Sending Timely Status Updates through Channel with Random Delay via Online Learning

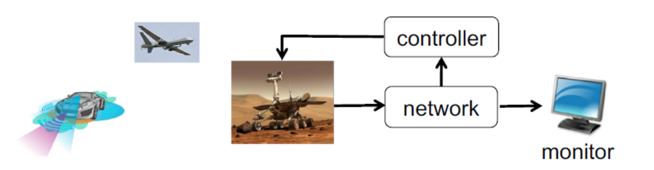
Haoyue Tang, Yuchao Chen, Jintao Wang, Jingzhou Sun, Jian Song

Department of Electronic Engineering, Tsinghua University

Contents

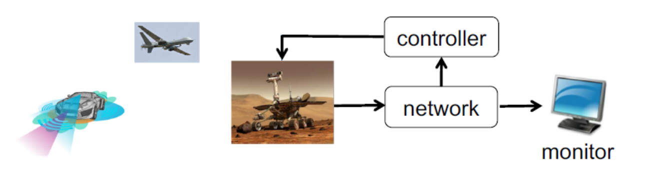
Background

Real-time service requires fresh data



Background

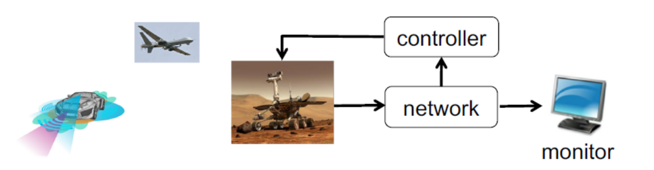
Real-time service requires fresh data



Real-time data analytics becomes important

Background

Real-time service requires fresh data

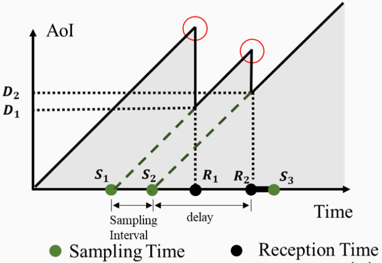


Real-time data analytics becomes important

Research Problem

How to measure and optimize information freshness?

Freshness Metric: Age of Information

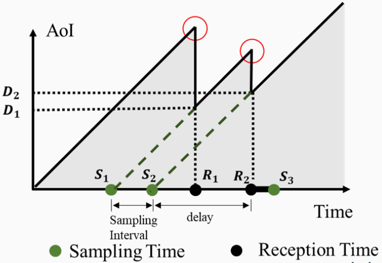


- By definition, the AoI at time t, denoted by $A(t)$

$$A(t) \triangleq t - S_{i(t)}, \tag{1}$$

where $i(t) := \arg\max\{i | R_i \leq t\}$ is the index of the recently received sample.

Freshness Metric: Age of Information

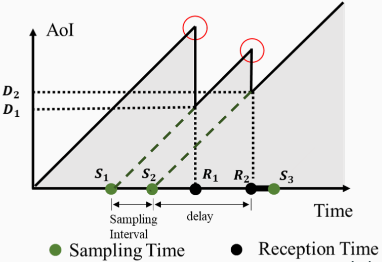


- By definition, the AoI at time t, denoted by $A(t)$

$$A(t) \triangleq t - S_{i(t)}, \quad (1)$$

  where $i(t) := \arg\max\{i | R_i \leq t\}$ is the index of the recently received sample.
- No sample received: grows linearly
- New sample received: drops to packet delay

Freshness Metric: Age of Information



- By definition, the AoI at time t, denoted by $A(t)$

$$A(t) \triangleq t - S_{i(t)}, \tag{1}$$

where $i(t) := \arg\max\{i | R_i \le t\}$ is the index of the recently received sample.

