An Analytical Approach for Minimizing the Age of Information in a Practical CSMA Network

Suyang Wang, Oluwaseun T. Ajayi, Yu Cheng
Department of Electrical and Computer Engineering, Illinois Institute of Technology, USA 60616
Email: swang133@hawk.iit.edu; oajayi6@hawk.iit.edu; cheng@iit.edu

IEEE INFOCOM 2024 - IEEE Conference on Computer Communications | 979-8-3503-8350-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/INFOCOM52122.2024.10621330

***Abstract*—Age of information (AoI) is a crucial metric in modern communication systems, quantifying the information freshness at the receiver side. This study proposes a novel and general approach utilizing stochastic hybrid systems (SHS) for AoI analysis and minimization in carrier sense multiple access (CSMA) networks. Specifically, we consider a practical networking scenario where multiple nodes contend for transmission through a standard CSMA-based medium access control (MAC) protocol, and the tagged node under consideration uses a small transmission buffer for a low AoI. We for the first time develop an SHS-based analytical model for this finite-buffer transmission system over the CSMA MAC. Moreover, we develop a creative method to incorporate the collision probability into the SHS model, with background nodes having heterogeneous traffic arrival rates. This new model enables us to analytically find the optimal sampling rate to minimize the AoI of the tagged node in a wide range of practical networking scenarios. Our analysis reveals insights into buffer size impacts when jointly optimizing throughput and AoI. The SHS model is cast over an 802.11-based MAC to examine the performance, with comparison to ns-based simulation results. The accuracy of the modeling and the efficiency of optimal sampling are convincingly demonstrated.**

***Index Terms*—Age of information, carrier sense multiple access, stochastic hybrid systems, queueing systems, performance optimization.**

I. Introduction

Information freshness has become increasingly crucial in timely updating systems, with the prosperous development of time-sensitive applications in the Internet of things (IoT) and cyber-physical systems (CPS). Consequently, the age of information (AoI) has recently been introduced as a metric to quantify information timeliness, which is defined as the elapsed time since the generation of the most recent update [1]–[3]. AoI captures the effects of source sampling rate and packet delivery latency on information freshness at the receiver side. For instance, a low sampling rate in a simple queuing system results in low packet latency since the queue is often empty, but infrequent information updates may lead to outdated knowledge at the receiver, and therefore cause a high AoI. On the other hand, increasing the sampling rate may not always benefit the AoI, as a high arrival rate of updating messages tends to cause congestion and a large queueing delay.

The emergence of the AoI concept immediately stimulates a large number of studies that re-examine the queueing theory for AoI analysis in different queuing and service settings [3]–[8]. Specifically, the work in [3] delved into the average age analysis in classic first-come-first-serve (FCFS) queues, including the $M/M/1$, $M/D/1$, and $D/M/1$ systems. The AoI research also involves the study of sampling strategies in sensor networks, the age metrics beyond average analysis, and the packet management mechanisms [9]–[16]. A stochastic hybrid system (SHS) provides an alternate approach for average AoI analysis when the communication system under consideration can be modeled as a continuous-time Markov chain (CTMC). A set of balance equations, incorporating the CTMC state distributions, the age-increasing process when residing in each state, and the age reset upon state transitions, can be established to solve the average AoI [4], [17]. In this paper, we are to leverage the SHS method to develop a new analytical model for AoI analysis considering both the carrier sense multiple access (CSMA)-based medium access control (MAC) and the queueing effect.

Recently, there has been significant attention on optimizing the AoI in wireless networks under centralized scheduling policies [18]–[28]. While these centralized approaches hold merit, it is essential to extend AoI-related studies to wireless networks running over distributed MAC protocols, given the prevalence of time-sensitive applications in such settings [1], [29]–[43]. The work in [1] proposes an FCFS system with a small buffer and a well-chosen broadcast period to achieve a low system age in CSMA-based vehicular networks without modifying the standard hardware. However, this study purely relies on simulations and lacks theoretical analysis. The work in [37] develops a CSMA system mimicking the nearly optimal max-weight policy to minimize the AoI in a single-hop wireless network. The work in [38] presents an age analysis of a p-persistent CSMA augmented with additional collision resolution mechanisms. The work in [39] leverages multi-agent reinforcement learning (MARL) for more efficient and situation-aware data transmissions in the CSMA system, outperforming the traditional MAC in terms of both average and peak AoI. Nevertheless, [37]–[39] cannot be directly applied over a standard off-the-shelf CSMA MAC. In [40], [41], analytical approaches to minimize the AoI of a tagged node in a practical 802.11-based network are proposed by service time approximation and arrival rate optimization, under the setting of infinite buffer size. The work in [42] develops an SHS-based model for AoI analysis in a collision-free CSMA network, which motivates this study. However, the model in [42] omits queueing for simplification, assuming the zero-delay sampling upon demand and the immediate service with preemption, which cannot be applied in a practical networking scenario with the off-the-shelf devices that normally adopt a

transmission buffer. Moreover, [42] approximates the collision effect by bounding the channel access time and lacks the performance evaluation in a practical collision-prone system.

This paper presents a novel SHS-based approach that provides a significant leap in AoI analysis for real-world wireless networks using CSMA-based medium access control. Unlike existing approaches that require hardware redesign or protocol modifications, our method enables a tagged node to set a small MAC queue for minimal AoI in a general network setting where background nodes may take heterogeneous updating rates. To the best of our knowledge, this work is the first to integrate the influence of packet collisions into the SHS model, ensuring its accuracy and practical applicability. Exploiting our model, we achieve analytical insights into the optimal sampling rate, enabling us to minimize the tagged node's AoI across diverse practical networking scenarios. Furthermore, our investigation delves into the impact of buffer size optimization when jointly considering throughput and AoI. To validate the SHS model's performance, we apply it to an 802.11-based MAC and demonstrate its accuracy and efficiency by comparing it against ns-based simulation results. The proposed approach holds significant promise for optimizing communication efficiency and enhancing network performance in real-world applications. Furthermore, we elucidate the trade-off between AoI and throughput, offering valuable insights and guiding the design of application-oriented sampling strategies. Remarkably, our method is generally applicable for AoI analysis over the family of CSMA-based MAC protocols without requiring modifications to standard hardware or protocol designs, making it highly deployable and practical. In summary, this paper has multi-fold contributions.

1) We develop a general and accurate SHS-based model for analyzing AoI in CSMA networks, capturing the interplay between queueing effects and channel contention effectively.
2) We propose a creative method to incorporate the collision effect into the SHS model. The impact of all the heterogeneous background nodes is jointly modeled with one aggregated state, which has the dual benefits of enabling a scalable model and integrating with the practical MAC analysis [44].
3) We quantitatively determine the optimal sampling rate to minimize AoI. Unlike existing finite-buffer system approaches that heavily rely on simulations, our rigorous modeling approach allows us to analytically determine the optimal sampling rate, enhancing the flexibility and applicability of the AoI minimization technique.
4) Our analytical model readily facilitates the joint optimization of throughput and AoI. It can find applications in many service scenarios where status monitoring applications generate small packets that can be piggybacked with data packets, we introduce a novel "throughput weighted age of information (TwAoI)" to guide the optimal sampling rate selection balancing the throughput and AoI requirements.
5) We conduct practical ns-3 simulations over the 802.11-based network to validate the effectiveness and accuracy of the proposed analytical model.

The remainder of this paper is structured as follows. Section II presents the system model, including the general CSMA networking model and the basics of AoI and SHS modeling. Section III demonstrates how to construct the SHS model for finite buffer transmission over the CSMA MAC. The model is further enhanced to incorporate collisions in Section IV. Numerical results and the throughput-AoI joint evaluation are presented in Section V. Section VI reviews more related work and Section VII concludes the paper.

II. System Model

This section first describes a generalized networking model underpinning the family of CSMA-based MAC protocols. We then briefly explain the AoI concept and present the basics of SHS modeling for average AoI analysis.

