# Age-of-Information-based Scheduling in Multiuser Uplinks with Stochastic Arrivals: A POMDP Approach

Aoyu Gong[1], Tong Zhang[2], He Chen[2], and Yijin Zhang[1]

[1]School of Electronic and Optical Engineering
Nanjing University of Science and Technology (NJUST)

[2]Department of Information Engineering
The Chinese University of Hong Kong (CUHK)

IEEE Global Communications Conference 2020

- The information freshness has become an increasingly important performance metric in this era of the Internet of Things (IoT).
- The concept of the age of information (AoI) has been proposed to measure the information freshness from the monitor's perspective.
- In this paper, we focus on scheduling problems of minimizing the network-wide AoI in multiuser uplinks with stochastic arrivals.

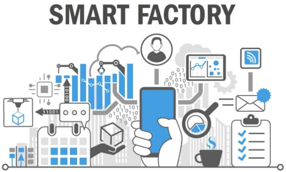



Figure: Smart Factory

Figure: Smart Healthcare

Figure: Smart Transport

# Motivation

## Multiuser Downlinks with Stochastic Arrivals

- There is no uncertainty about the status update arrivals from the monitor's perspective since status updates arrive at the monitor.

What about **multiuser uplinks with stochastic arrivals**?

## Weakness of Existing Work

- Most existing work assumed end nodes used extra feedback to report their status update arrivals to the monitor.
- Such feedback leads to considerable overhead and thus makes the corresponding scheduling policies hard to implement.

To combat this weakness, we assume that **there is no extra feedback**, which leads to a scheduling problem under **partial system information**.

# System Model - I

- We consider a multiuser uplink system where $K$ end nodes report their freshest status updates to a common monitor.
- In each time slot, a new status update arrives at end node $i \in \{1, 2, \ldots, K\}$ with probability $\lambda_i \in (0, 1]$.

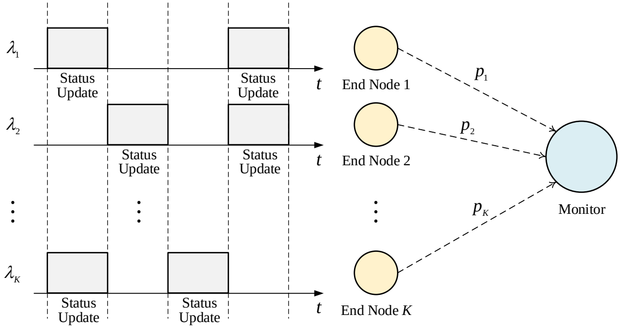



Figure: The multiuser uplink system with stochastic arrivals of status updates.

# System Model - II

- At the beginning of each time slot, the monitor schedules at most one end node to transmit its freshest status update.
- The transmission of end node $i$ to the monitor has a successful probability $p_i$ and an error probability $(1 - p_i)$.

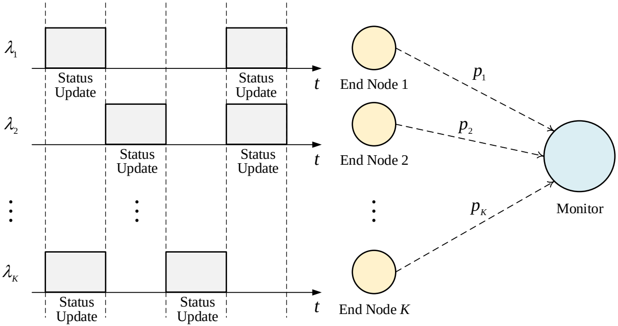


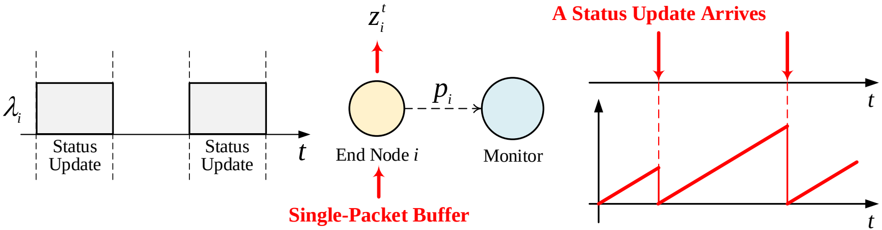

Figure: The multiuser uplink system with stochastic arrivals of status updates.