- No sample received: grows linearly
- New sample received: drops to packet delay

Previous Work

Minimizing AoI is different from max throughput/min delay

Challenges

Goal: min AoI under unknown communication statistics

Existing work

- Unknown AoI penalty functions Tripathi and Modiano (2021)
- Utility/delay optimization under an AoI constraint Li (2021)
- min AoI using RL/Bandits (Atay et al. (2021); Leng and Yener (2021), ...) Theoretic analysis are missing.

Challenges

Goal: min AoI under unknown communication statistics

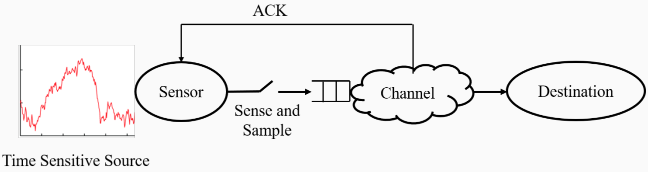
Existing work

- Unknown AoI penalty functions Tripathi and Modiano (2021)
- Utility/delay optimization under an AoI constraint Li (2021)
- min AoI using RL/Bandits (Atay et al. (2021); Leng and Yener (2021), ...) Theoretic analysis are missing.

Contributions

- Reformulate AoI minimization problem as a Renewal-Reward Process, then propose an online algorithm.
- Derive the convergence rate of the proposed online algorithm.
- Establish converse result for any causal sampling algorithm. Verify minimax order optimal.

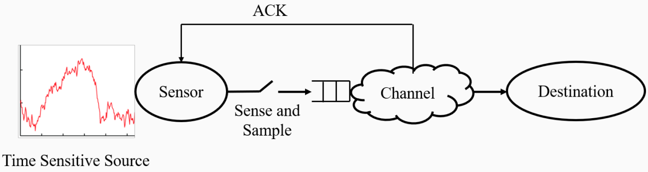
System Model



- Point-to-point link: **Sensor** senses and submits samples to the **Destination**
- **Channel**: FIFO queue with i.i.d. transmission times
- **Feedback**: zero-delay ACK
    - Busy/Idle state of the channel is known to the sensor.

The transmission delay D_k of sample k in the channel is i.i.d following distribution $\mathbb{P}_D$.

System Model



- Point-to-point link: **Sensor** senses and submits samples to the **Destination**
- **Channel**: FIFO queue with i.i.d. transmission times
- **Feedback**: zero-delay ACK
  - Busy/Idle state of the channel is known to the sensor.

The transmission delay D_k of sample k in the channel is i.i.d following distribution $\mathbb{P}_D$.

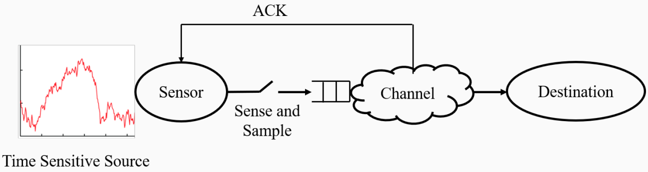
Assumption 1

$\mathbb{P}_D$ is absolutely continuous and is first and second order bounded, i.e.,

$$\overline{D}_{\text{lb}} \leq \overline{D} := \mathbb{E}[D] \leq \overline{D}_{\text{ub}}, M_{\text{lb}} \leq \mathbb{E}[D^2] \leq M_{\text{ub}}. \quad (2)$$

Problem Formulation

Minimizing the AoI under a sampling frequency constraint



Optimization Problem

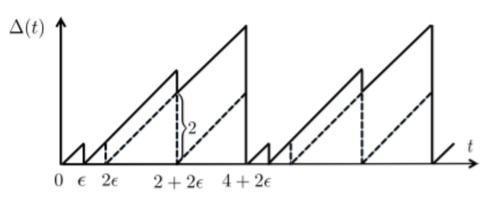
$$\inf_{\pi \in \Pi} \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}\left[\int_{t=0}^{T} A(t) \mathrm{d}t\right], \tag{3a}$$

$$\text{s.t.} \quad \liminf_{K \to \infty} \frac{1}{K} \mathbb{E}\left[\sum_{k=1}^{K} (S_{k+1} - S_k)\right] \geq \frac{1}{f_{\max}}. \tag{3b}$$

Observation: Samples waiting in the queues are not longer fresh.