A. A General CSMA Network Model

We consider a general CSMA network where each node communicates with the monitor residing at the access point (AP) via a shared wireless channel coordinated by the CSMA MAC protocol. Each node sends packets through its local tail-drop FCFS MAC queue, and the traffic arrival process is modeled as a Poisson process. Different nodes may have different traffic arrival rates based on their respective applications. To facilitate the SHS modeling, we first consider a generalized CSMA setting and then demonstrate how it is cast over an 802.11-based implementation. Generally, each node independently senses the channel before transmission. If the channel is sensed idle, the node then starts an exponentially distributed backoff period and transmits a packet upon the end of the backoff. Successful transmissions result in the removal of the packet from the MAC queue, while collisions trigger retransmissions, going through the same backoff process. We consider a constant channel capacity and exponentially distributed packet size to facilitate mathematical tractability. Note that such a CSMA setting is similar to the one used in [40], the most related work to this study; however, our model is much more practical, where we consider the tagged node uses a finite buffer, and collision is to be included into modeling. Note that the simulation study in [1] demonstrates that a small transmission buffer assisted with an optimal traffic arrival rate is a valid setting to minimize the average AoI. In this work, we are to develop an analytical model that can significantly extend the applicability and convenience of the finite-buffer system. Specifically, we are to develop an SHS-based model to compute the average AoI of the tagged node perceived by the monitor at the AP, incorporating the buffering effect at the transmission queue and the collision over the CSMA MAC.

B. SHS Modeling for AoI Analysis

The age of information specifies how much time has elapsed since the most recent update's generation. The update i is sent through the system and eventually arrives at the monitor when a source creates update i with timestamp $u_i(t)$. The monitor

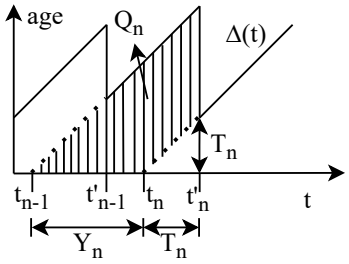


Fig. 1. Sawtooth age waveform.

observes its most recent update received at time t with an age of $t - u_i(t)$. Age increases linearly over time as the update becomes older without any new updates arriving. The average AoI is calculated as the time-average of the instantaneous age waveform $\Delta(t)$, as shown in Fig. 1, over a sufficiently large time duration τ: $\langle\Delta\rangle_\tau = \frac{1}{\tau}\int_0^\tau \Delta(t)dt$. The interarrival time and system time for the n^{th} update are denoted as $Y_n = t_n - t_{n-1}$ and $T_n = t'_n - t_n$, respectively. According to Fig. 1, the sum of each shaded area $Q_n = \frac{1}{2}(T_n + Y_n)^2 - \frac{1}{2}T_n^2$ is equal to the integral denoted by $\langle\Delta\rangle$. The average AoI is

$$\Delta = \lim_{\tau\to\infty}\langle\Delta\rangle_\tau = \frac{E[Q_n]}{E[Y_n]}. \tag{1}$$

Please refer to the tutorial article [45] for analytical details of applying equation (1) to different queuing systems. To compute the average age in complex systems like CSMA, traditional graphical approaches can be cumbersome and impractical in lossy systems where packets may be dropped or in complex environments where various events must be considered. To address these challenges, we employ the SHS method to calculate the overall average AoI. The SHS model considers the networking system as CTMC, facilitated by the assumption of exponential backoff and packet size. A set of balance equations, incorporating the CTMC state distributions, the age-increasing process when residing in each state, and the age reset upon state transitions, can be established to solve the average AoI [4], [17]. In the remainder of this paper, we adopt similar mathematical notations as that used in [4] to develop our SHS model.

In an SHS, our proposed method involves labeling the node of interest and utilizing the continuous state as a vector of age-related processes to track the age of their updates at a monitoring point. We model our network by the states $(\vec{q}(t), \mathbf{x}(t))$ where:

1) The discrete process $\vec{q}(t) \in \mathbb{Q}$ represents the state of the network at time t, where $\mathbb{Q}$ denotes the discrete set of possible values that $\vec{q}(t)$ can take. Note that we use $\vec{q}(t)$ to denote the state, as the states in our SHS model take the form of a pair of scalar values.
2) $\mathbf{x}(t) = [x_0(t), x_1(t), \cdots x_K(t)]$ is a vector that traces the age evolution of the interested updating packets at the monitor and in the queue. Specifically, $x_0(t)$ is the age of the interested node (or namely, the tagged node)'s latest update at the monitor. The tagged node's queue capacity is denoted as K, and packets in the queue are indexed from 1 to K, with packet-1 being the head of the queue. The age of a tagged node's packet in the i^{th} place of its MAC queue is represented by $x_i(t), (1 \le i \le K)$. In such systems, we calculate the average AoI for the tagged node's update process by determining the limit of the expected value of $x_0(t)$ as time approaches infinity.

A Markov process known as $\vec{q}(t)$ can be visually depicted as a Markov chain $(\mathbb{Q}, L)$. This chain comprises vertices, with each vertex being a state $\vec{q}(t) \in \mathbb{Q}$, and directed transition edges $l \in L$ with the departing and entering states of transition l denoted as $\vec{q}_l$ and $\vec{q}'_l$, respectively. The transition rate of each edge is denoted by $\lambda^{(l)}\delta_{\vec{q}_l,\vec{q}(t)}$, where the Kronecker delta function ensures that the transition l can only occur when the discrete process $\vec{q}(t)$ equals $\vec{q}_l$. The outgoing and incoming transitions sets for each state $\vec{q}$ are defined as follows:

$$L_{\vec{q}} = \{l \in L : \vec{q}_l = \vec{q}\}, \quad L'_{\vec{q}} = \{l \in L : \vec{q}'_l = \vec{q}\}. \tag{2}$$

The Markov chain part in SHS handles the complex relevant events in the network. In AoI analysis, another aspect SHS should handle is the age evolution of the tagged node. Intuitively, the transitions in the discrete process would lead to a reset in the continuous process. In other words, when a transition l occurs, the discrete process moves to a different state $\vec{q}'_l$, and a sudden drop in the continuous process $\mathbf{x}' = \mathbf{x}\mathbf{A}_l$ occurs when a packet is successfully delivered at the monitor. The transition reset maps are represented by the matrix $\mathbf{A}_l$ and play a crucial role in modeling the evolution of the age process. As the age process can only increase linearly with time in each state $\vec{q} \in \mathbb{Q}$, to account for this aspect, the continuous process $\mathbf{x}$ satisfies the following first-order differential equation:

$$\dot{\mathbf{x}} = \mathbf{b}_{\vec{q}}. \tag{3}$$

In this equation, $\mathbf{b}_{\vec{q}} = [b^0_{\vec{q}}, b^1_{\vec{q}}, b^2_{\vec{q}}, \cdots b^K_{\vec{q}}]$ is a binary vector in which $b^i_{\vec{q}}$ equals 1 when the age $x_i(t)$ increases at a unit rate in state $\vec{q}$, and equals 0 when the age remains constant in that state. For instance, the age of each update at the monitor always grows linearly with time, so $b^0_{\vec{q}} = 1$ for all $\vec{q}$. In order to determine the average age of the system using SHS, it is necessary to define the following quantities for each state $\vec{q}$:

$$\pi_{\vec{q}}(t) = E\left[\delta_{\vec{q},\vec{q}(t)}\right] = P(\vec{q}(t) = \vec{q}), \tag{4}$$

$$\mathbf{v}_{\vec{q}}(t) = [v_{\vec{q}0}(t), v_{\vec{q}1}(t), \cdots, v_{\vec{q}K}(t)] = E\left[\mathbf{x}(t)\delta_{\vec{q},\vec{q}(t)}\right], \tag{5}$$

where $\pi_{\vec{q}}(t)$ is the stationary probability distribution of the Markov chain and $\mathbf{v}_{\vec{q}}(t)$ is the correlation between the age process $\mathbf{x}(t)$ and the discrete state $\vec{q}(t)$. In order to establish the existence and uniqueness of a steady-state distribution for the Markov chain, it is necessary to make the assumption that the Markov chain $\vec{q}(t)$ is ergodic. This requirement is satisfied by defining the steady-state probability vector $\bar{\pi} = [\bar{\pi}_{\vec{q}}]_{\vec{q}\in\mathbb{Q}}$ as the solution to the general balance equations:

