Abstract—Delay and Age-of-Information (AoI) are two crucial performance metrics for emerging time-sensitive applications in Industrial Internet of Things (IIoT). In order to achieve optimal performance, studying the inherent interplay between these two parameters in non-trivial task. In this work, we consider a Device-to-Device (D2D)-based heterogeneous IIoT network that supports two types of traffic flows, namely Aoi-orientated. First, we introduce a distributed backlog-aware random access protocol that allows the Aoi-orientated nodes to opportunistically access the channel based on the queue occupancy of the delay-oriented node. Then, we develop an analytical framework to evaluate the average delay and the average Aoi, and formulate an optimization problem to minimize the Aoi under a given delay constraint. Finally, we provide numerical results to demonstrate the impact of different network parameters on the performance in terms of the average delay and the average Aoi. We also give the numerical solutions of the optimal parameters that minimize the Aoi subject to a delay constraint.

Index Terms—Industrial IoT, AoI, D2D communication

I. INTRODUCTION

The emerging technology of Internet of Things (IoT) enables ubiquitous connectivity and innovative services spanning a wide range of applications, such as industrial automation, healthcare and intelligent transportation [1]. The realization of such applications implies the interconnection of a large number of heterogeneous and pervasive IoT devices continuously generating and transmitting sensor data to their respective destinations. According to the Global System for Mobile Communications Association (GSMA), it is expected that the number of connected IoT devices will grow up to 75 billion by 2025 [2]. As a subset of the IoT, Industrial IoT (IIoT) has gained significant attention as a key technology supporting machine type communication for time-sensitive industrial applications [3]. For a vast number of IIoT applications, the field network comprises a high number of sensors and actuators that sense the physical environment and perform appropriate actions, respectively [4]. Different from the conventional star-network configuration with a static sink, Device-to-device (D2D) communications provide a promising transmission paradigm for sensor-actuator pairs which are in propinquity [5]. In such setups, sensors directly transmit data to actuators without involving a central node, resulting in improved throughput and energy efficiency.

Typically, the field network in IIoT applications supports transmissions of heterogeneous traffic flows which are characterized by different performance objectives [6]. In process monitoring and control scenarios, the distributed sensor nodes either report special events to their corresponding actuators that perform actions in response, or transmit either status updates about a certain process. In the former case, it is essential to maintain low latency communication of event reporting from sensors to actuators in order to avoid production inefficiency or safety-critical situations. An event could be any incident happening, e.g., fire or leakage of gas, that should be delivered to the corresponding actuator within a predefined deadline. In the latter case, the freshness of the status updates is crucial for proper functioning of the system as the actuation decisions are mainly influenced by the freshest information available [7]. For instance, in oil refineries, valve actuators should acquire timely monitoring of oil level to avoid spilling of oil tanks [8].

Classical time-related performance metrics, such as throughput and delay, have a transmitter-side view of the channel and are insufficient to reflect information freshness at the receiver side. Age of Information (AoI) has been proposed to characterize information freshness which measures the time elapsed since the generation of the most recently received packet [9]. Although D2D communication can improve throughput and energy efficiency [10], the high number of deployed sensor-actuator pairs and the uncoordinated channel access pose a challenge on guaranteeing the performance objectives (low latency and low AoI) of heterogeneous traffic streams.

In this work, we consider a D2D-based heterogeneous IIoT network that supports two traffic flows, delay-oriented traffic and Aoi-oriented traffic. The delay-oriented traffic is transmitted by a single node and is subject to a predefined delay constraint, while the Aoi-oriented traffic is transmitted by a set of nodes distributed according to a stochastic point process. Our contributions in this work can be summarized as follows:

- We introduce a distributed backlog-aware random access protocol, in which the Aoi-oriented nodes access the shared channel based on the backlog status of the delay-oriented node.
- Using stochastic geometry, Discrete-Time Markov Chain (DTMC) and queuing theory, we derive closed-form expressions for the average AoI, average delay, and the average queue size of the delay-oriented node, and formulate an optimization problem to minimize the average...
AoI under a given delay constraint.

- We evaluate the network performance using numerical results in order to investigate the impact of different network parameters on the average AoI and the average delay.
- We solve the formulated optimization via a 2-D exhaustive search under different network parameters.

II. RELATED WORK

AoI has been receiving a growing attention as a promising metric for time-sensitive applications. Several works were conducted to study and analyze AoI utilizing queuing theory [11], [12]. The AoI minimization problem has been addressed through different approaches including packet management schemes [13], [14], adjusting sampling rate [9] and introducing packet deadlines [15]. The queuing analysis in these works considers only the temporal traffic dynamics associated with the transmitters and receivers based on the assumption of an error-free channel, i.e., without considering transmission failures due to interference and collisions that are inevitable in the industrial environment. In [16]–[19], the authors introduced a number of scheduling protocols with the goal to minimize the average AoI considering wireless networks with unreliable channels. All these works are based on a centralized scheduling policy which is not applicable to the D2D communication scenario in IIoT, as it would incur high communication overhead and cannot scale with the network size.

The main research on scheduling in wireless D2D networks focuses on the analysis and optimization of resource allocation [20], traffic density [21], or user fairness [22], however, there are few works that focus on the study and minimization of AoI. The authors in [23] statistically analyzed the peak AoI and the conditional success probability in a D2D communication scenario under preemptive and non-preemptive queues. Based on a stochastic geometry approach, the authors in [24] proposed a distributed scheduling policy that utilizes local observations, encapsulated in the concept of stopping sets to make scheduling decisions that minimize the network-wide average AoI. The work in [25] introduced a joint optimization problem of the transmit probability and transmit power to maximize the energy efficiency of a covert D2D communication under AoI constraint. The authors in [26] proposed an age-aware scheduling algorithm where edge controllers assist each D2D node to select a proper resource block and transmit power level to satisfy the AoI constraint. All these works consider that the network supports only a single traffic flow, which is age-oriented traffic and aim to minimize the average AoI without considering the coexistence of other traffic flows. The works in [27], [28] proposed scheduling policies to optimize the AoI while satisfying throughput constraints. The authors in [29] proposed a link scheduling policy for a joint optimization of AoI and throughput based on a deep learning approach. These works assume that, at each timeslot, the same user transmit either AoI-oriented packet or a throughput-oriented packet. The works in [30], [31] study the AoI-delay tradeoff for multi-server queuing systems considering error-free transmissions which is not applicable to real IIoT systems. In [32], the authors studied the delay and AoI violation probabilities considering finite-block length packets under noisy channels, however, the study was conducted separately for the two parameters.

III. SYSTEM MODEL

A. Network Topology and Communication Model

We consider a heterogeneous IIoT network comprising a set of sensor-actuators (D2D) pairs that are randomly distributed in a circular region $C$ with a radius $R$ as depicted in Fig. 1. The network supports two traffic flows, namely AoI-sensitive flow ($T_{AoI}$) and delay-sensitive flow ($T_D$). The $T_{AoI}$ traffic represents the status updates of an underlying time-varying process, where the $T_D$ traffic denotes event reporting about a certain incident.

