



Analysis of Algorithms on Growing Networks

--- Can you **correct** all of *an increasing number of coupons*?

*Shuji Kijima (Shiga Univ.)



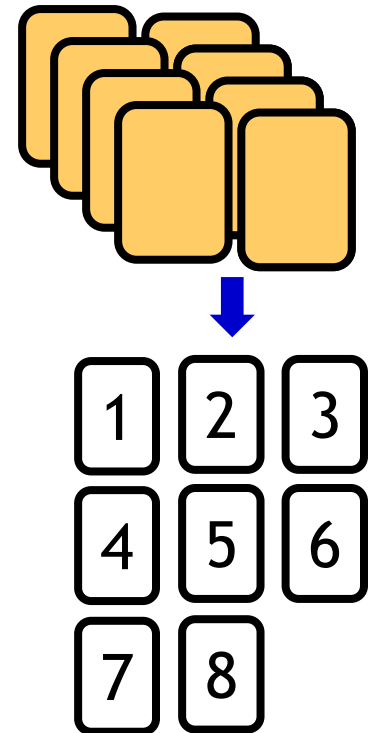
1. Coupon Collector's Problem

Collecting Coupons (a.k.a. Coupon collector's problem)

- ✓ n cards are laid face-down on a desk, where the faces are marked with distinguishing motifs.
- ✓ You will repeat trials, each of which consists of choosing a card, checking its face, putting it back face-down, and mixing cards completely (i.e., sampling with replacement).
- ✓ How many times do you need to repeat the trials to see all n motifs?

Ex.

- P●KEM●N trading cards
- Bikkuriman
- Capsule toys



Collecting Coupons

- ✓ $X_t \in \{1, \dots, n\}$ ($t = 1, 2, \dots$) independent, $\Pr[X_t = k] = \frac{1}{n}$ for $k \in \{1, \dots, n\}$.
- ✓ $T = \min\{t' \mid \{X_1, \dots, X_{t'}\} = \{1, \dots, n\}\}$: **completion time**.

Q. Find $E[T]$.

Collecting Coupons

- ✓ $X_t \in \{1, \dots, n\}$ ($t = 1, 2, \dots$) independent, $\Pr[X_t = k] = \frac{1}{n}$ for $k \in \{1, \dots, n\}$.
- ✓ $T = \min\{t' \mid \{X_1, \dots, X_{t'}\} = \{1, \dots, n\}\}$: **completion time**.

Q. Find $E[T]$.

Let $T_k := \min\{t \mid |\{X_1, \dots, X_t\}| = k\}$ for $k = 1, 2, \dots, n$.

Let $S_k := T_k - T_{k-1}$.

Claim. $E[S_k] = \frac{n}{n-k+1}$.

✓ Prob. on cond. $k-1$: $p_k = \frac{n-(k-1)}{n}$.

✓ Thus $E[S_k] = \frac{1}{p_k} = \frac{n}{n-k+1}$ (Geom. Distr.).

Since $T = T_n = \sum_{k=1}^n S_k$,

$$E[T] = E[T_n] = \sum_{k=1}^n E[S_k] = \sum_{k=1}^n \frac{n}{n-k+1} = n \sum_{k'=1}^n \frac{1}{k'} \approx n(\ln n + 0.577).$$

Collecting Coupons

- ✓ $X_t \in \{1, \dots, n\}$ ($t = 1, 2, \dots$) independent, $\Pr[X_t = k] = \frac{1}{n}$ for $k \in \{1, \dots, n\}$.
- ✓ $T = \min\{t' \mid \{X_1, \dots, X_{t'}\} = \{1, \dots, n\}\}$: **completion time**.

Q. Find $E[T]$.

Q. Find $\Pr[T \geq 2n \ln n]$.

Collecting Coupons

- ✓ $X_t \in \{1, \dots, n\}$ ($t = 1, 2, \dots$) independent, $\Pr[X_t = k] = \frac{1}{n}$ for $k \in \{1, \dots, n\}$.
- ✓ $T = \min\{t' \mid \{X_1, \dots, X_{t'}\} = \{1, \dots, n\}\}$: **completion time**.

