Peak Age of Information Optimization of Slotted Aloha: FCFS Versus LCFS

Wen Zhan, *Member, IEEE*, Dewei Wu, *Student Member, IEEE*, Xinghua Sun, *Member, IEEE*, Ziyang Guo, *Member, IEEE*, Peng Liu, *Member, IEEE*, and Jingjing Liu, *Member, IEEE*

***Abstract*—This article aims at optimizing the Peak Age of Information (PAoI) performance of slotted Aloha networks with two representative data queue service disciplines, i.e., first-come first-served (FCFS) and last-come first-served (LCFS). Specifically, by assuming a unit-size buffer and Bernoulli packet arrivals, the average PAoI are derived and minimized by properly tuning the packet arrival rate and the channel access probability of each sensor. Depending on whether the packet arrival rate can be tuned or not, the individual optimization and joint optimization are both considered. The optimal system parameters and corresponding minimum PAoI are explicitly characterized. The analysis shows that in the joint optimization case, the PAoI linearly increases with the network scale and the minimum PAoI with LCFS is 16.8% lower than that with FCFS in the massive access scenario. Yet, for achieving such performance gain over FCFS, each sensor with LCFS should generate a new sample in each time slot, indicating soaring energy consumption due to the sampling operation. The analysis shows that the energy efficiency with LCFS is always lower than that with FCFS, and the gap keeps increasing with the number of sensors, which reveals a clear age-energy tradeoff in terms of service disciplines.**

***Index Terms*—Age of information, Aloha networks, channel access probability, packet arrival rate.**

I. Introduction

THE information freshness is crucial for many real-time energy-limited Internet of Things (IoT) applications, such as vehicle tracking and e-health, where fast and accurate system responses are vital. However, in large-scale IoT networks, due to the broadcast nature of the wireless channel, the fierce channel competition may lead to frequent packet transmission failures that deteriorate the information freshness. Considering the explosive growth of the time-sensitive IoT application market [1], [2], [3], [4], it is urgent to study how to optimize information freshness performance of the large-scale distributed IoT networks.

To characterize the timeliness of information, Age-of-Information (AoI) was proposed in [5], which measures the time elapsed since the last received packet was generated. AoI captures the impact of both transmission delay and data generation frequency, and has attained significant attention in recent years, where average AoI and Peak AoI (PAoI) are the two most popular AoI performance metrics. The latter represents the maximum value of AoI achieved just before an update is received. Meanwhile, PAoI is of interest in applications with timeliness threshold restrictions, and therefore will be the focus of this article.

A. AoI Optimization in Slotted Aloha

For large-scale IoT systems, where hundreds of sensors are deployed over a large region, the centralized access control strategy is impractical due to intolerably high coordination costs. Random access schemes, such as slotted Aloha, have proven to be a simple yet elegant solution, where each sensor independently determines when to transmit. The minimum coordination and distributed control enable Aloha to become one of the most widely deployed access schemes in IoT-originated wireless networks such as LoRa and NB-IoT [6], [7].

For decades, extensive works have been done on the throughput and delay analysis of Aloha networks [8], [9], [10], [11], [12], [13]. It was found that the performance optimization of the network lies in the proper selection of the channel access probability and packet arrival rate (i.e., the sampling rate in the context of information update system [14]). A large channel access probability (resp. packet arrival rate) leads to mounting channel contention and poor performance, while a small channel access probability (resp. packet arrival rate) results in an insufficient use of channel resource. However, the search for the optimal parameter configuration has long been known as notoriously difficult. This is because the service rate of each sensor's data queue depends on the aggregate activities of all sensors as the packet transmission of a sensor can be successful if and only if all other sensors keep silent, and consequently each sensor's data queue is coupled with each other. The performance optimization becomes even more complicated when the inherent bi-stable behavior problem is considered [10]. That is, the Aloha network has two steady-state points (i.e., the probability of successful

Manuscript received 31 July 2022; revised 1 March 2023; accepted 24 April 2023. Date of publication 2 May 2023; date of current version 25 October 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFA0716600, in part by the Shenzhen Science and Technology Innovation Program under Grant RCBS20210706092408010, and in part by the National Natural Science Foundation of China under Grants 62001524 and 62174181. Recommended for acceptance by Dr. Xiaoli Chu. *(Corresponding author: Xinghua Sun.)*

Wen Zhan, Dewei Wu, Xinghua Sun, and Jingjing Liu are with the School of Electronics and Communication Engineering, Shenzhen Campus of Sun Yat-sen University, Shenzhen 518063, China (e-mail: zhanw6@mail.sysu.edu.cn; wudw5@mail2.sysu.edu.cn; sunxinghua@mail.sysu.edu.cn; liujj77@mail.sysu.edu.cn).

Ziyang Guo and Peng Liu are with the Huawei Technologies Co., Ltd., Shenzhen 518063, China (e-mail: guoziyang@huawei.com; jeremy.liupeng@huawei.com).

Digital Object Identifier 10.1109/TNSE.2023.3272360

transmissions of packets) and suffers the risk of dropping to the lower one, on which the network performance becomes poor. As such, numerous works have been done on characterizing the bi-stable region of Aloha networks, which is found to be determined by various system parameters including the packet arrival rate and the channel access probability [15], [16].

The AoI optimization of Aloha networks gains significant attention recently. To push the age performance to the limit, the generate-at-will traffic model, where sensors generate new samples whenever they decide to transmit, has been widely used [17], [18], [19], [20]. In [17], the average AoI of slotted Aloha was characterized and optimized by properly choosing the channel access probability. [18] considered a multi-channel Aloha network, and optimized the age performance via joint access control and resource allocation. To make the best use of each update transmission, threshold-based random access schemes were proposed, in which channel access request of a sensor is permitted only if the information age exceeds a predetermined threshold [19], [20].

Although the generate-at-will traffic model benefits the age performance by generating updates immediately before transmissions, it is not in line with many practical IoT services, such as smart grid [21], where the updates come into the data queue periodically or stochastically, instead of the on-demanding manner. For instance, in smart grid, the reporting period of devices ranges from 5 mins to 24 hours depending on the practical needs [22]. In [23] and [24], the optimal transmission probabilities were searched based on an iterative algorithm assuming Bernoulli arrival with a fixed packet arrival rate. However, this approach is too demanding for practical low-end IoT devices. Note that besides the channel access probability, the packet arrival rate can also be a tunable system parameter. A question naturally arises: How to jointly select the packet arrival rate and channel access probability of each sensor for the age performance optimization of slotted Aloha? Although the above works have focused on optimally choosing the channel access probability, few have considered joint tuning when the bi-stable behavior of Aloha networks is included.

B. Age Performance in Terms of Service Disciplines: FCFS Versus LCFS

The analysis of the classical queueing systems, such as M/M/c, M/M/∞ and Geo/Geo/c [25], [26], [27], [28], [29], have shown that the age performance of the network crucially depends on the service disciplines of the data queue, which mainly contains two categories: 1) First-come first-served (FCFS), i.e., all data packets are served in the order of arrival; 2) Last-come first-served (LCFS), i.e., newly arrived packets have priority and will be served first. It is revealed in [30], [31] that the system could achieve better age performance with LCFS than that with FCFS, especially when the packet arrival rate is large.

Yet, for FCFS, a large packet arrival rate does not necessarily contribute to age optimization, as the queue would be backlogged and queueing delay is enlarged [29]. On the other hand, with LCFS, although increasing the packet arrival rate will always reduce the average age, frequent information updates induce excessive energy consumption. Towards this issue, lots of endeavors have been devoted to searching for the optimal packet arrival rate in battery-limited networks [32], [33].

For slotted Aloha, the age performance with different service disciplines has been investigated and compared. In [34], with the FCFS service discipline, the PAoI was characterized based on a spatiotemporal model, and the optimal channel access probability was derived as a function of local topology information, by leveraging the dominant system approach. [35] extended the analysis to the random walk case, and obtained the optimal channel access probability and the packet arrival rate for PAoI minimization, and the age-energy tradeoff is analyzed in [36]. In [37], the average AoI and PAoI were evaluated for both FCFS and LCFS service disciplines, where the effectiveness of LCFS over FCFS in reducing AoI was confirmed.

Note that the studies in [34], [35], [36], [37] consider the ad-hoc scenario, where multiple transmitter-receiver pairs coexist and compete with each other. While, little research has been done on the optimal age performance comparison with different service disciplines in the multiple access scenario, where multiple nodes transmit to a common receiver. It is thus unknown whether in the multiple access scenario, the age performance with LCFS is better than that with FCFS, and how much performance gain that the network can achieve. Additionally, what is the price to pay, particularly from the energy efficiency point of view?

C. Contribution

To address the above open issues, in this article, we consider a slotted Aloha network where the packet arrival process follows Bernoulli distribution and each sensor is equipped with a unit-size buffer. By extending the analysis in [15], we take the bi-stable behavior of slotted Aloha into account and obtain the explicit expression of average PAoI. Based on the bi-stable region, the channel access probability and the packet arrival rate for PAoI optimization are carefully selected such that the network would not fall into the bi-stable region, avoiding the risk of dropping to the lower steady-state point. In specific, our contributions are summarized as follows:

- *Individual optimization:* Given the packet arrival rate, we obtain the minimum PAoI and corresponding optimal channel access probabilities in both FCFS and LCFS cases. The analysis shows that the optimal channel access probability for PAoI minimization is independent of the service disciplines. In contrast, the minimum PAoI with LCFS would be smaller than that with FCFS, and the gap grows with the number of sensors and the packet arrival rate.
- *Joint optimization:* When the packet arrival rate can be tuned, the minimum PAoI and corresponding optimal packet arrival rate and channel access probability are derived as explicit functions of the number of sensors in both FCFS and LCFS cases. It is found that with FCFS, the optimal packet arrival rate and channel access probability decrease with the number of sensors. Yet, with LCFS, the optimal packet arrival rate remains constant regardless of the network scale. On the other hand, the minimum PAoI linearly increases with the number of sensors, and the

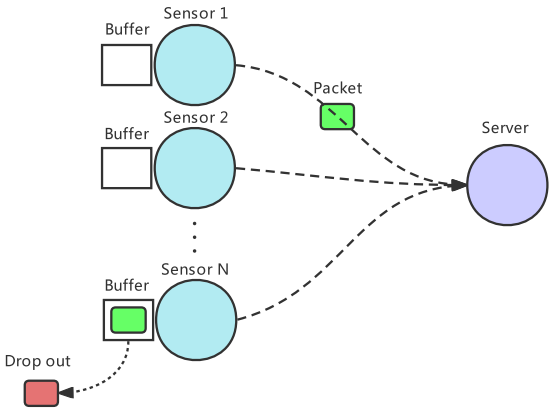

Fig. 1. An Aloha network with multiple sensors and one server.

minimum PAoI with LCFS is 16.8% lower than the PAoI with FCFS in the massive access scenario.