Solution: Focus on policy that waits for $W_k \geq 0$ to submit sample $(k+1)$ after the ACK of the k-th sample is received.

A Counter-Intuitive Example



Suppose the transmission delay sequence is $\{0, 0, 2, 2, 0, 0, 2, 2, \cdots\}$.

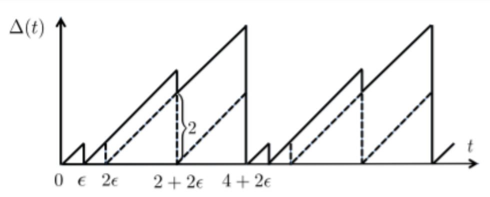
Sampling policy:

- Delay=0, wait for ϵ to take the next sample
- Delay=2, take the next sample immediately

$$\overline{A} = \frac{8 + 2\epsilon + \epsilon^2}{4 + 2\epsilon}. \tag{4}$$

Zero Wait $\epsilon = 0$, $\overline{A} = 2$; $\epsilon = 1$, $\overline{A} = 11/6$.

A Counter-Intuitive Example



Suppose the transmission delay sequence is $\{0, 0, 2, 2, 0, 0, 2, 2, \cdots\}$.
Sampling policy:

- Delay=0, wait for ϵ to take the next sample
- Delay=2, take the next sample immediately

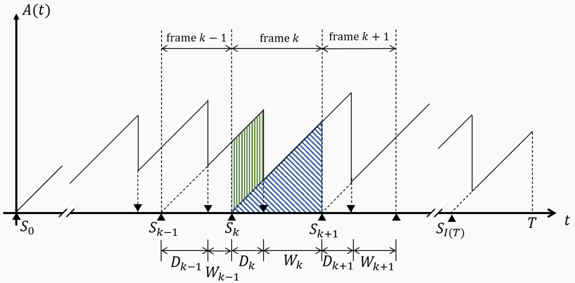
$$\overline{A} = \frac{8 + 2\epsilon + \epsilon^2}{4 + 2\epsilon}. \tag{4}$$

Zero Wait $\epsilon = 0$, $\overline{A} = 2$; $\epsilon = 1$, $\overline{A} = 11/6$.

Take away message

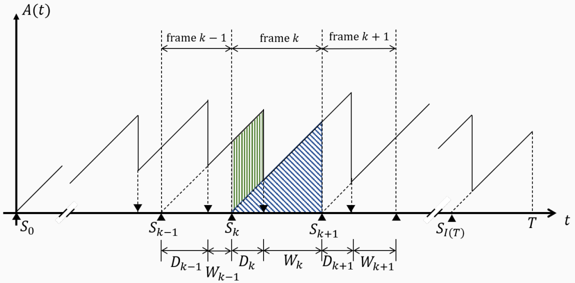
- Zero-wait is not AoI minimum.
- When delay is zero, the new sample taken is wasted!

Problem Resolution



- Frame length k: $L_k = D_k + W_k$.
- Cumulative Age in frame k: $R_k = \frac{1}{2}(D_k + W_k)^2 + D_k(D_{k-1} + W_{k-1})$.

Problem Resolution

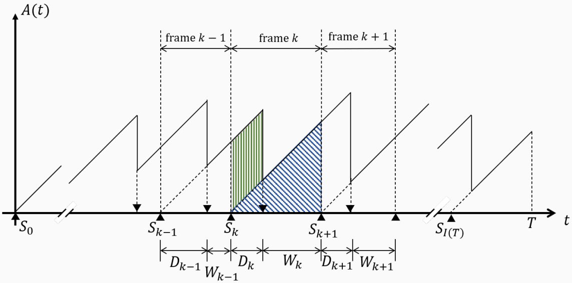


- Frame length k: $L_k = D_k + W_k$.
- Cumulative Age in frame k: $R_k = \frac{1}{2}(D_k + W_k)^2 + D_k(D_{k-1} + W_{k-1})$.

