X-armed Bandits for Optimizing Information Freshness in Robots Communication

Thanh Le, Yusheng Ji, Thanh-Trung Nguyen, John C.S. Lui

The Graduate University for Advanced Studies (SOKENDAI) - Japan

24^{th} Oct. 2024

Acknowledgement: This research is supported in part by the Japanese Ministry of Internal Affairs and Communications (MIC) SCOPE project No.JP235006102.

Agenda

1. Introduction
2. System model, metrics, and formulation
3. Sampling policies
4. Experimental results
5. Conclusion

Agenda

1. Introduction
2. System model, metrics, and formulation
3. Sampling policies
4. Experimental results
5. Conclusion

Application of Autonomous Mobile Robots (AMRs)

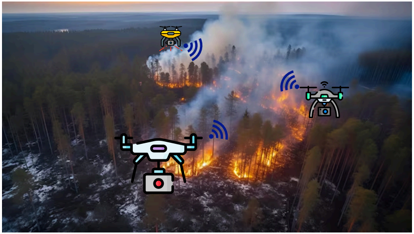


Figure: Drones observe their nearby environment and share their observations.

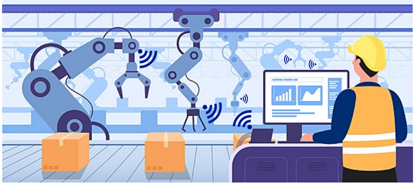


Figure: Robots and sensors in factory observe the manufacturing process and share the status to TV visualization.

AMRs: Tasks and Requirements

Tasks

- A group of AMRs that colaboratively solving a common task need to share their local observations
- An AMR that monitors a certain manufacturing process and send its observation to a remote monitor

Requirements

- Require latest observations from neighboring robots
- Need most recent status of a manufacturing process

$\rightarrow$ Require fresh information, characterized by a metric called Age of Information (AoI)

Optimize AMRs communication by minimize AoI

Age of Information (AoI)

- AoI is the metric to quantify information freshness
- Information usually has highest value when it is fresh
- The AoI at the monitor **increases linearly** in time when there is no update and is **reset to the delay** when an update is received
- Constraint: the communication channel can be congested when the sampling rate is too high

Contribution

- Study how to choose the sampling rate
- Implement a Proof of Concept (PoC) using 1-hop ad-hoc WiFi link
- Choose continuous sampling rate
- Visualize AoI for different sampling policies

Agenda

1 Introduction

2 System model, metrics, and formulation

3 Sampling policies

4 Experimental results

5 Conclusion

System model

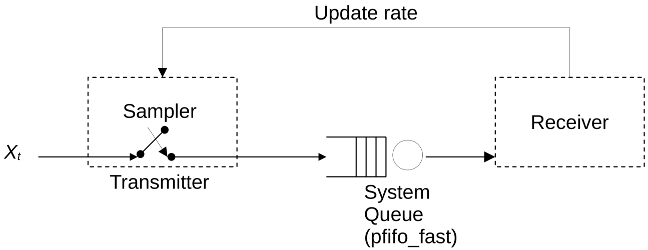


Figure: System model

- A source process X_t at **transmitter**
- Sample and send to a **receiver**
- 1-hop, ad-hoc Wifi 2.4GHz communication channel
- Sampled packets (with size 1kB) are sent via UDP
- Background traffic consumes an unknown amount of Wifi bandwidth
- Queueing discipline: single-server FIFO (linux kernel default)

Age of Information (AoI)

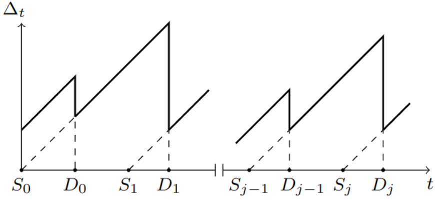


Figure: The evolution of AoI over time

Notations

The i^{th} sample is

- Sample at the **transmitter** at time S_i
- Deliver to the **receiver** at D_i

AoI

Let $U_t = max\{S_i : D_i \leq t\}$, the generation time of the freshest sample that has been delivered by time t. The AoI is defined by:

$$\Delta_t = t - U_t \tag{1}$$

Goal

$$\underset{\lambda}{\text{Minimize}} \quad \frac{1}{T}\sum_{t=1}^{T} \Delta_t \tag{2}$$

(3)

Agenda

1 Introduction

2 System model, metrics, and formulation

3 Sampling policies

4 Experimental results

5 Conclusion

Zero-wait policy

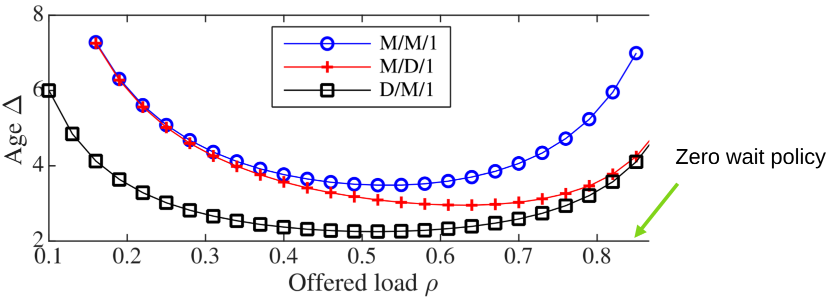


Figure: Theoretical AoI as a function of offered load for M/M/1, M/D/1, or D/M/1 queue[1].