- *Age-energy tradeoff:* We evaluate the energy efficiency of each sensor in the joint optimization case with different service disciplines, and find that the energy efficiency with LCFS is always lower than that with FCFS. For instance, by adopting LCFS discipline to achieve 16.8% PAoI gain compared with FCFS, approximately 100 times more energy is consumed if the sampling energy consumption is half of the transmission energy consumption in a 200-node network, and the energy efficiency with LCFS is approximate 1.05% of that with FCFS. Moreover, the gap in energy efficiency between FCFS and LCFS keeps increasing with the growth of the network scale, indicating a clear tradeoff between the age performance and the energy efficiency in terms of service disciplines.

The rest of this article is organized as follows. The preliminary analysis of steady-state points and PAoI are presented in Section II. Both the individual optimization and the joint optimization for PAoI minimization are addressed in Section III with FCFS discipline and Section IV with LCFS discipline, respectively. Section V compares the minimum PAoI in FCFS case and that in LCFS case. The age-energy tradeoff is characterized in Section VI. Concluding remarks are summarized in Section VII.

II. System Model and Preliminary Analysis

Consider a slotted Aloha network containing n sensors and one common server, and each sensor is equipped with a unit-size buffer, as shown in Fig. 1. Assume that all the sensors are synchronized and can start a transmission only at the beginning of a time slot, and each transmission occupies one time slot. Similar to [20], [24], we consider the collision model where the packets generated by sensors are transmitted over a noiseless channel, and each packet can be successfully transmitted when there is no concurrent transmission from other sensors.

Assume that with probability $\lambda \in (0, 1]$, each sensor performs the sampling operation at the beginning of each time slot for generating a new packet (i.e., update). When the buffer is non-empty, the sensor will access the channel in each time slot with probability q until the packet is successfully transmitted. We assume the ACK/NACK transmission from the destination is instantaneous and error-free [37], [38], [39]. In this article, the FCFS discipline and the LCFS discipline are both considered:

- With FCFS, a newly arrived packet will be dropped out if the buffer is full.
- With LCFS, a newly arrived packet replaces the Head-of-Line (HoL) packet of the buffer if the buffer is full[1].

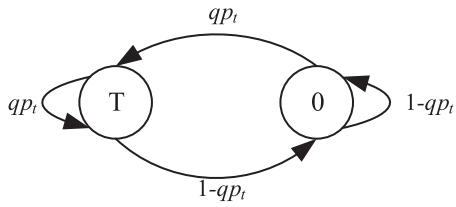

Fig. 2. State transition diagram of each individual HoL packet.

A. Steady-State Points

Note that a discrete-time Markov chain has been established in [15] for characterizing the behavior of the HoL packet of each sensor, as shown in Fig. 2. Let p_t denote the probability of successful transmission of HoL packets at time slot $t = 1, 2, \ldots$. A fresh packet is initially in State T, and remains in State T if it is successfully transmitted with probability qp_t. If it is transmitted but encounters a collision, then it goes to State 0. In State 0, if it is successfully transmitted with probability qp_t, then it shifts to State T. The Markov chain is uniformly strongly ergodic if and only if the following limit exists [40]:

$$\lim_{t\to\infty} p_t = p. \tag{1}$$

We can derive the steady-state probability distribution of the Markov chain: $\pi_T = pq$ and $\pi_0 = 1 - pq$, based on which the fixed-point equation of the steady-state probability of successful transmission of each packet p is obtained in [15] as

$$p = \exp\left(\frac{-\hat{\lambda}q}{\lambda + pq}\right), \tag{2}$$

where $\hat{\lambda} = n\lambda$ is the aggregate packet arrival rate.

Equation (2) may have more than one roots. The analysis in [15] reveals that the network has either two steady-state points, i.e., the desired steady-state point p_L and the undesired steady-state point p_A with $p_L > p_A$, or one steady-state point p_L. Both p_L and p_A are the roots of (2). The number of roots in (2) is determined by the number of sensors n, the packet arrival rate λ and the channel access probability q. Accordingly, we can define the following stable regions in terms of (n, q, λ):

- **Bi-stableregion** $\mathcal{B} = \{(n, q, \lambda) | n > \frac{4}{q}, \lambda_1 < \lambda < \lambda_2\}$, in which the network has two steady-state points p_A and

[1]Packet preemption is adopted here, i.e., even when the HoL packet is in service, the newly arrived packet can replace it.

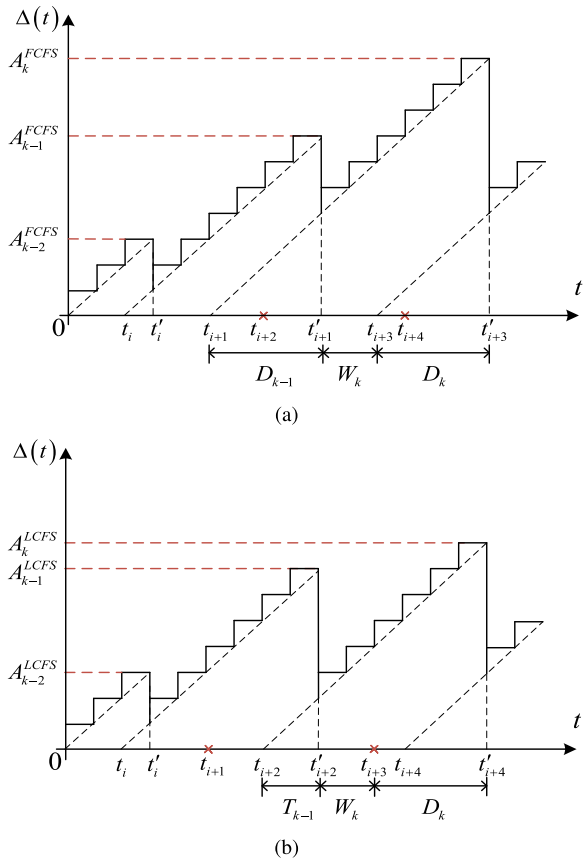

Fig. 3. AoI evolution of a sensor with (a) FCFS discipline, (b) LCFS discipline.

p_L. The boundaries λ_1 and λ_2 are given by (7) and (8) in [15], i.e.,

$$\begin{cases} \lambda_1 = \dfrac{2}{n\left(1-\frac{2}{nq}-\sqrt{1-\frac{4}{nq}}\right)\cdot\exp\left(\frac{2}{1-\sqrt{1-\frac{4}{nq}}}\right)} \\ \lambda_2 = \dfrac{2}{n\left(1-\frac{2}{nq}+\sqrt{1-\frac{4}{nq}}\right)\cdot\exp\left(\frac{2}{1+\sqrt{1-\frac{4}{nq}}}\right)}. \end{cases} \tag{3}$$

- **Mono-stableregion** $\mathcal{M} = \bar{\mathcal{B}}$, in which the network has only one steady-state point p_L. Note that the network stays in either the mono-stable region $\mathcal{M}$ or bi-stable region $\mathcal{B}$. Accordingly, $\mathcal{M}$ is the complementary set of bi-stable region $\mathcal{B}$.

B. Peak Age of Information

This article focuses on the peak age of information performance of the network. For illustration, Fig. 3(a) and (b) depict the evolution of AoI $\Delta(t)$ over time t with FCFS and LCFS, respectively, where t_i denotes the time instance at which ith packet arrives and t'_i denotes the time instance that i^{th} packet is successfully transmitted, $i \in \{1, 2, \ldots\}$. Note that due to finite buffer size, some packets are dropped. To avoid ambiguity, we use k to indicate the k^{th} successfully transmitted packets. As only successfully-transmitted packets contribute to the age performance, we refer to them as informative packets [41], [42].

As Fig. 3 illustrates, the AoI grows over time. When a packet is successfully transmitted, AoI reduces to the time elapsed since the generation of the delivered packet, i.e., $\Delta(t) = t - u(t), t \in \{1, 2, \cdots\}$, where $u(t) = \max\{t_i | 0 < t'_i < t\}$. The PAoI, denoted by A_k, is the AoI achieved immediately before receiving the k^{th} packet. It is clear that the PAoI crucially depends on the service disciplines.

With FCFS, we can see from Fig. 3(a) that the packet arrivals at t_{i+2} is dropped because the buffer is full. The age performance only relies on the packet which arrives when the buffer is empty, i.e., the i^{th} packet, the $i+1^{\text{th}}$ packet and the $i+3^{\text{th}}$ packet in Fig. 3(a). Accordingly, upon the successful reception of the k^{th} packet, the PAoI A_k^{FCFS} can be written as

$$A_k^{FCFS} = D_{k-1} + W_k + D_k, \tag{4}$$

where

- D_k denotes the access delay, i.e., the interval from the arrival of the first packet after the successful transmission of the $k-1^{\text{th}}$ informative packet until the successful transmission of the k^{th} informative packet.
- W_k denotes the idle period, which is the time interval from the successful transmission of $k-1^{\text{th}}$ informative packet until the arrival of the next packet.

With LCFS, we can see from Fig. 3(b) that the $i+1^{\text{th}}$ packet is dropped and replaced by a fresher packet, i.e., the $i+2^{\text{th}}$ packet. The age performance only relies on the most up-to-date packet, i.e., the i^{th} packet, the $i+2^{\text{th}}$ packet and the $i+4^{\text{th}}$ packet in Fig. 3(b). Accordingly, the PAoI A_k^{LCFS} can be written as

$$A_k^{LCFS} = T_{k-1} + W_k + D_k, \tag{5}$$

where

- T_k is the service time of the k^{th} informative packet, i.e., the interval from the arrival of k^{th} informative packet to its successful transmission.