In our model, we consider a single D2D pair for transmitting the $T_D$ traffic which characterized by a strict delay performance. The $T_D$ transmitter is fixed in predefined location where the $T_D$ receiver is located at the origin of considered network region with a distance $d_D$. The $T_D$ transmitter has an infinite output queue where the arrival process of the $T_D$ packets follows a Poisson process with average rate $\lambda_D$ packets/slot, which captures the scenario where such traffic can be triggered by the occurrence of some random incident [33]. The $T_D$ packets are transmitted in First-Come-First-Served (FCFS) basis, and a packet is removed from the buffer if it is successfully transmitted, which is acknowledged by the $T_D$ receiver through an ACK feedback. If no ACK is received within a predefined timeout, the packet is retransmitted. We assume that the ACK transmission is instantaneous and error-free, which is common assumption in the literature [34].

Further, we consider a set of $T_{AoI}$ transmitters randomly scattered according to a homogeneous Poisson Point Process (PPP) $\phi_A$ with spatial density $\lambda_A$, with $\phi_A = \{x_i = \mathbb{R}^2, \forall i \in \mathbb{N}^+, i \geq 1\}$ [35], where $x_i$ is the location of the $i$-th $T_{AoI}$ transmitter. Each $T_{AoI}$ transmitter has a dedicated receiver which is located at a distance $d_A$ in a random orientation. The $T_{AoI}$ nodes report status updates following the generate-at-will
model\textsuperscript{[9], [36]}, i.e., the T\textsubscript{Aol} transmitter is activated with a certain probability, and once becoming active, it samples a fresh packet and transmits it to the T\textsubscript{Aol} receiver\textsuperscript{[7]}. After a transmission attempt, the T\textsubscript{Aol} discards the packet regardless it is a successful or failed attempt\textsuperscript{[8]}. The network operates in a time-slotted fashion where the generation and the transmission of a packet is aligned with the boundary of a timeslot and a transmission attempt takes a constant time, which is assumed to be equal to the duration of one timeslot. The T\textsubscript{D} transmitter freely access the channel as long as there are packets in its queue. The communication between each D2D pair is performed via a shared communication channel in a slotted ALOHA-like access using a slot-access probability that is subject to optimization. In other words, in every slot there is a random number of interfering users that are active, comprising the T\textsubscript{D} and/or the activated T\textsubscript{Aol} transmitters. The collisions among simultaneously active users are not necessarily destructive, due to the capture effect\textsuperscript{[39]}. In that sense, a packet is decoded successfully at the end of a timeslot when the Signal-to-Interference plus-Noise ratio (SINR) at the corresponding receiver exceeds the capture ratio $\beta$. The received SINR at an arbitrary receiving node $i$ is given as

$$\text{SINR}_i = \frac{P_i |h_{i,i}|^2 d_{i,i}^{-\alpha}}{\sigma^2 + \sum_{j \in \mathcal{K}(i)} P_j |h_{j,i}|^2 d_{j,i}^{-\alpha}}, \quad (1)$$

where $P_i$ is the transmitting power of node $i$, $h_{j,i}$, $d_{j,i}$ represents the Raleigh fading of the channel between the transmitter $j$ and the receiver $i$ with $h_{j,i} \sim \mathcal{CN}(0,1)$, $d_{j,i}$ is the distance between the transmitter $j$ and the receiver $i$, $\alpha$ is the path loss exponent, $\mathcal{K}$ is the set of transmitting nodes at the same timeslot and $\sigma^2$ denotes additive white Gaussian noise (AWGN) power.

### B. Analysis of the Average AoI and the average Delay

In this section, we analyze the Average AoI of the T\textsubscript{Aol} traffic considering that the status updates from all T\textsubscript{Aol} nodes are equally important, and without loss of generality, we evaluate the average AoI at an arbitrary T\textsubscript{Aol} receiver in discrete time. In our model, the AoI represents the number of time slots elapsed since the last received T\textsubscript{Aol} packet was generated. If $\Delta(t)$ denotes the AoI at the end of time slot $t$, then we have

$$\Delta(t + 1) = \begin{cases} \Delta(t) + 1 & \text{unsuccessful transmission} \\ 1 & \text{successful transmission}. \end{cases} \quad (2)$$

Fig.\textsuperscript{[2]} shows the Discrete-Time Markov Chain (DTMC) model of the AoI correspond to an arbitrary T\textsubscript{Aol} node where each state represents the value of the AoI at the tagged T\textsubscript{Aol} receiver. The DTMC transits from any state $j$ to 1 only upon a successful reception of a T\textsubscript{Aol} packet, otherwise it transits to state $j + 1$. Let $X_t$ represent the value of $\Delta(t)$ at time slot $t$, then the transition probability from state $j$ to state $k$ is $P_{jk} = \Pr(X_{t+1} = k | X_t = j)$, and the transition matrix $P_{\text{Aol}}$ is written as

$$P_{\text{Aol}} = \begin{bmatrix} s_A & 1 - s_A & 0 & 0 & \ldots \\ s_A & 0 & 1 - s_A & 0 & \ldots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \end{bmatrix}, \quad (3)$$

where $s_A$ denotes the average successful update probability per timeslot, which is defined as the probability that a T\textsubscript{Aol} is delivered successfully to its corresponding receiver node. The row vector $\pi_{\text{AoI}} = [\tilde{\pi}_1, \tilde{\pi}_2, \ldots, \tilde{\pi}_{l-1}, \tilde{\pi}_l, \ldots]$ represents the steady-state probability vector of the DTMC in Fig.\textsuperscript{[2]} where $\tilde{\pi}_l = \lim_{t \to +\infty} P(X_t = l)$ denotes the probability that the AoI is equal to $l$ at the steady state.

Using the set of equations $\pi_{\text{Aol}} P_{\text{D}} = \pi_{\text{Aol}}$ and $\sum_l \pi_l = 1$, $\pi_l$ is obtained as

$$\tilde{\pi}_l = s_A (1 - s_A)^{(l-1)}, \quad \forall l. \quad (4)$$

Accordingly, the average AoI $\bar{\Delta}$ is calculated as

$$\bar{\Delta} = \sum_{l=1}^{\infty} \pi_l l = \sum_{l=1}^{\infty} l s_A (1 - s_A)^{(l-1)}$$

$$= \frac{s_A}{1 - s_A} \sum_{l=1}^{\infty} l (1 - s_A)^l. \quad (5)$$

Since we have $s_A < 1$,\textsuperscript{[12]} can be rewritten as

$$\bar{\Delta} = \frac{s_A}{1 - s_A} \frac{1 - s_A}{s_A} = \frac{1}{s_A}. \quad (6)$$

However, the average AoI cannot account for extreme AoI events occurring with very low probabilities. As mentioned in Section\textsuperscript{[22]}, the received updates of the T\textsubscript{Aol} traffic are used for control and actuation actions, which implies certain requirements on tolerated values of AoI. In that context, we also analyze the AoI violation probability, which is the probability that the AoI exceeds a certain constraint. Denote $c$ as the target AoI constraint, then we can express the AoI violation probability as

$$\Pr(\Delta > c) = 1 - \Pr(\Delta \leq c) = 1 - \sum_{l=1}^{c} \pi_l$$

$$= 1 - \frac{s_A}{1 - s_A} \sum_{l=1}^{c} (1 - s_A)^l \overset{(b)}{=} (1 - s_A)^c, \quad (7)$$

\textsuperscript{1}The generate-at-will model is commonly used for optimal AoI performance as it avoids unnecessary staleness to the sampled information due to queuing\textsuperscript{[9], [36]–[38]}.\textsuperscript{2}However, our analysis can be extended to other arrival models.

\textsuperscript{2}This scenario is based on the fact that retransmitting a stale packet might be useless for the receiver. In that sense, it would be better to transmit a fresh status update about the underlying industrial process instead of consuming the communication resources for retransmitting outdated information. Adopting retransmission-aided approach is out of the scope of this paper and left as a future work.