Q. Find $E[T]$.

Q. Find $\Pr[T \geq 2n \ln n]$.

By Chebyshev's inequality,

$$\Pr[X \geq 2E[X]] \leq \frac{\text{Var}[X]}{E[X]^2} \leq \frac{\frac{n^2 \pi^2}{6}}{(n \ln n)^2} = \frac{\pi^2}{6(\ln n)^2}$$

≈ 0.078 (for $n = 100$)

Improved by Hoeffding ineq., Chernoff ineq., etc.

an increasing number of

2. Collecting coupons

Collecting *an increasing number of* coupons

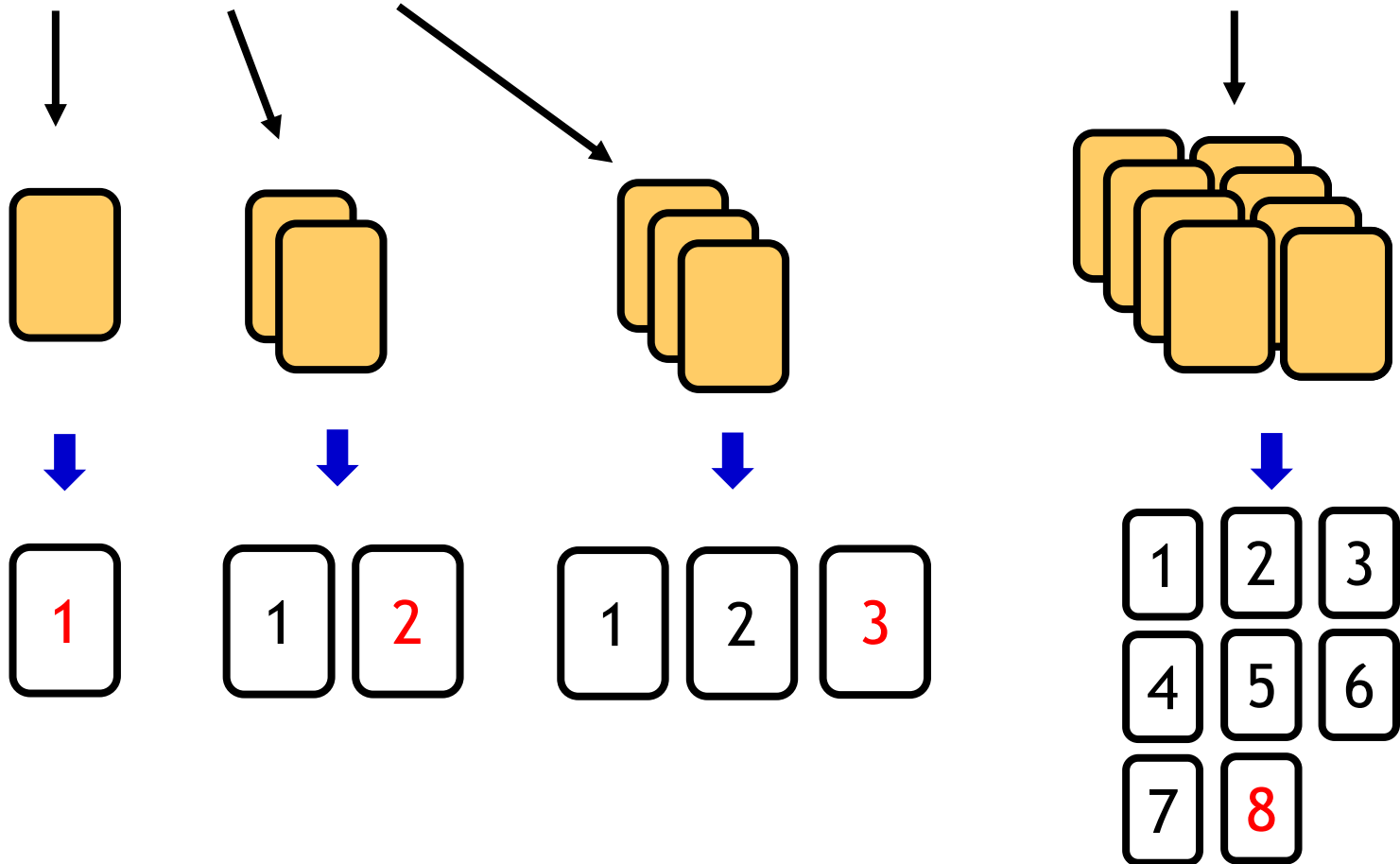
Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**

“complete”
= collect all types.

