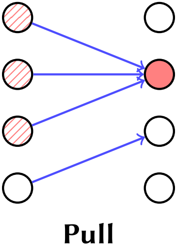
Push

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

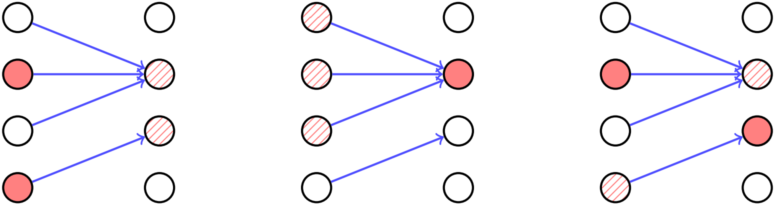




Random-Phone-Call Modell

- 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



Push **Pull** **Push & Pull**

Zeit- und Nachrichtenkomplexität

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.

Zeit- und Nachrichtenkomplexität

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.

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

Zeit- und Nachrichtenkomplexität

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.

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

Zeit- und Nachrichtenkomplexität

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.

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

Zeit- und Nachrichtenkomplexität

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.

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

Frage: Wie kann die Nachrichtenkomplexität reduziert werden?

Zeit- und Nachrichtenkomplexität

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.

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

Frage: Wie kann die Nachrichtenkomplexität reduziert werden?

Theorem ([Karp et al. '00])

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

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 uninformierten Knoten an und bezeichen die Knoten mit $1, \dots, U(t)$

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:** $n - \frac{n}{\ln n} > U(t) \geq \sqrt{3cn(\ln n)^5}$
   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:** $n - \frac{n}{\ln n} > U(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $U(t) < \sqrt{3cn(\ln n)^5}$
   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:** $n - \frac{n}{\ln n} > U(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $U(t) < \sqrt{3cn(\ln n)^5}$
   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:** $n - \frac{n}{\ln n} > U(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $U(t) < \sqrt{3cn(\ln n)^5}$
   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

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

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

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl informierter Knoten, die uninformierte Knoten anrufen (Push)
  - $B(t)$: Anzahl uninformierter Knoten, die informierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl informierter Knoten, die informierte Knoten anrufen (Push&Pull)

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

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

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl informierter Knoten, die uninformierte Knoten anrufen (Push)
  - $B(t)$: Anzahl uninformierter Knoten, die informierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl informierter Knoten, die informierte 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

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl informierter Knoten, die uninformierte Knoten anrufen (Push)
  - $B(t)$: Anzahl uninformierter Knoten, die informierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl informierter Knoten, die informierte 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 informiert 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

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl informierter Knoten, die uninformierte Knoten anrufen (Push)
  - $B(t)$: Anzahl uninformierter Knoten, die informierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl informierter Knoten, die informierte 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 informiert 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

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl informierter Knoten, die uninformierte Knoten anrufen (Push)
  - $B(t)$: Anzahl uninformierter Knoten, die informierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl informierter Knoten, die informierte 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 informiert 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

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl informierter Knoten, die uninformierte Knoten anrufen (Push)
  - $B(t)$: Anzahl uninformierter Knoten, die informierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl informierter Knoten, die informierte 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 informiert werden: $\sum_{t=1}^{t^*} B(t) \leq n$
- Anzahl der in der Wachstumsphase gesendeten Nachrichten:

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

$$\leq \frac{3n}{\ln n} \cdot t^* + n$$

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

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

Nachrichtenkomplexität:

- Sei $M(t)$ die Anzahl der in Runde t gesendeten Nachrichten
- Drei Arten relevanter Interaktionen:
  - $A(t)$: Anzahl informierter Knoten, die uninformierte Knoten anrufen (Push)
  - $B(t)$: Anzahl uninformierter Knoten, die informierte Knoten anrufen (Pull)
  - $C(t)$: Anzahl informierter Knoten, die informierte 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 informiert werden: $\sum_{t=1}^{t^*} B(t) \leq n$
- Anzahl der in der Wachstumsphase gesendeten Nachrichten:

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

$$\leq \frac{3n}{\ln n} \cdot t^* + n = O(n)$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft: $u(t)$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft: $u(t)$