\textsuperscript{3}The received SINR at an arbitrary receiving node $i$ is given as

$$\text{SINR}_i = \frac{P_i |h_{i,i}|^2 d_{i,i}^{-\alpha}}{\sigma^2 + \sum_{j \in \mathcal{K}(i)} P_j |h_{j,i}|^2 d_{j,i}^{-\alpha}}, \quad (1)$$

where $P_i$ is the transmitting power of node $i$, $h_{j,i}$, $d_{j,i}$ represents the Raleigh fading of the channel between the transmitter $j$ and the receiver $i$ with $h_{j,i} \sim \mathcal{CN}(0,1)$, $d_{j,i}$ is the distance between the transmitter $j$ and the receiver $i$, $\alpha$ is the path loss exponent, $\mathcal{K}$ is the set of transmitting nodes at the same timeslot and $\sigma^2$ denotes additive white Gaussian noise (AWGN) power.
where (b) is obtained by using \( \sum_{i=0}^{y} x^i = \frac{x^{y+1}-1}{x-1} \) with \( x \neq 1 \).

We define the average delay \( D_{av} \) of a \( T_D \) packet as the time elapsed from the instant the packet was generated until it is successfully delivered to its destination. Therefore, \( D_{av} \) consists of the queuing time of the packet \( T_D \) packet and its transmission time. According to Little’s law [40], the queuing time is calculated as \( Q_{av} \), where \( Q_{av} \) denotes the average queue size of the \( T_D \) source. Then, the average delay \( D_{av} \) can be expressed as

\[
D_{av} = \frac{Q_{av}}{\lambda_D} + \frac{1}{s_D}, \tag{8}
\]

where \( s_D \) is the average probability of a successful transmission of a \( T_D \) packet per timeslot (average service probability), which is inversely proportional to the transmission time component of \( D_{av} \) [41]. As we can see from (6) and (8), \( \Delta \) and \( D_{av} \) depend on the values \( s_A, s_D, Q_{av} \). In the next sections, we derive the expressions of \( s_A, s_D, Q_{av} \) based on our proposed distributed backlog-aware channel access.

IV. ANALYSIS OF THE PROPOSED DISTRIBUTED BACKLOG-AWARE CHANNEL ACCESS

In the conventional slotted Aloha-based D2D communication, all nodes compete blindly and equally to transmit their data to their respective receivers via the shared channel. For a network with dense \( T_{AoI} \) links (which is the scenario of our considered network model), this means that the \( T_D \) node will have a low chance to gain access to the channel. This in turn results in a congestion problem where the buffered \( T_D \) packets suffer from an extended queuing time and a poor delay performance. However, such performance would be unacceptable in IIoT applications where safety-critical events (\( T_D \) traffic), such as fire detection or gas leakage, must be reported with very low delay in order to perform the appropriate actuation in response.

On the other hand, low AoI is required for the \( T_{AoI} \) traffic in order maintain fresh knowledge of the underlying industrial process. In that case, it is inefficient to force the \( T_{AoI} \) nodes to remain totally silent when the \( T_D \) node is active, in particular as the channel offers a possibility for capture effect to take place. In this regard, we propose a backlog-aware channel access to achieve a proper delay-AoI tradeoff in dense IIoT networks. In our proposed scheme, the \( T_D \) node transmits a packet in each timeslot as long as it has a non-empty queue, while the \( T_{AoI} \) nodes access the channel in an opportunistic, slotted ALOHA fashion, with the slot-access probability optimized according to the status of the queue size of the \( T_D \) node. Denoting \( Q \) as the random variable representing the queue size of the \( T_D \) node, the introduced access model is defined by the following cases.

Case 1: Empty Queue (\( Q = 0 \))

When the queue of the \( T_D \) node is empty, i.e. \( Q = 0 \), the \( T_{AoI} \) nodes access the channel with a probability \( p_1 \). Let \( \phi_{A_0} \) represent the locations of the active \( T_{AoI} \) transmitters when the \( T_D \) node is silent. Leveraging the thinning property of Poisson processes [42], \( \phi_{A_0} \) follows a homogeneous PPP with intensity \( p_1 \lambda_A \) and we have \( K = \phi_{A_1} \). According to Slivnyak’s theorem [42] and without loss of generality, the following analysis focuses on a typical active \( T_{AoI} \) pair. For the considered channel model and conditioned on the spatial realization \( \phi_{A_0} \), the successful decoding probability \( p_{A_0} \) at the tagged \( T_{AoI} \) receiver when \( Q = 0 \) is obtained as

\[
p_{A_0} = \Pr(\text{SINR} > \beta | K = \phi_{A_0})
= \Pr \left( \frac{P_2 |h_{i,i}|^2 d_i^{-\alpha}}{\sigma^2 + \sum_{j \in \phi_{A_0} \setminus \{i\}} P_2 |h_{j,i}|^2 d_j^{-\alpha}} > \beta \right)
= \exp \left( -\frac{p_1 \lambda_A d_i^2 \beta^2}{\sin^2(\frac{\beta}{\alpha})} \right) \exp \left( -\beta \sigma^2 d_i^2 \frac{P_2}{P_1} \right), \tag{9}\]

where \( P_2 \) is the transmitting power of the \( T_{AoI} \) nodes. The result in (a) follows by leveraging the Rayleigh distribution of the channel gain, i.e., \( |h_{i,j}|^2 \sim \exp(1) \) and the the probability generating functional (PGFL) of the PPP [35].

Case 2: Moderate Backlog (\( 1 < Q < M \))

We denote \( M \) as the backlog threshold that will determine the activity limit for the \( T_{AoI} \) nodes. When the queue is non-empty, the \( T_D \) node transmits in each timeslot, while the \( T_{AoI} \) nodes attempt to access the channel with a probability \( p_2 \). We assume that \( p_2 < p_1 \), as the \( T_{AoI} \) nodes should cause, as well as experience a lower interference level on the channel. Let \( x_D \) represent the location of the \( T_D \) node and \( \phi_{A_1} \) denote the locations of the active \( T_{AoI} \) nodes that follow a homogeneous PPP with intensity \( p_2 \lambda_A \). Then we have \( K = (x_D \cup \phi_{A_1}) \). In this case, the successful decoding probability \( p_{D_1} \) at the \( T_D \) receiver, conditioned on the spatial realization \( \phi_{A_1} \), is given as

\[
p_{D_1} = \Pr(\text{SINR} > \beta | K = (x_D \cup \phi_{A_1}))
= \Pr \left( \frac{P_2 |h_{D,i}|^2 d_i^{-\alpha}}{\sigma^2 + \sum_{j \in \phi_{A_1}} P_2 |h_{j,i}|^2 d_j^{-\alpha}} > \beta \right)
= \exp \left( -\frac{p_2 \lambda_A d_i^2 \beta^2}{\sin^2(\frac{\beta}{\alpha})} \right) \exp \left( -\beta \sigma^2 d_i^2 \frac{P_2}{P_1} \right). \tag{10}\]

For the active \( T_{AoI} \) nodes, we derive the successful decoding probability at an arbitrary \( T_{AoI} \) receiver using the following proposition.