By comparing (4) and (5), we can see that the difference lies in the first item on the right side of the equations, i.e., D_k and T_k. Note that D_k is the time spent from the generation of one packet when the queue is empty until the queue is empty again. It should be pointed out that after the generation of the first packet when the queue is empty, many other packets may arrive. While, T_k only refers to the service time of the last packet (i.e., the packet to be successfully transmitted) to arrive in the queue. For instance, in Fig. 3(b), T_{k-1} is the interval $[t_{i+2}, t'_{i+2}]$ and D_{k-1} is the interval $[t_{i+1}, t'_{i+2}]$. Therefore, the physical meaning of D_k and that of T_k are distinct and $T_k \leq D_k$ holds. Moreover, as T_k, W_k and D_k for $k \in \{1, 2, \ldots\}$ are i.i.d. random variables, we drop the subscript k of T_k, W_k and D_k for simplicity.

C. Problem Definition

In this article, our objective is to minimize the average PAoI[2] A via optimally tuning the channel access probability q and the packet arrival rate λ, where $A = \lim_{N\to\infty} \frac{\sum_{k=1}^{N} A_k}{N}$. Note that

[2]The average AoI can also be characterized based on the analytical approach in this article. How to optimize the average AoI performance is another important issue that will be addressed in our future work.

the channel access probability is a MAC layer parameter that can be dynamically adjusted, while the packet arrival rate is usually an application-dependent parameter. Accordingly, we consider two optimization problems. The first one is the individual optimization problem, where the network aims at minimizing the PAoI A by individually tuning q given the packet arrival rate λ, i.e.,

$$A^*|_\lambda = \min_{0<q\le 1} A. \tag{6}$$

If the packet arrival rate λ can be tuned, then we have the joint optimization problem, where the PAoI is minimized by jointly tuning q and λ, i.e.,

$$A^* = \min_{0<q\le 1, 0<\lambda\le 1} A. \tag{7}$$

In the following sections, we consider the above optimization problems in FCFS and LCFS cases, respectively.

III. PAoI Optimization With FCFS

Let us start by deriving the expression of PAoI A^{FCFS} with FCFS. The analysis in [16] has revealed that the mean access delay is given by

$$E[D] = \frac{1}{qp}. \tag{8}$$

Since the departure of $i-1^{\text{th}}$ successfully transmitted packet and the arrival of a new packet can occur at the same time, we can have

$$E[W] = \frac{1}{\lambda} - 1, \tag{9}$$

Based on (4), (8)–(9), we can then obtain

$$A^{FCFS} = \frac{2}{qp} + \frac{1}{\lambda} - 1. \tag{10}$$

A. Individual Optimization

The following theorem presents the optimal channel access probability $q^*|_\lambda$ that minimizes the PAoI when the packet arrival rate and the number of sensors are given.

Theorem 1: With FCFS, given the aggregate packet arrival rate $\hat{\lambda}$, the minimum PAoI is given by

$$A^{FCFS*}|_\lambda = \begin{cases} 2e\cdot n - \frac{1}{\lambda} - 1 & \hat{\lambda} > \hat{\lambda}_0, \\ \frac{n\left(-2\mathbb{W}_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)-1\right)}{2\mathbb{W}^2_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)p^*_L} + \frac{1}{\lambda} - 1 & \text{otherwise}, \end{cases} \tag{11}$$

which is achieved when the channel access probability

$$q^*|_\lambda = \begin{cases} \frac{\lambda}{n\lambda - e^{-1}} & \hat{\lambda} > \hat{\lambda}_0, \\ \frac{4\mathbb{W}^2_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)}{n\left(-2\mathbb{W}_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)-1\right)} & \text{otherwise}, \end{cases} \tag{12}$$

where $\hat{\lambda}_0 \approx 0.48$, $W_{-1}(\cdot)$ is the secondary branch of the Lambert W function, and p^*_L is the non-zero root of the following equation

$$p^*_L = \exp\left(\frac{-4n\lambda\mathbb{W}^2_{-1}(-\frac{\sqrt{n\lambda}}{2})}{4p^*_L\mathbb{W}^2_{-1}(-\frac{\sqrt{n\lambda}}{2}) + n\lambda\left(-2\mathbb{W}_{-1}(-\frac{\sqrt{n\lambda}}{2})-1\right)}\right). \tag{13}$$

Proof: Given the aggregate packet arrival rate $\hat{\lambda}$, it can be seen from (8) and (10) that PAoI minimization is equivalent to mean access delay minimization. Therefore, the optimal channel access probability for minimizing the mean access delay can also minimize the PAoI. The optimal channel access probability has been obtained in [16], as shown by (12). The minimum PAoI can then be obtained by substituting (12) into (2) and (10). ■

Theorem 1 reveals that the optimal channel access probability $q^*|_{\hat{\lambda}>\hat{\lambda}_0}$ is $\frac{\lambda}{n\lambda-e^{-1}}$ when $\hat{\lambda} > \hat{\lambda}_0$, where $\hat{\lambda}_0 \approx 0.48$ is the single non-zero root of the equation $\hat{\lambda} - \hat{\lambda}(1+1/\mathbb{W}_{-1}(-\sqrt{\hat{\lambda}}/2))^2 = 4(\hat{\lambda} - e^{-1})$, which is derived by combining $q^*|_{\hat{\lambda}>\hat{\lambda}_0} = q_1$ and q_1 is given by (10) in [16]. $\hat{\lambda}_0$ is the threshold of the aggregate packet arrival rate $\hat{\lambda}$, above which the network is guaranteed to operate at the desired steady-state point p_L; otherwise, the network has two steady-state points. If $\hat{\lambda} < \hat{\lambda}_0$, then $q^*|_\lambda$ is set to be $\frac{4\mathbb{W}^2_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)}{n\left(-2\mathbb{W}_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)-1\right)}$, which is the root of $\lambda = \lambda_1$ for avoiding the network falls into the bi-stable region $\mathcal{B}$, where λ_1 is given in (3).

Fig. 4 depicts how the PAoI A^{FCFS} varies with the channel access probability for $n = 100$ with the aggregate packet arrival rate $\hat{\lambda} = 0.4$ or 0.8. The simulation setup follows Section II and each case runs for 10^7 time slots via a MATLAB-based simulator. With $\hat{\lambda} = 0.4 \le \hat{\lambda}_0$, it can be seen from Fig. 4(a) that the PAoI decreases as q increases when the network stays in mono-stable region $\mathcal{M}$. When $q > q^*|_\lambda \approx 0.047$, the network falls into bi-stable region $\mathcal{B}$, in which it has two steady-state points, i.e., the desired steady-state point p_L and the undesired steady-state point p_A. Once the network shifts to the undesired steady-state point p_A, the PAoI A rises sharply. To hedge against this case, the optimal channel access probability $q^*|_\lambda$ is set to the region boundary, i.e., the root of $\lambda = \lambda_1$. On the other hand, if the aggregate packet arrival rate increases to 0.8, as shown in Fig. 4(b), then the network always stays in mono-stable region $\mathcal{M}$ and operates at the desired steady-state point p_L. The minimum PAoI can be achieved when the channel access probability $q^*|_\lambda = \frac{\lambda}{n\lambda-e^{-1}} \approx 0.018$. Simulation results presented in Fig. 4 well agree with the analysis.

B. Joint Optimization

Note that if the packet arrival rate of each sensor λ can be adjusted, then the network can jointly tune q and λ for PAoI optimization. The following theorem presents the optimal channel access probability q^*_F and packet arrival rate λ^*_F for PAoI optimization.

Theorem 2: With FCFS, the minimum PAoI is given by

$$A^{FCFS*} \approx 3.27n - 1, \tag{14}$$

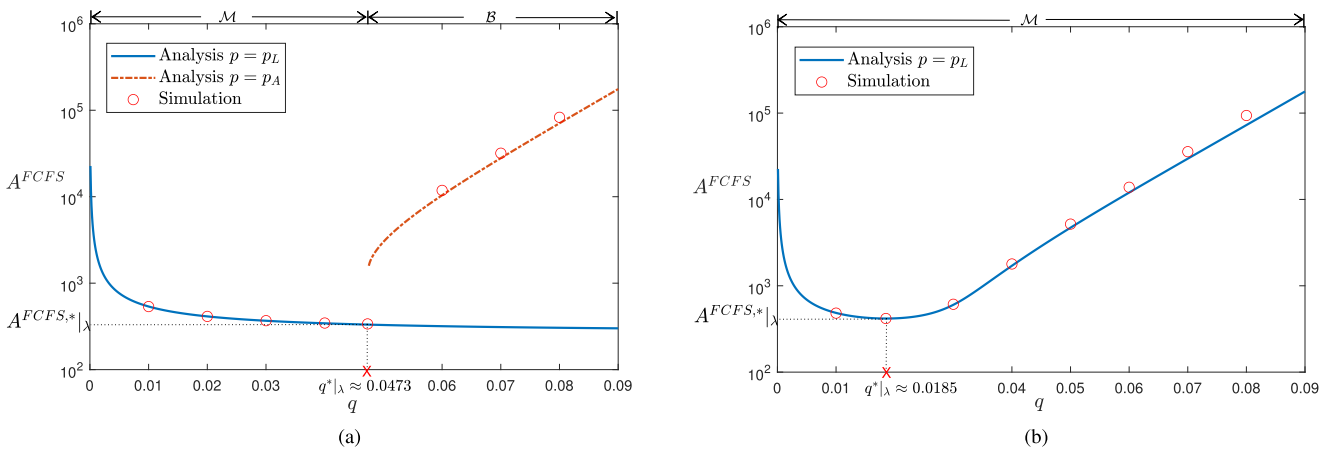

Fig. 4. PAoI A^{FCFS} versus the channel access probability q with FCFS discipline, $n = 100$. (a) $\hat{\lambda} = 0.4$. (b) $\hat{\lambda} = 0.8$.