$$\bar{\pi}_{\vec{q}}\left(\textstyle\sum_{l\in L_{\vec{q}}}\lambda^{(l)}\right) = \textstyle\sum_{l\in L'_{\vec{q}}}\lambda^{(l)}\bar{\pi}_{\vec{q}_l}, \quad \vec{q} \in \mathbb{Q}, \tag{6}$$

$$\textstyle\sum_{\vec{q}\in\mathbb{Q}}\bar{\pi}_{\vec{q}} = 1. \tag{7}$$

More detail shown in [4] justifies that $\mathbf{v}_{\vec{q}(t)}$ converges in this ergodic case satisfying:

$$\bar{\mathbf{v}}_{\vec{q}} \sum_{l \in L_{\vec{q}}} \lambda^{(l)} = \mathbf{b}_{\vec{q}} \bar{\pi}_{\vec{q}} + \sum_{l \in L'_{\vec{q}}} \lambda^{(l)} \bar{\mathbf{v}}_{\vec{q}_l} \mathbf{A}_l, \quad \vec{q} \in \mathbb{Q}. \tag{8}$$

Finally, the average age of the tagged node is:

$$\Delta = \sum_{\vec{q} \in \mathbb{Q}} \bar{v}_{\vec{q}0}. \tag{9}$$

The fundamental outcomes presented above constitute the foundation of the AoI analysis through SHS. These results will be utilized in the following section to examine the average AoI in various scenarios.

III. Finite Buffer Transmission over the CSMA

In our model, a certain node i generates status update packets for a certain application at a rate of λ_i, which follows a Poisson process. These packets are then added to the transmission queue. Node i takes an exponentially distributed backoff procedure with the average of $1/R_i$. Once node i finishes its backoff period, the packet at the head of its queue starts the transmission. The transmission time is exponentially distributed with an average of $1/H_i$; H_i can also be interpreted as the average transmission rate. Upon a successful transmission, the number of packets in the queue decreases by one.

We consider a CSMA network of N nodes, trying to send updating messages to the monitor at the AP. We are to develop the SHS model through the perspective of the tagged node, which uses a transmission queue of size K packets. The remaining $N-1$ nodes function as background traffic, transmitting with varying traffic rates. A significant feature of our model is to jointly describe the impact of the $N-1$ background nodes on the tagged node using one aggregated state. Specifically, the network's state at a certain time is denoted by a 2-tuple state, where the first element of the tuple indicates the state of the tagged node, and the second element is the state of the background nodes. The total set of possible states $\mathbb{Q} = \{(k, Q), (C_k, Q), (k, C)\}, k \in [0, K]$. Detailed specifications of the states are as follows.

- The states $(k, Q), k \in [1, K]$, represent that both the tagged node and the background traffic independently start the exponential backoff process to contend for the channel access, and the tagged node currently has k packets in the queue. Here, we use "Q" to indicate there is a certain background node (or, equivalently, the aggregated background node) tending to grab the channel. Note that the channel competition from any background node will cause the same impact on the tagged node; thus, our model does not differentiate the exact background node and just uses "Q" to indicate there is background traffic contending for the channel. In a typical practical scenario with multiple background nodes, the background traffic almost always exists. So even after one successful transmission from the background node, the state still goes back to "Q". The state $(0, Q)$ indicates that the tagged node has no data and only background traffic tries to use the channel.
- The states $(C_k, Q), k \in [1, K]$, represent that the tagged node grabbed the channel to transmit the head-of-queue packet, upon the end of the backoff, with other $k-1$ packets staying in the buffer. Since a node cannot be in the transmission state with an empty queue, C_0 doesn't exist. Also, for the system with $K = 1$, we denote C_k as C for simplification as illustrated in Fig. 2.
- The states $(k, C), k \in [0, K]$, represent that the background traffic finished the backoff earlier than the tagged node and grabbed the channel for transmission.
- Note that a transmission will freeze the backoff of other contending nodes due to the carrier sensing mechanism. Thus, the tagged node's backoff operation either successfully ends to transmit or is frozen by transmission from a background node. Accordingly, the states $(C_k, C), k \in [1, K]$ do not exist. Also, due to the memoryless property of exponential distribution, a newly started backoff period and a resumed backoff period follow the same distribution.

Figure 2 describes the SHS Markovian chain for the scenario where the tagged node has a transmission queue size of 1 packet (i.e., $K = 1$). State transitions occur when

- a new update enters the tagged node's queue, with arrival rate λ, or
- a node among all N nodes captures the channel, or
- retransmission occurs due to a collision (the analytical details about collisions are to be discussed in the next section) or
- a transmission is successfully completed.

To analyze the AoI associated with the tagged node, all elements in the age vector $\mathbf{x} = [x_0, x_1, \cdots, x_K]$ are to be involved. Here, x_0 signifies the age perceived at the monitor, while $x_1, x_2, \cdots, x_K$ represent the age progression of the packets placed at the $1^{st}, 2^{nd}, \cdots, K^{th}$ position in the queue, respectively. Leveraging the equations provided in Section II-B, we can efficiently compute the average age; the analytical details are to be elaborated in Sections IV and V.

IV. SHS Modeling with the Collision

This section introduces the proposed SHS-based model incorporating collisions. We will first elaborate on a CSMA network containing a tagged node with a queue size of $K = 1$ and updates arrival rate λ. The SHS model is developed with an assumption that each transmission will be successful with a probability $1-p$, or suffer from a collision with a probability p. Without loss of generality, we index the tagged node as node 1, and the aggregated background node as node 2. We then take the 802.11 distributed coordination function (DCF) as a concrete case to demonstrate the engineering method to obtain the collision probability in a practical setting. We then show in detail how to extend our model to the case where the tagged node has a queue size of 2 and discuss the practicability and challenges of SHS modeling for larger buffer scenarios.

A. Incorporating the Collisions

The SHS Markov chain for the discrete state is illustrated in Fig. 2. At time t, the system state $\vec{q}(t) \in \mathbb{Q}$, with

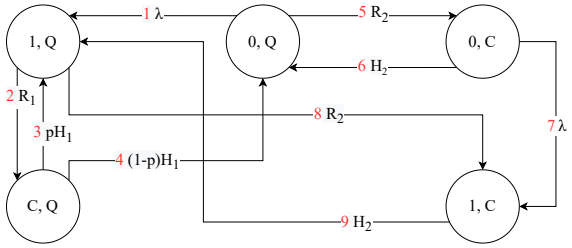


Fig. 2. The SHS Markov chain for the tagged node in the CSMA network with a queue size of 1 packet and collision incorporated.

TABLE I
SHS TRANSITIONS (TAGGED NODE'S QUEUE SIZE = 1)

l	$\vec{q}_l \to \vec{q}'_l$	$\lambda^{(l)}$	$\mathbf{x}\mathbf{A}_l$	$\mathbf{v}_{\vec{q}_l}\mathbf{A}_l$
1	$(0,Q) \to (1,Q)$	λ	$[x_0, 0]$	$[v_{(0,Q)0}, 0]$
2	$(1,Q) \to (C,Q)$	R_1	$[x_0, x_1]$	$[v_{(1,Q)0}, v_{(1,Q)1}]$
3	$(C,Q) \to (1,Q)$	pH_1	$[x_0, x_1]$	$[v_{(C,Q)0}, v_{(C,Q)1}]$
4	$(C,Q) \to (0,Q)$	$(1-p)H_1$	$[x_1, 0]$	$[v_{(C,Q)1}, 0]$
5	$(0,Q) \to (0,C)$	R_2	$[x_0, 0]$	$[v_{(0,Q)0}, 0]$
6	$(0,C) \to (0,Q)$	H_2	$[x_0, 0]$	$[v_{(0,C)0}, 0]$
7	$(0,C) \to (1,C)$	λ	$[x_0, 0]$	$[v_{(0,C)0}, 0]$
8	$(1,Q) \to (1,C)$	R_2	$[x_0, x_1]$	$[v_{(1,Q)0}, v_{(1,Q)1}]$
9	$(1,C) \to (1,Q)$	H_2	$[x_0, x_1]$	$[v_{(1,C)0}, v_{(1,C)1}]$