Proposition 1: When \( 1 \leq Q \leq M \), the successful decoding probability \( p_{A_1} \) at an arbitrary \( T_{AoI} \) node is given by

\[
p_{A_1} = \exp \left( -\frac{p_2 \lambda_A d_i^2 \beta^2}{\sin^2(\frac{\beta}{\alpha})} \right) \frac{\exp \left( -\beta \sigma^2 d_i^2 \frac{P_2}{P_1} \right)}{1 + \frac{d_i^2}{\pi \sigma^2}} \tag{11}\]

where we have

\[
\mathbb{E}[d_{D,i}] = \int_0^{2\pi} \frac{1}{2\pi} \int_0^R \frac{2r}{R^2} \sqrt{r^2 + d_{D,0}^2 - 2d_0 d_{D,0} \cos \theta} \, dr \, d\theta.
\]

Proof: See Appendix A.
Case 3: Congested Queue \((Q > M)\)

When the queue size of the \(T_D\) node exceeds the backlog threshold \(M\), i.e., \(Q > M\), then all the \(TAoL\) nodes remain silent and only the \(T_D\) node transmits. In this case, the successful decoding probability \(PD_0\) at the \(T_D\) receiver is given by

\[
PD_0 = Pr(SINR_D > \beta | K = x_D) = Pr\left( \frac{P_h|x_D|^{2} d_{D}^{\alpha}}{\sigma^2} > \beta \right) = \exp\left( -\beta^2 d_{D}^{\alpha} \right).
\]

Note that \(PD_0 > PD_1\). Exploiting \([9], [10], [11], \) and \([12]\), we can express \(s_A\) and \(s_D\) as follows

\[
s_A = P_1 \Pr(SINR_{i\in\Phi_\Lambda_0} > \beta | Q = 0) \Pr(Q = 0) + P_2 \Pr(SINR_{i\in\Phi_\Lambda_0} > \beta | 1 \leq Q \leq M) \Pr(1 \leq Q \leq M) = P_1 \Pr(Q = 0) + P_2 \Pr(Q = M),
\]

\[
s_D = \frac{\Pr(1 \leq Q \leq M)PD_1 + \Pr(Q > M)PD_0}{\Pr(1 \leq Q \leq M) + \Pr(Q > M)}.
\]

As indicated by \([13] \) and \([14]\), both \(s_A\) and \(s_D\) depend on the distribution of the random variable \(Q\), which is analyzed in the next section.

V. THE DISTRIBUTION OF THE QUEUE SIZE \(Q\)

In this section, we analyze the distribution of the queue size \(Q\) to derive \(\Pr(Q > 0)\), \(\Pr(1 \leq Q \leq M)\), and \(\Pr(Q > M)\). The evaluation of the queue size \(Q\) can be modelled by the DTMC shown in Fig. 3, where each state represents the queue size. The success probability is \(PD_1\), when \(1 \leq Q \leq M\), while it is \(PD_0\) when \(Q > M\).

Let \(\pi_D = \{\pi_0, \pi_1, \ldots\}\) represent the steady-state probability vector of the queue size of the \(T_D\) node where \(\pi_n = \lim_{n \to \infty} \Pr(Q = n)\) is the steady-state probability of having \(n\) packets at the queue. The stationary distribution of the queue size of the \(T_D\) node is given by the following lemma:

**Lemma 1:** According to the DTMC in Fig. 3, the steady-state probability \(\pi_n\) is given as

\[
\pi_n = \begin{cases} 
\frac{\lambda_D(p_{D_0} - \lambda_D)}{\lambda_D(p_{D_1} - \lambda_D)} \pi_0, & 1 \leq n \leq M \\
\frac{\lambda_D p_{D_1} (1 - \lambda_D)^{n - 1}}{P_D 1_{(1 - \lambda_D)^n}} \pi_0, & n > M 
\end{cases}
\]

where \(\pi_0\) is given as

\[
\pi_0 = \begin{cases} 
\frac{(PD_1 - \lambda_D)(PD_0 - \lambda_D)}{PD_1 - \lambda_D - \lambda_D - (1 - \lambda_D)p_{D_1}} \lambda_D & \lambda_D \neq PD_1 \\
\frac{PD_0 - PD_1}{PD_1 - \lambda_D} & \lambda_D = PD_1
\end{cases}
\]

**Proof:** See Appendix B.

Note that the queue of the \(T_D\) node is stable under the condition that \(\lambda_D < PD_0\) (See Appendix C).

Based on the results of the steady-state probabilities in Lemma 1, the distribution of the queue size \(Q\) can be given as follows:

**Lemma 2:** For a stable queue \((\lambda_D < PD_0)\), and if \(\lambda_D \neq PD_1\) then the steady-state probabilities for the queue size are

\[
\Pr(1 \leq Q \leq M) = \frac{\lambda_D(1 - \psi^M)(PD_0 - \lambda_D)}{PD_1PD_0 - \lambda_DPD_1 - \lambda_D\psi^M(PD_0 - PD_1)},
\]

and we also obtain

\[
\Pr(Q > M) = \frac{\lambda_D\psi^M(PD_1 - \lambda_D)}{PD_1PD_0 - \lambda_DPD_1 - \lambda_D\psi^M(PD_0 - PD_1)},
\]

where we have

\[
\psi = \left( \frac{\lambda_D(1 - PD_1)}{(1 - \lambda_D)PD_1} \right).
\]

**Proof:** See Appendix D.

Based on the results obtained from Lemma 1 and Lemma 2, the average queue size \(Q_{avg}\) can be obtained from the following theorem:

**Theorem 1:** The average queue size of the \(T_D\) node is given by

\[
Q_{avg} = \sum_{n=1}^{\infty} n \pi_n = Q_1 + Q_2
\]

\[
Q_1 = \frac{Q_1 + Q_2}{PD_1PD_0 - \lambda_DPD_1 - \lambda_D\psi^M(PD_0 - PD_1)},
\]

where we have

\[
Q_1 = \lambda_D(1 - \lambda_D)PD_1 \left( \frac{PD_0 - \lambda_D}{PD_1 - \lambda_D} \right) \left( M\psi^{M+1} - \psi^M(M + 1) + 1 \right),
\]

\[
Q_2 = \psi^M \lambda_D(PD_1 - \lambda_D) \left( M + \frac{PD_0 - \lambda_D}{PD_1 - \lambda_D} \right).
\]

**Proof:** See Appendix E.

Hence, based on \([17] \) and \([18] \), we can obtain \(s_A\) and \(s_D\) as shown in \([21] \) and \([22] \). Then, by substituting \(s_A\) in \([6] \), we can obtain \(\Delta\). Similarly, by substituting \(s_D\) and \(Q_{avg}\) in \([8] \), we can obtain \(D_{avg}\).

As we can see from \([13] \) and \([6] \), when the \(T_D\) link is inactive, i.e., \(Q = 0\), the age of the \(TAoL\) nodes mainly depends on the access probability \(p_1\) for a given value of the transmission power \(P_2\). In that case, if we consider \(\Delta\) as a function of \(p_1\), then there is an optimal value \(p_1^*\) that minimizes

\[\text{Our derived expressions also hold for } \lambda_D = PD_1 \text{ by substituting } \pi_0 \text{ with its corresponding expression in } [16].\]
The optimization problem in (23) can be solved via the first-order optimality condition [43], which yields

\[
s_D = \frac{1}{1 - \left(\frac{(pD_1 - \lambda_D)(pD_0 - \lambda_D)}{pD_1pD_0 - \lambda_DpD_1 - \lambda_D\psi^M(pD_0 - pD_1)}\right)} \exp \left(-\frac{\beta\sigma^2d_D^2}{P_1}\right) \left(\frac{\lambda_D(1 - \psi^M)(pD_0 - \lambda_D)}{pD_1pD_0 - \lambda_DpD_1 - \lambda_D\psi^M(pD_0 - pD_1)}\right).
\]

By setting the obtained optimal access probability \( p_1^* \) in (13), and by fixing the arrival rate \( \lambda_D \) and the transmission power \( P_1 \), we can see that the average AoI \( \Delta \) mainly depends on the access probability \( p_2 \) and the transmission power \( P_2 \).