# Local Age

## Local Age

- Define **the local age of end node** $i$ as the time elapsed since the generation of the freshest status update at end node $i$.
- The evolution of **the local age of end node** $i$, denoted by $z_i^t$, is

$$z_i^{t+1} = \begin{cases} z_i^t + 1, & \text{no status update arrives in slot } t, \\ 1, & \text{a new status update arrives in slot } t. \end{cases} \quad (1)$$

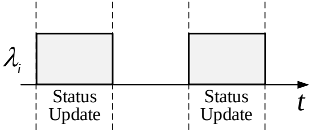


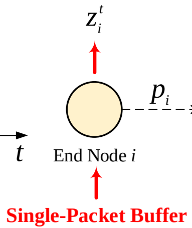
# Age of Information

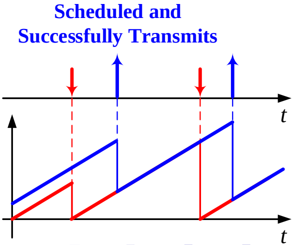
## Age of Information

- Define **the age of information (AoI) of end node** $i$ as the time elapsed since the generation (at end node $i$) of the latest received status update (at the monitor).
- The evolution of **the AoI of end node** $i$, denoted by $h_i^t$, is

$$h_i^{t+1} = \begin{cases} z_i^t + 1, & \text{scheduled and successfully transmits in slot } t \\ h_i^t + 1, & \text{otherwise.} \end{cases} \tag{2}$$



# Main Contributions

## Weakness of Existing Work

- Most existing work assumed end nodes used extra feedback to report their status update arrivals to the monitor.
- Such feedback leads to considerable overhead and thus makes the corresponding scheduling policies hard to implement.

To combat this weakness, we assume that **there is no extra feedback**, which leads to a scheduling problem under **partial system information**.

## Main Contributions

- Formulate a partially observable Markov decision process (POMDP) to characterize the dynamic behavior of such system.
- Use a dynamic programming (DP) algorithm to attain the optimal policy and devise a myopic policy with low computation complexity.

# POMDP Framework

A POMDP models an agent decision process in which it is assumed that

- the system dynamics are determined by an MDP, and
- the agent cannot directly observe the underlying state.

## Definition

A POMDP can be described as a tuple $< \mathcal{S}, \mathcal{A}, \mathcal{O}, T, R, Z >$.

- $\mathcal{S}$ is a finite set of states.
- $\mathcal{A}$ is a finite set of actions.
- $\mathcal{O}$ is a finite set of observations.
- $T$ is the state transition function, $T : \mathcal{S} \times \mathcal{A} \rightarrow \Pi(\mathcal{S})$.
- $R$ is the reward function, $R : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$.
- $Z$ is the observation function, $Z : \mathcal{S} \times \mathcal{A} \rightarrow \Pi(\mathcal{O})$.

# POMDP Formulation - I

## States

- Denote a state of end node $i$ in slot $t$ by $\mathbf{s}_i^t \triangleq [h_i^t, z_i^t]$.
  - $h_i^t \in \mathcal{T} \triangleq \{1, 2, 3, \ldots\}$ is its instantaneous AoI at the monitor
  - $z_i^t \in \mathcal{T}$ is its local age
- Denote a state of the POMDP in slot $t$ by $\mathbf{s}^t \triangleq [\mathbf{h}^t, \mathbf{z}^t]$.
  - $\mathbf{h}^t \triangleq [h_1^t, \ldots, h_K^t] \in \mathcal{H} = \mathcal{T}^K$ is the AoI of all end nodes
  - $\mathbf{z}^t \triangleq [z_1^t, \ldots, z_K^t] \in \mathcal{Z} = \mathcal{T}^K$ is the local age of all end nodes

## Actions

- Denote an action of end node $i$ in slot $t$ by $a_i^t \in \{0, 1\}$.
  - If end node $i$ is scheduled, $a_i^t = 1$; otherwise, $a_i^t = 0$.
- Denote an action of the POMDP in slot $t$ by $\mathbf{a}^t \triangleq [a_1^t, \ldots, a_K^t]$.
  - In the single-antenna system considered, we have $\sum_{i=1}^K a_i^t \leq 1$.

# POMDP Formulation - II

## Observations

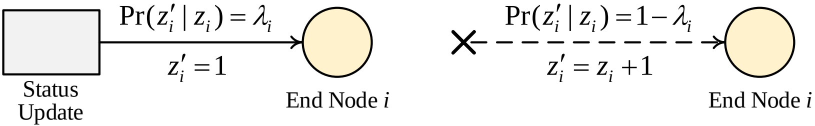
- Denote an observation of end node $i$ in slot $t$ by $\mathbf{o}_i^t \triangleq [h_i^t, \hat{z}_i^t]$.
  - $h_i^t$ is its fully observed AoI at the monitor
  - $\hat{z}_i^t \in \{\mathcal{T}, X\}$ is its partially observed local age, where $X$ means no observation of the local age of an end node
- Denote an observation of the POMDP in slot $t$ by $\mathbf{o}^t \triangleq [\mathbf{o}_1^t, \ldots, \mathbf{o}_K^t]$.

