Dynamic Random Access without Observation under Deadline-Constrained Periodic Traffic

Aoyu Gong, *Graduate Student Member, IEEE*, Yuan-Hsun Lo, *Member, IEEE*, Yan Lin, *Member, IEEE*, and Yijin Zhang, *Senior Member, IEEE*

***Abstract*—This paper focuses on random access in uplink systems under deadline-constrained periodic traffic, which is typical for many real-time Internet of Things scenarios. To achieve very low overhead in random access, we consider dynamic slotted ALOHA without observation where each active node adopts time-dependent but observation-independent transmission probabilities. Built on the theory of blind Markov decision processes, we develop an analytical framework of such dynamic access with a simplified version of information states, which leads to optimal time-dependent transmission probabilities. Further, based on this framework, we derive simple closed-form expressions for optimal time-dependent transmission probabilities and maximum long-run system throughput, which makes our scheme also enjoy very low complexity. Numerical results show that the proposed scheme outperforms other random access schemes without observation over a wide range of network configurations.**

***Index Terms*—Internet of Things, delivery deadline, random access, blind Markov decision processes**

I. Introduction

To enable timely monitoring and control, real-time Internet of Things (IoT) systems that impose a strict deadline on packet delivery have become prevalent in industry, environment, and transportation domains [1], [2]. Out of several such systems, this paper focuses on uplink IoT systems under deadline-constrained periodic traffic, which are typical for automation control loops and process control use cases [3]. For example, with a predefined delivery deadline, pressure sensors periodically send values to a machine PLC, so that the PLC can continuously monitor the status and take necessary actions.

Usually, an uplink IoT system involves a massive number of battery-powered, low-cost, and uncoordinated sensor nodes sending small packets to an access point (AP). So, it is essential to develop a low-complexity and low-overhead random access scheme to achieve excellent throughput/reliability performance under deadline-constrained periodic traffic. The design philosophy is expected to be different from what has been commonly used under deadline-unconstrained traffic.

Many recent studies have been devoted to this design issue. Under p-fixed slotted ALOHA where each active node adopts a fixed transmission probability p, Deng *et al.* [4] investigated the throughput performance for any p using a recursive algorithm, but the fixed p inevitably restricts the maximum achievable throughput performance. Under p-dynamic slotted ALOHA where each active node adopts time-dependent p, Zhao *et al.* [5] proposed to double or halve p according to the previous AP feedback, and analyzed the reliability performance relying on an absorbing Markov chain. However, this scheme was designed without any principle of optimality. Still under p-dynamic slotted ALOHA, Bae *et al.* [6] proposed to change p for maximizing the instantaneous throughput relying on a complete knowledge of the current number of active nodes, which is, however, unrealistic. This scheme was designed without a principle of dynamic programming optimality. To overcome the weakness in [4]–[6], built on the theory of partially observable Markov decision processes (POMDPs), Zhang *et al.* [7] proposed two dynamic schemes for maximizing the long-run throughput when each node only has an incomplete knowledge of the current number of active nodes. A common characteristic of [4]–[7] is that they all require instantaneous AP feedback to notify the transmission outcome and update the knowledge of the number of current active nodes, which introduces excessive overhead for small packets and resource-constrained nodes [8], [9]. To eliminate such overhead, another type of random access, called K-repetition [10], has been suggested in 3GPP R16. It requires each active node to randomly and uniformly choose K slots for transmissions before deadline without the need of observing the AP feedback and channel status (idle/busy). Gong *et al.* [11] focused on a similar system model and considered dynamic slotted ALOHA where each active node adopts both time- and observation-dependent transmission probabilities. However, the study therein requires each active node to constantly sense the channel status and maintain an activity belief, which may be impractical for low-cost nodes. Liu *et al.* [12] focused on centralized scheduling under Bernoulli traffic and optimized the age of information based on a POMDP framework. However, the centralized manner makes the technical difficulty in [12] quite different from that in random access [4]–[7], [10], [11].

Motivated by [4]–[7], [10], [11], this paper aims to design a p-dynamic slotted ALOHA scheme without observation.

This work was supported in part by the National Natural Science Foundation of China under Grants 62071236, 62001225, in part by the National Science and Technology Council of Taiwan under Grant NSTC 112-2115-M-153-MY2, and in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University under Grant 2022D07. (Corresponding author: Yijin Zhang.)

A. Gong is with the School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland (e-mail: aoyu.gong@epfl.ch).

Y.-H. Lo is with the Department of Applied Mathematics, National Pingtung University, Pingtung 90003, Taiwan (e-mail: yhlo@mail.nptu.edu.tw).

Y. Lin is with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: yanlin@njust.edu.cn).

Y. Zhang is with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: yijin.zhang@gmail.com).

Different from [4]–[6], [10], the designed scheme accounts for the dynamic programming optimality. Different from [7], [11], it has closed-form expressions and thus achieves very low complexity. Different from [4]–[7], it needs no observations and thus achieves very low overhead. In this scheme, each active node adopts time-dependent but observation-independent transmission probabilities. Our contributions are as follows.

(i) In Section III, based on the theory of blind Markov decision processes (MDPs) [13], we develop an analytical framework for p-dynamic slotted ALOHA without observation, which leads to optimal time-dependent transmission probabilities. Moreover, we prove a simplified version of information states to simplify this framework, which serves as a basis for subsequent derivations.

(ii) In Section IV, we derive closed-form expressions for optimal time-dependent transmission probabilities obtained from this framework, which makes our scheme enjoy very low complexity, thus being well suited for resource-constrained nodes. The explicit results help to easily understand how to optimally utilize channel and time resources under different urgencies.

(iii) In Section V, we present numerical results to evaluate the performance advantage of the proposed scheme and validate our theoretical findings.

To the best of our knowledge, this paper is the first to use the theory of blind MDPs in deadline-constrained random access.