Collecting *an increasing number of* coupons

Day (t)	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9
# types	1	2	3	4	5	6	7	8	9
$\Pr[X_t = k]$	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{9}$



Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**

Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**
 - **Impossible** (once you draw a duplication).

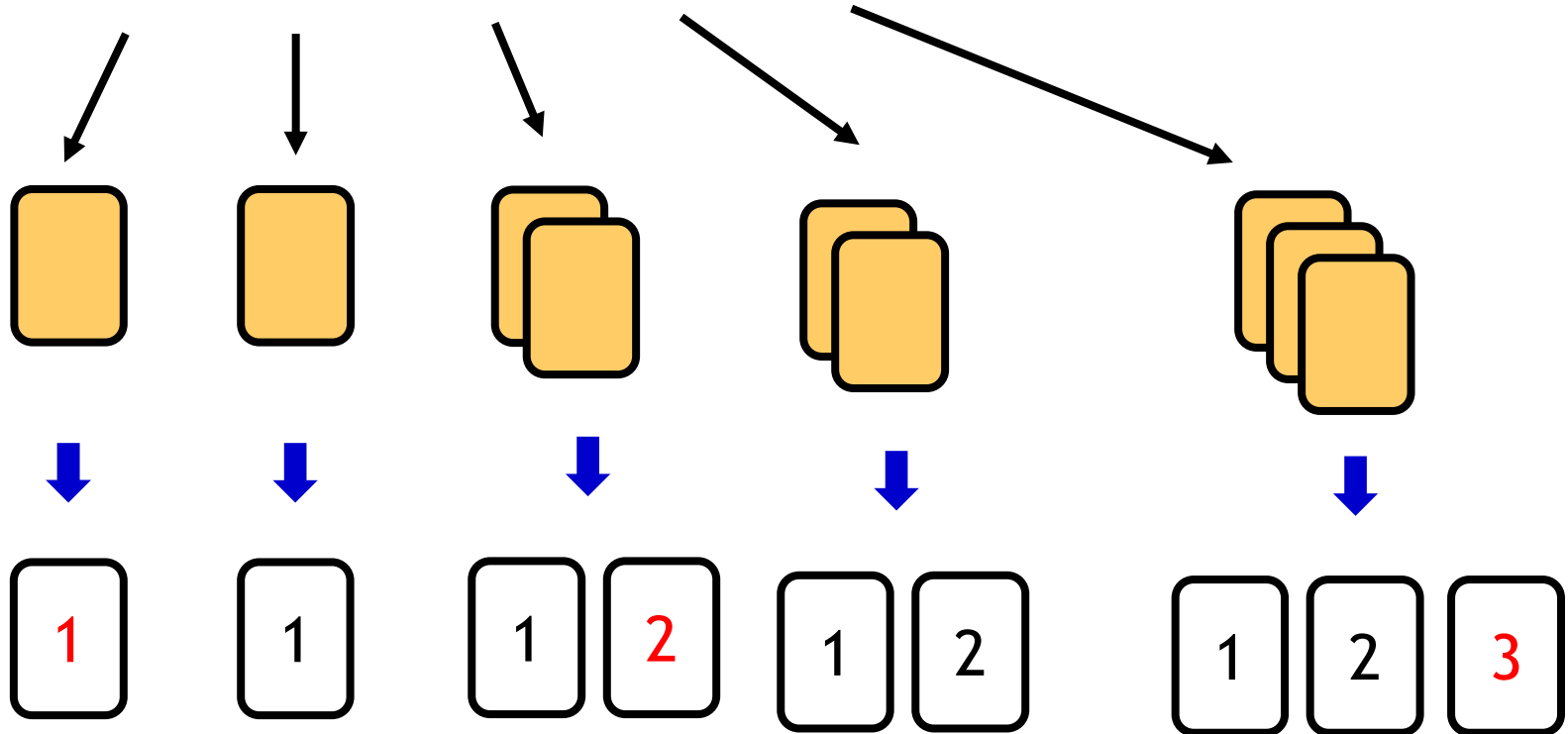
Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**
 - **Impossible** (once you draw a duplication).
- A new type is released **every two days**. I.e., you draw from $\left\lceil \frac{d}{2} \right\rceil$ on the d^{th} day. **When can you complete?**