Delay D_k is i.i.d, for stationary policy π:

$$\overline{A}_\pi = \limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^{K} R_k\right]}{\mathbb{E}\left[\sum_{k=1}^{K} L_k\right]} = \limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^{K} \frac{1}{2}(D_k + W_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^{K} L_k\right]} + \overline{D}. \quad (5)$$

Problem Resolution



- Frame length k: $L_k = D_k + W_k$.
- Cumulative Age in frame k: $R_k = \frac{1}{2}(D_k + W_k)^2 + D_k(D_{k-1} + W_{k-1})$.

Delay D_k is i.i.d, for stationary policy π:

$$\overline{A}_\pi = \limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^K R_k\right]}{\mathbb{E}\left[\sum_{k=1}^K L_k\right]} = \limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^K \frac{1}{2}(D_k + W_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^K L_k\right]} + \overline{D}. \quad (5)$$

$Q_k := \frac{1}{2}(W_k + D_k)^2$ and L_k are i.i.d $\Rightarrow$

Renewal-Reward Process Optimization

Algorithm Design (1): Offline Policy

Assuming $\mathbb{P}_D$ is known, computing $\pi^\star$:

$$\gamma^\star := \min_{\pi} \limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^{K} \frac{1}{2}(D_k + W_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^{K}(D_k + W_k)\right]}, \text{s.t., } \mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K}(D_k + W_k)\right] \geq \frac{1}{f_{\max}}.$$

- A Constrained Markov Process (CMDP) in a continuous space
  - State: observed delay $D_k \in \mathbb{R}$ / Action: select $W_k \in \mathbb{R}$.

Algorithm Design (1): Offline Policy

Assuming $\mathbb{P}_D$ is known, computing $\pi^\star$:

$$\gamma^\star := \min_{\pi} \limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^{K} \frac{1}{2}(D_k + W_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^{K}(D_k + W_k)\right]}, \text{s.t., } \mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K}(D_k + W_k)\right] \geq \frac{1}{f_{\max}}.$$

- A Constrained Markov Process (CMDP) in a continuous space
  - State: observed delay $D_k \in \mathbb{R}$ / Action: select $W_k \in \mathbb{R}$.

Property 1: $\forall \pi$ that satisfies the frequency constraint (Π_{cons}),

$$\limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^{K} \frac{1}{2}(D_k + W_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^{K}(D_k + W_k)\right]} \geq \gamma^\star.$$

- If $\gamma^\star$ is known

$$\min_{\pi} \left(\frac{1}{2}\mathbb{E}\left[(D + \pi(D))^2\right] - \gamma^\star \mathbb{E}\left[(D + \pi(D))\right]\right), \text{s.t.} \mathbb{E}[D + \pi(D)] \geq \frac{1}{f_{\max}}. \quad (6)$$

Algorithm Design (1): Offline Policy

Assuming $\mathbb{P}_D$ is known, computing $\pi^\star$:

$$\gamma^\star := \min_{\pi} \limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^{K} \frac{1}{2}(D_k + W_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^{K}(D_k + W_k)\right]}, \text{s.t., } \mathbb{E}\left[\frac{1}{K}\sum_{k=1}^{K}(D_k + W_k)\right] \geq \frac{1}{f_{\max}}.$$

- A Constrained Markov Process (CMDP) in a continuous space
  - State: observed delay $D_k \in \mathbb{R}$ / Action: select $W_k \in \mathbb{R}$.