II. System Model and Problem Formulation

We assume a wireless network consisting of a finite number, $N \geq 2$ of nodes which contend to transmit packets to a common AP. All the nodes are inside the AP's reception range and can interfere with each other's transmissions. The global time axis is structured frame-by-frame. Each frame is partitioned into $D \geq 1$ equal-duration time slots indexed by $t \in \mathcal{T} \triangleq \{1, 2, \ldots, D\}$. The nodes are aware of the slot boundaries. With probability $\lambda \in (0, 1]$, a single-slot packet is generated by each node independently at the start of each frame. A packet generated in a frame has a strict delivery deadline D slots, i.e., it will be dropped when this frame finishes. We assume that a packet is successfully received by the AP with probability $\sigma \in (0, 1]$ if and only if its transmission does not overlap with other transmissions.

As in [14], [15], we assume that each packet is not acknowledged and will not be retransmitted based on the following considerations. First, waiting for the AP feedback introduces additional latency, which is undesirable for real-time services. Second, feedback reception and retransmissions are both power-consuming, which are costly for resource-constrained nodes. Third, built-in redundancy in sensor coverage is usually applied in IoT systems to ensure reliability instead of retransmissions [16]. For the sake of low overhead, we also assume that all the nodes always perform no channel observation.

At the beginning of slot t in an arbitrary frame, we consider that a node is active if a packet is generated by this node at the start of this frame and has not been sent yet, and each active node sends its packet with probability $p_t \in [0, 1]$. Let $n_t \in \mathcal{N} \triangleq \{0, 1, \ldots, N\}$ represent the number of active nodes.

Based on the above assumptions, a p-dynamic slotted ALOHA scheme without observation can be defined by a sequence of transmission probabilities $\boldsymbol{\pi} \triangleq (p_1, p_2, \ldots, p_D) \in [0, 1]^D$. We want to find an optimal scheme $\boldsymbol{\pi}^* \triangleq (p_1^*, p_2^*, \ldots, p_D^*)$ that maximizes the long-run system throughput, i.e.,

$$\Theta^{\boldsymbol{\pi}} \triangleq \mathbb{E}^{\boldsymbol{\pi}}[\sigma n_t p_t (1 - p_t)^{n_t - 1}]/D,$$
$$\boldsymbol{\pi}^* \in \arg\max_{\boldsymbol{\pi} \in [0,1]^D} \Theta^{\boldsymbol{\pi}}.$$

Denote by $\mathbf{B}(M, q)$ a binomial distribution with parameters M and q where the k-th ($0 \leq k \leq M$) component is equal to $\binom{M}{k} q^k (1-q)^{M-k}$.

Two practical scenarios can be found as follows.

- In a periodic event-triggered control application [17], a fixed number of sensors associated with a process are deployed to periodically measure plant outputs and send their measurements to a controller when pre-designed conditions are met.
- In a group-based event detection application [18], a fixed number of nodes are deployed at deterministic locations to periodically monitor the same area of interest and send their reports to a controller for detecting physical events.

III. Blind MDP Framework

We formulate a blind MDP based on the definitions below.

(i) <u>*States, actions:*</u> We view n_t as the state at slot t, and view p_t as the action at slot t.

(ii) <u>*State Transition Function:*</u> The state transition function $T_t(n', n, p_t)$ is defined to be the probability of starting in $n_t = n$ and ending in $n_{t+1} = n'$ under the transmission probability p_t. So, for each $n, n' \in \mathcal{N}$ and each $p_t \in [0, 1]$, we have

$$T_t(n', n, p_t) = \begin{cases} \binom{n}{n-n'} p_t^{n-n'} (1-p_t)^{n'}, & \text{if } n' \leq n, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

(iii) <u>*Information State:*</u> The information state at slot t is denoted by $\mathbf{I}_t \triangleq [I_t(0), I_t(1), \ldots, I_t(N)]$, where $I_t(n)$ is the conditional probability (given all past transmission probabilities) that $n_t = n$. For each $t \in \mathcal{T} \setminus \{D\}$, given $\mathbf{I}_t$ and p_t, $\mathbf{I}_{t+1}$ can be computed by

$$I_{t+1}(n') = \sum_{n \in \mathcal{N}} T_t(n', n, p_t) I_t(n), \ \forall n' \in \mathcal{N}. \tag{2}$$

Further, we prove the following simplified version of $\mathbf{I}_t$ that is equivalent to the original one. The proof is relegated to Appendix A.

Lemma 1. *For $t \in \mathcal{T}$, $\mathbf{I}_t = \mathbf{B}(N, \alpha_t)$ where*

$$\alpha_t = \begin{cases} \lambda, & \textit{if } t = 1, \\ \alpha_{t-1}(1 - p_{t-1}), & \textit{otherwise.} \end{cases} \tag{3}$$

Let $\mathcal{A}_t$ represent the set consisting of all possible values of α_t.

(iv) *Reward Function:* The reward $R_t(\alpha_t, p_t)$ is defined to be the mean number of packets successfully received by the AP at slot t under $\mathbf{I}_t = \mathbf{B}(N, \alpha_t)$ and p_t. So,

$$R_t(\alpha_t, p_t) = \sum_{n=1}^{N} I_t(n)\sigma n p_t(1-p_t)^{n-1} \tag{4}$$

$$\stackrel{(*)}{=} \sigma N \alpha_t p_t (1-\alpha_t p_t)^{N-1}, \tag{5}$$

where the proof of $(*)$ is relegated to Appendix B.