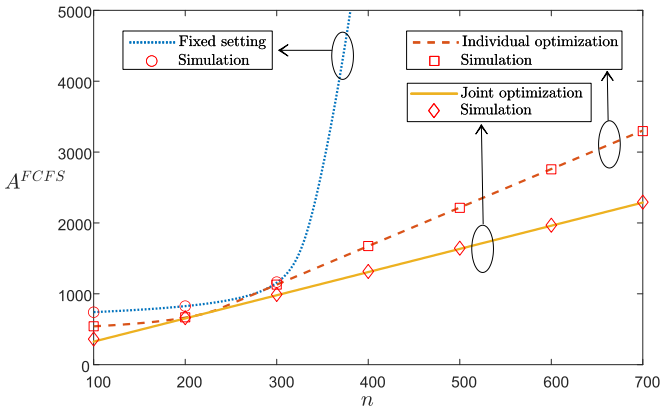

Fig. 5. PAoI A^{FCFS} versus n in three different cases. $q = 0.01$. $\lambda = 0.002$.

which is be achieved when the channel access probability and the packet arrival rate are set to be

$$\begin{cases} q_F^* \approx \frac{4.543}{n}, \\ \lambda_F^* \approx \frac{0.4395}{n}. \end{cases} \tag{15}$$

Proof: See Appendix A. ■

Theorem 2 reveals that as the number of sensors n increases, both the optimal channel access probability q_F^* and packet arrival rate λ_F^* should be properly reduced in order to alleviate the contention. Meanwhile, the minimum PAoI A^{FCFS*} linearly increases with n.

To evaluate the performance gain brought by the optimal tuning, Fig. 5 demonstrates how the PAoI A^{FCFS} varies with n in three cases: 1) fixed setting with $\lambda = 0.002$ and $q = 0.01$, 2) individual optimization with $\lambda = 0.002$ and 3) joint optimization. It can be seen from Fig. 5 that with fixed setting, the PAoI exponentially increases with n, indicating that the age becomes intolerably high in the massive access scenarios. In contrast, with either individual optimization or joint optimization, the PAoI linearly grows with n. The performance gain becomes significant, especially when n is large or both the channel access probability q and the packet arrival rate λ can be dynamically tuned.

IV. PAoI Optimization With LCFS

Let us now consider the LCFS discipline. Note that different from the FCFS case, with LCFS, the i^{th} successfully transmitted packet may not be the first packet arrives after the successful transmission of the $i-1^{\text{th}}$ informative packet, since packets are discarded once a new packet arrives. Accordingly, by comparing Fig. 3(b) with Fig. 3(a), we can see that $T_k \leq D_k$. T_k denotes the service time of k^{th} successfully transmitted packet and D_k denotes the time interval from the arrival of a packet when the buffer is empty until the successful transmission of the k^{th} informative packet.

To derive the PAoI under the LCFS discipline A^{LCFS}, let us first focus on the access delay D and the service time T. Specifically, the probability mass function (PMF) of D can be obtained based on (1) in [16] as

$$P\{D = j\} = \begin{cases} qp, & j = 1, \\ (1-qp)^{j-1}qp, & j \geq 2. \end{cases} \tag{16}$$

Given the access delay D, the service time T should be lower than or equal to D. Assume that the access delay $D = j$ and the service time $T = i$, there are following two cases: 1) $i = j$, indicating that the informative packet is the packet arrives at the queue when the queue is empty. Since no packet arrives in the following $i-1$ time slots, the probability of this case is given by $(1-\lambda)^{i-1}$; 2) $i < j$, indicating that the informative packet arrives after the first packet arrives at the queue when the queue is empty. The probability of this case is given by $\lambda(1-\lambda)^{i-1}$ as no packet is generated after the arrival of the informative packet. Accordingly, the conditional probability of the service time $T = i$ given the access delay $D = j$, $P\{T = i|D = j\}$, can be expressed as

$$P\{T = i|D = j\} = \begin{cases} (1-\lambda)^{i-1}, & i = j \\ \lambda(1-\lambda)^{i-1}, & 1 \leq i < j, \end{cases} \tag{17}$$

where $i, j \in \{1, 2, \ldots\}$. By combining (16) and (17), the PMF of the service time T can be obtained as

$$P\{T=i\} = \sum_{j=i}^{+\infty} P\{T=i|D=j\}P\{D=j\}$$
$$= ((1-\lambda)(1-qp))^{i-1}(qp+\lambda(1-qp)). \quad (18)$$

Then the mean service time $E[T]$ can be obtained as

$$E[T] = \sum_{i=1}^{+\infty} P\{T=i\}\cdot i = \frac{1}{qp+(1-qp)\lambda}. \quad (19)$$

Finally by combing (5), (8), and (19), the PAoI under the LCFS discipline can be written as

$$A^{LCFS} = \frac{1}{qp} + \frac{1}{qp+(1-qp)\lambda} + \frac{1}{\lambda} - 1. \quad (20)$$

Comparison of A^{FCFS} in (10) with A^{LCFS} in (20) brings $A^{FCFS} - A^{LCFS} = \frac{1}{qp} - \frac{1}{qp+(1-qp)\lambda} > 0$, implying that the network can always obtain better age performance with LCFS discipline. However, as we will show in Section VI, such performance gain is achieved at the cost of additional energy consumption, which could be the key concern in battery-limited scenarios.

A. Individual Optimization

Similar to the PAoI optimization in FCFS case, in LCFS case, we are also interested in tuning the system parameters, i.e., q and λ, for minimizing A^{LCFS}. The following theorem presents the optimal channel access probability $q^*|_\lambda$ that minimizes the PAoI when the packet arrival rate λ and the number of sensors n are given.

Theorem 3: With LCFS, given the aggregate packet arrival rate $\hat{\lambda}$, the minimum PAoI is given by

$$A^{LCFS*}|_\lambda = \begin{cases} e\cdot n + \frac{n\lambda - e^{-1}}{\lambda^2(n-e^{-1})} - 1 & \hat{\lambda} > \hat{\lambda}_0, \\ \frac{\alpha}{1+\lambda(\alpha-1)} + \alpha + \frac{1}{\lambda} - 1 & \text{otherwise}, \end{cases} \quad (21)$$

which is achieved when the channel access probability is set to be $q^*|_\lambda$ in (12), where $\alpha = \frac{n\left(-2\mathbb{W}_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)-1\right)}{4\mathbb{W}_{-1}^2\left(-\frac{\sqrt{n\lambda}}{2}\right)p_L^*}$.

Proof: As shown in (8) and (20), the PAoI is minimized when the access delay is minimized. Accordingly, similar to the proof of Theorem 1, the optimal channel access probability is given by (12), based on which (21) can be derived. ■

It is interesting to see that the optimal channel access probability $q = q^*|_\lambda$ for minimizing A^{FCFS} and that for minimizing A^{LCFS} are the same. Specifically, according to (10) and (20), we can observe that given the aggregate packet arrival rate $\hat{\lambda}$, the PAoI is determined by the access delay $E[D] = 1/qp$, which is the time length from the generation of a packet when the buffer is empty until the buffer becomes empty again. The access delay is independent of whether the HoL packet is replaced by the newly arrived packet or not, but solely depends on the channel access probability q. Therefore, $q^*|_\lambda$ for minimizing the access delay can minimize A^{FCFS} as well as A^{LCFS}.

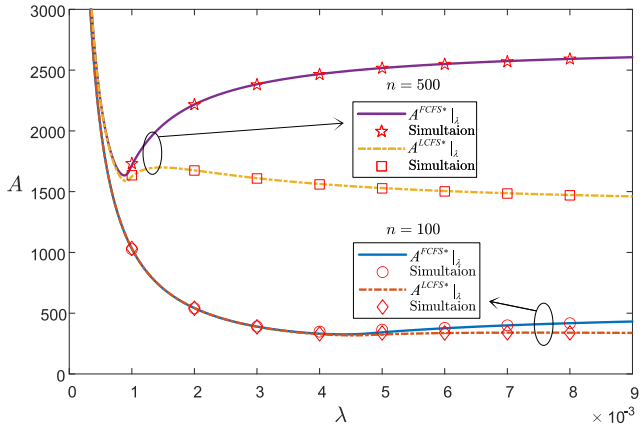

Fig. 6. PAoI $A^{FCFS*}|_\lambda$ and $A^{LCFS*}|_\lambda$ versus the packet arrival rate λ, $n \in \{100, 500\}$, $q = q^*|_\lambda$.

B. Joint Optimization

With joint tuning, the following theorem presents the optimal channel access probability q_L^* and packet arrival rate λ_L^* for PAoI optimization.

Theorem 4: With LCFS, the minimum PAoI is given by

$$A^{LCFS*} = e\cdot n + 1, \quad (22)$$

which is achieved when the channel access probability and the packet arrival rate are set to be

$$\begin{cases} q_L^* = \frac{1}{n-e^{-1}}, \\ \lambda_L^* = 1. \end{cases} \quad (23)$$

Proof: See Appendix B. ■

Theorem 4 reveals that to achieve the minimum PAoI with LCFS discipline, the packet arrival rate $\lambda_L^* = 1$, that is, each sensor receives a new update at each time slot. It is in sharp contrast to the FCFS case, where the optimal packet arrival rate $\lambda_F^* \approx \frac{0.4395}{n} \ll 1$. The reason stems from packet preemption under LCFS discipline, where a larger packet arrival rate makes the HoL packets contain fresher information. While, with FCFS, the packet arrival rate has to be reduced for reliving the channel contention.

V. Optimal PAoI Performance Comparison

So far, we have obtained the minimum PAoI and the corresponding optimal settings of channel access probability and the packet arrival rate in both FCFS and LCFS cases. There is a general consensus that the network achieves better age performance with LCFS than that with FCFS. Yet, how much the age performance gain can be achieved needs further study.

