Age of Information for Short-Packet Covert Communication

Weiwei Yang, Xingbo Lu, Shihao Yan, Member, IEEE, Feng Shu, Member, IEEE, and Zan Li, Senior Member, IEEE

Abstract—In this letter, we develop a new framework to jointly characterize covertness and timeliness of short-packet communications, in which a new metric named covert age of information (CAoI) is first proposed and then a closed-form expression for the average CAoI is derived. Our examination explicitly reveals the tradeoff between communication covertness and timeliness affected by the block-length, transmit power and prior transmission probability. Multiple transmission designs are tackled in order to minimize the average CAoI subject to covertness constraint, where the resultant differences relative to the designs with traditional metrics (e.g., effective covert rate, AoI) as objective functions are clarified. Our examination demonstrates that the optimal block-length is not the largest one, which is optimal in delay-constrained covert communication without considering the communication timeliness, and the optimal prior transmission probability may not be 1/2, which has been widely assumed in the literature of covert communications.

Index Terms—Covert communication, short-packet communication, age of information, finite block-length, timeliness.

I. INTRODUCTION

COVERT communication, which enables reliable communication between a transmitter (Alice) and a receiver (Bob) while maintaining a low probability of this communication being detected by a warden (Willie), has recently received ever-increasing research interests [1]–[7]. As a pioneering work, [2] first proved the information-theoretic limits of covert communication rate named square root law, in which Alice could reliably and covertly deliver at most $\sqrt{n}$ bits over $n$ additive white Gaussian noise (AWGN) channel uses. However, its achievable covert rate is zero as it is an asymptotic argument for a large enough number of channel uses. Fortunately, subsequent studies showed that Willie’s observation uncertainties suffered from background noise [4], detection channel [5], or interference [6] can assist to achieve a strictly positive covert transmission rate. Furthermore, [7] first investigated the impact of finite channel uses on Willie’s detection, which demonstrates that the finite channel uses result in an observation uncertainty at Willie while satisfying the delay constraint at Bob. We note that the aforementioned works mainly focused on guaranteeing covertness of wireless transmission. However, for some wireless military communications, such as real-time battlefield state sensing, monitoring and controlling, both communication covertness and timeliness are of significant importance. Although the finite number of channel uses characterizes the transmission delay, the timeliness of the data packet has not been well investigated in the context of covert communications.

Age of information (AoI), which is defined as the time elapsed since the last successfully decoded packet was generated at a transmitter [8], has proposed as a new metric to characterize the timeliness of a data packet [9]. Various AoI metrics that are applicable to a wide variety of systems were reviewed in [10]. Recently, some works have focused the AoI with security requirements, where the average AoI under the passive eavesdropping [11] and active jamming attacks [12] were studied. In order to achieve a high level of security, the authors of [13] conducted a preliminary work on examining AoI in covert communications, where Willie’s detection error probability was maximized subject to the average AoI constraint. However, this work considered that a full-duplex receiver transmits artificial noise to confuse Willie and assumed that the data block-length is infinite, which limit the achievable AoI due to self-interference and transmission delay. To the best of our knowledge, a framework that jointly considers covertness and AoI in short-packet communications is still missing, where a larger number of channel uses help achieve communication covertness and reliability, but minimizing transmission delay desires a smaller number of channel uses. Furthermore, the optimal transmission design that jointly guarantees communication covertness and timeliness of short-packet communications has never been tackled.

Motivated by the above discussions, in this letter, a new framework to jointly characterize communication covertness and timeliness of short-packet communications is proposed, where a new metric named covert age of information (CAoI) is first introduced. Considering AWGN channels, a closed-form expression for the average CAoI is derived. It is shown that the average CAoI can jointly characterize the effective covert rate and transmission delay, which allows us to examine the tradeoff between communication covertness and timeliness by optimizing the block-length, the transmit power, and the prior transmission probability. In addition, the differences between the average CAoI and traditional metrics, such as the effective covert rate, the transmission delay, and the average AoI, are also clearly revealed. Our examination shows that the optimal block-length to minimize the CAoI is balanced between maximizing the effective covert rate and minimizing the transmission delay, which is not the largest one, and the optimal prior transmission probability is not always 1/2.
II. SYSTEM MODEL

In this letter, we consider a scenario where Alice periodically generates a short packet and opportunistically transmits it to Bob as soon as possible, while there is a warden Willie who tries to detect this transmission based on his observations. We assume that each of Alice, Bob, and Willie is equipped with a single antenna. Since we focus on jointly investigating the covertness and timeliness of short-packet communications, without loss of generality, we assume that the wireless channels from Alice to Bob and from Alice to Willie are subject to AWGN, and the communication time is divided into slots, each of which is with the duration $T$. In order to simultaneously ensure communication covertness and timeliness, we assume that a short data packet containing $D$ nats information is transmitted by Alice with the prior transmission probability $\rho_1$ at each time slot. Then, when Alice is transmitting, the received signals at Bob can be expressed as