Let $J^{\boldsymbol{\pi}}$ represent the expected total reward obtained at slots $1, 2, \ldots, D$ under the policy $\boldsymbol{\pi}$, i.e.,

$$J^{\boldsymbol{\pi}} \triangleq \mathbb{E}^{\boldsymbol{\pi}}\left[\sum_{t=1}^{D} R_t(\alpha_t, p_t)\right] = \Theta^{\boldsymbol{\pi}} D.$$

Let $U_t^*(\alpha_t)$ represent the maximum total expected reward obtained at slots $t, t+1, \ldots, D$ when $\mathbf{I}_t = \mathbf{B}(N, \alpha_t)$. Obviously, $U_1^*(\alpha_1) = \Theta^{\boldsymbol{\pi}^*} D$. Then, we arrive at the following Bellman equation:

$$\begin{aligned} U_D^*(\alpha_D) &= \max_{p_D \in [0,1]} R_D(\alpha_D, p_D),\ \forall \alpha_D \in \mathcal{A}_D, \\ U_t^*(\alpha_t) &= \max_{p_t \in [0,1]} R_t(\alpha_t, p_t) + U_{t+1}^*(\alpha_t - \alpha_t p_t), \\ &\qquad \forall \alpha_t \in \mathcal{A}_t, \forall t \in \mathcal{T} \setminus \{D\}. \end{aligned} \tag{6}$$

Solving (6) formally leads to $\boldsymbol{\pi}^*$ and $\Theta^{\boldsymbol{\pi}^*}$. However, getting $\boldsymbol{\pi}^*$ is computationally heavy as $\cup_{t \in \mathcal{T}} \mathcal{A}_t$ and the action space $[0, 1]$ are both infinite.

IV. Closed-Form Optimal Scheme

In this section, to reduce the computational complexity of solving $\boldsymbol{\pi}^*$, we derive simple closed-form expressions for $\boldsymbol{\pi}^*$. A closed-form expression for $\Theta^{\boldsymbol{\pi}^*}$ is also provided.

First, we provide key properties of $R_t(\alpha_t, p_t)$ and $U_t^*(\alpha_t)$ in Lemma 2. Its proof is relegated to Appendix C.

Lemma 2. *For $t \in \mathcal{T}$, if $\alpha_t > 0$, the followings hold.*

(i) When $N\alpha_t \geq 1$, $R_t(\alpha_t, p_t)$ achieves its maximum $\sigma(1-\frac{1}{N})^{N-1}$ only at $p_t = \frac{1}{N\alpha_t}$,
(ii) When $0 < N\alpha_t < 1$, $R_t(\alpha_t, p_t)$ achieves its maximum $\sigma N\alpha_t(1-\alpha_t)^{N-1}$ only at $p_t = 1$, and the maximum value is strictly smaller than $\sigma(1-\frac{1}{N})^{N-1}$, and
(iii) $U_t^*(\alpha_t) \leq (D-t+1)\sigma(1-\frac{1}{N})^{N-1}$.

Based on Lemma 2, we derive closed-form p_t^* and $U_t^*(\alpha_t)$ when $\alpha_t > 0$. Its proof is relegated to Appendix D.

Lemma 3. *For $t \in \mathcal{T}$,*

(i) if $N\alpha_t \geq D-t+1$, then $p_t^* = \frac{1}{N\alpha_t}$ and $U_t^*(\alpha_t) = (D-t+1)\sigma(1-\frac{1}{N})^{N-1}$, and
(ii) if $0 < N\alpha_t < D-t+1$, then $p_t^* = \frac{1}{D-t+1}$ and $U_t^*(\alpha_t) = \sigma N\alpha_t(1-\frac{\alpha_t}{D-t+1})^{N-1}$.

We are ready for deriving closed-form $\boldsymbol{\pi}^*$ and $\Theta^{\boldsymbol{\pi}^*}$.

Theorem 4. *For $N \geq 2$, $\lambda \in (0, 1]$, and $D \geq 1$, we have*

$$p_t^* = \begin{cases} \frac{1}{N\lambda - t + 1}, & \text{if } N\lambda \geq D, \\ \frac{1}{D-t+1}, & \text{otherwise}, \end{cases}$$

for each $t \in \mathcal{T}$, and

$$\Theta^{\boldsymbol{\pi}^*} = \begin{cases} \sigma(1-\frac{1}{N})^{N-1}, & \text{if } N\lambda \geq D, \\ \sigma\frac{N\lambda}{D}(1-\frac{\lambda}{D})^{N-1}, & \text{otherwise}. \end{cases}$$

Proof. For the case $N\lambda \geq D$, by Lemma 3 (i), we have $U_1^*(\lambda) = D\sigma(1-\frac{1}{N})^{N-1}$. So, we obtain $\Theta^{\boldsymbol{\pi}^*} = \frac{U_1^*(\lambda)}{D} = \sigma(1-\frac{1}{N})^{N-1}$. When $t = 1$, by Lemma 1, we know $\alpha_1 = \lambda$. As $N\lambda \geq D$, by Lemma 3 (i), we have $p_1^* = \frac{1}{N\lambda}$. When $t = 2$, by Lemma 1, we know $\alpha_2 = \lambda - \frac{1}{N}$. As $N(\lambda - \frac{1}{N}) \geq D-1$, by Lemma 3 (i), we further have $p_2^* = \frac{1}{N\lambda-1}$. For $t = 3, 4, \ldots, D$, we can obtain the optimal scheme for this case by repeating the aforementioned reasoning iteratively.

For the case $N\lambda < D$, by Lemma 3 (ii), we know $U_1^*(\lambda) = \sigma N\lambda(1-\frac{\lambda}{D})^{N-1}$. So, we obtain $\Theta^{\boldsymbol{\pi}^*} = \sigma\frac{N\lambda}{D}(1-\frac{\lambda}{D})^{N-1}$. For each $t \in \mathcal{T}$, by Lemma 3 (ii), we have $p_t^* = \frac{1}{D-t+1}$. ☐

Remark 1: If $N\lambda \geq D$, each active node will adopt p_t^* maximizing the instantaneous throughput. Otherwise, they will adopt p_t^* transmitting all the backlogged packets uniformly in D slots. Note that the latter case is equivalent to 1-repetition [10].