$\mathbb{Q} = \{(k,Q), (C_k,Q), (k,C)\}$, where $k \in [0,K]$ with $K = 1$. The continuous process $\mathbf{x}(t) = [x_0(t), x_1(t)]$ where $x_0(t)$ traces the age of information perceived at the monitor, and $x_1(t)$ tracks the age process of the update in the tagged node's queue, which is the next update going to be delivered to the monitor. Upon a new update delivery, the age at the monitor will be updated to the age of the just-received message; mathematically presented, x_0 would drop to the value of x_1, reflected by $\mathbf{x}' = \mathbf{x}\mathbf{A}_l$. With this being said, using $(0,*)$ to denote the states when there is no update in the tagged node's queue, $x_1(t)$ is irrelevant in state $(0,*)$ and is thus set to 0. The SHS transitions are enumerated in Table I. Each edge in Fig. 2 from node $\vec{q}_l$ to $\vec{q}'_l$ represents state transition occurs at exponential rate $\lambda^{(l)}$ from state $\vec{q}_l$ to $\vec{q}'_l$. The resultant new vector of the age process is $\mathbf{x}\mathbf{A}_l$, and $\mathbf{v}_{\vec{q}_l}\mathbf{A}_l$ is the correlation between the age process $\mathbf{x}(t)$ and the discrete state $\vec{q}(t)$, which is a crucial part for AoI calculation using equation (8). We would like to highlight transitions $l = 3$ and 4, which are exclusively introduced in this work to handle the collision effect on the AoI in the CSMA networks. The explanations of each transition in Table I are provided below.

- $l = 1$ A new update arrives at the tagged node's empty MAC queue with rate λ, causing the queue size to increase from zero to one. However, the state transition from $(0,Q)$ to $(1,Q)$ does not cause age reduction since no new update is delivered to the monitor, that is, $x'_0 = x_0$. Also, the new arrival is fresh with age zero, reflecting as $x'_1 = 0$. Regarding the $\mathbf{x}\mathbf{A}_l = [x_0, 0]$, the $\mathbf{v}_{\vec{q}_l}\mathbf{A}_l$ is thus written as $[v_{(0,Q)0}, 0]$, accordingly.
- $l = 2$ The tagged node, with an update in the queue, captures the channel with the backoff rate R_1. Since no update delivery occurs, the age of the update at the monitor remains the same, i.e., $x'_0 = x_0$. Also, since the update has been in the queue for a while, $x'_1 = x_1$.
- $l = 3$ Upon capturing the channel, the transmission fails due to a collision with probability p at a rate of pH_1, where $1/H_1$ is the channel holding time for each transmission. Subsequently, the tagged node enters the retransmission procedure after the failed transmission, leading to state transitions back to $(1,Q)$.
- $l = 4$ The update departs the queue and is received by the monitor at rate $(1-p)H_1$, leading to the age at the monitor drops to $x'_0 = x_1$.
- $l = 5$ One of the $N-1$ background nodes captures the channel, and the state transitions at rate R_2. x_1 is irrelevant to this transition and thus makes no change.
- $l = 6$ As the behaviors of the $N-1$ background nodes do not affect the tagged node's $\mathbf{x}$, transition $l = 6$ simply unifies the successful and collided transmission into one single transition at rate H_2, indicating a background node's update transmission has ended and the channel is released and free for new contention.
- $l = 7$ During the background node's transmission phase, a new update arrives at the tagged node's queue at rate λ, and there is no reset to the tagged node's $\mathbf{x}$.
- $l = 8$ One of the $N-1$ background nodes captures the channel. Since there is no new tagged node's update delivery, it does not incur any reset in the tagged node's $\mathbf{x}$. Also, because the update in the tagged node's queue has already been in it for a while, $x'_1 = x_1$.
- $l = 9$ A successful or collided transmission happens at rate H_2. This is again not relevant to any reset in $\mathbf{x}$.

The first-order differential equation of the continuous process $\mathbf{x}$ satisfies $\dot{\mathbf{x}} = \mathbf{b}_{\vec{q}} = \begin{cases} [1 \quad 0], & \vec{q} = (0,*), \\ [1 \quad 1], & \text{otherwise}, \end{cases}$ indicating the age-increasing process when residing in each state. Specifically, the age x_0 always increases at a unit rate, and the age x_1 increases at a unit rate only in states other than $\vec{q} = (0,*)$. To leverage equation (8) for AoI calculation, firstly we are going to obtain the stationary probability $\bar{\pi} = [\bar{\pi}_{\vec{q}}]_{\vec{q} \in \mathbb{Q}}$ of the CTMC via equations (6) and (7). Here we rewrite equation (6) as $\bar{\pi}\mathbf{D} = \bar{\pi}\mathbf{B}$, where $\mathbf{D} = \text{diag}[\lambda + R_2, R_1 + R_2, H_1, \lambda + H_2, H_2]$,

$$\mathbf{B} = \begin{bmatrix} 0 & \lambda & 0 & R_2 & 0 \\ 0 & 0 & R_1 & 0 & R_2 \\ (1-p)H_1 & pH_1 & 0 & 0 & 0 \\ H_2 & 0 & 0 & 0 & \lambda \\ 0 & H_2 & 0 & 0 & 0 \end{bmatrix}. \tag{10}$$

Thus, we have stationary probabilities. Next, we can use equation (8) to solve

$$\begin{aligned} \bar{\mathbf{v}} &= [\bar{\mathbf{v}}_{(0,Q)} \quad \bar{\mathbf{v}}_{(1,Q)} \quad \bar{\mathbf{v}}_{(C,Q)} \quad \bar{\mathbf{v}}_{(0,C)} \quad \bar{\mathbf{v}}_{(1,C)}] \\ &= [\bar{v}_{(0,Q)0} \quad \bar{v}_{(0,Q)1} \quad \bar{v}_{(1,Q)0} \quad \bar{v}_{(1,Q)1} \quad \bar{v}_{(C,Q)0} \\ &\quad \bar{v}_{(C,Q)1} \quad \bar{v}_{(0,C)0} \quad \bar{v}_{(0,C)1} \quad \bar{v}_{(1,C)0} \quad \bar{v}_{(1,C)1}]. \end{aligned} \tag{11}$$

Specifically, at state $\vec{q} = (0,Q)$, the equation obtained through equation (8) is

$$\begin{aligned} (\lambda + R_2)[\bar{v}_{(0,Q)0} \quad \bar{v}_{(0,Q)1}] &= [1 \quad 0]\bar{\pi}_{(0,Q)} \\ &+ (1-p)H_1[\bar{v}_{(C,Q)1} \quad 0] + H_2[\bar{v}_{(0,C)0} \quad 0]. \end{aligned} \tag{12}$$

Similarly, the equations regarding states $\vec{q} = (1,Q)$, (C,Q), $(0,C)$, $(1,C)$ are listed below, respectively.

$$\begin{aligned} (R_1 + R_2)[\bar{v}_{(1,Q)0} \quad \bar{v}_{(1,Q)1}] &= [1 \quad 1]\bar{\pi}_{(1,Q)} + \lambda[\bar{v}_{(0,Q)0} \quad 0] \\ &+ pH_1[\bar{v}_{(C,Q)0} \quad \bar{v}_{(C,Q)1}] + H_2[\bar{v}_{(1,C)0} \quad \bar{v}_{(1,C)1}]. \end{aligned} \tag{13}$$

$$H_1[\bar{v}_{(C,Q)0} \quad \bar{v}_{(C,Q)1}] = [1 \quad 1]\bar{\pi}_{(C,Q)} + R_1[\bar{v}_{(1,Q)0} \quad \bar{v}_{(1,Q)1}]. \quad (14)$$

$$(H_2 + \lambda)[\bar{v}_{(0,C)0} \quad \bar{v}_{(0,C)1}] = [1 \quad 0]\bar{\pi}_{(0,C)} + R_2[\bar{v}_{(0,Q)0} \quad 0]. \quad (15)$$

$$\begin{aligned} H_2[\bar{v}_{(1,C)0} \quad \bar{v}_{(1,C)1}] &= [1 \quad 1]\bar{\pi}_{(1,C)} \\ &+ \lambda[\bar{v}_{(0,C)0} \quad 0] + R_2[\bar{v}_{(1,Q)0} \quad \bar{v}_{(1,Q)1}]. \end{aligned} \quad (16)$$