Property 1: $\forall \pi$ that satisfies the frequency constraint (Π_{cons}),

$$\limsup_{K\to\infty} \frac{\mathbb{E}\left[\sum_{k=1}^{K} \frac{1}{2}(D_k + W_k)^2\right]}{\mathbb{E}\left[\sum_{k=1}^{K}(D_k + W_k)\right]} \geq \gamma^\star.$$

- If $\gamma^\star$ is known

$$\min_{\pi} \left(\frac{1}{2}\mathbb{E}\left[(D + \pi(D))^2\right] - \gamma^\star \mathbb{E}\left[(D + \pi(D))\right]\right), \text{s.t.} \mathbb{E}[D + \pi(D)] \geq \frac{1}{f_{\max}}. \quad (6)$$

Property 2: For given γ, equation (6) can be solved using Lagrange
$\mathcal{L}(\pi, \gamma, \nu) := \frac{1}{2}\mathbb{E}\left[(D + \pi(D))^2\right] - (\gamma + \nu)\mathbb{E}[(D + \pi(D))] + \nu\frac{1}{f_{\max}}$.

$$\pi^\star_{\gamma,\nu}(d) = (\gamma + \nu - d)^+ . \quad (7)$$

Offline Design: Compute $\gamma^\star + \nu^\star$ via Bi-section search Sun et al. (2017)

Algorithm Design (2): Online Policy

How to obtain $\gamma^\star, \nu^\star$ when $\mathbb{P}_D$ is unknown?

- $\gamma^\star = \frac{\mathbb{E}\left[\frac{1}{2}(D+\pi(D))^2\right]}{\mathbb{E}[D+\pi(D)]} = \arg\min_\gamma (\mathbb{E}[Q] - \gamma\mathbb{E}[L])^2$: Robbins-Monro (SGD) Neely (2021)
- Dual optimizer $\nu^\star$: Virtual queue to satisfy sampling frequency constraint.

In frame k, observe transmission delay D_k and then:

- **Generate Sample/Batch k:** Wait for $W_k = (\gamma_k + \nu_k - D_k)^+$ to take sample $k+1$. The reward/length of frame k:
$$Q_k = \frac{1}{2}(D_k + W_k)^2, L_k = D_k + W_k. \tag{8}$$

Algorithm Design (2): Online Policy

How to obtain $\gamma^\star, \nu^\star$ when $\mathbb{P}_D$ is unknown?

- $\gamma^\star = \frac{\mathbb{E}\left[\frac{1}{2}(D+\pi(D))^2\right]}{\mathbb{E}[D+\pi(D)]} = \arg\min_\gamma (\mathbb{E}[Q] - \gamma \mathbb{E}[L])^2$: Robbins-Monro (SGD) Neely (2021)
- Dual optimizer $\nu^\star$: Virtual queue to satisfy sampling frequency constraint.

In frame k, observe transmission delay D_k and then:

- **Generate Sample/Batch k:** Wait for $W_k = (\gamma_k + \nu_k - D_k)^+$ to take sample $k+1$. The reward/length of frame k:
$$Q_k = \frac{1}{2}(D_k + W_k)^2, L_k = D_k + W_k. \tag{8}$$
- **Approximate γ_k via SGD:** Goal: $\min_\gamma \left(\mathbb{E}[Q] - \gamma \mathbb{E}[L]\right)^2$, $\partial\gamma = -L_k(Q_k - \gamma L_k)$.
$$\gamma_{k+1} = \left[\gamma_k + \eta_k (Q_k - \gamma_k L_k)\right]_{\gamma_{\text{lb}}}^{\gamma_{\text{ub}}}. \tag{9}$$

($\eta_k = \frac{1}{k\bar{D}_{\text{lb}}}$, Recall SGD in ML, step-sizes should be diminishing, γ_{ub} and γ_{lb} can be estimated from transmission delay D upper and lower bounds.)

Algorithm Design (2): Online Policy

How to obtain $\gamma^\star, \nu^\star$ when $\mathbb{P}_D$ is unknown?