Remark 2: The analytical framework developed in this paper is based on the model-based blind MDP formulation. If the number of nodes is unknown or varies with time, model-free algorithms in reinforcement learning can be utilized as they do not require learning a model of the environment [19]. However, such algorithms often exhibit high computational complexity and slow convergence due to the need for a large number of interactions with the environment, which may pose challenges for low-cost nodes.

V. Numerical Evaluation

In this section, we compare the proposed scheme $\boldsymbol{\pi}^*$, the myopic scheme $\boldsymbol{\pi}^{\text{myo}}$ [6], the optimal static scheme $\boldsymbol{\pi}^{\text{sta}}$ [4], and 1-repetition [10]. Here, $\boldsymbol{\pi}^{\text{sta}} \in \arg\max_{p \in [0,1]} \Theta^{\boldsymbol{\pi}}$ s.t. $p_t = p$, $\forall t \in \mathcal{T}$. We set up the numerical experiments according to the system model specified in Section II.

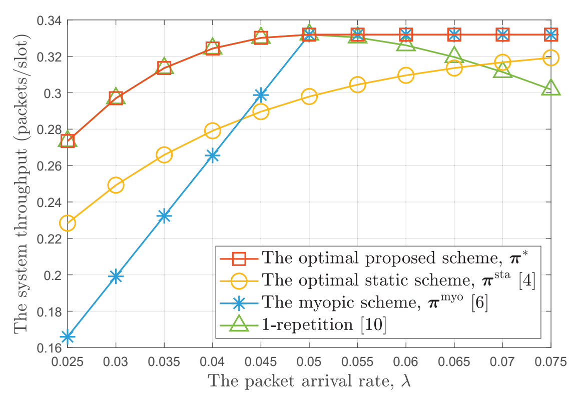

Fig. 1: The system throughput (packets/slot) versus the packet arrival rate λ when $N = 200$, $D = 10$, $\sigma = 0.9$.

Figs. 1–2 show the throughput performance as functions of the packet arrival rate λ and the delivery deadline D, respectively. We observe that $\boldsymbol{\pi}^*$ significantly outperforms $\boldsymbol{\pi}^{\text{myo}}$ with 9.51%–64.66% improvement when $N\lambda < D$, and attains an equivalent performance as π^{myo} when $N\lambda \geq D$. We also observe that π^* significantly outperforms π^{sta} with 1.79%–19.75% improvement. As expected, since π^* includes π^{myo} and π^{sta} as particular cases, π^* provides an upper bound on the throughput performance of them. Meanwhile, we observe that π^* attains an equivalent performance as 1-repetition when $N\lambda \leq D$, and outperforms 1-repetition with 0.48%–36.26% improvement when $N\lambda > D$. These comparisons confirm Lemma 3 and Theorem 4. We note from Figs. 1–2 that the throughput advantage of π^* over π^{myo} becomes noticeable as λ decreases or D increases. This is because the waste of time resources becomes severer in π^{myo}. We also note that the throughput advantage over π^{sta} becomes noticeable as λ decreases or D increases. This is because dynamic access plays a stronger role in improving the throughput under less delivery urgency or less contention intensity. We further note that the throughput advantage over 1-repetition becomes noticeable as λ increases or D decreases. This is because the waste of channel resources becomes severer in 1-repetition.

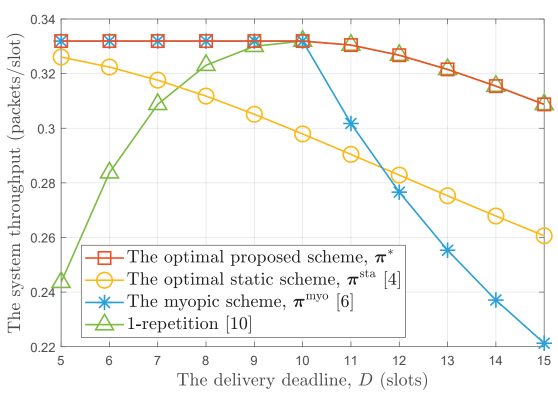

Fig. 2: The system throughput (packets/slot) versus the delivery deadline D when $N = 200$, $\lambda = 0.05$, $\sigma = 0.9$.

The above comparisons indicate that π^* always performs best, which can be attributed to the benefit of π^* combining the advantages of other schemes, that is, fully utilizing the channel resources when $N\lambda \geq D$ as in π^{myo}, and fully utilizing the time resources when $N\lambda \leq D$ as in 1-repetition.

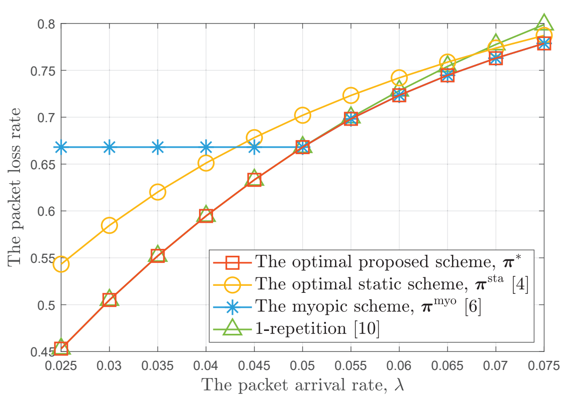

Fig. 3: The packet loss rate versus the packet arrival rate λ when $N = 200$, $D = 10$, $\sigma = 0.9$.

Fig. 3 shows the packet loss rate versus the packet arrival rate λ. We observe that π^* outperforms π^{myo} with 5.51%–47.45% reduction when $N\lambda < D$, 1-repetition with 0.20%–2.58% reduction when $N\lambda > D$, and π^{sta} with 1.09%–19.90% reduction in all the cases. We also observe that π^* attains an equivalent performance as π^{myo} when $N\lambda \geq D$ and as 1-repetition when $N\lambda \leq D$. Once again, these results indicate that π^* always performs best, i.e., it provides a lower bound on the packet loss rate. As λ increases, all the four schemes suffer high packet loss rates due to lack of retransmissions and observations. Nevertheless, this issue can be mitigated by employing the built-in redundancy. For example, in group-based detection [16], an event is considered to be detected as soon as a portion of users have successfully sent their reports for this event.