Collecting *an increasing number of* coupons

Day	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9
# types	1	1	2	2	3	3	4	4	5
$\Pr[X_t = k]$	$\frac{1}{1}$	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{5}$



Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**
 - **Impossible** (once you draw a duplication).
- A new type is released **every two days**. I.e., you draw from $\left\lceil \frac{d}{2} \right\rceil$ on the d^{th} day. **When can you complete?**

Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**
 - **Impossible** (once you draw a duplication).
- A new type is released **every two days**. I.e., you draw from $\left\lceil \frac{d}{2} \right\rceil$ on the d^{th} day. **When can you complete?**
 - **Impossible**, maybe.

Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**
 - **Impossible** (once you draw a duplication).
- A new type is released **every two days**. I.e., you draw from $\left\lfloor \frac{d}{2} \right\rfloor$ on the d^{th} day. **When can you complete?**
 - **Impossible**, maybe.
- A new type is released **every 10 days**. I.e., you draw from $\left\lfloor \frac{d}{10} \right\rfloor$ on the d^{th} day. **When can you complete?**

Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**
 - **Impossible** (once you draw a duplication).
- A new type is released **every two days**. I.e., you draw from $\left\lceil \frac{d}{2} \right\rceil$ on the d^{th} day. **When can you complete?**
 - **Impossible**, maybe.
- A new type is released **every 10 days**. I.e., you draw from $\left\lceil \frac{d}{10} \right\rceil$ on the d^{th} day. **When can you complete?**
 - **Impossible**. (Consider the 1000th day. Can you draw the new one from 100 types in 10 days?)

Collecting *an increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- A new type of coupons is released **everyday**. I.e., you draw a coupon from d types on the d^{th} day. **When can you complete?**
 - **Impossible** (once you draw a duplication).
- A new type is released **every two days**. I.e., you draw from $\left\lfloor \frac{d}{2} \right\rfloor$ on the d^{th} day. **When can you complete?**
 - **Impossible**, maybe.
- A new type is released **every 10 days**. I.e., you draw from $\left\lfloor \frac{d}{10} \right\rfloor$ on the d^{th} day. **When can you complete?**
 - **Impossible**. (Consider the 1000th day. Can you draw the new one from 100 types in 100 days?)

Not very fun...

Collecting *a* moderately *increasing number of* coupons

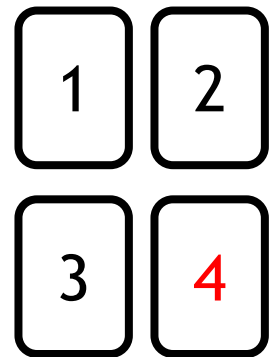
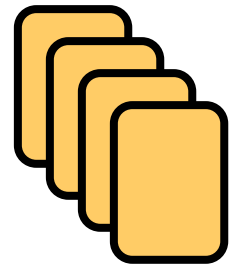
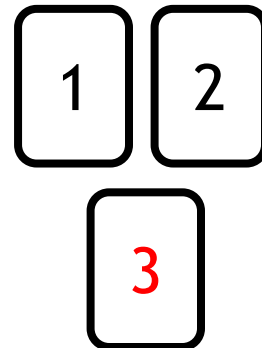
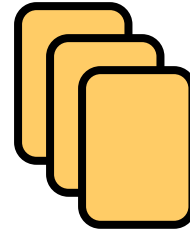
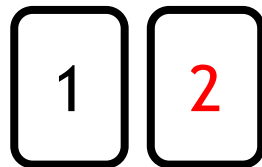
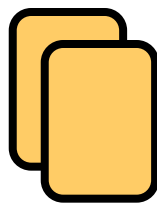
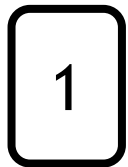
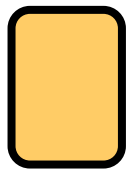
Draw a coupon everyday, u.a.r. from a finite #types.