Let us start with the individual optimization case when only the channel access probability is optimally tuned. Fig. 6 illustrates how PAoI $A^*|_\lambda$ varies with the packet arrival rate λ when $q = q^*|_\lambda$ and $n \in \{100, 500\}$. It can be seen that when the packet arrival rate λ is small, the age performance is similar in FCFS and LCFS. Intuitively, with small λ and optimal channel access probability, the successful packet transmission probability would be

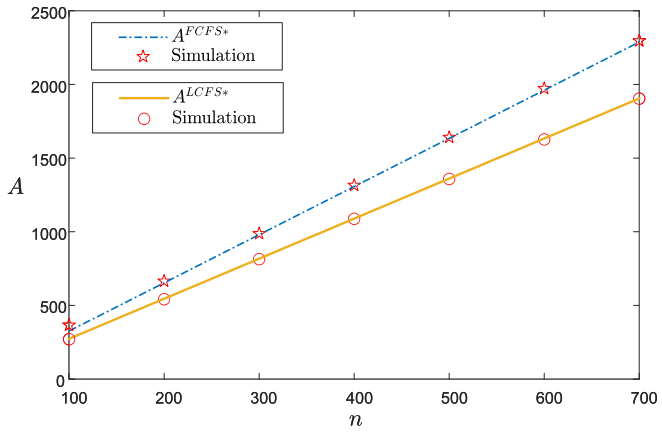

Fig. 7. The minimum PAoI A^* versus n with FCFS and LCFS discipline in the joint optimization case.

high such that the packets that arrived at the empty queue can always be successfully transmitted before a new one arrives. Therefore, the difference in service disciplines has limited effect on age performance.

In sharp contrast, when λ or n is large, the age performance gain that LCFS can achieve over FCFS becomes clear. We can see from Fig. 6 that the PAoI with FCFS $A^{FCFS*}|_\lambda$ increases with λ even with $q = q^*|_\lambda$ because of the mounting channel contention along with large access delay. Meanwhile, the PAoI with LCFS $A^{LCFS*}|_\lambda$ is insensitive to the variation of the packet arrival rate λ since the HoL packet is often replaced by a new packet which carries fresher information.

For joint optimization, Fig. 7 demonstrates the minimum PAoI with FCFS and that with LCFS discipline in the joint optimization case. It can be seen that the PAoI increases linearly with the number of sensors n. The minimum PAoI with FCFS A^{FCFS*} is always larger than that with LCFS A^{LCFS*}. In the massive access case, we can get the asymptotic ratio between A^{FCFS*} and A^{LCFS*} as

$$\lim_{n\to\infty} \frac{A^{FCFS*} - A^{LCFS*}}{A^{FCFS*}} = 16.8\%, \tag{24}$$

according to (14) and (22). Equation (24) reveals that if the network uses the LCFS discipline, rather than the FCFS discipline, the minimum PAoI can be reduced by 16.8%.

Although significant age performance gain can be obtained with LCFS in the massive access scenarios, the packet arrival rate $\lambda_L^* = 1$, implying that each sensor in the status update system should generate a new sample in each time slot. However, frequent sampling operations consume a large proportion of the energy budget, which will shorten the lifetime of sensors with limited battery capacity. The age-energy tradeoff deserves further study.

VI. Age-Energy Tradeoff Between FCFS and LCFS

In this section, we will evaluate the energy-efficiency in FCFS and LCFS cases, and then discuss the age-energy tradeoff between the FCFS and LCFS disciplines[3].

A. Energy Efficiency

The energy efficiency is defined as the ratio of the number of successfully transmitted information bits to the total energy consumption [43], which can be written as

$$\eta = \lim_{z\to\infty} \frac{z k_b}{\sum_{k=1}^{z} \bar{E} \cdot Y_k}. \tag{25}$$

The numerator of (25) is the number of successfully transmitted information bits, where z is the number of successfully transmitted packets, and k_b is the number of information bits in each packet. The denominator of (25) is the energy consumption, where $\bar{E}$ is the average energy consumption of each sensor in each time slot, and $Y_k = W_k + D_k$ is the inter-departure time between the k^{th} and $k-1^{\text{th}}$ successfully transmitted packet. By combing (8) and (9), we have

$$E[Y_k] = \frac{1}{qp} + \frac{1}{\lambda} - 1. \tag{26}$$

Let us now derive the average energy consumption of each sensor in each time slot $\bar{E}$. Specifically, for each sensor in each time slot, it may perform the sampling operation for generating a new packet, or transmit a packet, or keep silent. Accordingly, the energy consumption in each time slot contains the following parts: 1) E_T: the energy consumption of the interface during packet transmission; 2) E_S: the energy consumption of sampling operation at the beginning of the time slot; 3) E_B: the baseline energy consumption when the sensor keeps silent, i.e., no data transmission. Combining the above three parts, the average energy consumption of each sensor $\bar{E}$ in each time slot can be written as

$$\bar{E} = \lambda_e E_S + \rho q E_T + (1 - \rho q) E_B, \tag{27}$$

where λ_e is the effective arrival rate [29], and

$$\rho = \frac{\lambda}{\lambda + qp}, \tag{28}$$

is the offered load [15], i.e., the probability that the sensor has a packet in the buffer for transmission. In particular, at the beginning of each time slot, each sensor will decide whether to sample or not, for which the average energy consumption is $\lambda_e E_S$. After the sampling decision-making process, the sensor with a packet in its buffer will transmit the packet during the rest of one time slot with probability q; otherwise, it will keep silent till the end of this time slot. Accordingly, the average energy consumption is $\rho q E_T + (1 - \rho q) E_B$.

Note that we are interested in the energy efficiency in the joint optimization case, where the PAoI is optimized. By further combining (2), (15), (23), and (28), the offered load in FCFS case and that in LCFS case can be derived as

$$\rho|_{q=q_F^*,\lambda=\lambda_F^*} \approx 0.1793, \text{ and } \rho|_{q=q_L^*,\lambda=\lambda_L^*} \approx 1, \tag{29}$$

[3]As the focus of this article is on the age performance limit, the age-energy tradeoff is discussed based on the premise that both the packet arrival rate and the channel access probability have been optimally configured for PAoI minimization. The analysis can be extended to evaluate the age-energy tradeoff with the general setting of system parameters.

respectively.

For the effective arrival rate, note that with FCFS, in order to save energy, sensors would not perform the sampling operation when it has packets in the buffer[4]. While, with LCFS, sampling operations are performed even when the buffer is non-empty. Accordingly, the effective arrival rate in FCFS case and that in LCFS case can be derived as

$$\lambda_e^{FCFS} = (1-\rho)\lambda, \text{ and } \lambda_e^{LCFS} = \lambda, \tag{30}$$

respectively.

By combining (15), (23), (27), (29), and (30), the average energy consumption in each time slot with FCFS discipline $\bar{E}^{FCFS}$ can be expressed as

$$\bar{E}^{FCFS} = \frac{0.3607}{n}E_S + \frac{0.8146}{n}E_T + (1-\frac{4.543}{n})E_B, \tag{31}$$

and with LCFS discipline $\bar{E}^{LCFS}$ as

$$\bar{E}^{LCFS} = E_S + \frac{1}{n}E_T + (1-\frac{1}{n})E_B, \tag{32}$$

with an approximation $\frac{1}{n-e^{-1}} \approx \frac{1}{n}$ when n is large. By comparing (31) and (32), it can be obtained that $\bar{E}^{LCFS} - \bar{E}^{FCFS} > 0$, indicating that the energy consumption of the system with LCFS is always higher than that with FCFS. Finally, by combining (15), (23), (25)–(32), the energy efficiency with FCFS discipline can be obtained as

$$\eta^{FCFS} = \frac{k_b}{\left(\frac{0.3607}{n}E_S + \frac{0.8146}{n}E_T + \left(1-\frac{4.543}{n}\right)E_B\right)\cdot(2.772n-1)}, \tag{33}$$

and with LCFS discipline η^{LCFS} as

$$\eta^{LCFS} = \frac{k_b}{\left(E_S + \frac{1}{n}E_T + \left(1-\frac{1}{n}\right)E_B\right)\cdot 2.718n}. \tag{34}$$

when the channel access probability q and the packet arrival rate λ are configured for PAoI minimization.

B. Discussion

For discussion, let us consider the scenario in which the transmission energy E_T is 24.75 mJ per time slot and the average baseline energy consumption E_B is 15 μJ per time slot [44]. The information bits of each data packet is $k_b = 20$ kbits, and the energy consumption for data sampling E_S is assumed to be dependent on the application.

Fig. 8 illustrates how the energy efficiency η^{FCFS} and η^{LCFS} vary with the number of sensors n with $E_S \in \{0.5E_T, 2E_T\}$. Specifically, as n grows, the channel access probability of each sensor is reduced according to (15) and (23), in which case the sensor keeps silent often. Thus, in both FCFS and LCFS cases, the time-average energy consumption of each sensor drops, so does the energy efficiency. Note that η^{LCFS} is more sensitive to the variation of n, because a large body of the energy consumption comes from the sampling operation, which is required to be done in each time slot (i.e., $\lambda_L^* = 1$) for refreshing the packet

[4]This contradicts to the assumption in Section II that each sensor performs the sampling operation at the beginning of each time slot with probability λ. Note that since the buffer size is one, with FCFS, the newly incoming packet is discarded when the buffer is non-empty, and has no effect on the age performance and the channel contention process. Therefore, the above analysis still holds in the case that no sampling operation is performed when the buffer is full.

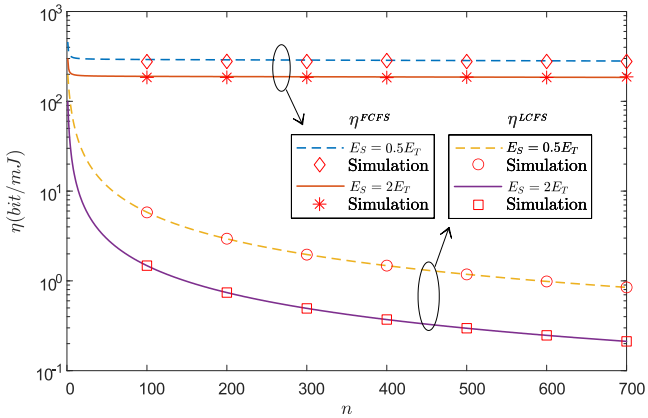

Fig. 8. The energy efficiency η^{LCFS} and η^{FCFS} versus n in the joint optimization case. $E_S \in \{0.5E_T, 2E_T\}$.

even when each sensor does not perform data transmission. This also leads to $\lim_{n\to\infty} \bar{E}^{LCFS} - \bar{E}^{FCFS} = E_S$, indicating that in the massive access scenario, the energy consumption due to sampling operation is the major factor in determining the gap between $\bar{E}^{LCFS}$ and $\bar{E}^{FCFS}$. As a result, it can be seen from Fig. 8 that the energy efficiency gap between LCFS and FCFS grows with n and E_S.