**Remark 1:** The average delay \( D_{\text{avg}} \) is an independent function of \( p_1 \) given a certain value of \( \Delta_D \). On the other hand, increasing \( p_2 \) and/or \( P_2 \) would in turn improve \( \Delta \) for the T AoI nodes but at the cost of increased delay \( D_{\text{avg}} \) for the T AoI node due to the increased level of interference. This is because the probability \( p_D_1 \) is decreased accordingly leading to increased queuing time at the T AoI node. Thus, the average delay \( D_{\text{avg}} \) is an increasing function of \( p_2 \) and \( P_2 \).

For a given delay constraint \( D_{\text{max}} \) of the T AoI traffic, there exists an optimal pair \((p_2^*, P_2^*)\) that minimizes the average AoI \( \Delta \). Considering the delay constraint \( D_{\text{max}} \) and the queue stability condition \( (\lambda_D < p_D) \) of the T AoI node, the optimal pair \((p_2^*, P_2^*)\) lies within a feasible region \( F_R \) that can be defined as

\[
F_R = \{(p_2, P_2) : D_{\text{avg}} < D_{\text{max}}, \lambda_D < p_D\}.
\]

In order to find the optimal pair \((p_2^*, P_2^*)\), we formulate the following optimization problem

\[
\min_{(p_2, P_2)} \Delta \quad \text{subject to} \quad D_{\text{avg}}(p_2, P_2) < D_{\text{max}} \quad 0 < p_2 \leq 1 \quad 0 < P_2 \leq P_{\text{max}},
\]

where \( P_{\text{max}} \) denotes the maximum transmission power of the T AoI nodes. Deriving closed form solutions for the optimal values \((p_2^*, P_2^*)\) is a hard problem due to the complexity of the corresponding expressions of \( s_D, s_A \) and \( D_{\text{avg}} \). Therefore, in the next section we do a numerical analysis of the solution to the optimization problem in (26).

### Table 1

**Evaluation Parameters.**

| Parameter | Value |
|-----------|-------|
| \( \lambda_D \) | \( 2 \times 10^{-4} \) |
| \( d_D \) | 100 m |
| \( d_A \) | 50 m |
| \( R \) | 300 m |
| \( \alpha \) | 3 |
| \( \beta \) | 0 dB |
| \( \sigma^2 \) | \(-90 \text{ dBm}\) |
| \( P_1 \) | 100 mW |
| \( P_{\text{max}} \) | 0.02 mW |
| \( D_{\text{max}} \) | 5 time slots |
VI. RESULTS AND DISCUSSION

In this section, we present numerical results to evaluate the performance of the proposed backlog-aware method in terms of the average delay ($D_{avg}$) and the average AoI ($\Delta$). We also introduce numerical solutions to the optimization problem (26) under different values of $\lambda_D$ and $M$. Unless stated otherwise, we adopt the evaluation parameters described in Table I.

Fig. 4 shows the successful decoding probabilities in (9), (10), (11) and (12) as a function of the access probabilities $p_1$ and $p_2$. The value of $p_{D0}$ is independent of the access probabilities $p_1$ and $p_2$, thus showing a constant value of 0.996. The successful decoding probability $p_{A0}$ depends only on $p_1$ while the values of $p_{D1}$ and $p_{A1}$ depend on $p_2$. Increasing $p_2$ implies higher collision probability between the T_AoI nodes, hence decreasing $p_{A0}$. Also, when the T_D node is active ($1 \leq Q \leq M$), increasing $p_1$ would in turn decrease both $p_{A1}$ and $p_{D1}$. In order to satisfy the queue stability condition $\lambda_D < p_{D0}$, we utilize $\lambda_D < 0.996$ for all the following results.

Fig. 5 presents the average AoI $\Delta$ versus the access probability $p_2$ under different values of $\lambda_D \in \{0.2, 0.6\}$ and $M \in \{1, 3\}$. Besides the numerical results obtained from (6), we also plot the simulation results obtained via discrete-time simulations in MATLAB to support our analytical analysis. Each simulation run lasts for $10^4$ timeslots, and we plot the average results over 100 runs. First, we can see that the simulation results match well with the numerical results, which validates our analytical analysis presented in Section *****. Moreover, we note that $\Delta$ is not a monotonic function of $p_2$. At very low values of $p_2$, the tagged T_AoI node is likely to be inactive for excessive period of time. For a certain values of $\lambda_D$ and $M$, the average AoI first decreases as $p_2$ increases where the tagged T_AoI node tends to transmit updates to its corresponding actuator more frequently. However, after a certain value of $p_2$, the average AoI at the corresponding actuator starts to significantly increase due to consecutive failed transmission attempts from the tagged source as a result of increased level of contention among the T_AoI nodes. We also observe that increasing $M$ would improve the AoI as it implies a higher chance for the T_AoI source to be active. For the same $M$, higher $\lambda_D$ degrades the AoI performance as the probability $Pr(Q > M)$ increases, which implies that the T_AoI nodes remain silent (Case 3 in Section IV).

Fig. 6 shows the evaluation of the average delay $D_{avg}$ against $p_2$ under different values of $\lambda_D \in \{0.2, 0.6, 0.8\}$ and $M \in \{1, 3, 9\}$. For a certain $M$, we can observe that the average delay increases rapidly as $p_2$ increases and tends to saturate at a certain point at high values of $\lambda_D$, which can be explained as follows. Increasing $p_2$ will in turn decrease $p_{D1}$ due to the increased interference from the T_AoI nodes, which leads to extended average queueing time at the T_D node. When $\lambda_D$ is high, it is likely that the T_AoI nodes remain inactive as the probability $Pr(Q > M)$ is high, hence the T_D node would probably have interference-free transmission with a stable delay. This also explains the counter-intuitive observation in Fig. 6 where the T_D node may experience a lower delay at high value of $\lambda_D$ than the one at low value of $\lambda_D$. The figure also shows that, for the same $\lambda_D$, increasing $M$ degrades the average delay, because the backlog-aware strategy becomes weaker leading to lower $s_D$ and higher $D_{avg}$ accordingly.

Fig. 7 plots the boundary of the feasible region $F_R$ defined in (25) with $\lambda_D \in \{0.2, 0.6\}$ and $M \in \{1, 3\}$. The area below each curve comprises the possible pairs $(p_2, p_3)$ that solve the optimization problem in (26) and satisfy the delay constraint $D_{max}$. As we can see from the figure, higher $M$ implies narrower $F_R$ where the average delay becomes higher, hence narrowing down the feasible region of $(p_2, p_3)$ that satisfies the delay constraint $D_{max}$.