VI. Conclusion

To provide very low-overhead random access for deadline-constrained periodic traffic, this paper considered p-dynamic slotted ALOHA without observation where each active node adopts time-dependent but observation-independent transmission probabilities. To also achieve very low complexity, based on a blind MDP framework, simple closed-form expressions for optimal time-dependent transmission probabilities have been derived. Our future work would focus on the dynamic optimization of more advanced random access schemes, such as non-orthogonal-multiple-access-based grant-free access [20] and successive-interference-cancellation-based grant-free access [21].

Appendix A
Proof of Lemma 1

We shall prove this result by induction on t from $t = 1$ up to D. First, given N and λ, we have $\mathbf{I}_1 = \mathbf{B}(N, \lambda)$, forming the induction basis. Next, when $t \in \mathcal{T} \setminus \{1\}$, assume $\mathbf{I}_{t-1} = \mathbf{B}(N, \alpha_{t-1})$. For the case $p_t = 1$, we have $\mathbf{I}_t = \mathbf{B}(N, 0)$. For the case $p_t < 1$, by (1) and (2), for each $n' \in \mathcal{N}$, we have

$$
\begin{aligned}
&I_t(n') = \\
&\sum_{n=n'}^{N} \binom{n}{n-n'} p_{t-1}^{n-n'} (1-p_{t-1})^{n'} \binom{N}{n} \alpha_{t-1}^{n} (1-\alpha_{t-1})^{N-n}, \\
&= \binom{N}{n'} (\alpha_{t-1} - \alpha_{t-1}p_{t-1})^{n'} (1 - \alpha_{t-1} + \alpha_{t-1}p_{t-1})^{N-n'} \\
&\quad \times \sum_{m=0}^{N-n'} \binom{N-n'}{m} \left(\frac{\alpha_{t-1}p_{t-1}}{1-\alpha_{t-1}+\alpha_{t-1}p_{t-1}}\right)^{m} \\
&\quad \times \left(1 - \frac{\alpha_{t-1}p_{t-1}}{1-\alpha_{t-1}+\alpha_{t-1}p_{t-1}}\right)^{N-n'-m} \\
&= \binom{N}{n'} (\alpha_{t-1} - \alpha_{t-1}p_{t-1})^{n'} (1 - \alpha_{t-1} + \alpha_{t-1}p_{t-1})^{N-n'}.
\end{aligned}
$$

So, we have $\mathbf{I}_t = \mathbf{B}(N, \alpha_t)$ where $\alpha_t = \alpha_{t-1}(1 - p_{t-1})$.

Appendix B
Proof of (*) in (5)

By plugging the expression of $I_t(n)$ into (4) yields

$$
R_t(\alpha_t, p_t) = \sum_{n=1}^{N} \binom{N}{n} \alpha_t^n (1-\alpha_t)^{N-n} \sigma n p_t (1-p_t)^{n-1}.
$$

So, it is easy to see that $R_t(\alpha_t, p_t) = 0$ when $\alpha_t p_t = 1$. For the case $\alpha_t p_t < 1$, we further have

$$R_t(\alpha_t, p_t) = \sigma N \alpha_t p_t (1-\alpha_t p_t)^{N-1} \sum_{m=0}^{N-1} \binom{N-1}{m} \times \Big(\frac{\alpha_t - \alpha_t p_t}{1-\alpha_t p_t}\Big)^m \Big(1 - \frac{\alpha_t - \alpha_t p_t}{1-\alpha_t p_t}\Big)^{N-1-m} = \sigma N \alpha_t p_t (1-\alpha_t p_t)^{N-1}.$$

Appendix C
Proof of Lemma 2

When α_t is fixed, $R_t(\alpha_t, p_t)$ can be viewed as a function of one variable p_t. By (5), define

$$f(x) \triangleq \sigma N \alpha_t x (1-\alpha_t x)^{N-1}, \quad x \in [0, +\infty). \tag{7}$$

Differentiating $f(x)$ with respect to x yields that

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \sigma N \alpha_t (1 - N\alpha_t x)(1-\alpha_t x)^{N-2}. \tag{8}$$

From (8), for $x \in [0,1]$, we know $f(x)$ achieves its maximum $\sigma(1-\frac{1}{N})^{N-1}$ only at $x = \frac{1}{N\alpha_t}$ if $N\alpha_t \geq 1$, and achieves its maximum $\sigma N \alpha_t (1-\alpha_t)^{N-1}$ only at $x = 1$ otherwise. Further, when $0 < N\alpha_t < 1$, since $f(x)$ is continuous on $[0, \frac{1}{N\alpha_t}]$ and $\frac{\mathrm{d}}{\mathrm{d}x} f(x) > 0$ for $x \in (0, \frac{1}{N\alpha_t})$, we obtain that $f(x)$ is strictly increasing on $[0, \frac{1}{N\alpha_t}]$, implying $f(1) < f(\frac{1}{N\alpha_t}) = \sigma(1-\frac{1}{N})^{N-1}$. The proof of Lemma 2 (i), (ii) is thus completed.