$$y_b[i] = x[i] + n_b[i],$$

where $i = 1, \ldots, L$ denotes the index over $L$ channel uses, $x[i] \sim C\mathcal{N}(0, P)$ denotes the Gaussian signaling, $P$ denotes the transmit power at Alice, and $n_b[i] \sim C\mathcal{N}(0, \sigma_w^2)$ denotes the AWGN at Bob with zero mean and variance $\sigma_w^2$. Likewise, the observed signals at Willie are given by

$$\begin{align*}
\mathcal{H}_0 : y_w[i] &= n_w[i], \\
\mathcal{H}_1 : y_w[i] &= x[i] + n_w[i],
\end{align*}$$

where $\mathcal{H}_0$ is the null hypotheses and represents that Alice is silent, $\mathcal{H}_1$ is the alternative hypotheses and represents that Alice is transmitting, and $n_w[i] \sim C\mathcal{N}(0, \sigma_w^2)$ denotes the AWGN at Willie with zero mean and variance $\sigma_w^2$.

A. Detection Performance at Willie

We consider a worst-case scenario, where the complete knowledge on the transmit power $P$, the block-length $L$, and the prior transmission probability $\rho_1$ is known by Willie. As such, Willie can adopt an optimal statistical hypothesis test (i.e., likelihood ratio test, LRT) to distinguish between the null and alternative hypotheses. Accordingly, the total detection error probability can be expressed as

$$P_{we} = \rho_0 P_{FA} + \rho_1 P_{MD},$$

where $\rho_0 = 1 - \rho_1$ denotes the prior probability that Alice does not transmit, $P_{FA}$ denotes the false alarm probability that Willie agrees $\mathcal{H}_1$ while $\mathcal{H}_0$ is true, respectively, and $P_{MD}$ denotes the missed detection probability that Willie agrees $\mathcal{H}_0$ but $\mathcal{H}_1$ is true. Achieving covert communication requires to guarantee the constraint $P_{we} \geq \min(\rho_0, \rho_1) - \delta$ [2], where $P_{we}$ and $\delta$ denote the minimum total detection error probability and the predetermined covertness tolerance level, respectively.

Based on the observed signal sequences given in (2), LRT with the optimal threshold $\mu = \rho_0/\rho_1$ can be expressed as

$$\Lambda(y_w) = \frac{P_1}{P_0} \frac{\prod_{i=1}^{L} f(y_w[i]|\mathcal{H}_1)}{\prod_{i=1}^{L} f(y_w[i]|\mathcal{H}_0)} \geq \frac{D_1}{D_0},$$

where $f(y_w[i]|\mathcal{H}_0)$ and $f(y_w[i]|\mathcal{H}_1)$ respectively denote the likelihood functions under $\mathcal{H}_0$ and $\mathcal{H}_1$, i.e., $y_w[i]|\mathcal{H}_0 \sim C\mathcal{N}(0, \sigma_w^2)$ and $y_w[i]|\mathcal{H}_1 \sim C\mathcal{N}(0, P + \sigma_w^2)$, and $D_1$ and $D_0$ represent Willie’s decisions that the covert communication happens or not, respectively.

Following (4), the false alarm probability and missed detection probability can be derived, of which the expressions involve lower incomplete Gamma functions and are given in [7]. However, the lower incomplete Gamma function is hard to be used for further analysis. Therefore, in this work, we use a generalized bound on the total variation distance to guarantee the covertness constraint, i.e., $P_{we} \geq \min(\rho_0, \rho_1) - \max(\rho_0, \rho_1) V_T(P_0, P_1)$ where $V_T(P_0, P_1)$ denotes the total variation distance between $P_0$ and $P_1$. Accordingly, the covertness constraint can be further bounded by

$$D(P_0, P_1) \leq 2 \left( \frac{\delta}{\max(\rho_0, \rho_1)} \right)^2,$$

where $D(P_0, P_1) = L(\ln(1 + \frac{P}{\sigma_w^2}) - \frac{P}{P + \sigma_w^2})$ denotes the relative entropy (Kullback-Leibler divergence) from $P_0$ to $P_1$.

We note that the covertness constraint can be also interpreted as $P_{cop} \leq \delta$, where $P_{cop}$ denotes the covert communication outage probability, which is the probability that the covertness constraint cannot be satisfied. Following (5), we have an upper bound on $P_{cop}$, which is given by

$$P_{cop} = \max(\rho_0, \rho_1) \left( \frac{L}{2} \ln \left(1 + \frac{P}{\sigma_w^2} \right) - \frac{P}{P + \sigma_w^2} \right) \leq \delta.$$

B. Effective Covert Rate and Communication Delay

For the considered short-packet covert communication systems, the packet error probability cannot be ignored [14]. With the fixed rate $R_b = D/\lambda$, the packet error probability can be given by [15]

$$P_{be} \approx Q \left( \frac{\sqrt{L}(1 + \gamma_b)(\ln(1 + \gamma_b) - R_b)}{\sqrt{\gamma_b}(1 + \gamma_b)} \right),$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)/dt$, and $\gamma_b = P/\sigma_b^2$ denotes the signal-to-noise ratio (SNR).