Then we can obtain $\bar{v}_{(0,Q)0}$, $\bar{v}_{(1,Q)0}$, $\bar{v}_{(C,Q)0}$, $\bar{v}_{(0,C)0}$, $\bar{v}_{(1,C)0}$ by solving equation (8). Finally, the average AoI at the monitor for the tagged node with queue size $K = 1$ is

$$\Delta_{K=1} = \sum_{\vec{q}\in\mathbb{Q}} \bar{v}_{\vec{q}0} = \bar{v}_{(0,Q)0} + \bar{v}_{(1,Q)0} + \bar{v}_{(C,Q)0} + \bar{v}_{(0,C)0} + \bar{v}_{(1,C)0}. \quad (17)$$

B. SHS Instantiation over the 802.11 DCF

In this part, we outline the process of obtaining the backoff-related parameter R and the collision probability p when instantiating the SHS model over an 802.11 DCF MAC. We leverage the results in [46], where both the protocol specifications and analytical details for 802.11 DCF are presented.

The work in [46] considers the saturated transmission scenario, where each node always has packets in the buffer contending for the channel for transmission. We can still leverage the saturated analysis in [46] to facilitate our analysis in the heterogeneous setting. Our approach involves modeling all the queueing and transmission details of the tagged node over a finite buffer with a finite arrival rate, using the SHS modeling as described in Section IV-A. In a highly utilized channel with multiple nodes, the tagged node nearly always experiences channel competition from background traffic. Thus, we can use the saturated approximation to obtain a collision probability. Note that, our SHS model is designed for the tagged node, but it can be applied to any node of interest in a heterogeneous setting for accurate performance evaluation. Thus, our methodology of integrating detailed SHS modeling for the tagged node and saturated approximation for collision probability represents a scalable engineering approach that accommodates various numbers of background nodes and heterogeneous traffic arrival processes.

According to 802.11 DCF, each node accesses an idle channel by an exponential backoff procedure, where the contention window is dynamically adjusted depending on the number of unsuccessful retransmissions. The channel access can be described by a transmission probability τ, which may thus suffer a conditional collision probability p, as demonstrated in [46]. Let $CW_{\min}$ denote the initial contention window size, m denote the maximum backoff stage, and a denote the retransmission limit. The values of p and τ can be obtained by solving the following coupled equations for a network with N nodes: $p = 1 - (1-\tau)^{N-1}$, $\tau = \frac{2(1-p^{a+1})}{1-p^{a+1}+p\cdot CW_{\min}\sum_{i=0}^{m-1}(2p)^i + CW_{\min}(1-2^m p^{a+1})}$. Note that p is the approximate collision probability to be used in our SHS model as illustrated in Fig. 2. The exponential backoff period in Fig. 2 can be obtained by computing the average backoff time amortized over all the backoff stages. The average contention window $\bar{W}$, in terms of time slots is $\bar{W} = \sum_{k=1}^{a+1} p^{k-1}(1-p)^{I\{k<a+1\}} \sum_{j=1}^{k} \frac{CW(j)-1}{2}$, where $CW(k) = \min\left(2^m CW_{\min}, 2^{k-1} CW_{\min}\right), \quad k = 1, \ldots, a+1$,

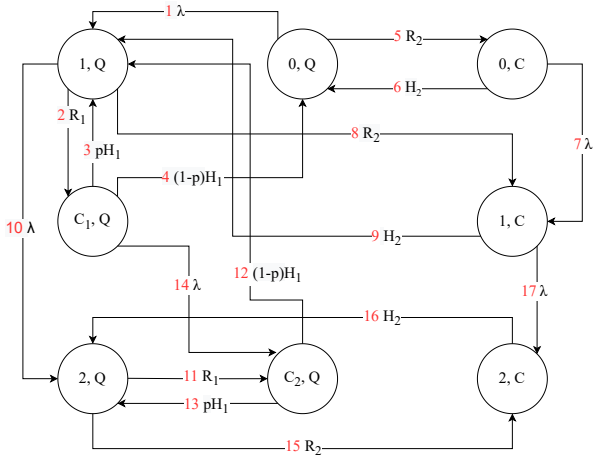


Fig. 3. The SHS Markov chain for the tagged node in the CSMA network, with the queue size of 2 packets and collision incorporated.

and the indicator $I\{A\}$ is set to 1 if A is true, and to 0 otherwise [47]. Let t_{slot} denote the slot time, the average backoff period of the tagged node can be set as $R_1 = 1/(t_{slot} \cdot \bar{W})$. Node 2 in the SHS model in Fig. 2 is aggregated over the $N-1$ background nodes, and thus the channel access time is the minimum of $N-1$ exponential variables. Therefore, $R_2 = (N-1)/(t_{slot} \cdot \bar{W})$. With these values, the SHS model can be computed for performance analysis, and for other CSMA-based protocols, appropriate MAC analysis is needed to determine p, R_1, and R_2. For example, the MAC analysis of 802.15.4 can be found in [48] and the references therein.

C. Age Analysis for Larger Queue Size

In order to analytically evaluate the impact of buffer size on the AoI performance, we need to extend the SHS model to the cases of $K > 1$, which is not a trivial task. Where the SHS Markov chain is readily available, referring to the descriptions in Section III, the balance equations need to be established for the age vector $\mathbf{x} = [x_0, x_1, \cdots, x_K]$, as the perceived age at the monitor (x_0) depends on age values of all the packets in the queue (x_1 to x_K). In this part, we give the details of extending the SHS model to the case of $K = 2$. The operations for further extension to $K > 2$ cases are not difficult to figure out following the same principle.

The SHS Markov chain is updated shown in Fig. 3. The state space has increased to depict the network dynamics as the queue size increases by 1. Specifically, three states $(2, Q)$, (C_2, Q), and $(2, C)$ are added. In this case, the ages of three updates are of our concern, say, the newest update at the monitor, the update in the first place (head) of the tagged node's queue (if any), and the update in the second place of the tagged node's queue (if any), traced by the continuous process $\mathbf{x}(t) = [x_0(t), x_1(t), x_2(t)]$, respectively. Once the update departs the FCFS queue and is successfully delivered to the monitor, x_0 would drop to the value of x_1, reflected by $\mathbf{x}' = \mathbf{x}\mathbf{A}_l$. And a new age is given to x_1 with the value of x_2, represented as $x_1' = x_2$. The SHS transitions are enumerated in Table II. Transitions $l = 3, 4, 12$, and 13 tackle the collision