We shall prove Lemma 2 (iii) by induction on t from $t = D$ down to 1. When $t = D$, by (6) and Lemma 2 (i), (ii), we have $U_D^*(\alpha_D) = \max_{p_D \in [0,1]} R_D(\alpha_D, p_D) \leq \sigma(1-\frac{1}{N})^{N-1}$ for $\alpha_D \in \mathcal{A}_D \setminus \{0\}$, forming the induction basis. Next, when $t \in \mathcal{T} \setminus \{D\}$, assume $U_{t+1}^*(\alpha_{t+1}) \leq (D-t)\sigma(1-\frac{1}{N})^{N-1}$. By Lemma 1, Lemma 2 (i), (ii), and (6), we have

$$\begin{aligned} U_t^*(\alpha_t) &= \max_{p_t \in [0,1]} R_t(\alpha_t, p_t) + U_{t+1}^*(\alpha_t - \alpha_t p_t) \\ &\leq \sigma\Big(1-\frac{1}{N}\Big)^{N-1} + (D-t)\sigma\Big(1-\frac{1}{N}\Big)^{N-1} \\ &= (D-t+1)\sigma\Big(1-\frac{1}{N}\Big)^{N-1}, \end{aligned}$$

for $\alpha_t \in \mathcal{A}_t \setminus \{0\}$. This completes the proof.

Appendix D
Proof of Lemma 3

The following notation and definition are useful throughout this proof. Let $U_t^\diamond(\alpha_t, p_t)$ represent the expected total reward obtained at slots $t, t+1, \ldots, D$ under $\mathbf{I}_t = \mathbf{B}(N, \alpha_t)$ when each active node adopts p_t and $(p_{t+1}^*, \ldots, p_D^*)$. So, we have

$$\begin{aligned} &U_D^\diamond(\alpha_t, p_t) = R_D(\alpha_t, p_t), \quad \forall \alpha_t \in \mathcal{A}_D, \\ &U_t^\diamond(\alpha_t, p_t) = R_t(\alpha_t, p_t) + U_{t+1}^*(\alpha_t - \alpha_t p_t), \\ &\qquad\qquad \forall \alpha_t \in \mathcal{A}_t, \forall t \in \mathcal{T} \setminus \{D\}. \end{aligned} \tag{9}$$

We proceed by induction on t from $t = D$ down to 1. First, consider $t = D$. If $N\alpha_D \geq 1$, it follows from (6) and Lemma 2 (i) that $p_D^* = \frac{1}{N\alpha_D}$ and $U_D^*(\alpha_D) = \sigma(1-\frac{1}{N})^{N-1}$. If $0 < N\alpha_D < 1$, it follows from (6) and Lemma 2 (ii) that $p_D^* = 1$ and $U_D^*(\alpha_D) = \sigma N \alpha_D (1-\alpha_D)^{N-1}$. So, the assertion holds for $t = D$.

Assume the assertion is true for $t+1$ with $t \leq D-1$, i.e.,

(i) if $N\alpha_{t+1} \geq D-t$, then $p_{t+1}^* = \frac{1}{N\alpha_{t+1}}$ and $U_{t+1}^*(\alpha_{t+1}) = (D-t)\sigma(1-\frac{1}{N})^{N-1}$; and

(ii) if $0 < N\alpha_{t+1} < D-t$, then $p_{t+1}^* = \frac{1}{D-t}$ and $U_{t+1}^*(\alpha_{t+1}) = \sigma N \alpha_{t+1} (1 - \frac{\alpha_{t+1}}{D-t})^{N-1}$.

First, consider $N\alpha_t \geq D-t+1$. In this case, $N\alpha_t \geq 1$. By Lemma 2 (i), $R_t(\alpha_t, p_t) \leq \sigma(1-\frac{1}{N})^{N-1}$ and the equality occurs only when $p_t = \frac{1}{N\alpha_t}$. When p_t is chosen to be $\frac{1}{N\alpha_t}$, by (3), $N\alpha_{t+1} = N\alpha_t - N\alpha_t p_t \geq D-t$. By (9) and the hypothesis (i), we have $U_t^\diamond(\alpha_t, \frac{1}{N\alpha_t}) = \sigma(1-\frac{1}{N})^{N-1} + (D-t)\sigma(1-\frac{1}{N})^{N-1} = (D-t+1)\sigma(1-\frac{1}{N})^{N-1}$, which attains the optimal value $U_t^*(\alpha_t)$ by Lemma 2 (iii). Therefore, $p_t^* = \frac{1}{N\alpha_t}$, and the result follows.

Now, consider $0 < N\alpha_t < D-t+1$ that can be divided into the following two subcases.

Subcase 1: $0 < N\alpha_t < D-t$. In this subcase, by (3), $N\alpha_{t+1} = N\alpha_t - N\alpha_t p_t < D-t$. It follows from (5), (9), and the hypothesis (ii) that

$$\begin{aligned} U_t^\diamond(\alpha_t, p_t) &= R_t(\alpha_t, p_t) + U_{t+1}^*(\alpha_t - \alpha_t p_t) \\ &= \sigma N \alpha_t p_t (1-\alpha_t p_t)^{N-1} \\ &\quad + \sigma N (\alpha_t - \alpha_t p_t)\Big(1 - \frac{\alpha_t - \alpha_t p_t}{D-t}\Big)^{N-1}. \end{aligned} \tag{10}$$

By viewing (10) as a function of one variable p_t, its derivative is given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}p_t} U_t^\diamond(\alpha_t, p_t) &= \sigma N \alpha_t (1 - N\alpha_t p_t)(1-\alpha_t p_t)^{N-2} \\ &- \sigma N \alpha_t \Big(1 - \frac{N\alpha_t(1-p_t)}{D-t}\Big)\Big(1 - \frac{\alpha_t(1-p_t)}{D-t}\Big)^{N-2}. \end{aligned} \tag{11}$$

Notice that $U_t^\diamond(\alpha_t, p_t)$ is continuous on $p_t \in [0,1]$. Since $D-t \geq 1$, one has $p_t < \frac{1-p_t}{D-t}$ if $p_t < \frac{1}{D-t+1}$ and $p_t > \frac{1-p_t}{D-t}$ if $\frac{1}{D-t+1} < p_t$. It follows from (11) that $U_t^\diamond(\alpha_t, p_t)$ is strictly increasing on $p_t \in [0, \frac{1}{D-t+1}]$ and strictly decreasing on $p_t \in [\frac{1}{D-t+1}, 1]$. Hence, we have $p_t^* = \frac{1}{D-t+1}$, and thus $U_t^*(\alpha_t) = \sigma N \alpha_t (1 - \frac{\alpha_t}{D-t+1})^{N-1}$ by plugging $p_t = \frac{1}{D-t+1}$ into (10).