Zero-wait policy

`Zero-wait` policy always sends update immediately after previous packet has been sent and leads to highest offered load.

[1] Sanjit Kaul, Roy Yates, and Marco Gruteser. "Real-time status: How often should one update?" In: *2012 Proceedings IEEE INFOCOM*. IEEE. 2012, pp. 2731–2735.

Finite difference policy (online learning)

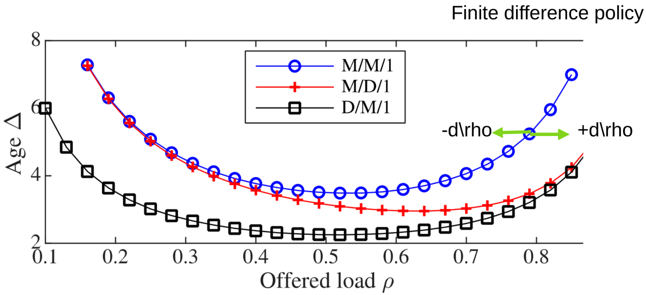


Figure: Theoretical AoI as function of offered load for M/M/1, M/D/1, or D/M/1 queue.

Finite diference policy

- Assume that the function of AoI w.r.t sampling rate (offered load) is convex
- Approximate the derivatives of AoI w.r.t sampling rate using nearby function values
- Use the derivatives to iteratively find the minimum

Problem - The AoI function in PoC is stochastic and has hard local optima.

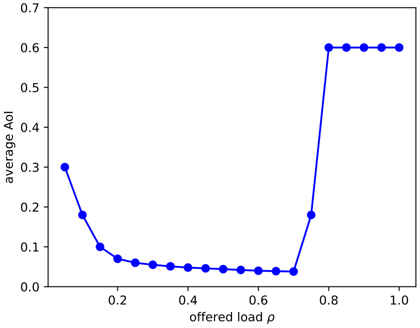


Figure: Average AoI as a function offered load on our PoC system.

- $\rho > 0.7$ is the critical point when the system becomes unstable
- Using finite difference method $\rightarrow$ cannot escape local optima within $[0.8, 1.0]$
- Note that

$$\rho = \frac{\text{throughput}}{\text{assumed maximum bandwidth}}$$

Solution X-armed bandit (XAB) policy (online learning)

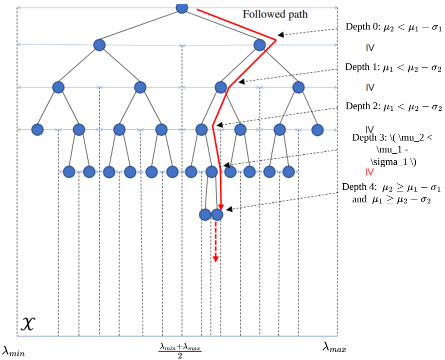


Figure: Illustration of X-armed bandits (XAB).

Hierarchical Parititioning XAB for Solving Blackbox Minimization Problem

- At depth 0, a single node cover the entire search space
- Split search space of each node into two equal partitions
- Randomly draw values from leaf nodes, until the **average** objective value of one branch **smaller** than the **average plus uncertainty** of objective value of another branch
- Iteratively traverse path with highest estimated objective value to evaluate

2

[2] Wenjie Li, Chi-Hua Wang, Guang Cheng, and Qifan Song. "Optimum-statistical collaboration towards general and efficient black-box optimization". In: *arXiv preprint arXiv:2106.09215* (2021).

Agenda

1 Introduction

2 System model, metrics, and formulation

3 Sampling policies

4 Experimental results

5 Conclusion

Experimental setup

Devices

- Two Rasberry Pi 4 devices placed two corners in our laboratory
- Testing duration per policy: 60s
- Search space: 5(ms/sample) to 50(ms/sample)
- Update packet size: 1kbyte
- Average over 10 independent trials

Traffic control to limit bandwidth

$ sudo tc qdisc add dev wlan0 root handle 1:0 tbf rate 1mbit burst 25kb limit 250kb