Somit:

$$\mathrm{Ex}[U(t+1)] = u(t) \cdot U(t)$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft: $u(t)$

Somit:

$$\mathrm{Ex}[U(t+1)] = u(t) \cdot U(t)$$

$$\mathrm{Ex}[u(t+1)] = \mathrm{Ex}[\frac{U(t+1)}{n}] = (u(t))^2$$

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft: $u(t)$

Somit:

$$\mathrm{Ex}[U(t+1)] = u(t) \cdot U(t)$$

$$\mathrm{Ex}[u(t+1)] = \mathrm{Ex}[\frac{U(t+1)}{n}] = (u(t))^2$$

Iteriert:

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

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft: $u(t)$

Somit:

$$\mathrm{Ex}[U(t+1)] = u(t) \cdot U(t)$$

$$\mathrm{Ex}[u(t+1)] = \mathrm{Ex}[\frac{U(t+1)}{n}] = (u(t))^2$$

Iteriert:

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

Beobachtung

Für $k \geq 2\log_2(\ln n)$ und $U(t) < n - \frac{n}{\ln n}$ (also $u(t) < 1 - \frac{1}{\ln n}$) gilt:

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

Pull – Quadratisches Schrumpfen in Erwartung

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft: $u(t)$

Somit:

$$\mathrm{Ex}[U(t+1)] = u(t) \cdot U(t)$$

$$\mathrm{Ex}[u(t+1)] = \mathrm{Ex}[\frac{U(t+1)}{n}] = (u(t))^2$$

Iteriert:

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

Beobachtung

Für $k \geq 2\log_2(\ln n)$ und $U(t) < n - \frac{n}{\ln n}$ (also $u(t) < 1 - \frac{1}{\ln n}$) gilt:

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

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

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

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

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $U(t+1) = \sum_{j=1}^{U(t)} X_j(t)$
- $\Pr[X_j(t) = 1] = u(t)$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $U(t+1) = \sum_{j=1}^{U(t)} X_j(t)$
- $\Pr[X_j(t) = 1] = u(t)$
- Setze $p := u(t)$ und somit $\mu = u(t) \cdot U(t)$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $U(t+1) = \sum_{j=1}^{U(t)} X_j(t)$
- $\Pr[X_j(t) = 1] = u(t)$
- Setze $p := u(t)$ und somit $\mu = u(t) \cdot U(t)$
- Setze $\delta := \frac{1}{(\ln n)^2}$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $U(t+1) = \sum_{j=1}^{U(t)} X_j(t)$
- $\Pr[X_j(t) = 1] = u(t)$
- Setze $p := u(t)$ und somit $\mu = u(t) \cdot U(t)$
- Setze $\delta := \frac{1}{(\ln n)^2}$

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

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $U(t+1) = \sum_{j=1}^{U(t)} X_j(t)$
- $\Pr[X_j(t) = 1] = u(t)$
- Setze $p := u(t)$ und somit $\mu = u(t) \cdot U(t)$
- Setze $\delta := \frac{1}{(\ln n)^2}$

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

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $U(t+1) = \sum_{j=1}^{U(t)} X_j(t)$
- $\Pr[X_j(t) = 1] = u(t)$
- Setze $p := u(t)$ und somit $\mu = u(t) \cdot U(t)$
- Setze $\delta := \frac{1}{(\ln n)^2}$

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

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

Schrumpfungsphase: $U(t) \geq \sqrt{3cn(\ln n)^5}$

Anwendung der Chernoff Bound

- Zufallsvariable für jeden uninformierten Knoten in Runde t

$$X_j(t) := \begin{cases} 1 & \text{falls uninformierter Knoten } j \text{ uninformierten Knoten anruft} \\ 0 & \text{andernfalls} \end{cases}$$

- $U(t+1) = \sum_{j=1}^{U(t)} X_j(t)$
- $\Pr[X_j(t) = 1] = u(t)$
- Setze $p := u(t)$ und somit $\mu = u(t) \cdot U(t)$
- Setze $\delta := \frac{1}{(\ln n)^2}$

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

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

Schrumpfungsphase: $U(t) \geq \sqrt{3cn(\ln n)^5}$

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

Quadratisches Schrumpfen:

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

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

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

Quadratisches Schrumpfen:

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

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Iteriert:

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

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

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

Quadratisches Schrumpfen:

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

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Iteriert:

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

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

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

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

Quadratisches Schrumpfen:

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

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Iteriert:

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

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

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

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

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

Quadratisches Schrumpfen:

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

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Iteriert:

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

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

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

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

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

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

Quadratisches Schrumpfen:

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

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Iteriert:

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

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

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

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

(Wir dürfen annehmen, dass $\sqrt{3cn(\ln n)^5} \geq e^2$)

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

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

Quadratisches Schrumpfen:

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

mit hoher Wahrscheinlichkeit in jeder Runde der Schrumpfungsphase

Iteriert:

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

mit hoher Wahrscheinlichkeit nach k Runden der Schrumpfungsphase

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

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

(Wir dürfen annehmen, dass $\sqrt{3cn(\ln n)^5} \geq e^2$)

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

Wahrscheinlichkeit, dass uninformierter Knoten in $2c + 3$ Runden nur uninformierte Knoten anruft:

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

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

Wahrscheinlichkeit, dass uninformierter Knoten in $2c + 3$ Runden nur uninformierte Knoten anruft:

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

Wahrscheinlichkeit, dass in $2c + 3$ Runden mindestens einer der uninformierten Knoten nur uninformierte Knoten anruft (und damit uninformiert bleibt):

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

Wahrscheinlichkeit, dass uninformierter Knoten in $2c+3$ Runden nur uninformierte Knoten anruft:

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

Wahrscheinlichkeit, dass in $2c+3$ Runden mindestens einer der uninformierten Knoten nur uninformierte Knoten anruft (und damit uninformiert bleibt):

(Union Bound)

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

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

Wahrscheinlichkeit, dass uninformierter Knoten in $2c+3$ Runden nur uninformierte Knoten anruft:

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

Wahrscheinlichkeit, dass in $2c+3$ Runden mindestens einer der uninformierten Knoten nur uninformierte Knoten anruft (und damit uninformiert bleibt):

(Union Bound)

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

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

Wahrscheinlichkeit, dass uninformierter Knoten in $2c+3$ Runden nur uninformierte Knoten anruft:

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

Wahrscheinlichkeit, dass in $2c+3$ Runden mindestens einer der uninformierten Knoten nur uninformierte Knoten anruft (und damit uninformiert bleibt):

(Union Bound)

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

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

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

Wahrscheinlichkeit, dass uninformierter Knoten in $2c+3$ Runden nur uninformierte Knoten anruft:

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

Wahrscheinlichkeit, dass in $2c+3$ Runden mindestens einer der uninformierten Knoten nur uninformierte Knoten anruft (und damit uninformiert bleibt):

(Union Bound)

$$\begin{aligned}
\leq \sqrt{3cn(\ln n)^5} \cdot q^{2c+3} &= \frac{\left(\sqrt{3cn(\ln n)^5}\right)^{2c+4}}{n^{2c+3}} \\
&= \frac{(3c(\ln n)^5)^{c+2} \cdot n^{c+2}}{n^{2c+3}} \quad = \frac{(3c(\ln n)^5)^{c+2}}{n^{c+1}}
\end{aligned}$$

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

Wahrscheinlichkeit, dass uninformierter Knoten uninformierten Knoten anruft:

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

Wahrscheinlichkeit, dass uninformierter Knoten in $2c+3$ Runden nur uninformierte Knoten anruft:

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

Wahrscheinlichkeit, dass in $2c+3$ Runden mindestens einer der uninformierten Knoten nur uninformierte Knoten anruft (und damit uninformiert bleibt):

(Union Bound)

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

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

(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:** $n - \frac{n}{\ln n} > U(t) \geq \sqrt{3cn(\ln n)^5}$
   Dauer: $O(\log \log n)$ Runden
   #Nachrichten: $O(n \log \log n)$
3. **Schluss:** $U(t) < \sqrt{3cn(\ln n)^5}$
   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
  - $(1 - \frac{1}{x})^x \leq \frac{1}{e}$
  - $(1 + \frac{1}{x})^x \leq e$
- Analyse „kippt“ sobald gewisser Anteil an Knoten informiert 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