Subcase 2: $D-t \leq N\alpha_t < D-t+1$. We consider two situations according to the possible values of p_t.

Subcase 2.1: $p_t > 1 - \frac{D-t}{N\alpha_t}$. Since $p_t > 1 - \frac{D-t}{N\alpha_t}$ implies that $N\alpha_t - N\alpha_t p_t < D-t$, we have $N\alpha_{t+1} < D-t$ by (3). It follows from (5), (9), and the hypothesis (ii) that

$$\begin{aligned} U_t^\diamond(\alpha_t, p_t) &= \sigma N \alpha_t p_t (1-\alpha_t p_t)^{N-1} \\ &\quad + \sigma N (\alpha_t - \alpha_t p_t)\Big(1 - \frac{\alpha_t - \alpha_t p_t}{D-t}\Big)^{N-1}. \end{aligned}$$

Note that $1 - \frac{D-t}{N\alpha_t} < \frac{1}{D-t+1}$ due to the assumption $N\alpha_t < D-t+1$. By the same argument in *Subcase 1*, we have

$$U_t^\diamond(\alpha_t, p_t) \leq \sigma N \alpha_t \Big(1 - \frac{\alpha_t}{D-t+1}\Big)^{N-1}, \tag{12}$$

and the equality holds when $p_t = \frac{1}{D-t+1}$.

Subcase 2.2: $p_t \leq 1 - \frac{D-t}{N\alpha_t}$. Since $p_t \leq 1 - \frac{D-t}{N\alpha_t}$ implies that $N\alpha_t - N\alpha_t p_t \geq D-t$, we have $N\alpha_{t+1} \geq D-t$ by (3). It follows from (5), (9), and the hypothesis (i) that

$$\begin{aligned} U_t^\diamond(\alpha_t, p_t) &= R_t(\alpha_t, p_t) + U_{t+1}^*(\alpha_t - \alpha_t p_t) \\ &= \sigma N \alpha_t p_t (1-\alpha_t p_t)^{N-1} + (D-t)\sigma\Big(1-\frac{1}{N}\Big)^{N-1}. \end{aligned}$$

It suffices to maximize $\sigma N\alpha_t p_t(1-\alpha_t p_t)^{N-1}$ to compute $U_t^*(\alpha_t)$. Recall the function $f(x)$ given in (7), which is defined for the maximization of $\sigma N\alpha_t p_t(1-\alpha_t p_t)^{N-1}$. By the argument in the proof of Lemma 2, $f(x)$ is strictly increasing on $[0, \frac{1}{N\alpha_t}]$. Since $1-\frac{D-t}{N\alpha_t} < \frac{1}{N\alpha_t}$ due to the assumption that $N\alpha_t < D-t+1$, we have

$$\begin{aligned} U_t^\diamond(\alpha_t, p_t) &\leq U_t^\diamond\Big(\alpha_t, 1-\frac{D-t}{N\alpha_t}\Big) \\ &= \sigma N\alpha_t\Big(1-\frac{D-t}{N\alpha_t}\Big)\Big(1-\alpha_t\Big(1-\frac{D-t}{N\alpha_t}\Big)\Big)^{N-1} \\ &\quad + (D-t)\sigma\Big(1-\frac{1}{N}\Big)^{N-1}. \end{aligned} \tag{13}$$

In what follows, we shall claim that the upper bound in (12) is strictly larger than the upper bound in (13). If this assertion is true, the maximum value of $U_t^*(\alpha_t)$ occurs in *Subcase 2.1*, and thus the proof is completed.

Observe that $N\alpha_t \geq 1$ due to the assumption $D-t \leq N\alpha_t$ and $t \leq D-1$, and observe that $1-\frac{D-t}{N\alpha_t} < \frac{1}{D-t+1} < \frac{1}{N\alpha_t}$ due to the assumption $N\alpha_t < D-t+1$. Define

$$h(x) \triangleq x(1-\alpha_t x)^{N-1}, \quad \text{for } x \in \Big[0, \frac{1}{N\alpha_t}\Big].$$

The derivative of $h(x)$ is $h'(x) = (1-N\alpha_t x)(1-\alpha_t x)^{N-2}$. For the case $N=2$, we have $h''(x) = -2x < 0$. For the case $N > 2$, we have

$$h''(x) = -\alpha_t(2N-2-N^2\alpha_t x + N\alpha_t x)(1-\alpha_t x)^{N-3}.$$

Let $g(x) \triangleq 2N-2-N^2\alpha_t x + N\alpha_t x$ for $x \in [0, \frac{1}{N\alpha_t}]$. We have $g'(x) = N\alpha_t(1-N) < 0$ for $x \in (0, \frac{1}{N\alpha_t})$. So, $g(x)$ is strictly decreasing on $[0, \frac{1}{N\alpha_t}]$ and $g(x) \geq g(\frac{1}{N\alpha_t}) = N-1 > 0$ for $x \in [0, \frac{1}{N\alpha}]$. We further have $h''(x) = -\alpha_t g(x)(1-\alpha_t x)^{N-3} < 0$. Therefore, $h(x)$ is strictly concave on $[0, \frac{1}{N\alpha_t}]$.