- The n^{th} type of coupon is released n days after $n - 1^{\text{th}}$ was released. **When can you complete?**

Collecting a moderately increasing number of coupons

Day	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9
# types	1	2	2	3	3	3	4	4	4
$\Pr[X_t = k]$	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

1st period 2nd period 3rd period 4th period



Collecting *a* moderately *increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- The n^{th} type of coupon is released n days after $n - 1^{\text{th}}$ was released. **When can you complete?**

Collecting *a* moderately *increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- The n^{th} type of coupon is released n days after $n - 1^{\text{th}}$ was released. **When can you complete?**
 - **Impossible.** (because a new type appears forever!)

Collecting *a moderately increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- The n^{th} type of coupon is released n days after $n - 1^{\text{th}}$ was released. **When can you complete?**
 - **Impossible.** (because a new type appears forever!)
 - Q. How many types are collected in the end of n^{th} period?

Collecting *a moderately increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- The n^{th} type of coupon is released n days after $n - 1^{\text{th}}$ was released. **When can you complete?**

➤ **Impossible.** (because a new type appears forever!)

➤ Q. How many types are collected in the end of n^{th} period?

1. $o(n)$

2. cn

3. $n - c\sqrt{n}$

4. $n - c \log n$

5. $n - c$



c is some constant.

Collecting *a* moderately *increasing number of* coupons

Draw a coupon everyday, u.a.r. from a finite #types.

- The n^{th} type of coupon is released n days after $n - 1^{\text{th}}$ was released. **When can you complete?**

➤ **Impossible.** (because a new type appears forever!)

➤ Q. How many types are collected in the end of n^{th} period?

1. $o(n)$

2. cn

3. $n - c\sqrt{n}$

4. $n - c \log n$

5. $n - c$ (right answer)



c is some constant.

Collecting *a moderately increasing number*

Draw a coupon everyday
 $\delta(n)$: #days of the n^{th} period
 U_n : #items uncollected
 in the end of n^{th} period

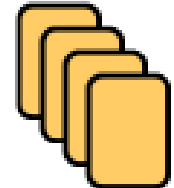
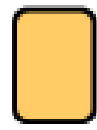
Prop.

If $\delta(n) = n$ then $E[U_n] < 1$.

Collecting *a moderately increasing number of coupons*

Day	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9
# types	1	2	2	3	3	3	4	4	4
$\Pr[X_i = k]$	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

1st period 2nd period 3rd period 4th period



Collecting a moderately increasing number

Draw a coupon everyday
 $\delta(n)$: #days of the n^{th} period
 U_n : #items uncollected
in the end of n^{th} period

Prop.

If $\delta(n) = n$ then $E[U_n] < 1$.

Proof.

✓ $\mathcal{E}_{i,n} := \begin{cases} 1 & \text{(item } i \text{ is uncollected in the end of the } n^{\text{th}} \text{ period)} \\ 0 & \text{(item } i \text{ is collected by the end of the } n^{\text{th}} \text{ period)} \end{cases}$

for $i = 1, 2, \dots, n$.

✓ $U_n = \sum_{i=1}^n \mathcal{E}_{i,n}$

✓ Prob. that item n is uncollected in the end of the n^{th} period:

$$\Pr[\mathcal{E}_{n,n} = 1] = \left(1 - \frac{1}{n}\right)^n < e^{-1}$$

✓ Prob. that item i ($i \leq n$) is uncollected in the end of the n^{th} period:

$$\Pr[\mathcal{E}_{i,n} = 1] = \left(1 - \frac{1}{i}\right)^i \left(1 - \frac{1}{i+1}\right)^{i+1} \dots \left(1 - \frac{1}{n}\right)^n < \left(\frac{1}{e}\right)^{n+1-i}$$

✓ $E[U_n] = \sum_{i=1}^n \Pr[\mathcal{E}_{i,n}] < \sum_{i=1}^n \left(\frac{1}{e}\right)^{n+1-i} = \frac{1}{e} + \frac{1}{e^2} + \dots + \frac{1}{e^n} < \frac{\frac{1}{e}}{1 - \frac{1}{e}} = \frac{1}{e-1} < 0.582$.