TABLE II
SHS TRANSITIONS (TAGGED NODE'S QUEUE SIZE = 2)

l	$\vec{q}_l \to \vec{q}'_l$	$\lambda^{(l)}$	$\mathbf{x}\mathbf{A}_l$	$\mathbf{v}_{\vec{q}_l}\mathbf{A}_l$
1	$(0,Q) \to (1,Q)$	λ	$[x_0, 0, 0]$	$[v_{(0,Q)0}, 0, 0]$
2	$(1,Q) \to (C_1,Q)$	R_1	$[x_0, x_1, 0]$	$[v_{(1,Q)0}, v_{(1,Q)1}, 0]$
3	$(C_1,Q) \to (1,Q)$	pH_1	$[x_0, x_1, 0]$	$[v_{(C_1,Q)0}, v_{(C_1,Q)1}, 0]$
4	$(C_1,Q) \to (0,Q)$	$(1-p)H_1$	$[x_1, 0, 0]$	$[v_{(C_1,Q)1}, 0, 0]$
5	$(0,Q) \to (0,C)$	R_2	$[x_0, 0, 0]$	$[v_{(0,Q)0}, 0, 0]$
6	$(0,C) \to (0,Q)$	H_2	$[x_0, 0, 0]$	$[v_{(0,C)0}, 0, 0]$
7	$(0,C) \to (1,C)$	λ	$[x_0, 0, 0]$	$[v_{(0,C)0}, 0, 0]$
8	$(1,Q) \to (1,C)$	R_2	$[x_0, x_1, 0]$	$[v_{(1,Q)0}, v_{(1,Q)1}, 0]$
9	$(1,C) \to (1,Q)$	H_2	$[x_0, x_1, 0]$	$[v_{(1,C)0}, v_{(1,C)1}, 0]$
10	$(1,Q) \to (2,Q)$	λ	$[x_0, x_1, 0]$	$[v_{(1,Q)0}, v_{(1,Q)1}, 0]$
11	$(2,Q) \to (C_2,Q)$	R_1	$[x_0, x_1, x_2]$	$[v_{(2,Q)0}, v_{(2,Q)1}, v_{(2,Q)2}]$
12	$(C_2,Q) \to (1,Q)$	$(1-p)H_1$	$[x_1, x_2, 0]$	$[v_{(C_2,Q)1}, v_{(C_2,Q)2}, 0]$
13	$(C_2,Q) \to (2,Q)$	pH_1	$[x_0, x_1, x_2]$	$[v_{(C_2,Q)0}, v_{(C_2,Q)1}, v_{(C_2,Q)2}]$
14	$(C_1,Q) \to (C_2,Q)$	λ	$[x_0, x_1, 0]$	$[v_{(C_1,Q)0}, v_{(C_1,Q)1}, 0]$
15	$(2,Q) \to (2,C)$	R_2	$[x_0, x_1, x_2]$	$[v_{(2,Q)0}, v_{(2,Q)1}, v_{(2,Q)2}]$
16	$(2,C) \to (2,Q)$	H_2	$[x_0, x_1, x_2]$	$[v_{(2,C)0}, v_{(2,C)1}, v_{(2,C)2}]$
17	$(1,C) \to (2,C)$	λ	$[x_0, x_1, 0]$	$[v_{(1,C)0}, v_{(1,C)1}, 0]$

effect on the age of information in the CSMA networks. Especially, right before transition $l = 12$, the tagged node had two updates in its queue, and the age vector was x_0, x_1, x_2. As transition $l = 12$ occurs, the tagged node's update departs the queue and gets to the monitor, causing age reduction $x'_0 = x_1$ at the monitor. Thus, the second update in the queue moves to the first place, leading to $x'_1 = x_2$. Consequently, the age map resets to $\mathbf{x}\mathbf{A}_l = [x_1, x_2, 0]$. The rest of the transitions are straightforward and the explanation will be omitted due to the space limit.

The first-order differential equation of the continuous process $\mathbf{x}$ satisfies $\dot{\mathbf{x}} = \mathbf{b}_{\vec{q}} = \begin{cases} [1 \quad 0 \quad 0], & \vec{q} = (0,*), \\ [1 \quad 1 \quad 0], & \vec{q} = (1,*), (C_1,*), \\ [1 \quad 1 \quad 1], & \vec{q} = (2,*), (C_2,*), \end{cases}$ indicating the age-increasing process when residing in each state. Specifically, the age x_0 always increases at a unit rate, and the age x_1 increases at a unit rate in states other than $\vec{q} = (0,*)$, when there is an update waiting at the first place of the queue. The age x_2 increases at a unit rate in states $\vec{q} = (2,*)$ and $(C_2,*)$, when there is an update waiting at the second place of the queue. Similarly, we obtain the stationary probability with equation $\bar{\pi}\mathbf{D} = \bar{\pi}\mathbf{B}$, and use equation (8) to solve $\bar{\mathbf{v}}$.

The average AoI at the monitor for the tagged node with queue size $K = 2$ is

$$\Delta_{K=2} = \sum_{\vec{q} \in \mathbb{Q}} \bar{v}_{\vec{q}0} = \bar{v}_{(0,Q)0} + \bar{v}_{(1,Q)0} + \bar{v}_{(C_1,Q)0} + \bar{v}_{(0,C)0} + \bar{v}_{(1,C)0} + \bar{v}_{(2,Q)0} + \bar{v}_{(C_2,Q)0} + \bar{v}_{(2,C)0}. \tag{18}$$

The modeling efforts above also demonstrate the principles for further extension: the perceived age at the monitor, x_0, in fact, depends on the age values of all the packets in the queue, x_1 to x_k, as every buffered packet will ultimately move to the head of the queue and their age values then impact x_0. Such a principle guides the analysis of the age reset matrix $\mathbf{A}_l$ and the establishment of the appropriated balance equations. We also construct and compute the SHS model with $K = 3$, and related numerical results are given in Section V. We, however, omit the modeling details for $K = 3$ due to the page limit.

TABLE III
IEEE 802.11 CONFIGURATIONS

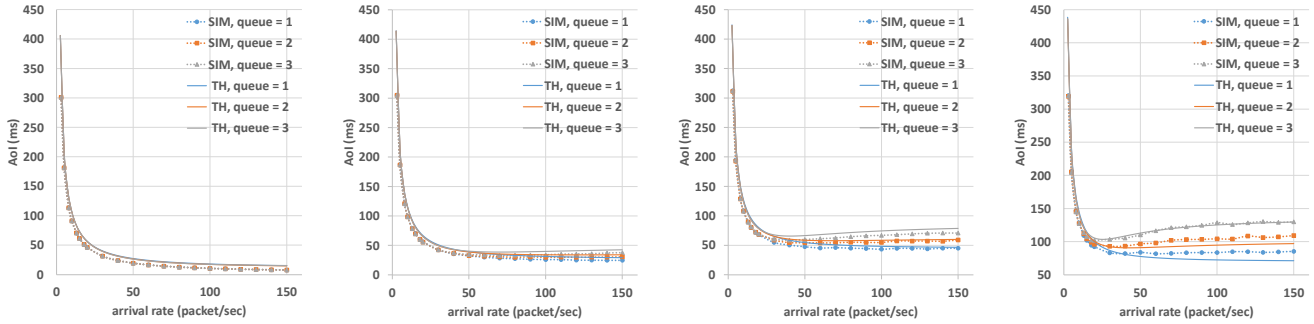
Bit rate for DATA frame	11 Mbps
Bit rate for ACK frame	1 Mbps
Bit rate for PLCP & Preamble	1 Mbps
Slot time	$20\,\mu s$
DIFS	$50\,\mu s$
SIFS	$10\,\mu s$
PHY header	192 bits
MAC header	224 bits
IP header	160 bits
Packet payload size	8000 bits
ACK	112 bits + PHY header
Initial contention window size	31
Maximum backoff stages	5
Maximum retransmission limit	7

V. NUMERICAL RESULTS

In this section, we apply our SHS model for AoI analysis and optimization over the practical IEEE 802.11 DCF protocol. Note that, to the best of our knowledge, this work is the first analytical AoI study from a tagged node's perspective over a collision-prone CSMA with saturated and unsaturated background nodes; we thus could not get comparable existing results in the same setting for performance benchmark. Therefore, we resort to ns-3 simulations to examine the accuracy of our SHS-based analysis. Furthermore, we consider that in practice, status monitoring normally generates small packets and can be piggybacked with data packets. We thus also apply our model to numerically assist the joint throughput and AoI optimization. With a slight abuse of the notation for convenience, we use "n" to indicate the number of background nodes in all the results.