As $1-\frac{D-t}{N\alpha_t} < \frac{1}{D-t+1} < \frac{1}{N\alpha_t}$, by the concavity of $h(x)$ and Jensen's inequality, we have

$$\begin{aligned} &\frac{1}{D-t+1}h\Big(1-\frac{D-t}{N\alpha_t}\Big) + \frac{D-t}{D-t+1}h\Big(\frac{1}{N\alpha_t}\Big) \\ &< h\Big(\frac{1}{D-t+1}\Big(1-\frac{D-t}{N\alpha_t}\Big) + \frac{D-t}{D-t+1}\Big(\frac{1}{N\alpha_t}\Big)\Big) \\ &= h\Big(\frac{1}{D-t+1}\Big), \end{aligned}$$

and thus

$$\begin{aligned} &\frac{1}{D-t+1}\Big(1-\alpha_t\Big(\frac{1}{D-t+1}\Big)\Big)^{N-1} \\ &> \frac{1}{D-t+1}\Big(1-\frac{D-t}{N\alpha_t}\Big)\Big(1-\alpha_t\Big(1-\frac{D-t}{N\alpha_t}\Big)\Big)^{N-1} \\ &\quad + \frac{D-t}{D-t+1}\Big(\frac{1}{N\alpha_t}\Big)\Big(1-\alpha_t\Big(\frac{1}{N\alpha_t}\Big)\Big)^{N-1}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \Big(1-\frac{\alpha_t}{D-t+1}\Big)^{N-1} &> \frac{D-t}{N\alpha_t}\Big(1-\frac{1}{N}\Big)^{N-1} \\ &\quad + \Big(1-\frac{D-t}{N\alpha_t}\Big)\Big(1-\alpha_t\Big(1-\frac{D-t}{N\alpha_t}\Big)\Big)^{N-1}, \end{aligned}$$

and then the result follows.

References

[1] E. Fadel, V. C. Gungor, L. Nassef, N. Akkari, M. A. Malik, S. Almasri, and I. F. Akyildiz, "A survey on wireless sensor networks for smart grid," *Comput. Commun.*, vol. 71, pp. 22–33, 2015.

[2] K. S. Adu-Manu, C. Tapparello, W. Heinzelman, F. A. Katsriku, and J.-D. Abdulai, "Water quality monitoring using wireless sensor networks: Current trends and future research directions," *ACM Trans. Sens. Netw.*, vol. 13, no. 1, pp. 1–41, 2017.

[3] *A 5G traffic model for industrial use cases*. White Paper, 5G Alliance for Connected Industries and Automation, 2019.

[4] L. Deng, J. Deng, P.-N. Chen, and Y. S. Han, "On the asymptotic performance of delay-constrained slotted ALOHA," in *Proc. IEEE ICCCN*, 2018, pp. 1–8.

[5] L. Zhao, X. Chi, L. Qian, and W. Chen, "Analysis on latency-bounded reliability for adaptive grant-free access with multipackets reception (MPR) in URLLCs," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 892–895, 2019.

[6] Y. H. Bae and J. W. Baek, "Age of information and throughput in random access-based IoT systems with periodic updating," *IEEE Wirel. Commun. Lett.*, vol. 11, no. 4, pp. 821–825, 2022.

[7] Y. Zhang, A. Gong, L. Deng, Y.-H. Lo, Y. Lin, and J. Li, "Achieving maximum urgency-dependent throughput in random access," *IEEE Trans. Commun.*, early access, 2022, doi: 10.1109/TCOMM.2023.3305465.

[8] Y. Liu, Y. Deng, M. Elkashlan, A. Nallanathan, and G. K. Karagiannidis, "Analyzing grant-free access for URLLC service," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 741–755, 2020.

[9] J. Ding, M. Nemati, S. R. Pokhrel, O.-S. Park, J. Choi, and F. Adachi, "Enabling grant-free URLLC: An overview of principle and enhancements by massive MIMO," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 384–400, 2022.

[10] S. E. Elayoubi, P. Brown, M. Deghel, and A. Galindo-Serrano, "Radio resource allocation and retransmission schemes for URLLC over 5G networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 896–904, 2019.

[11] A. Gong, Y. Zhang, L. Deng, F. Liu, J. Li, and F. Shu, "Dynamic optimization of random access in deadline-constrained broadcasting," *IEEE Trans. Netw. Sci. Eng.*, 2023.

[12] J. Liu, R. Zhang, A. Gong, and H. Chen, "Optimizing age of information in wireless uplink networks with partial observations," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 4105–4118, 2023.

[13] K. Chatterjee, R. Saona, and B. Ziliotto, "Finite-memory strategies in POMDPs with long-run average objectives," *Math. Oper. Res.*, 2021.

[14] X. Cao, J. Chen, Y. Cheng, X. S. Shen, and Y. Sun, "An analytical MAC model for IEEE 802.15. 4 enabled wireless networks with periodic traffic," *IEEE Trans. Wirel. Commun.*, vol. 14, no. 10, pp. 5261–5273, 2015.

[15] B. Wu, M. D. Lemmon, and H. Lin, "Formal methods for stability analysis of networked control systems with IEEE 802.15. 4 protocol," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 5, pp. 1635–1645, 2017.

[16] Y. Wang, M. C. Vuran, and S. Goddard, "Analysis of event detection delay in wireless sensor networks," in *Proc. IEEE INFOCOM*, 2011, pp. 1296–1304.

[17] A. Fu and M. Mazo, "Traffic models of periodic event-triggered control systems," *IEEE Trans. Autom. Control*, vol. 64, no. 8, pp. 3453–3460, 2018.

[18] Y. Wang, M. C. Vuran, and S. Goddard, "Cross-layer analysis of the end-to-end delay distribution in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 305–318, 2011.

[19] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[20] C.-S. Chang, D.-S. Lee, and C. Wang, "Asynchronous grant-free uplink transmissions in multichannel wireless networks with heterogeneous QoS guarantees," *IEEE/ACM Trans. Netw.*, vol. 27, no. 4, pp. 1584–1597, 2019.

[21] L. Valentini, M. Chiani, and E. Paolini, "Interference cancellation algorithms for grant-free multiple access with massive MIMO," *IEEE Trans. Commun.*, 2023.