In Table II we present the numerical solutions $(p_2^*, p_3^*)$ to the optimization problem in (26) under different values of $\lambda_D$ and $M$ and give the minimum achievable average AoI $\Delta^*$. The optimal solutions presented in this table are obtained via a 2-D numerical exhaustive search. As we can see, for the same value of $\lambda_D$, the minimum average AoI ($\Delta^*$) can be improved by increasing $M$ as the probability that the T_AoI node to be active is increased accordingly, i.e., higher $p_2^*$. However, higher $M$ increases the average delay as shown by Fig. 6 and tighten the feasible region $F_R$ in order to satisfy the constraint $D_{max}$. For instance, when $\lambda_D = 0.2$, increasing $M$ from 1 to 3 would improve $\Delta^*$ when the $p_2^*$ is increased from 0.275 to 0.303. However, increasing $M$ further to 6 would imply decreasing $p_2^*$ to 0.282 in order to keep $p_{D1}$ sufficiently high to comply with the delay constraint of $D_{max} = 5$ timeslots. In addition, for a given value of $M$, the optimal access probability $p_2^*$ and

TABLE II

| $\lambda_D$ | $M$ | $p_2^*$ | $p_3^*$ (mW) | $\Delta^*$ |
|-------------|-----|--------|-------------|-----------|
| 0.2         | 1   | 0.275  | 0.0122      | 8.22      |
|             | 3   | 0.303  | 0.0185      | 6.65      |
|             | 6   | 0.282  | 0.0161      | 6.01      |
| 0.6         | 1   | 0.145  | 0.0091      | 13.28     |
|             | 3   | 0.192  | 0.0132      | 11.22     |
|             | 6   | 0.184  | 0.0121      | 10.08     |
| 0.8         | 1   | 0.125  | 0.0063      | 16.81     |
|             | 3   | 0.132  | 0.0072      | 15.01     |
|             | 6   | 0.108  | 0.0058      | 14.09     |
the optimal transmission power $P_2^*$ are lower for higher $\lambda_D$ in order to reduce the interference from the $T_{AoI}$ nodes and satisfy $D_{max}$.

VII. CONCLUSION AND FUTURE WORK

In this paper, the inherent interplay between delay and AoI is investigated for a D2D-based heterogeneous IIoT network. We proposed a distributed backlog-aware access protocol and developed closed-form expressions for the average AoI and average delay using stochastic geometry and queuing theory. We evaluated the network performance using numerical results under different parameters and obtained optimal solutions that minimize the AoI via 2-D exhaustive numerical search. As a future work, this work can be extended to adopt queue management approaches and adaptive retransmission to introduce extra enhancement to the AoI. Moreover, we consider using machine learning-based techniques to set the optimal network parameters to adapt to the dynamic industrial environment.

ACKNOWLEDGEMENT

This paper has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 883315.

APPENDIX

A. Proof of Proposition 1

The successful update probability $p_{A_1}$ of an arbitrary $T_{AoI}$ pair when the $T_D$ node is active is given as
and an arbitrary φ can be rewritten as

\[ \exp \left( -s \sum_{j \in \phi, \phi \{i \}} |h_{j,i}|^2 d_{j,i}^{-\alpha} \right) = \exp \left[ \frac{\pi \rho_\lambda d_A^2 \beta \frac{x}{\alpha} \sin(\frac{\pi}{\alpha})}{\beta B} \right]. \]  \tag{29}

Based on (45), we can also have

\[ \mathbb{E} \left[ \frac{1}{1 + \frac{p_1}{\rho_\lambda} \beta d_A^2 d_{D,i}^{-\alpha}} \right] \approx \frac{1}{1 + \frac{d_A^2 \rho_\lambda}{\beta B}} \left( \frac{\beta B}{\pi} \right)^2. \]  \tag{30}

Then, with (29) and (30), \( p_{A_1} \) can be expressed as

\[ p_{A_1} \approx \exp \left[ -\frac{\pi \rho_\lambda d_A^2 \beta \frac{x}{\alpha} \sin(\frac{\pi}{\alpha})}{\beta B} \right] \frac{\exp \left( \frac{\beta x d_{A_1}}{d_A^2} \right)}{1 + \frac{d_A^2 \rho_\lambda}{\beta B} \left( \frac{\beta B}{\pi} \right)^2}. \]  \tag{31}

The level of interference experienced by an arbitrary \( T_{AoI} \) receiver from the \( T_D \) transmitter depends on the distance \( d_{D,i} \). Therefore, the term \( \mathbb{E}[d_{D,i}] \) is calculated by taking the expectation over all the possible locations of that \( T_{AoI} \) receiver within the considered region \( C \). Since the locations of the \( T_{AoI} \) transmitters follow a homogeneous PPP model, we can consider that the locations of the \( T_{AoI} \) receivers approximately follow a uniform distribution in the region \( C \) with a radius \( R \). Accordingly, the distance \( d_0 \) of an arbitrary \( T_{AoI} \) receiver from origin of \( C \) follow the probability distribution function \( f_{d_0}(r) \) which is given by

\[ f_{d_0}(r) = \begin{cases} \frac{2r}{R^2} & \text{if } 0 \leq r \leq R \\ 0 & \text{otherwise}. \end{cases} \]  \tag{32}

If we consider the region \( C \) depicted in Fig. 8, then we have

\[ d_{D,i} = \sqrt{d_{D,0}^2 + d_{D,0}^2 - 2d_0 d_{D,0} \cos \theta}, \]  \tag{33}

where \( d_{D,0} \) is the distance from the \( T_D \) transmitter to the origin of the region \( C \), and \( \theta \in \{0, 2\pi\} \) is a uniformly distributed random variable. Then, the expectation \( \mathbb{E}[d_{D,i}] \) can be obtained as

\[ \mathbb{E}[d_{D,i}] = \int_0^{2\pi} \frac{1}{2\pi} \int_0^R \frac{2r}{R^2} \sqrt{r^2 + d_{D,0}^2 - 2d_0 d_{D,0} \cos \theta} \, r \, dr \, d\theta. \]  \tag{34}

Based on (31) and (32), we can obtain the probability \( p_{A_1} \), which concludes the proof.

\section*{B. Proof of Lemma 1}

The balance equations for the DTMC in Fig. 3 can be obtained as

\[ \lambda_D \pi_0 = (1 - \lambda_D) p_{D_1} \pi_1 \Leftrightarrow \pi_1 = \pi_0 \frac{\lambda_D}{1 - \lambda_D} p_{D_1}, \]

\[ \pi_1 (\lambda_D (1 - p_{D_1}) + (1 - \lambda_D) p_{D_1}) = \lambda_D \pi_0 + \pi_2 p_{D_1}, \]

\[ \Leftrightarrow \pi_2 = \pi_0 \frac{\lambda_D^2 (1 - p_{D_1})}{(1 - \lambda_D)^2 p_{D_1}}. \]  \tag{35}

Fig. 7. The feasible region \( F_R \) under different values of \( \lambda_D \) and \( M \).