## Belief State

- Denote a belief state of the POMDP in slot $t$ by $\mathbf{I}^t \triangleq [\mathbf{h}^t, \mathbf{b}^t]$.
  - $\mathbf{b}^t$ is a probability distribution over $\mathcal{Z}$, where $b(\mathbf{z})$ is the probability assigned to $\mathbf{z}^t = \mathbf{z}$ by $\mathbf{b}^t = \mathbf{b}$, satisfying $\sum_{\mathbf{z}\in\mathcal{Z}} b(\mathbf{z}) = 1$.
  - Note that $\mathbf{h}^t$ is fully observable, i.e., its belief update is always deterministic given $\mathbf{h}^{t-1}$, $\mathbf{a}^{t-1}$ and $\mathbf{o}^{t-1}$.

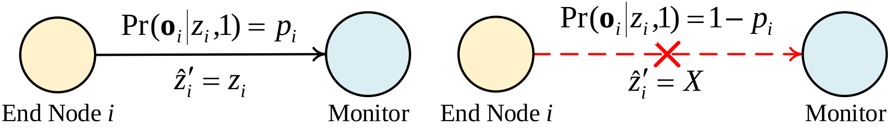
## State Transition Function

- Let $\Pr(\mathbf{z}'|\mathbf{z}) = \prod_{i=1}^{K} \Pr(z_i'|z_i)$ denote the state transition function.




## Observation Function

- Let $\Pr(\mathbf{o}|\mathbf{z}, \mathbf{a}) = \prod_{i=1}^{K} \Pr(\mathbf{o}_i|z_i, a_i)$ denote the observation function.



# POMDP Formulation - IV

## Belief Update

- When given $\mathbf{I}^t = \mathbf{I}$, $\mathbf{a}^t = \mathbf{a}$ and $\mathbf{o}^t = \mathbf{o}$, for $\forall i$, $h_i^{t+1}$ can be updated as follows:

$$h_i^{t+1} = \begin{cases} \hat{z}_i + 1, & \text{if } \hat{z}_i \neq X, \\ h_i + 1, & \text{if } \hat{z}_i = X. \end{cases} \tag{3}$$

- When given the same condition, for $\forall \mathbf{z}' \in \mathcal{Z}$, $b^{t+1}(\mathbf{z}')$ can be updated via the Bayes' theorem:

$$b^{t+1}(\mathbf{z}') = \Big\{ \sum_{\mathbf{z} \in \mathcal{Z}} \Pr(\mathbf{z}'|\mathbf{z}) \Pr(\mathbf{o}|\mathbf{z}, \mathbf{a}) b(\mathbf{z}) \Big\} / \Pr(\mathbf{o}|\mathbf{I}, \mathbf{a}), \tag{4}$$

  where $\Pr(\mathbf{o}|\mathbf{I}, \mathbf{a}) = \sum_{\mathbf{z} \in \mathcal{Z}} \Pr(\mathbf{o}|\mathbf{z}, \mathbf{a}) b(\mathbf{z})$.
- We denote the process by the update function $\mathbf{I}^{t+1} = f(\mathbf{I}, \mathbf{a}, \mathbf{o})$.

# POMDP Formulation - V

## Reward

- Define the reward at belief state $\mathbf{I}^t = \mathbf{I}$ as the weighted sum of the instantaneous AoI of all end nodes, i.e., $R(\mathbf{I}) \triangleq \sum_{i=1}^{K} \omega_i h_i$.

## Policy

- Define a decision rule as a mapping from the belief space $\mathcal{I}$ into the action space $\mathcal{A}$, i.e., $d_t : \mathcal{I} \to \mathcal{A}$.
- Define a policy as a sequence of decision rules, i.e., $\pi = \{d_1, \ldots, d_T\}$.