Collecting *a moderately increasing number*

Prop.

If $\delta(n) = \frac{n}{2}$ then $E[U_n] < ???$.

Draw a coupon everyday

$\delta(n)$: #days of the n^{th} period

U_n : #items uncollected

in the end of n^{th} period

Collecting *a moderately increasing number*

Draw a coupon everyday

$\delta(n)$: #days of the n^{th} period

U_n : #items uncollected

in the end of n^{th} period

Prop.

If $\delta(n) = \frac{n}{2}$ then $E[U_n] < 2$.

Proof.

✓ $\mathcal{E}_{i,n} := \begin{cases} 1 & \text{(item } i \text{ is uncollected in the end of the } n^{\text{th}} \text{ period)} \\ 0 & \text{(item } i \text{ is collected by the end of the } n^{\text{th}} \text{ period)} \end{cases}$

for $i = 1, 2, \dots, n$.

✓ $U_n = \sum_{i=1}^n \mathcal{E}_{i,n}$

✓ Prob. that item n is uncollected in the end of the n^{th} period:

$$\Pr[\mathcal{E}_{n,n} = 1] = \left(1 - \frac{1}{n}\right)^{\frac{n}{2}} < e^{-\frac{1}{2}}$$

✓ Prob. that item i ($i \leq n$) is uncollected in the end of the n^{th} period:

$$\Pr[\mathcal{E}_{i,n} = 1] = \left(1 - \frac{1}{i}\right)^{\frac{i}{2}} \left(1 - \frac{1}{i+1}\right)^{\frac{i+1}{2}} \dots \left(1 - \frac{1}{n}\right)^{\frac{n}{2}} < \left(\frac{1}{e}\right)^{\frac{n+1-i}{2}}$$

✓ $E[U_n] = \sum_{i=1}^n \Pr[\mathcal{E}_{i,n}] < \sum_{i=1}^n \left(\frac{1}{e}\right)^{\frac{n+1-i}{2}} = \frac{1}{e^{\frac{1}{2}}} + \frac{1}{e^{\frac{1}{2}}} + \dots + \frac{1}{e^{\frac{1}{2}}} < \frac{\frac{1}{e^{\frac{1}{2}}}}{1 - \frac{1}{e}} = \frac{1}{e^{\frac{1}{2}} - 1} < 1.542$.

Collecting a moderately increasing number

Draw a coupon everyday

$d(n)$: #days of the n^{th} period

U_n : #items uncollected

in the end of n^{th} period

Prop.

If $d(n) = cn$ then $E[U_n] < \frac{1}{e^c - 1}$.

Proof.

✓ Prob. that item i ($i \leq n$) is uncollected in the end of the n^{th} period:

$$\Pr[\mathcal{E}_{i,n} = 1] = \left(1 - \frac{1}{i}\right)^{ci} \left(1 - \frac{1}{i+1}\right)^{c(i+1)} \dots \left(1 - \frac{1}{n}\right)^{cn} \leq \left(\frac{1}{e}\right)^{c(n+1-i)}$$

$$\begin{aligned} \checkmark E[U_n] &= \sum_{i=1}^n \Pr[\mathcal{E}_{i,n}] \leq \sum_{i=1}^n \left(\frac{1}{e}\right)^{c(n+1-i)} = \frac{1}{e^c} + \frac{1}{e^{2c}} + \dots + \frac{1}{e^{nc}} < \frac{\frac{1}{e^c}}{1 - \frac{1}{e^c}} = \frac{1}{e^c - 1} \end{aligned}$$

Collecting a moderately increasing number

Draw a coupon everyday

$\delta(n)$: #days of the n^{th} period

U_n : #items uncollected

in the end of n^{th} period

Prop. (lower bound)

If $\delta(n) = cn$ then $E[U_n] \geq \frac{n}{cn+1} \left(1 - \frac{1}{n}\right)^{cn} \approx \frac{1}{ce^c}$.