- $\gamma^\star = \frac{\mathbb{E}\left[\frac{1}{2}(D+\pi(D))^2\right]}{\mathbb{E}[D+\pi(D)]} = \arg\min_\gamma (\mathbb{E}[Q] - \gamma\mathbb{E}[L])^2$: Robbins-Monro (SGD) Neely (2021)
- Dual optimizer $\nu^\star$: Virtual queue to satisfy sampling frequency constraint.

In frame k, observe transmission delay D_k and then:

- **Generate Sample/Batch k:** Wait for $W_k = (\gamma_k + \nu_k - D_k)^+$ to take sample $k+1$. The reward/length of frame k:
$$Q_k = \frac{1}{2}(D_k + W_k)^2, L_k = D_k + W_k. \quad (8)$$
- **Approximate γ_k via SGD:** Goal: $\min_\gamma (\mathbb{E}[Q] - \gamma\mathbb{E}[L])^2$, $\partial\gamma = -L_k(Q_k - \gamma L_k)$.
$$\gamma_{k+1} = [\gamma_k + \eta_k(Q_k - \gamma_k L_k)]_{\gamma_{\text{lb}}}^{\gamma_{\text{ub}}}. \quad (9)$$

($\eta_k = \frac{1}{k\bar{D}_{\text{lb}}}$, Recall SGD in ML, step-sizes should be diminishing, γ_{ub} and γ_{lb} can be estimated from transmission delay D upper and lower bounds.)

- **Update dual optimizer ν_k:**
$$\nu_{k+1} = \left(\nu_k + \frac{1}{V}(L_k - \frac{1}{f_{\max}})\right)^+. \quad (10)$$

Theoretic Analysis (1)

Constraint Satisfaction

U_k records the sampling constraint violation up to the k-th sample.
The proposed algorithm satisfies the sampling frequency constraint in the sense that:

$$\liminf_{K\to\infty} \frac{1}{K}\mathbb{E}\left[\sum_{k=1}^{K} U_k\right] < \infty. \tag{11}$$

Proof: Lyapunov Drift Plus Penalty

Theoretic Analysis (2)–Minimax Order Optimal

Learning Rate

If the delay D is bounded, the time-average AoI of the proposed algorithm converges to $\overline{A}_{\pi^\star}$ almost surely, i.e.,

$$\lim_{K\to\infty} \frac{\int_{t=0}^{S_{K+1}} A(t)\mathrm{d}t}{S_{K+1}} = \overline{A}_{\pi^\star}, \text{w.p.1}, \tag{12}$$

and the convergence rate:

$$\frac{\mathbb{E}\left[\int_{t=0}^{S_{K+1}} A(t)\mathrm{d}t\right]}{\mathbb{E}[S_{K+1}]} - \overline{A}_{\pi^\star} = \mathcal{O}\left(\frac{\ln K}{K}\right). \tag{13}$$

Converse Result

Let $\mathcal{P}_w$ be the set of probabilities so that, if delay $D \sim \mathbb{P} \in \mathcal{P}_w$, the age optimum sampling policy is not zero-wait, for any causal policy π:

$$\inf_{\pi} \sup_{\mathbb{P}\in\mathcal{P}_w} \left(\frac{\mathbb{E}\left[\int_{t=0}^{S_{K+1}} A(t)\mathrm{d}t\right]}{\mathbb{E}[S_{K+1}]} - \overline{A}_{\pi^\star} \right) = \Omega\left(\frac{\ln K}{K}\right). \tag{14}$$

Convergence Rate Analysis: perturbed ODE method;

Converse Result: Le Cam's two point method.

Simulations (1)

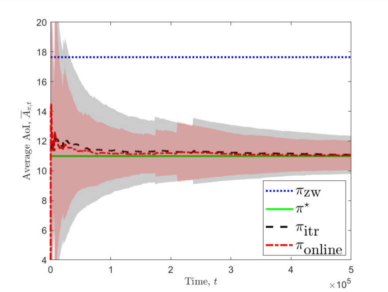


Figure 1: AoI evolution with time (red denotes the confidence interval).