Consider one special case where $n = 200$ and $E_S = 0.5E_T$, the energy consumption with LCFS $\bar{E}^{LCFS}$ is approximately 100 times higher than that with FCFS $\bar{E}^{FCFS}$ according to (31) and (32), meanwhile the energy efficiency with LCFS η^{LCFS} is approximately 1.05% of that with FCFS η^{FCFS} according to (33) and (34). If the number of sensors n further increase to $n = 600$, the energy consumption gap will grow, such that $\bar{E}^{LCFS}$ is approximately 300 times higher than that with FCFS $\bar{E}^{FCFS}$, meanwhile the energy efficiency η^{LCFS} is approximately 0.35% of that with FCFS η^{FCFS}. This reveals that although the network obtains 16.8% age performance gain with LCFS, the energy budget might be intolerably high in large-scale IoT networks, which should be the key concern for practical systems that needs further study.

VII. Conclusion and Future Work

In this article, we focus on the PAoI optimization in slotted Aloha networks with FCFS and LCFS service disciplines by optimally tuning the channel access probability and packet arrival rate of each sensor. Specifically, depending on whether the packet arrival rate can be tuned or not, the individual optimization and the joint optimization are studied, where the optimal system parameters are explicitly characterized with the consideration of the bi-stable behavior of Aloha networks. The analysis shows that regardless of the individual optimization or joint optimization, the minimum PAoI with LCFS is lower than that with FCFS. Yet, in the joint optimization case, the age performance gain of LCFS over FCFS is achieved via a high packet arrival rate, which, nevertheless, results in soaring power consumption due to frequent sampling operations. A closer look is then cast on the age-energy tradeoff in slotted Aloha networks,

which reveals that the energy efficiency gap between those two disciplines increases with the network scale.

Note that this article assumes the classical collision model, where at most one packet can be successfully decoded in each time slot. The collision model is overly pessimistic if there exists a large difference in received power among packets. Meanwhile, the packet error rate is significant when the packets are short [45]. Therefore, it is of great importance to extend the analysis to incorporate the effect of advanced receiver structures and the short packet transmission [46]. Moreover, the homogeneous scenario is considered in this article, where all sensors have the same packet arrival rate and channel access probability. In practice, many different IoT applications may coexist together, each of which contains a large number of sensors with distinct traffic characteristics and quality-of-service requirements. How to optimize the age performance of Aloha networks in the heterogeneous scenario is another interesting issue that deserves much attention in the future study.

Appendix A
Proof of Theorem 2

To derive the optimal channel access probability and packet arrival rate for joint optimization, let us first obtain the optimal packet arrival rate $\lambda^*|_q$ when the access probability is fixed. According to (2) and (10), we have

$$\frac{\partial A^{FCFS}}{\partial \lambda} = \frac{2nq}{(pq+\lambda)^2 - n\lambda pq^2} - \frac{1}{\lambda^2}, \tag{35}$$

$\lim_{\lambda\to 0} \frac{\partial A^{FCFS}}{\partial \lambda} < 0$, and $\lim_{\lambda\to 1} \frac{\partial A^{FCFS}}{\partial \lambda} \approx 2nq - 1$. With $q^*|_\lambda > 1/2n$ according to (12), we have $\lim_{\lambda\to 1} \frac{\partial A^{FCFS}}{\partial \lambda} > 0$, implying that $\frac{\partial A^{FCFS}}{\partial \lambda} = 0$ has at least one non-zero root in (0,1). In the following, we only consider $q > 1/2n$ case.

By assuming the network always operates at the desired steady-state point p_L, Lemma 1 presents the optimal packet arrival rate $\lambda^*_{p_L}$.

Lemma 1: With $1/2n < q \le 1$ and $p = p_L$, the minimum PAoI

$$A^{FCFS*}|_{q,p=p_L} = \frac{nq\left(1+\sqrt{1+\frac{4}{nq}}\right)+2}{2q\exp\left(-\frac{2}{1+\sqrt{1+\frac{4}{nq}}}\right)} - 1, \tag{36}$$

is achieved when the packet arrival rate

$$\lambda^*_{p_L} = \frac{2q\exp\left(-\frac{2}{1+\sqrt{1+\frac{4}{nq}}}\right)}{nq\left(1+\sqrt{1+\frac{4}{nq}}\right)-2}. \tag{37}$$

Proof: By combining $\frac{\partial A^{FCFS}}{\partial \lambda} = 0$ and (35), we have $(pq+\lambda)^2 - nq\lambda(pq+\lambda) - nq\lambda^2 = 0$, for which the single positive solution gives (37). ■

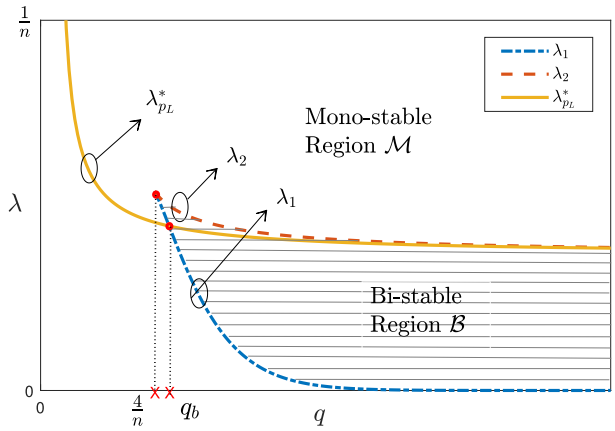

Fig. 9. Bi-stable region $\mathcal{B}$ and mono-stable region $\mathcal{M}$.

Based on Lemmas 1 and 2 further presents the optimal packet arrival rate in the general case, i.e., without the assumption of $p = p_L$.

Lemma 2: With $1/2n < q \le 1$, the optimal packet arrival rate for minimizing the PAoI is given by

$$\lambda^*|_q = \begin{cases} \frac{2q\cdot\exp\left(-\frac{2}{1+\sqrt{1+\frac{4}{nq}}}\right)}{nq\left(1+\sqrt{1+\frac{4}{nq}}\right)-2} & 1/2n < q \le q_b, \\ \frac{2}{n\left(m-\frac{2}{nq}\right)\cdot\exp\left(\frac{2}{m}\right)} & q_b < q \le 1, \end{cases} \tag{38}$$

where $m = 1 - \sqrt{1-\frac{4}{nq}}$ and $q_b \approx \frac{4.51}{n}$.

Proof: Note that with $1/2n < q \le 1$ and $\lambda = \lambda^*_{p_L}$, the network may shift into the bi-stable region $\mathcal{B}$ and therefore not operate at the desired steady-state point p_L. Fig. 9 illustrates the bi-stable region $\mathcal{B}$ and mono-stable region $\mathcal{M}$. It can be seen that $\lambda^*_{p_L}$ is included in mono-stable region $\mathcal{M}$ when the channel access probability is lower than the threshold $q_b \approx \frac{4.51}{n}$, which is solved by combining $\lambda_1 = \lambda^*_{p_L}$, (3) and (37).

When the channel access probability $1/2n < q \le q_b$, the network is guaranteed to stay in mono-stable region $\mathcal{M}$ and operate at the desired steady-state point p_L with the packet arrival rate $\lambda^*_{p_L}$, i.e.,

$$\arg\min_{\lambda} A^{FCFS}|_{1/2n<q\le q_b} = \lambda^*_{p_L}. \tag{39}$$

On the other hand, if $q > q_b$, then $(n, q, \lambda^*_{p_L}) \in \mathcal{B}$, implying that the network suffers from the risk of dropping to the undesired steady-state point p_A if $\lambda = \lambda^*_{p_L}$. To obtain the optimal packet arrival rate in this case, Fig. 10 depicts how the PAoI varies with λ in bi-stable region $\mathcal{B}$ and mono-stable region $\mathcal{M}$. It can be seen that when $\lambda \le \lambda_1$, the network operates at the desired steady-state point p_L and the PAoI decreases as λ increases. As λ increases to $\lambda > \lambda_1$, the network goes into the bi-stable region $\mathcal{B}$, and the PAoI will increase sharply if the network shifts to the undesired steady-state point p_A. Since a larger λ will increase the risk of the network shifting from p_L to p_A, the packet arrival rate λ should be set to the boundary λ_1 to minimize the chance

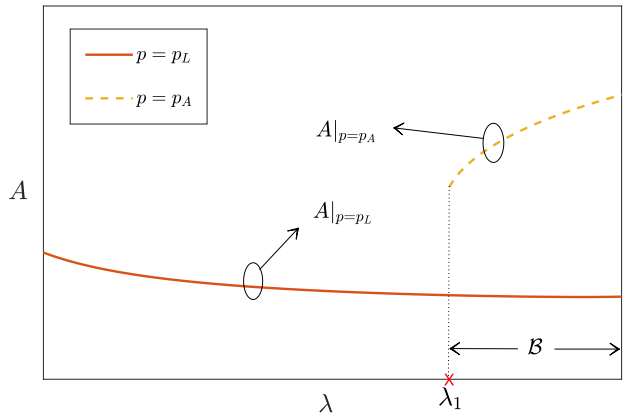

Fig. 10. PAoI A versus λ in bi-stable region $\mathcal{B}$ and mono-stable region $\mathcal{M}$.

of operating at p_A, i.e.,

$$\arg\min_{\lambda} A^{FCFS}|_{q\geq q_b} = \lambda_1. \tag{40}$$

Finally, the optimal packet arrival rate $\lambda^*|_q$ in (38) can be obtained by combining Lemma 1, (3), (39), and (40). ■

For joint optimization, the channel access probability and the packet arrival rate should satisfy (12) and (38) both. The following Lemma 3 shows the optimal setting of $\{q^*, \lambda^*\}$ in the joint optimization case.