Proof.

✓ $S(n) := \sum_{k=1}^n \prod_{i=k}^n \left(1 - \frac{1}{i}\right)^{ci} \quad (= E[U_n])$

✓ Proposition: $S(n) \geq \frac{n}{cn+1} \left(1 - \frac{1}{n}\right)^{cn}$.

- Claim 1: $S(n+1) = \left(1 - \frac{1}{n+1}\right)^{c(n+1)} (S(n) + 1)$.
- Claim 2: $S(n) \geq \frac{n}{cn+1} \left(1 - \frac{1}{n}\right)^{cn} \Rightarrow S(n) + 1 \geq \frac{n+1}{c(n+1)+1}$.
- Inductively,

$$\begin{aligned} S(n+1) &= \left(1 - \frac{1}{n+1}\right)^{c(n+1)} (S(n) + 1) \\ &\geq \left(1 - \frac{1}{n+1}\right)^{c(n+1)} \frac{n+1}{c(n+1)+1} \end{aligned}$$

Collecting a moderately increasing number

Draw a coupon everyday

$\delta(n)$: #days of the n^{th} period

U_n : #items uncollected

in the end of n^{th} period

Prop. (lower bound)

$$\text{If } \delta(n) = cn \text{ then } E[U_n] \geq \frac{n}{cn+1} \left(1 - \frac{1}{n}\right)^{cn} \simeq \frac{1}{ce^c}.$$

Proof.

$$\checkmark S(n) := \sum_{k=1}^n \prod_{i=k}^n \left(1 - \frac{1}{i}\right)^{ci} \quad (= E[U_n])$$

$$\checkmark \text{ Proposition: } S(n) \geq \frac{n}{cn+1} \left(1 - \frac{1}{n}\right)^{cn}.$$

- Claim 1: $S(n+1) = \left(1 - \frac{1}{n+1}\right)^{c(n+1)} (S(n) + 1)$.
- Claim 2: $S(n) \geq \frac{n}{cn+1} \left(1 - \frac{1}{n}\right)^{cn} \Rightarrow S(n) + 1 \geq \frac{n+1}{c(n+1)+1}$.
- Inductively,

$$\begin{aligned} S(n+1) &= \left(1 - \frac{1}{n+1}\right)^{c(n+1)} (S(n) + 1) \\ &\geq \left(1 - \frac{1}{n+1}\right)^{c(n+1)} \frac{n+1}{c(n+1)+1} \end{aligned}$$

Collecting *a moderately increasing number of* coupons

Lemma (about S_n)

$$\delta: \mathbb{N} \rightarrow \mathbb{N}, \quad S(n) := \sum_{k=1}^n \prod_{i=k}^n \left(1 - \frac{1}{i}\right)^{\delta(i)}$$

(i) If $\delta(i) \geq ci$ then $S(n) = O(1)$.

(ii) If δ is monotone nonincreasing ($\delta(i) \leq \delta(i+1)$) then

$$S(n) \geq \frac{n}{\delta(n)+1} \left(1 - \frac{1}{n}\right)^{\delta(n)}.$$

(iii) If δ is sublinear ($\frac{\delta(i)}{i} \geq \frac{\delta(i+1)}{i+1}$) then $S(n) \leq \frac{n}{\delta(n)}$.

(iv) If $\delta(i) = c$ (const.) then $S(n) \leq \frac{n}{c+1}$.

Collecting *an increasing number of* coupons

Draw a coupon everyday
 $\delta(n)$: #days of the n^{th} period
 U_n : #items uncollected
in the end of n^{th} period

Thm.

$\delta: \mathbb{N} \rightarrow \mathbb{N}$.

(i) If $\delta(i) \geq ci$ then $E[U_n] = O(1)$.

- If $\frac{\delta(i)}{i} \xrightarrow{i \rightarrow \infty} \infty$ then $E[U_n] = 0$.