Let $V^\pi(\mathbf{I})$ denote the expected total reward from slot 1 to slot $T$ when $\mathbf{I}^1 = \mathbf{I}$ and the policy $\pi$ is used, which can be defined by

$$V^\pi(\mathbf{I}) \triangleq \mathbb{E}\Big[\sum_{t=1}^{T} R(\mathbf{I}^t) | \mathbf{I}^1 = \mathbf{I}, \pi\Big] \;\rightarrow\; \text{(P): } \pi^* = \arg\min_{\pi} \frac{1}{TK} V^\pi(\mathbf{I}).$$

# The Optimal Policy

## The Total Expected Reward

- Let $U_t^\pi(\mathbf{I})$ denote the total expected reward from slot $t$ to slot $T$ when $\mathbf{I}^t = \mathbf{I}$ and the policy $\pi$ is used, which can be defined by

$$U_t^\pi(\mathbf{I}) \triangleq \mathbb{E}\left[\sum_{\tau=t}^{T} R(\mathbf{I}^\tau)|\mathbf{I}^t = \mathbf{I}, \pi\right] \rightarrow \text{(P): } \pi^* = \arg\min_{\pi} \frac{1}{TK} U_1^\pi(\mathbf{I}).$$

## Dynamic Programming

- For problem (P), we have the following bellman equation:

$$U_t^*(\mathbf{I}) = R(\mathbf{I}) + \max_{\mathbf{a}\in\mathcal{A}} \sum_{\mathbf{o}\in\mathcal{O}} \Pr(\mathbf{o}|\mathbf{I}, \mathbf{a}) U_{t+1}^*\big(f(\mathbf{I}, \mathbf{a}, \mathbf{o})\big), \tag{5}$$

  for each $t \in \{1, \ldots, T-1\}$ and $\mathbf{I} \in \mathcal{I}_t$.
- The minimal EWSAoI given $\mathbf{I}^1 = \mathbf{I}$ can be computed by $U_1^*(\mathbf{I})/TK$.

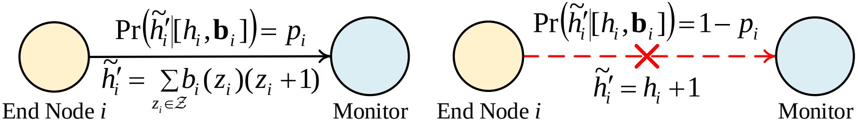
# A Myopic Policy - I

## Belief State (Sufficient Statistics)

- The monitor can only maintain probability distributions of the local age of each end node, denoted by $\mathbb{B}^t = [\mathbf{b}_1^t, \mathbf{b}_2^t \ldots, \mathbf{b}_K^t]$.
- The belief state of the POMDP can be redefined as $\mathbb{I}^t \triangleq [\mathbf{h}^t, \mathbb{B}^t]$.

## The One-Step Expected AoI

- Denote the expected AoI in the next slot of end node $i$ by $\mathbb{E}(h_i')$.
  - $\mathbb{E}(h_i') = (1 - a_i)(h_i + 1) + a_i^t\left(p_i \sum_{z_i \in \mathcal{T}} b_i(z_i)(z_i + 1) + (1 - p_i)(h_i + 1)\right)$



# A Myopic Policy - II

## The One-Step Expected Reward

- Given belief state $\mathbb{I}^t = \mathbb{I}$, if action $\mathbf{a}^t = \mathbf{a}$ is chosen in slot $t$, the one-step expected reward of the POMDP is given by

$$\hat{R}(\mathbb{I}, \mathbf{a}) = \sum_{i=1}^{K} \omega_i \left[ (1 - a_i)(h_i + 1) + a_i^t \left( p_i \sum_{z_i \in \mathcal{T}} b_i(z_i)(z_i + 1) + (1 - p_i)(h_i + 1) \right) \right]$$

## The Myopic Policy

- Given belief state $\mathbb{I}^t = \mathbb{I}$, the myopic policy can be given by

$$d_t(\mathbb{I}^t) = \arg\min_{\mathbf{a} \in \mathcal{A}} \hat{R}(\mathbb{I}, \mathbf{a}). \tag{6}$$

# Simulation - I

## Successful Transmission Probability

- The successful transmission probability of end node $i$ is computed by

$$p_i = 1 - \Pr(r_i < r_{th}) \overset{(a)}{=} \exp\left(-\frac{d_i^{\tau}(2^{r_{th}} - 1)}{\mathrm{SNR}}\right). \quad (7)$$

## Benchmarking Policies

- **Myopic policy (full knowledge), MPF**: Given $\mathbf{s}^t = \mathbf{s}$, the monitor schedules end node $i^*$ satisfying $i^* = \arg\max_{\forall i} \omega_i p_i (h_i - z_i)$.
- **MaxAoI policy, MAP**: Given $\mathbf{h}^t = \mathbf{h}$, the monitor schedules end node $i^*$ satisfying $i^* = \arg\max_{\forall i} \omega_i p_i h_i$.
- **Randomized policy, RDP**: The monitor schedules end node $i$ with probability $\mu_i = \sqrt{\omega_i / p_i} / (\sum_{j=1}^{K} \sqrt{\omega_j / p_j})$, $\forall i$.