Lemma 3: To minimize the PAoI with FCFS discipline, the channel access probability q^* and the packet arrival rate λ^* together satisfy

$$\begin{cases} q^* = \dfrac{\frac{2n\lambda^{*2}}{(p^*q^*+\lambda^*)^2-n\lambda^*p^*q^{*2}}-\frac{2}{q^*}}{p^*\exp\left(\frac{-2}{m}\right)\left(\frac{2-m}{m}\right)^2\left(\frac{2nq^*}{(p^*q^*+\lambda^*)^2-n\lambda^*p^*q^{*2}}-\frac{1}{\lambda^{*2}}\right)}, & (41) \\ \lambda^* = \dfrac{2}{n\left(m-\frac{2}{nq^*}\right)\cdot\exp\left(\frac{2}{m}\right)}, & (42) \end{cases}$$

where $m = 1-\sqrt{1-\frac{4}{nq^*}}$, and p^* denotes the successful transmission probability in (2) with $q = q^*$, $\lambda = \lambda^*$.

Proof: To ensure both (12) and (38) are satisfied, we have $q = \frac{4\mathbb{W}_{-1}^2\left(-\frac{\sqrt{n\lambda}}{2}\right)}{n\left(-2\mathbb{W}_{-1}\left(-\frac{\sqrt{n\lambda}}{2}\right)-1\right)}$ and $\lambda = \lambda_1$ by combining (3), (12) and (38). Let us derive the optimal channel access probability q^* first. By substituting $\lambda = \lambda_1$ into (11), the PAoI can be regarded as a function of channel access probability q, i.e.,

$$A^{FCFS}|_{\lambda=\lambda_1} = \frac{2}{qp'} + \frac{1}{\lambda_1} - 1. \tag{43}$$

where p' denotes the successful transmission probability in (2) with $\lambda = \lambda_1$, i.e., $p|_{\lambda=\lambda_1}$. By combining (2), (3), and (43), we have

$$\frac{\partial A^{FCFS}|_{\lambda=\lambda_1}}{\partial q} = -\exp\left(\frac{-2}{m}\right)\left(\frac{2-m}{m}\right)^2\left(\frac{2nq}{(p'q+\lambda_1)^2-n\lambda_1p'q^2}-\frac{1}{\lambda_1^2}\right) + \frac{2}{p'q}\cdot\left(\frac{n\lambda_1^2}{(p'q+\lambda_1)^2-n\lambda_1p'q^2}-\frac{1}{q}\right). \tag{44}$$

Recall that λ_1 in (3) exists if and only if $q \in [\frac{4}{n}, 1]$ [15]. We then have $\lim_{q\to\frac{4}{n}} \frac{\partial A^{FCFS}|_{\lambda=\lambda_1}}{\partial q} < 0$ and $\lim_{q\to 1} \frac{\partial A^{FCFS}|_{\lambda=\lambda_1}}{\partial q} > 0$. Numerical analysis shows that the equation $\frac{\partial A}{\partial q} = 0$ has one single root, i.e., q^*. (41) can be obtained by rewritten (44). The optimal packet arrival rate λ^* can be calculated by substituting q^* into (38). ■

The implicit nature of $\{q^*, \lambda^*\}$ sheds little light on the effect of network scale on the optimal setting. The difficulty stems from the joint calculation of (41)–(42). To simplify the calculation, the following corollary denotes $u = nq^*$ and merges (41)–(42) into (45).

Corollary 1: Let $u = nq^*$, (41) can be simplified as

$$2g_2(u)g_3(u)u^2 - 2g_1^2(u) = \left(\frac{g_2(u)g_3(u)u^3}{g_1^2(u)} - 2u\right)\left(\left(g_3(u)+\frac{g_1(u)}{u}\right)^2 - g_1(u)g_3(u)\right), \tag{45}$$

where $g_1(u)$, $g_2(u)$ and $g_3(u)$ are the function of u and given by

$$g_1(u) = \frac{2}{\left(1-\frac{2}{u}-\sqrt{1-\frac{4}{u}}\right)\cdot\exp\left(\frac{2}{1-\sqrt{1-\frac{4}{u}}}\right)}, \tag{46}$$

$$g_2(u) = \exp\left(\frac{-2}{1-\sqrt{1-\frac{4}{u}}}\right)\left(\frac{1+\sqrt{1-\frac{4}{u}}}{1-\sqrt{1-\frac{4}{u}}}\right)^2, \tag{47}$$

and

$$g_3(u) = \exp\left(-\frac{g_1(u)u}{g_1(u)+g_3(u)u}\right), \tag{48}$$

respectively.

Proof: According to (42), the optimal aggregate packet arrival rate $n\lambda^*$ is determined by u and can be written as $g_1(u)$ in (46). On the other hand, for the optimal channel access probability q^*, (41) can be rewritten as

$$\frac{2g_2(u)p^*nq^{*2}-2n\lambda^{*2}}{(p^*q^*+\lambda^*)^2-n\lambda^*p^*q^{*2}} - \frac{g_2(u)p^*q^{*2}-2\lambda^{*2}}{\lambda^{*2}q^*} = 0. \tag{49}$$

By substituting (46) into the fixed-point (2), the successful transmission probability p^* can be expressed as $g_3(u)$ in (48), which is an implicit function and also determined by u. Finally by combining (46)–(49), (45) can be obtained. ■

The solution of (45) can be obtained as $u \approx 4.543$ via brute-force searching. Therefore the optimal channel access probability $q^* = \frac{u}{n} \approx \frac{4.543}{n}$ and the optimal packet arrival rate $\lambda^* = \frac{g_1(u)}{n} \approx \frac{0.4395}{n}$, based on which the minimum PAoI can be obtained by combining (10) and (15).

Appendix B
Proof of Theorem 4

According to (2) and (20), we have

$$\frac{\partial A^{LCFS}}{\partial\lambda} = \frac{np^2q^2}{(pq+\lambda)^2-n\lambda pq^2}\cdot\left(\frac{1}{p^2q}+\frac{(1-\lambda)q}{(\lambda+(1-\lambda)pq)^2}\right) - \left(\frac{1-pq}{(\lambda+(1-\lambda)pq)^2}+\frac{1}{\lambda^2}\right). \tag{50}$$

By assuming $p = p_L$, the optimal channel access probability $q = \frac{\lambda}{n\lambda-e^{-1}}$ [16]. In this case, the successful transmission probability

$p_L = e^{-1}$, and (50) can be further derived as

$$\frac{\partial A^{LCFS}}{\partial \lambda} = \frac{g(\lambda)}{((nq\lambda+(1-nq)\lambda^2)(n\lambda-e^{-1}))^2}, \tag{51}$$

where $g(\lambda) = n(e^{-1}-n)\lambda^2 + 2e^{-1}(n-e^{-1})\lambda$. It can be seen that $g(\lambda)$ is the quadratic function of λ and $g(\lambda)$ has the same roots as (50) which can be calculated that $\lambda_a = 0$ and $\lambda_b = \frac{2e^{-1}}{n}$. Since $q = \frac{\lambda}{n\lambda - e^{-1}} \in (0,1]$, we have $\lambda \geq \frac{e^{-1}}{n-1}$. Let us consider the following two scenarios.

1) When $\frac{e^{-1}}{n-1} \leq \lambda < \lambda_b$, we have $g(p) > 0$ since $e^{-1} < n$ and $q \in (0,1]$. Therefore $\frac{\partial A^{LCFS}}{\partial \lambda} > 0$, the PAoI increases monotonically as λ increases. The minimum PAoI is obtained when $\lambda = \frac{e^{-1}}{n-1}$.
2) When $\lambda_b \leq \lambda \leq 1$, we have $g(p) < 0$ according to $e^{-1} < n$ and $q \in (0,1]$, so $\frac{\partial A^{LCFS}}{\partial \lambda} < 0$. The PAoI decreases monotonically with the increase of λ. The minimum PAoI is obtained when $\lambda = 1$.

By combining the above two scenarios, the minimum PAoI can be obtained when $\lambda = \frac{e^{-1}}{n-1}$ or $\lambda = 1$. When the packet arrival rate $\lambda = \frac{e^{-1}}{n-1}$, the optimal channel access probability $q^*|_\lambda = 1$ and it can be further verified that the network falls in the bi-stable region $\mathcal{B}$ and is quite possible not operating at p_L, the minimum PAoI cannot be achieved. On the other hand, when $\lambda = 1$, the optimal channel access probability can be obtained by substituting $\lambda = 1$ in (12) and it can be verified that the network falls in the mono-stable region $\mathcal{M}$ and therefore $p = p_L$. The corresponding optimal PAoI can be obtained by combining $\lambda = 1$, (12) and (20).

REFERENCES

[1] L. Lyu, C. Chen, S. Zhu, and X. Guan, "5G enabled codesign of energy-efficient transmission and estimation for industrial IoT systems," *IEEE Trans. Ind. Informat.*, vol. 14, no. 6, pp. 2690–2704, Jun. 2018.

[2] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. IEEE 8th Annu. Conf. Sensor, Mesh Ad Hoc Commun. Netw.*, 2011, pp. 350–358.

[3] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore, "Environmental wireless sensor networks," *Proc. IEEE*, vol. 98, no. 11, pp. 1903–1917, Nov. 2010.

[4] M. Chiang and T. Zhang, "Fog and IoT: An overview of research opportunities," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 854–864, Dec. 2016.

[5] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?," in *Proc. IEEE Int. Conf. Comput. Commun.*, 2012, pp. 2731–2735.

[6] LoRa Alliance, "LoRaWAN 1.1 specification," LoRa Alliance Tech. Committee, USA, Oct. 2017. [Online]. Available: https://lora-alliance.org/resource_hub/lorawan-specification-v1-1/

[7] Y.-P. E. Wang et al., "A primer on 3GPP narrowbandInternet of Things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 117–123, Mar. 2017.