(ii) If δ is monotone nondecreasing, unbounded and sublinear ($\frac{\delta(i)}{i} \geq \frac{\delta(i+1)}{i+1}$) then

$$E[U_n] = (1 - o(1)) \frac{n}{\delta(n)+1}.$$

(iii) If $\delta(i) = c$ (const.) then

$$E[U_n] = \left(1 - O\left(\frac{1}{n}\right)\right) \frac{n}{c+1}.$$

If $\delta(i) = o(i)$ then $E[U_n] \xrightarrow{n \rightarrow \infty} \infty$.
(e.g., $\delta(i) = \lfloor \sqrt{i} \rfloor$,
 $\delta(i) = \lfloor \log i \rfloor$ etc.)

From **collecting coupons**
to **random walks** on finite graphs

How many vertices does a **random walk**
miss in a network *with a moderately
increasing number of vertices?*

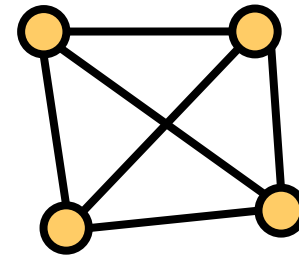
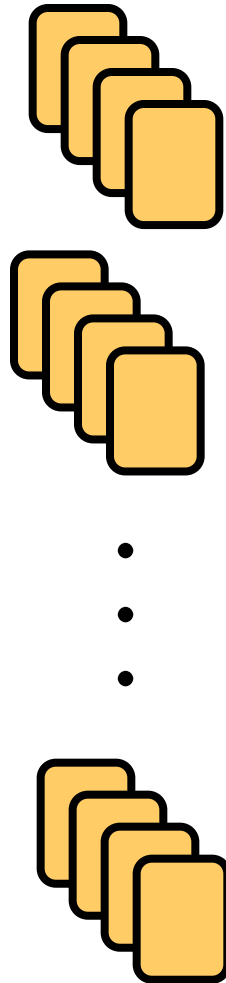
***Shuji Kijima** (Shiga Univ.)

Nobutaka Shimizu (Tokyo Tech)

Takeharu Shiraga (Chuo Univ.)

in Proc. **SODA 2021**
(open access)

(Classic) collecting coupons and RW on a complete graph



completion time of **coupons** = **cover time** of **RW**
on a complete graph

Random walks on growing graphs

RWoGG: $(\delta, (G^{(i)})_{i=1}^{\infty}, (P^{(i)})_{i=1}^{\infty})$

Thm. (general upper bound)

If $\delta(i) \geq ct_{\text{hit}}(i)$ ($c \geq 1$) then $E[U] = O(1)$.

Particularly, if $\frac{\delta(i)}{t_{\text{hit}}(i)} \xrightarrow{i \rightarrow \infty} \infty$ then $E[U_n] \xrightarrow{n \rightarrow \infty} 0$.

Thm. (upper bound for lazy and reversible walk)

Suppose $P^{(i)}$ is lazy and reversible.

If $\frac{t_{\text{hit}}(i)}{t_{\text{mix}}(i)} \geq \frac{i^\gamma}{c}$ and $\delta(i) \geq \frac{3ct_{\text{hit}}(i)}{i^\gamma}$ ($c > 0$) then $E[U_n] \leq \frac{8n^\gamma}{c} + 32$.

S. Kijima, N. Shimizu, T. Shiraga, How many vertices does a random walk miss in a network with moderately increasing the number of vertices?, in Proc. SODA 2021, 106–122.

Summary

- ✓ **Collecting an increasing number of coupons.**
- ✓ Cover times of random walks on growing networks.
 - $E[U_n]$ with respect to Duration δ .

Future works

- ✓ **Does $E[U_n] = O(1)$ require $\delta(i) = \Omega(t_{\text{hit}}(i))$?**
- ✓ Other than “cover time”?
 - Coming soon.
- ✓ From **random walks** to **optimization**:



Welcome!

□ **“Optimization in Dynamic Environment”**

Cf. calculus of variations

‘Many traditional approaches and measures for static networks are not adequate for dynamic networks. There is already strong evidence that there is room for the development of a rich theory.’

--- Othon Michail & Paul G. Spirakis CACM 2018

Fin.



The end

Thank you for the attention.