A. Simulations Setup

For the illustration purpose without losing generality, we construct a WiFi network following the MAC protocol's IEEE 802.11 distributed coordination function with detailed configurations shown in Table III. To focus on our objective, we eliminate uncertain influence caused by hidden terminal



(a) $n = 2$ (b) $n = 6$ (c) $n = 10$ (d) $n = 15$

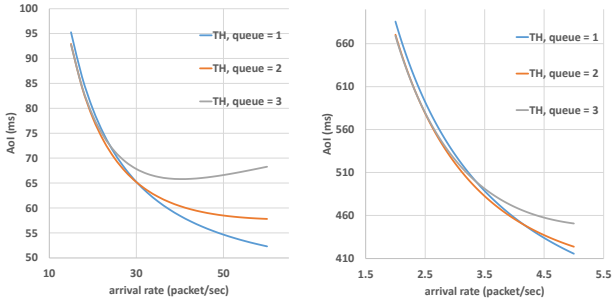
Fig. 4. The average AoI of the tagged node versus its traffic arrival rate (the sampling rate), with various numbers of background nodes.

problems by carefully arranging the location, power, and gain of each node.

We consider the WiFi network with each node equipped with a traffic generator following Poisson distribution. Each successful data transmission only contains one data frame. The average AoI for each experimental setting was computed over 2000 simulation runs, with each run lasting 300 seconds. Also, for each run, the samples in the initial 10% of the simulation period are discarded. Our setting adopts the heterogeneous setting with high channel utilization, where the traffic arrival rate of each background node is uniformly selected from the range $[50, 500]$, and default buffer size is adopted.

B. Analytical Accuracy Validation and Optimal Sampling Rate

We here conduct performance analysis for the scenarios where the queue sizes of the tagged node are $1, 2$, and 3. Fig. 4 compares the theoretical (TH) and simulation (SIM) results of the "AoI versus arrival rate" of the tagged node, obtained from varying numbers of background nodes ($n = 2, 6, 10, 15$), respectively. The analytical results and simulation results closely match in all the settings. When the network expands to a larger scale, specifically when $n = 15$, we can tell the small gap between the analytical results and the simulation results in the range of larger traffic arrival rate (say, > 50 packets/s), while the shapes of the curves still match well. One main reason leading to the gap is that the SHS model incorporates collision effect with approximations from the saturated MAC analysis as given in Section IV-B. Thus, when there are more background nodes, which accordingly contain more unsaturated nodes in our heterogeneous setting, the analytical deviation due to saturated approximation will be more obvious, as shown in Fig. 4 for $n = 10$ and 15. It is noteworthy that the analytical results are robustly accurate for all scenarios with a buffer size of 3; the small gap to simulation results even tends to diminish when there are more background nodes ($n = 15$). This is due to the fact that a larger buffer under a high arrival rate facilitates the network nodes to become saturated, matching the approximation technique adopted in our model. Also, SHS assumes exponential distributions, which might deviate from the true situation where many nodes interact through CSMA. Such inaccuracy due to exponential approximation might be



(a) $n = 10$ (b) $n = 50$

Fig. 5. Tagged node's AoI with queue size $K = 1$, 2, and 3, respectively, for the scenarios of (a) $n = 10$ and (b) $n = 50$ background nodes.

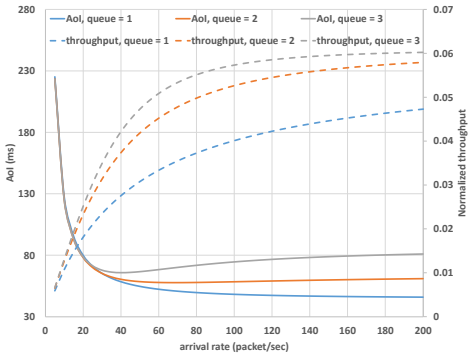
the main reason contributing to the phenomenon that the AoI curve has a larger gap in the range of a larger traffic arrival rate. The good news is that the optimal traffic arrival rate (i.e., the optimal sampling rate) minimizing the AoI suggested by the analytical results is robustly close to that indicated by the simulation curves in all the settings.

C. Queueing Benefit for Small Sampling Rates

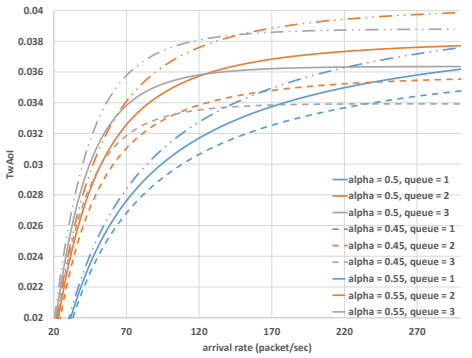
We further scrutinize the queueing effect on AoI as illustrated in Fig. 5. While the common sense is that buffering may negatively impact the AoI, the results in Fig. 5 reveal interesting insights when the tagged node works in a range of small sampling rates. In both scenarios with $n = 10$ and 50 nodes, we observe that employing a queue of size two yields a more favorable AoI compared to a queue of size one, especially when the traffic arrival rate is low. Such a finding is closely related to the transmission strategy considering collision. In a collision-prone networking context, the transmission opportunity upon an idle channel is a precious resource. A ready packet in the buffer can utilize such a resource more effectively and thus benefit the AoI. This finding suggests that an energy-constraint node, favoring a low sampling rate, can leverage the strategy of properly increasing its queue size to enhance the AoI performance.

D. Joint AoI and Throughput Optimization

Under practical conditions, status monitoring typically generates small-sized packets that can be conveniently integrated



(a) AoI and Throughput, $n = 10$



(b) TwAoI with various weights, $n = 10$

Fig. 6. Illustration of joint performance of AoI and throughput when $n = 10$.

with the transmission of data packets. With this in mind, we delve into performance optimization strategies that encompass both data throughput and the AoI. High throughput and low AoI are generally desirable traits, however, as illustrated in Fig. 6a, these features tend to have a negative correlation. To jointly optimize the throughput and AoI, we define "throughput weighted age of information (TwAoI)" denoted by γ as

$$\gamma = \alpha S + (1 - \alpha)\frac{1}{\Delta}. \tag{19}$$

Here, S represents the normalized throughput and $1/\Delta$ physically represents the age updating rate, making TwAoI intuitively balance the information transmission rate and the age updating rate in a certain application. The weight α can be adjusted to cater to different application requirements.

Fig. 6b showcases the TwAoI for a tagged node under various weight and queue size configurations. The results illustrate the delicate interplay between the weight distributions, queue sizes, and the resultant TwAoI. The important finding is that the setting with a buffer size $K = 2$ robustly performs well (although not always the largest) in terms of TwAoI under different weighting factor values or traffic arrival rates. By manipulating queue sizes and adjusting sampling rates, nodes can finetune their operations to find an optimal balance between throughput and AoI. Such adjustments could involve increasing queue size for a throughput-oriented strategy or shifting to a more aggressive sampling rate for AoI-focused scenarios.

VI. Related Work

The essence of AoI studies on queueing theory in work [3]–[8] is to determine the optimal queue arrival rate, striking a balance between the update frequency and queue congestion. Additionally, the lossy queue has been studied in the context of packet management, where arriving updates may be discarded while the server is busy or preempted (replaced) by fresher arrivals for older updates [9], [11]. The zero-wait policy [45] ensures that a new status update arrives just as the previous update packet departs the queue, leading to optimal throughput and delay performance. In contrast, works [6], [7] reveal that the zero-wait policy may not be optimal when channel resources are limited. Additionally, SHS [4], [17] has emerged as a powerful tool for AoI analysis, discovering in multiple sources real-time status updating scenarios where LCFS with preemption is considered. In this paper, resorting to the SHS modeling for AoI and optimal sampling rate analysis over the distributed MAC protocol is an innovative contribution.

The AoI optimization in CSMA networks is still an under-explored area with limited literature. Some works propose age-efficient transmission policies [33], [34] optimizing AoI in slotted ALOHA and CSMA networks using packet age-gain. In the work [42], AoI minimization in CSMA networks is investigated without considering the queueing effect. The focus is on minimizing the AoI through the optimization of backoff rates, which are upper-bounded to account for practical collision considerations. In [43], average AoI is analyzed in random access networks with stochastic packet generation considering collision effects but with the sources adopting the queueing discipline that always keeps only the freshest update. However, all the studies are limited to over-simplified models, simulation-oriented, collision-free, or requiring changes to the standard MAC protocol. The works [40], [41] make efforts to analyze AoI over a practical WiFi network by approximating the tagged node as a large FCFS server. However, the small AoI normally favors a small buffer system, this study moves a significant step forward to fill the gap.