# Simulation - II

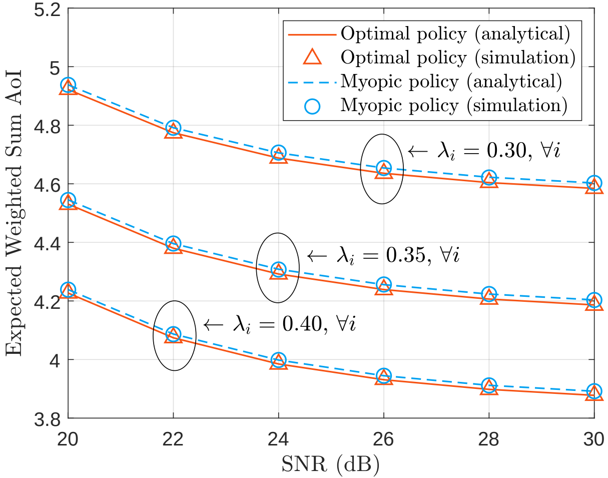



Figure: The EWSAoI as a function of the SNR for different configurations when $K = 2$, $T = 50$, $D = 15$, and $\omega_i = 1$, $\forall i$.

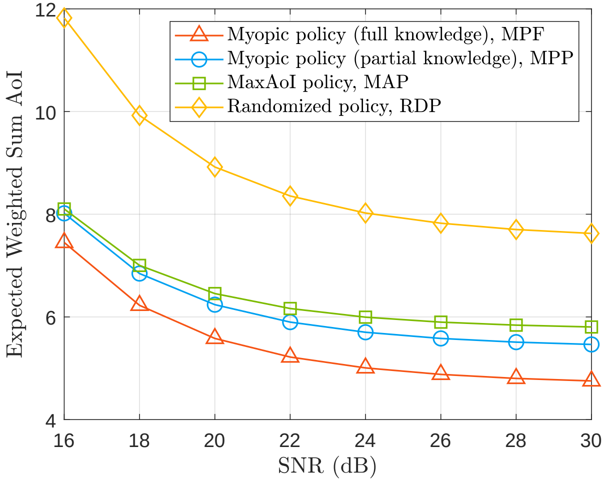



Figure: The EWSAoI as a function of the SNR when $K = 5$, $T = 10^6$, $D = 50$, $\lambda_i = 0.4$ and $\omega_i = 1$, $\forall i$.

# Simulation - IV

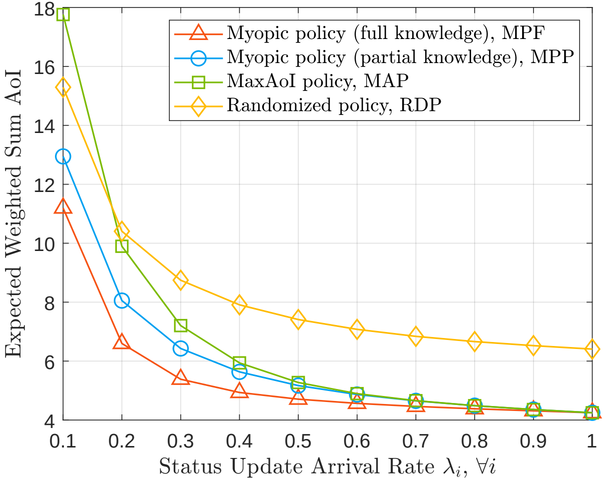



Figure: The EWSAoI as a function of the status update arrival rate when $K = 5$, $T = 10^6$, $D = 50$, SNR $= 25\,$dB and $\omega_i = 1$, $\forall i$.

# Conclusion

## System Model

In this paper, we have investigated **age-of-information-based scheduling in multiuser uplinks with stochastic arrivals**.

## POMDP Formulation and Solutions

- Formulate a partially observable Markov decision process (POMDP) to characterize the dynamic behavior of such system.
- Use a dynamic programming (DP) algorithm to attain the optimal policy and devise a myopic policy with low computation complexity.

## Simulation

The performance of the myopic policy approaches that of the optimal policy and is superior to that of the benchmarking policy.