Fig. 8. Geometric representation of the distribution of the \( T_D \) transmitter and an arbitrary \( T_{AoI} \) receiver.

\[
p_{A_1} = \Pr(\text{SINR}_i > \beta | K = \phi_{A_1} \cup x_D) = \Pr \left( \frac{P_2 |h_{i,i}|^2 d_{A_1}^{-\alpha}}{\sigma^2 + \sum_{j \in \phi_{A_1} \setminus \{i\}} P_2 |h_{j,i}|^2 d_{j,i}^{-\alpha} + P_1 |h_{D,i}|^2 d_{D,i}^{-\alpha}} > \beta \right) \]  \tag{27}

Since we have \( |h_{i,i}|^2 \sim \exp(1) \) and \( |h_{D,i}|^2 \sim \exp(1), \) so (27) can be rewritten as

\[
p_{A_1} = \exp \left( -\frac{\beta \sigma^2 d_{A_1}^2}{P_2} \right) \mathbb{E} \left[ \int_0^\infty \exp \left( -\frac{P_1}{P_2} d_{D,i}^{-\alpha} x \right) e^{-x} \, dx \right] \]

\[
\cdot \mathbb{E} \left[ \exp \left( -s \sum_{j \in \phi_{A_1} \setminus \{i\}} |h_{j,i}|^2 d_{j,i}^{-\alpha} \right) \right] \]

\[
= \exp \left( -\frac{\beta \sigma^2 d_{A_1}^2}{P_2} \right) \mathbb{E} \left[ \int_0^\infty \frac{1}{1 + \frac{P_1}{P_2} \beta d_A^2 d_{D,i}^{-\alpha}} \right] \]

\[
\cdot \mathbb{E} \left[ \exp \left( -s \sum_{j \in \phi_{A_1} \setminus \{i\}} |h_{j,i}|^2 d_{j,i}^{-\alpha} \right) \right], \]  \tag{28}

where the term \( L_S \) denotes the Laplace transform of the interference component of the nodes in \( \phi_{A_1} \) which can be expressed as [44].
Hence, for $1 \leq n \leq M$ we have
\[ \pi_n = \frac{\lambda_D^n (1 - p_{D_1})^{(n-1)}}{p_{D_1}^n (1 - \lambda_D)^n} \pi_0, \tag{36} \]
while for $n > M$ we get
\[ \pi_n = \frac{\lambda_D^n (1 - p_{D_1})^M (1 - p_{D_0})^{(n-M-1)}}{p_{D_1}^M p_{D_0}^{(n-M)} (1 - \lambda_D)^n} \pi_0. \tag{37} \]

Since we have $\sum_{n=0}^{\infty} \pi_n = 1$, with $\lambda_D \neq p_{D_1}$, the empty queue probability $\pi_0$ is expressed as
\[ \pi_0 = \frac{(p_{D_1} - \lambda_D)(p_{D_0} - \lambda_D)}{p_{D_1} p_{D_0} - \lambda_D p_{D_1} - \lambda_D \left[ \frac{\lambda_D (1 - p_{D_1})}{(1 - \lambda_D p_{D_1})} \right]^M (p_{D_0} - p_{D_1})}. \tag{38} \]

When $\lambda_D = p_{D_1}$, (38) is no longer valid ($\pi_0 = \frac{0}{\theta}$). In this case we apply L'Hôpital's rule and get $\pi_0$ as
\[ \pi_0 = \lim_{\lambda_D \to p_{D_1}} \frac{f(\lambda_D)}{h(\lambda_D)} = \frac{p_{D_0} - p_{D_1}}{p_{D_1} + (p_{D_0} - p_{D_1}) M + 1 - p_{D_1}}, \tag{39} \]
where $f(\lambda_D)$ and $h(\lambda_D)$ represent the nominator and denominator of (38), respectively. Combining (38) and (39), we get (16) which concludes the proof of Lemma 1.

C. Stability Condition of the $T_{AoI}$ Node Queue

Our basic definition for the queue stability is based on Loynes’ theorem [46]. The theorem states that, a queue is stable if the average service probability is greater than average arrival probability with the condition that the arrival process and the service process are strictly jointly stationary and. Hence, in order for the queue of the $T_D$ to be stable, we need to prove that the DTMC in Fig. [3] irreducible Markov chain. Then, the average $\lambda_D$ in (38) must satisfy the condition $0 < \pi_0 < 1$. We consider the following cases of the values of $\lambda_D$ relative to the values of $p_{D_1}$ and $p_{D_0}$:

- When $\lambda_D < p_{D_1}$, we have $\frac{\lambda_D (1 - p_{D_1})}{(1 - \lambda_D p_{D_1})} < 1$, then $h(\lambda_D) > p_{D_0} (p_{D_1} - \lambda_D)$. Hence, we have $\pi_0 < \frac{p_{D_0} - \lambda_D}{p_{D_0}} = 1 - \frac{\lambda_D}{p_{D_0}} < 1$.

- When $\lambda_D = p_{D_1}$, based on (39), it is obvious that $0 < \pi_0 < 1$.

- When $p_{D_1} < \lambda_D < p_{D_0}$ we have $\frac{\lambda_D (1 - p_{D_1})}{(1 - \lambda_D p_{D_1})} > 1$, then we have $f(\lambda_D) < 0$. Also, as from the following inequality
\[ \left( \frac{1 - \lambda_D}{1 - p_{D_1}} \right)^M \frac{p_{D_0} - \lambda_D}{p_{D_0} - p_{D_1}} < 1 < \left( \frac{\lambda_D}{p_{D_1}} \right)^{M+1} \]
\[ \Rightarrow p_{D_1} (p_{D_0} - \lambda_D) < \lambda_D \left( \frac{\lambda_D (1 - p_{D_1})}{(1 - \lambda_D p_{D_1})} \right)^M (p_{D_0} - p_{D_1}) \]
\[ \Rightarrow h(\lambda_D) < 0. \]

Thus, we have $\pi_0 = \frac{f(\lambda_D)}{h(\lambda_D)} > 0$. Since $\frac{\lambda_D (1 - p_{D_1})}{(1 - \lambda_D p_{D_1})} > 1$, we also have $h(\lambda_D) < p_{D_0} (p_{D_1} - \lambda_D) < 0$, so we obtain $\pi_0 < 1 - \frac{\lambda_D}{p_{D_0}} < 1$.

From the discussion of the above cases, we can see that the condition $0 < \pi_0 < 1$ is verified, and we can conclude that the queue of $T_D$ node is stable if $\lambda_D < p_{D_0}$.

D. Proof of Lemma 2

The probability $\Pr(1 \leq Q \leq M)$ is obtained as
\[ \Pr(1 \leq Q \leq M) = \sum_{n=1}^{M} \pi_n. \tag{40} \]

Based on (15) and when $\lambda_D < p_{D_0}$ and $\psi \neq 1$, we get
\[ \Pr(1 \leq Q \leq M) = \frac{\pi_0 M}{1 - p_{D_1}} \sum_{n=1}^{M} \left( \frac{\lambda_D (1 - p_{D_1})}{1 - \lambda_D p_{D_1}} \right)^n \]
\[ = \frac{\pi_0 M}{1 - p_{D_1}} \frac{\lambda_D (1 - p_{D_1})}{1 - \lambda_D p_{D_1}} \]
\[ = \frac{\lambda_D (1 - \psi^M)}{p_{D_1} p_{D_0} - \lambda_D p_{D_1} - \lambda_D \psi^M (p_{D_0} - p_{D_1})}. \tag{41} \]

For $n > M$, we obtain $\Pr(Q > M)$ as
\[ \Pr(Q > M) = \sum_{n=M+1}^{\infty} \pi_n = 1 - \pi_0 - \sum_{n=1}^{M} \pi_n \]
\[ = \frac{\lambda_D \psi^M (p_{D_1} - \lambda_D)}{p_{D_1} p_{D_0} - \lambda_D p_{D_1} - \lambda_D \psi^M (p_{D_0} - p_{D_1})}. \tag{42} \]