[8] Y. Li and L. Dai, "Maximum sum rate of slotted Aloha with capture," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 690–705, Feb. 2016.

[9] A. Carleial and M. Hellman, "Bistable behavior of ALOHA-type systems," *IEEE Trans. Commun.*, vol. 23, no. 4, pp. 401–410, Apr. 1975.

[10] L. Dai, "Stability and delay analysis of buffered Aloha networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2707–2719, Aug. 2012.

[11] R. R. Rao and A. Ephremides, "On the stability of interacting queues in a multiple-access system," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 918–930, Sep. 1988.

[12] H. Wu, C. Zhu, R. J. La, X. Liu, and Y. Zhang, "FASA: Accelerated S-ALOHA using access history for event-driven M2M communications," *IEEE/ACM Trans. Netw.*, vol. 21, no. 6, pp. 1904–1917, Dec. 2013.

[13] Y. Yang and T.-S. P. Yum, "Delay distributions of slotted ALOHA and CSMA," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1846–1857, Nov. 2003.

[14] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *J. Commun. Netw.*, vol. 21, no. 3, pp. 204–219, Jun. 2019.

[15] W. Zhan and L. Dai, "Massive random access of machine-to-machine communications in LTE networks: Modeling and throughput optimization," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2771–2785, Apr. 2018.

[16] W. Zhan and L. Dai, "Access delay optimization of M2M communications in LTE networks," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1675–1678, Dec. 2019.

[17] R. D. Yates and S. K. Kaul, "Status updates over unreliable multiaccess channels," in *Proc. IEEE Int. Symp. Inf. Theory*, 2017, pp. 331–335.

[18] B. Yu, Y. Cai, and D. Wu, "Joint access control and resource allocation for short-packet-based mMTC in status update systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 851–865, Mar. 2021.

[19] D. C. Atabay, E. Uysal, and O. Kaya, "Improving age of information in random access channels," in *Proc. IEEE Conf. Comput. Commun. Workshops*, 2020, pp. 912–917.

[20] O. T. Yavascan and E Uysal, "Analysis of slotted ALOHA with an age threshold," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1456–1470, May 2021.

[21] Z. M. Fadlullah, M. M. Fouda, N. Kato, A. Takeuchi, N. Iwasaki, and Y. Nozaki, "Toward intelligent machine-to-machine communications in smart grid," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 60–65, Apr. 2011.

[22] Verizon, "Smart grid traffic behaviour discussion," 3GPP R2-102340, RAN WG2 Meeting #69b, Apr. 2020. [Online]. Available: https://www.3gpp.org/dynareport?code=Meetings-R2.htm

[23] Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Can decentralized status update achieve universally near-optimal age-of-information in wireless multiaccess channels?," in *Proc. IEEE 30th Int. Teletraffic Congr.*, 2018, vol. 01, pp. 144–152.

[24] X. Chen, K. Gatsis, H. Hassani, and S. S. Bidokhti, "Age of information in random access channels," *IEEE Trans. Inf. Theory*, vol. 68, no. 10, pp. 6548–6568, Oct. 2022.

[25] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Proc. IEEE 46th Annu. Conf. Inf. Sci. Syst.*, 2012, pp. 1–6.

[26] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1360–1374, Mar. 2016.

[27] J.-B. Seo and J. Choi, "On the outage probability of peak age-of-information for D/G/1 queuing systems," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 1021–1024, Jun. 2019.

[28] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "On the age of information with packet deadlines," *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6419–6428, Sep. 2018.

[29] M. Costa, M. Codreanu, and A. Ephremides, "On the age of information in status update systems with packet management," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1897–1910, Apr. 2016.

[30] Y. Inoue, H. Masuyama, T. Takine, and T. Tanaka, "A general formula for the stationary distribution of the age of information and its application to single-server queues," *IEEE Trans. Inf. Theory*, vol. 65, no. 12, pp. 8305–8324, Dec. 2019.

[31] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "The age of information in a discrete time queue: Stationary distribution and non-linear age mean analysis," *IEEE J. Sel. Area Commun.*, vol. 39, no. 5, pp. 1352–1364, May 2021.

[32] S. Farazi, A. G. Klein, and D. R. Brown, "Average age of information for status update systems with an energy harvesting server," in *Proc. IEEE Int. Conf. Comput. Commun. Workshops*, 2018, pp. 112–117.

[33] S. Farazi, A. G. Klein, and D. R. Brown, "Age of information in energy harvesting status update systems: When to preempt in service?," in *Proc. IEEE Int. Symp. Inf. Theory*, 2018, pp. 2436–2440.

[34] H. H. Yang, A. Arafa, T. Q. S. Quek, and H. V. Poor, "Optimizing information freshness in wireless networks: A stochastic geometry approach," *IEEE Trans. Mobile Comput.*, vol. 20, no. 6, pp. 2269–2280, Jun. 2021.

[35] X. Sun, F. Zhao, H. H. Yang, W. Zhan, X. Wang, and T. Q. S. Quek, "Optimizing age of information in random-access Poisson networks," *IEEE Internet Things J.*, vol. 9, no. 9, pp. 6816–6829, May 2022.

[36] F. Zhao, X. Sun, W. Zhan, X. Wang, J. Gong, and X. Chen, "Age-energy tradeoff in random-access Poisson networks," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 4, pp. 2055–2072, Dec. 2022.

[37] H. H. Yang, C. Xu, X. Wang, D. Feng, and T. Q. S. Quek, "Understanding age of information in large-scale wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3196–3210, May 2021.

[38] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, Dec. 2018.

[39] P. D. Mankar, M. A. Abd-Elmagid, and H. S. Dhillon, "Spatial distribution of the mean peak age of information in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4465–4479, Jul. 2021.
[40] M. Iosifescu, *Finite Markov Processes and Their Applications*. New York, NY, USA:Dover, 2007.
[41] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Trans. Netw.*, vol. 28, no. 5, pp. 2215–2228, Oct. 2020.
[42] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.
[43] M. Xie, J. Gong, X. Jia, and X. Ma, "Age and energy tradeoff for multicast networks with short packet transmissions," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 6106–6119, Sep. 2021.
[44] K. F. Ramadan, M. Dessouky, M. Abd-Elnaby, and F. E. Abd El-Samie, "Energy-efficient dual-layer MAC protocol with adaptive layer duration for WSNs," in *Proc. IEEE 11th Int. Conf. Comput. Eng. Syst.*, 2016, pp. 47–52.
[45] X. Sun, W. Zhan, W. Liu, Y. Li, and Q. Liu, "Sum rate and access delay optimization of short-packet aloha," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 1501–1514, 2022.
[46] B. Yu, X. Chen, and Y. Cai, "Age of information for the cellular Internet of Things: Challenges, key techniques, and future trends," *IEEE Commun. Mag.*, vol. 60, no. 12, pp. 20–26, Dec. 2022.

Wen Zhan (Member, IEEE) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2012 and 2015, respectively, and the Ph.D. degree from the City University of Hong Kong, Hong Kong, in 2019. He was a Research Assistant and a Postdoctoral Fellow with the City University of Hong Kong. Since 2020, he has been with the School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen, China, where he is currently an Assistant Professor. His research interests include Internet of Things, modeling, and performance optimization of next-generation mobile communication systems.

Dewei Wu (Student Member, IEEE) received the B.E. degree in communication engineering from the School of Electronics and Communication Engineering, Sun Yat-sen University (Shenzhen Campus), Shenzhen, China, in 2021. He is currently working toward the M.E. degree in information and communication engineering with the School of Electronics and Communication Engineering, Sun Yat-sen University. His research interests include age of information, Internet of Things, and stochastic modeling of wireless network.

Xinghua Sun (Member, IEEE) received the B.S. degree from the Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China, in 2008, and the Ph.D. degree from the City University of Hong Kong (CityU), Hong Kong, in 2013. He was a Visiting Student with the National Institute for Research in Digital Science and Technology, France, in 2010, and a Postdoctoral Fellow with CityU, in 2013. From 2015 to 2016, he was a Postdoctoral Fellow with the University of British Columbia, Vancouver, BC, Canada. From July to August 2019, he was a Visiting Scholar with the Singapore University of Technology and Design, Singapore. From 2014 to 2018, he was an Associate Professor with NJUPT. Since 2018, he has been an Associate Professor with Sun Yat-sen University, Shenzhen, China. His research interests include stochastic modeling of wireless networks and machine learning for networking. Dr. Sun was the Technical Program Committee Member and Organizing Committee Member for numerous conferences.

Ziyang Guo (Member, IEEE) received the B.Eng. degree (Hons.) from the College of Control Science and Engineering, Zhejiang University, Hangzhou, China, in 2014, and the Ph.D. degree from the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, in 2018. From September 2016 to December 2016, she was a Visiting Student with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. Since September 2018, she has been a Senior Engineer with Huawei Technologies Co., Ltd. Her research interests include short-range wireless communication, multi-agent reinforcement learning, cyber-physical system security, and network estimation and control.

Peng Liu (Member, IEEE) received the B.S. degree and the Ph.D. degree in telecommunication engineering from Xidian University, Xi'an, China, in 2010 and 2015, respectively. From 2013 to 2014, he was a Visiting Scholar with Columbia University, New York, NY, USA, under the supervision of Prof. Wang. He is currently a Principal Engineer with Huawei Technologies Co., Ltd., Shenzhen, China. His research interests include short-range wireless communications, AI for Wi-Fi, and wireless physical layer security.

Jingjing Liu (Member, IEEE) received the B.Eng. (1st class Hons.) and M.Eng. degrees from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2003 and 2005, respectively, and the D.Phil. degree from the Department of Engineering Science, University of Oxford, Oxford, U.K., in 2009. From 2009 to 2010, he was a Postdoctoral Research with the University of Oxford. He was a Researcher with Peking University Shenzhen Research Institution for a few years. He is currently an Associate Professor with Sun-Yat Sen University, Shenzhen, China. His research interests include low power smart micro-sensor integrated circuit design, energy harvesting circuit, and BCI circuit.