VII. Conclusion

This research introduces a novel approach for AoI analysis and minimization in CSMA networks using SHS. We contribute original techniques that can incorporate finite buffer states, collision in transmission, and heterogeneous background nodes (through state aggregation) into the SHS model. Our model successfully optimizes the sampling rate for AoI minimization under various practical networking scenarios. Additionally, the study unveils valuable insights into the interplay between buffer size, throughput, and AoI. Simulations confirm the model's accuracy and the optimal sampling rate's effectiveness. As a next step, it will be interesting to apply our model for AoI analysis in other types of networks in the CSMA family, such as the IEEE 802.15.4 networks.

VIII. Acknowledgment

This work was supported in part by the U.S. National Science Foundation (NSF) under Grants CNS-2008092 and CNS-1816908.

REFERENCES

[1] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. of IEEE SECON*, 2011, pp. 350–358.

[2] S. Kaul, R. Yates, and M. Gruteser, "On piggybacking in vehicular networks," in *Proc. of IEEE GLOBECOM*, 2011, pp. 1–5.

[3] ——, "Real-time status: How often should one update?" in *Proc. of IEEE INFOCOM*, 2012, pp. 2731–2735.

[4] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1807–1827, 2019.

[5] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. of IEEE ISIT*, 2012, pp. 2666–2670.

[6] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[7] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. of IEEE ISIT*, 2015, pp. 3008–3012.

[8] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case," in *Proc. of IEEE ISIT*, 2017, pp. 326–330.

[9] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *Proc. of IEEE ISIT*, 2014, pp. 1583–1587.

[10] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *Proc. of IEEE ISIT*, 2015, pp. 1681–1685.

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

[12] J. Zhong, R. D. Yates, and E. Soljanin, "Two freshness metrics for local cache refresh," in *Proc. of IEEE ISIT*, 2018, pp. 1924–1928.

[13] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2215–2228, 2020.

[14] Y. Sun and B. Cyr, "Information aging through queues: A mutual information perspective," in *Proc. of IEEE SPAWC*, 2018, pp. 1–5.

[15] B. Yin, S. Zhang, Y. Cheng, L. X. Cai, Z. Jiang, S. Zhou, and Z. Niu, "Only those requested count: Proactive scheduling policies for minimizing effective age-of-information," in *Proc. of IEEE INFOCOM*, 2019, pp. 109–117.

[16] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 204–219, 2019.

[17] J. P. Hespanha, "Modelling and analysis of stochastic hybrid systems," *IEE Proceedings-Control Theory and Applications*, vol. 153, no. 5, pp. 520–535, 2006.

[18] Q. He, D. Yuan, and A. Ephremides, "Optimal link scheduling for age minimization in wireless systems," *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 5381–5394, 2018.

[19] Z. Jiang, B. Krishnamachari, X. Zheng, S. Zhou, and Z. Niu, "Timely status update in wireless uplinks: Analytical solutions with asymptotic optimality," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3885–3898, 2019.

[20] 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 Transactions on Networking*, vol. 26, no. 6, pp. 2637–2650, 2018.

[21] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. of IEEE INFOCOM*, 2018, pp. 1844–1852.

[22] ——, "Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints," *IEEE/ACM Transactions on Networking*, vol. 27, no. 4, pp. 1359–1372, 2019.

[23] C. Joo and A. Eryilmaz, "Wireless scheduling for information freshness and synchrony: Drift-based design and heavy-traffic analysis," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2556–2568, 2018.

[24] N. Lu, B. Ji, and B. Li, "Age-based scheduling: Improving data freshness for wireless real-time traffic," in *Proc. of ACM MobiHoc*, 2018, pp. 191–200.

[25] Y.-P. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," in *Proc. of IEEE ISIT*, 2017, pp. 561–565.

[26] Y.-P. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *Proc. of IEEE ISIT*, 2018, pp. 2634–2638.

[27] B. Zhou and W. Saad, "Minimizing age of information in the internet of things with non-uniform status packet sizes," in *Proc. of IEEE ICC*, 2019, pp. 1–6.

[28] Q. He, G. Dán, and V. Fodor, "Joint assignment and scheduling for minimizing age of correlated information," *IEEE/ACM Transactions on Networking*, vol. 27, no. 5, pp. 1887–1900, 2019.

[29] 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. of ITC*, vol. 01, 2018, pp. 144–152.

[30] R. Talak, S. Karaman, and E. Modiano, "Distributed scheduling algorithms for optimizing information freshness in wireless networks," in *Proc. of IEEE SPAWC*, 2018, pp. 1–5.

[31] Z. Jiang, S. Zhou, Z. Niu, and C. Yu, "A unified sampling and scheduling approach for status update in multiaccess wireless networks," in *Proc. of IEEE INFOCOM*, 2019, pp. 208–216.

[32] H. Chen, Y. Gu, and S.-C. Liew, "Age-of-information dependent random access for massive IoT networks," in *Proc. of IEEE INFOCOM WKSHPS*, 2020, pp. 930–935.

[33] X. Chen, K. Gatsis, H. Hassani, and S. S. Bidokhti, "Age of information in random access channels," in *Proc. of IEEE ISIT*, 2020, pp. 1770–1775.

[34] ——, "Age of information in random access channels," *IEEE Transactions on Information Theory*, pp. 1–1, 2022.

[35] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age of information performance of multiaccess strategies with packet management," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 244–255, 2019.

[36] A. M. Bedewy, Y. Sun, R. Singh, and N. B. Shroff, "Optimizing information freshness using low-power status updates via sleep-wake scheduling," in *Proc. of ACM MobiHoc*, 2020, p. 51–60.

[37] V. Tripathi, N. Jones, and E. Modiano, "Fresh-CSMA: A distributed protocol for minimizing age of information," in *Proc. of IEEE INFOCOM*, 2023, pp. 1–10.

[38] H. Pan, T.-T. Chan, J. Li, and V. C. M. Leung, "Age of information with collision-resolution random access," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 10, pp. 11 295–11 300, 2022.

[39] Z. Jiang, Y. Liu, J. Hribar, L. A. DaSilva, S. Zhou, and Z. Niu, "Smart: Situationally-aware multi-agent reinforcement learning-based transmissions," *IEEE Transactions on Cognitive Communications and Networking*, vol. 7, no. 4, pp. 1430–1443, 2021.

[40] S. Wang, Y. Cheng, L. X. Cai, and X. Cao, "Minimizing the age of information for monitoring over a WiFi network," in *Proc. of IEEE GLOBECOM*, 2022, pp. 383–388.

[41] S. Wang and Y. Cheng, "A deep learning assisted approach for minimizing the age of information in a WiFi network," in *Proc. of IEEE MASS*, 2022, pp. 58–66.

[42] A. Maatouk, M. Assaad, and A. Ephremides, "On the age of information in a CSMA environment," *IEEE/ACM Transactions on Networking*, vol. 28, no. 2, pp. 818–831, 2020.

[43] I. Kadota and E. Modiano, "Age of information in random access networks with stochastic arrivals," in *Proc. of IEEE INFOCOM*, 2021, pp. 1–10.

[44] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, 2000.

[45] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021.

[46] H. Zhai, Y. Kwon, and Y. Fang, "Performance analysis of IEEE 802.11 MAC protocols in wireless LANs," *Wireless communications and mobile computing*, vol. 4, no. 8, pp. 917–931, 2004.

[47] Y. Cheng, X. Ling, W. Song, L. X. Cai, W. Zhuang, and X. Shen, "A cross-layer approach for WLAN voice capacity planning," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 4, pp. 678–688, 2007.

[48] X. Ling, Y. Cheng, J. W. Mark, and X. Shen, "A renewal theory based analytical model for the contention access period of IEEE 802.15.4 MAC," *IEEE Transactions on Wireless Communications*, vol. 7, no. 6, pp. 2340–2349, 2008.