E. Proof of Theorem 1

The average queue size $Q_{avg}$ of the $T_{AoI}$ node can be given as
\[ Q_{avg} = \sum_{n=1}^{\infty} n \pi_n = \sum_{n=1}^{M} n \pi_n + \sum_{n=M+1}^{\infty} n \pi_n \]
\[ = \sum_{n=1}^{M} n \pi_n + \sum_{n=M+1}^{\infty} (M + n) \pi_{M+n} \]
\[ = \sum_{n=1}^{M} n \pi_n + M \sum_{n=1}^{\infty} \pi_{M+n} + \sum_{n=1}^{\infty} n \pi_{M+n} \]
\[ = R_1 + R_2 + R_3. \tag{43} \]

With $\lambda_D < p_{D_0}$ and $\psi \neq 1$, the first term $R_1$ can be obtained as
\[
\sum_{n=1}^{M} n \pi_n = \sum_{n=1}^{M} n \frac{\lambda_D (1 - p_{D_1})}{\lambda_D (1 - \lambda_D)^{n}} \pi_0
\]
\[
= \frac{\pi_0 \lambda_D}{(1 - \lambda_D) p_{D_1}} \sum_{n=1}^{M} \left[ \frac{\lambda_D (1 - p_{D_1})}{p_{D_1} (1 - \lambda_D)} \right]^{n-1}
\]
\[
= \frac{\pi_0 \lambda_D}{(1 - \lambda_D) p_{D_1}} \sum_{n=1}^{M} \left[ \frac{\lambda_D (1 - p_{D_1})}{p_{D_1} (1 - \lambda_D)} \right]^{n-1}
\]
\[
= \frac{(1 - \lambda_D) \lambda_D}{M \psi^M (M + 1) - \psi^M (M + 1) + 1}
\]
\[
\frac{1}{\lambda_D (1 - \lambda_D) p_{D_1} p_{D_0} - \lambda_D p_{D_1} - \lambda_D \psi^M (p_{D_0} - p_{D_1})}
\]
(44)

Based on the results of (18), the term \( R_2 \) is obtained as

\[ M \sum_{n=1}^{\infty} \pi_{M+n} = M \Pr (Q > M) \]
\[ = \frac{\lambda_D \psi^M (p_{D_1} - \lambda_D)}{p_{D_1} p_{D_0} - \lambda_D p_{D_1} - \lambda_D \psi^M (p_{D_0} - p_{D_1})} \]
(45)

Finally, the term \( R_3 \) can be obtained using the results from (44), (45) and (46) in (43), we get

\[ Q_{avg} = \frac{Q_1 + Q_2}{p_{D_0} p_{D_1} - \lambda_D p_{D_1} - \lambda_D \psi^M (p_{D_0} - p_{D_1})}, \]
(47)

where we have

\[ Q_1 = \lambda_D (1 - \lambda_D) \frac{p_{D_0} - \lambda_D}{p_{D_1} - \lambda_D} \]
\[ \cdot \left( M \psi^{M+1} - \psi^M (M + 1) + 1 \right), \]
(48)

\[ Q_2 = \psi^M \lambda_D (p_{D_1} - \lambda_D) \left( M + \frac{p_{D_0} - \lambda_D}{p_{D_0} - \lambda_D} \right). \]

REFERENCES

[1] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, “A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications,” IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1125–1142, 2017.
[2] G. Intelligence. (2021) The mobile economy 2021. [Online]. Available: https://www.gsma.com/mobileeconomy/
[3] E. Sisinni, A. Saifullah, S. Han, U. Jennehag, and M. Gidlund, “Industrial internet of things: Challenges, opportunities, and directions,” IEEE Transactions on Industrial Informatics, vol. 14, no. 11, pp. 4724–4734, 2018.
[4] F. Ademaj, M. Rzymowski, H.-P. Bernhard, K. Nyka, and L. Kulas, “Relay-aided wireless sensor network discovery algorithm for dense industrial iot utilizing espar antennas,” IEEE Internet of Things Journal, vol. 8, no. 22, pp. 16653–16665, 2021.
[5] L. Liu and W. Yu, “A d2d-based protocol for ultra-reliable wireless communications for industrial automation,” IEEE Transactions on Wireless Communications, vol. 17, no. 8, pp. 5045–5058, 2018.
[6] H. Farag, E. Sisinni, M. Gidlund, and P. Österberg, “Priority-aware wireless fieldbus protocol for mixed-criticality industrial wireless sensor networks,” IEEE Sensors Journal, vol. 19, no. 7, pp. 2767–2780, 2019.
[7] M. Li, C. Chen, C. Hua, and X. Guan, “Learning-based autonomous scheduling for aoi-aware industrial wireless networks,” IEEE Internet of Things Journal, vol. 7, no. 9, pp. 9175–9188, 2020.
[8] P. Gil, A. Santos, and A. Cardoso, “Dealing with outliers in wireless sensor networks: An oil refinery application,” IEEE Transactions on Control Systems Technology, vol. 22, no. 4, pp. 1589–1596, 2014.
[9] 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.
[10] R. Li, P. Hong, K. Xue, M. Zhang, and T. Yang, “Resource allocation for uplink noma-based d2d communication in energy harvesting scenario: A two-stage game approach,” IEEE Transactions on Wireless Communications, vol. 21, no. 2, pp. 976–990, 2022.
[11] H. H. Yang, C. Xu, X. Wang, D. Feng, and T. Q. S. Quek, “Understanding age of information in large-scale wireless networks,” IEEE Transactions on Wireless Communications, vol. 20, no. 5, pp. 3190–3210, 2021.
[12] R. D. Yates, “The age of information in networks: Moments, distributions, and sampling,” IEEE Transactions on Information Theory, vol. 66, no. 9, pp. 5712–5728, 2020.
[13] 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.
[14] 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.
[15] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, “On the age of information with packet deadlines,” IEEE Transactions on Information Theory, vol. 64, no. 9, pp. 6419–6428, 2018.
[16] 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.
[17] R. Talak, S. Karaman, and E. Modiano, “Improving age of information in wireless networks with perfect channel state information,” IEEE/ACM Transactions on Networking, vol. 28, no. 4, pp. 1765–1778, 2020.
[18] C. Li, Q. Liu, S. Li, Y. Chen, Y. T. Hou, W. Lou, and S. Kompella, “Scheduling with age of information guarantee,” IEEE/ACM Transactions on Networking, pp. 1–14, 2022.
[19] H. Farag, M. Gidlund, and Č. Stefanović, “A deep reinforcement learning approach for improving age of information in mission-critical iot,” in 2021 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT), 2021, pp. 14–18.
[20] M. Elnourani, S. Deshmukh, and B. Beferull-Lozano, “Distributed resource allocation in underlay multicast d2d communications,” IEEE Transactions on Communications, vol. 69, no. 5, pp. 3409–3422, 2021.
[21] G. Chisci, H. Elsayw, A. Conti, M.-S. Alouini, and M. Z. Win, “Uncoordinated massive wireless networks: Spatiotemporal models and multiaccess strategies,” IEEE/ACM Transactions on Networking, vol. 27, no. 3, pp. 918–931, 2019.
[22] M. Liu and L. Zhang, “Resource allocation for d2d underlay communications with proportional fairness using iterative-based approach,” IEEE Access, vol. 8, pp. 143 787–143 801, 2020.
[23] P. D. Mankar, M. A. Abd-Elmagid, and H. S. Dhillon, “Spatial distribution of the mean peak age of information in wireless networks,” IEEE Transactions on Wireless Communications, vol. 20, no. 7, pp. 4465–4479, 2021.