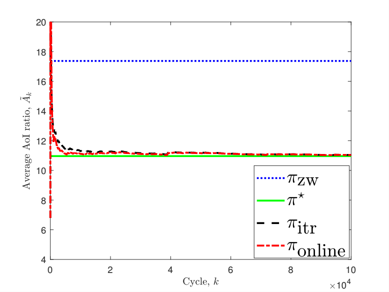


Figure 2: AoI evolution with frame $\mathbb{E}[\int_{t=0}^{S_{K+1}} A(t)\mathrm{d}t]/\mathbb{E}[S_{K+1}]$

The proposed algorithm learns the AoI minimum sampling policy adaptively when $T \to \infty$, the learning rate is faster than previous method.

Simulations (2)

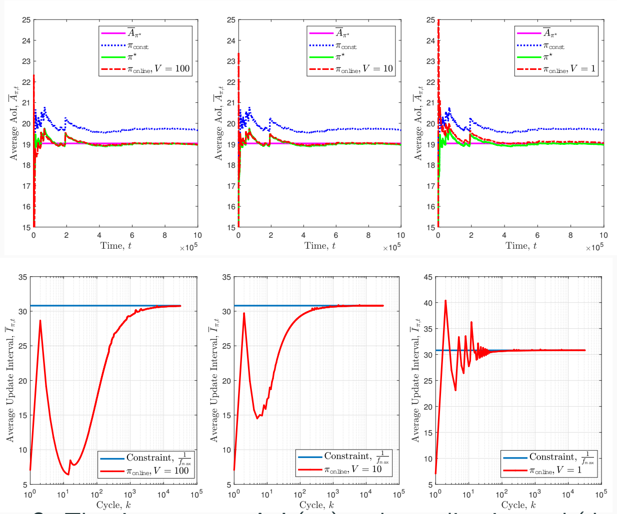


Figure 3: The time average AoI (up) and sampling interval (down).

π_{online} can satisfy the sampling frequency constraint, smaller V requires less time to satisfy the sampling frequency constraint.

Conclusions

- Contribution:
  Propose an online sampling strategy for AoI minimization.
- Theoretic Highlights:
  - The first work that adapt a Robbins-Monro algorithm Neely (2021) to an online AoI minimization problem.
  - Analyze the convergence behaviour and regret performance of the online learning algorithm.
  - Apply the Le Cam's two point method from statistics to the stochastic network community for proving converse.
- Full paper submitted to IEEE Transactions on Information Theory:
  H. Tang, Y. Chen, J. Wang, P. Yang and L. Tassiulas, "Age Optimal Sampling under Unknown Delay Statistics",
  https://arxiv.org/abs/2202.13367

Thank you! Questions?

References

E. U. Atay, I. Kadota, and E. H. Modiano. Aging wireless bandits: Regret analysis and order-optimal learning algorithm. In *WiOpt*, pages 57–64. IFIP, 2021.

S. Leng and A. Yener. An actor-critic reinforcement learning approach to minimum age of information scheduling in energy harvesting networks. In *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8128–8132, 2021. doi: 10.1109/ICASSP39728.2021.9415110.

B. Li. Efficient learning-based scheduling for information freshness in wireless networks. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, pages 1–10, 2021. doi: 10.1109/INFOCOM42981.2021.9488709.

M. J. Neely. Fast learning for renewal optimization in online task scheduling. *Journal of Machine Learning Research*, 22(279):1–44, 2021.

Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff. Update or wait: How to keep your data fresh. *IEEE Transactions on Information Theory*, 63(11):7492–7508, 2017.

V. Tripathi and E. Modiano. An online learning approach to optimizing time-varying costs of aoi. In *Proceedings of the Twenty-Second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, MobiHoc '21, page 241–250, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450385589. doi: 10.1145/3466772.3467053. URL https://doi.org/10.1145/3466772.3467053.