- Small bandwidth 1mbit/sec
- Small burst size and limit size

Result - Peak AoI

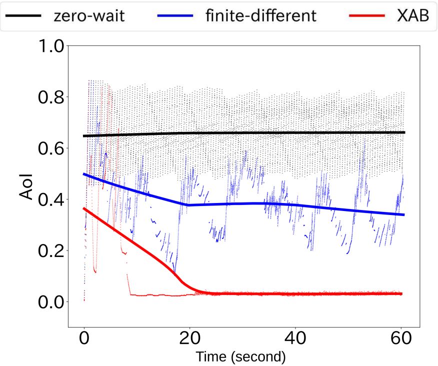


Figure: Average Peak AoI

Comments

- Zero-wait policy makes the network congested.
- Finite-difference policy was able to reduce AoI, but unable to converge
- XAB policy manages to gradually reduce AoI, and converge

Result - Number of Transmitted Packets

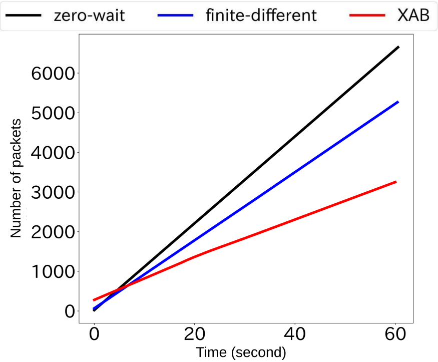
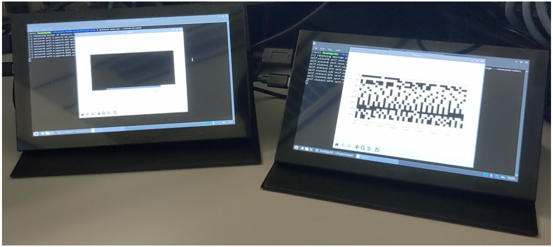


Figure: Total number of transmitted packets

- In average, `XAB` policy only sampled nearly 3000 packets within 60s
- `XAB` reduce about 50% of transmitted packets, in compared with `zero-wait` and `finite-difference`

Visualization of AoI



Left screen: Connected to Sender

- 2D image each black dot represents a packet
- Draw sequentially from left to right, up to down
- Draw when packet is sampled at sender

Right screen: Connected to Receiver

- Draw a black dot, represent the sampled timestamp, when packet is received
- Goal: Minimize AoI (reduce long consecutive white space)

Agenda

1 Introduction

2 System model, metrics, and formulation

3 Sampling policies

4 Experimental results

5 Conclusion

Conclusion

Summary

- Implemented a PoC status update system for AMRs communication via ad-hoc Wifi
- Applied `XAB` to reduce about 50% of transmitted packets, avoid network congestion, and minimize AoI

Future works

- Compare with other policies with discritized actions
- Scale to status update with multiple sources
- Incorporate data semantic to the sampling policies

Agenda

6 Related works

Optimal update rate to minimize M/M/1, D/M/1, and M/D/1 queues

Paper: Sanjit Kaul, Roy Yates, and Marco Gruteser. "Real-time status: How often should one update?" In: *2012 Proceedings IEEE INFOCOM*. IEEE. 2012, pp. 2731–2735

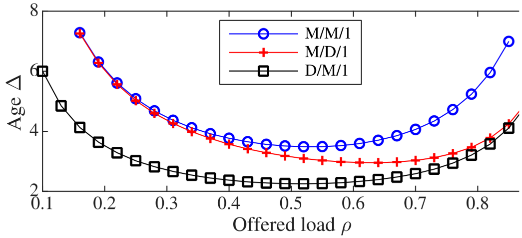


Figure: Compare AoI of different queueing models

Comments

- AoI is a deterministic convex function w.r.t offered load
- Easy to derive closed-form optimal sampling rate
- In our PoC, AoI is a stochastic non-convex function w.r.t offered load

Related works

Paper: Clement Kam, Sastry Kompella, and Anthony Ephremides. "Learning to sample a signal through an unknown system for minimum AoI". In: *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE. 2019, pp. 177–182

Comments

- Using SALSA algorithm with tile coding to learn the sampling rate under lossless TCP/IP connection
- Discretize the sampling rate.

Related works

Paper: Vishrant Tripathi, Igor Kadota, Ezra Tal, Muhammad Shahir Rahman, Alexander Warren, Sertac Karaman, and Eytan Modiano. "Wiswarm: Age-of-information-based wireless networking for collaborative teams of uavs". In: *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE. 2023, pp. 1–10

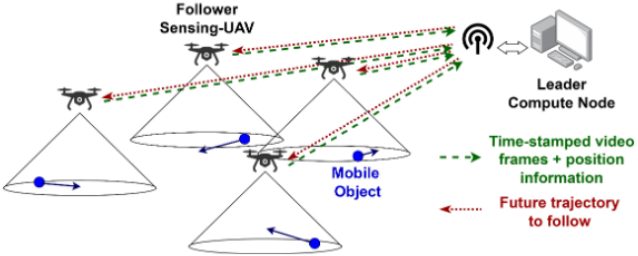


Figure: Leader node solve optimal sampling rate from multiple follower nodes to minimize AoI via WiFi. Using Whittle Index policy, time is synchronized.