Then, the effective covert rate with system bandwidth $W$ can be written as

$$R_c = \rho_1 W R_b (1 - P_{cop})(1 - P_{be}).$$

Besides, with the short-packet covert communications, the coding delay and propagation delay are typically much smaller than the transmission duration. Hereby, in this letter, we mainly focus on the transmission delay, which is equal to a time slot duration $T$ (i.e., $T = L/W$).

C. Cover Age of Information

We consider that Alice periodically generates a short data packet within the period $T$, but only opportunistically transmits the packet to guarantee the covertness of his wireless transmission behavior. As a result, some real-time data packets may not be timely transmitted. Furthermore, the covert outage probability $P_{cop}$ and packet error probability $P_{be}$ also potentially lead to invalid data packets at Bob, where a valid packet is that covertly transmitted and successfully decoded.

In order to jointly characterize the covertness and timeliness of short-packet communications, we introduce a new metric named CAnoI, which is defined as the time elapsed since the most recently valid data packet was originally generated at
Alice. Specifically, assuming the most recently valid packet was generated at the time $U(t)$, the CAoI at the time $t$ is

$$\Delta_{CAoI}(t) = t - U(t).$$  \hspace{1cm} (9)

We note that the CAoI linearly increases as time goes by, and it is reset as the age of valid packet only when a valid data packet is successfully received at Bob.

The average CAoI is the area under the CAoI curve averaged over time, which denotes the average time that Bob can receive a packet is successfully received at Bob. For convenience, we assume that the $(i-1)$-th valid data packet was generated at epoch $a_{i-1}$, and the departure time for the $(i-1)$-th and the $i$-th data packets are $a'_{i-1}$ and $a'_i$, respectively. Accordingly, the interval of departure time between the $(i-1)$-th and the $i$-th valid data packets can be denoted as $\Delta_i = a'_i - a'_{i-1}$. Furthermore, the sojourn time of the $(i-1)$-th valid data packet can be expressed as $Z_{i-1} = a'_{i-1} - a_{i-1}$, which is fixed and equal to the original age of valid data packet at Bob (i.e., $Z_{i-1} = T$). Assuming that there are $N(\tau)$ valid data packets over an interval $(0, \tau)$, the average arrival rate of valid data packet can be denoted as $\bar{\lambda}_b = \lim_{\tau \to \infty} \frac{N(\tau)}{\tau}$. We denote $S_i$ as the area of width $\bar{\lambda}_b \Delta_i$ over the departure time interval between the $(i-1)$-th and the $i$-th valid data packets, as shown in Fig. 1. Then, the average CAoI can be expressed as

$$\bar{\Delta}_{CAoI} = \lim_{\tau \to \infty} \frac{1}{\bar{\lambda}_b} \sum_{i=1}^{N(\tau)} S_i = \bar{\lambda}_b \mathbb{E}(S_i),$$  \hspace{1cm} (11)

of which a closed-form expression is derived in the following theorem.

**Theorem 1:** In the considered system, the closed-form expression for the average CAoI is derived as

$$\bar{\Delta}_{CAoI} = \frac{L \rho_1 W(1 - \bar{\rho}_c) (1 - P_{be})}{2W} + \frac{L}{2W}.$$  \hspace{1cm} (12)

**Proof:** Please find the detailed proof in Appendix.

Based on Theorem 1, we have the following insights. 1) The block-length $L$, transmit power $P$, and prior transmission probability $\rho_1$ have significative impact on the average CAoI. 2) The average CAoI jointly characterizes the traditional metrics including the effective covert rate $R_c$ and communication delay $T$, in which the average CAoI can be further written as $\bar{\Delta}_{CAoI} = D/R_c + T/2$. 3) The average CAoI can be simplified as the average AoI without covertness constraint, which is denoted as $\Delta_{AoI} = \frac{L(1 - P_{be})}{2W}$.

### III. Transmission Design to Minimize Average Covert Age of Information

#### A. Optimal $L$ to Minimize the Average CAoI

The block-length has a critical impact on the average CAoI, since it is a function of packet error probability, communication delay, and covert communication outage probability. Then, given the transmit power $P$ and prior transmission probability $\rho_1$, the optimization problem to minimize the average CAoI can be formulated as

$$\min_{L} \frac{L \rho_1 W(1 - \delta)(1 - P_{be})}{2W} + \frac{L}{2W}$$

subject to:

1) $D(P_0 \| P_1) \leq 2\left(\frac{\delta}{\max(\rho_0, \rho_1)}\right)^2$,

2) $\bar{L}_\min \leq L \leq \bar{L}_\max$, $L \in \mathbb{N}^+$.

where $c_2$ requires that the block-length $L$ is a positive integer, and is limited by the minimum value $\bar{L}_\min$ and the maximum value $\bar{L}_\max$. We note that the minimum value $\bar{L}_\min$, the given transmit power $P$, and the prior transmission probability $\rho_1$ should jointly satisfy the constraint $c_1$. Otherwise, the covert communication is not achievable.

To make the optimization problem (13) mathematically tractable, we adopt the following approximations. Firstly, the transmit power is typically small in covert communications, which results a low SNR. Accordingly, the covert constraint $c_1$ can be simplified as

$$c_1': \frac{L P^2}{\sigma_w^2 (\sigma_w^2 + P)} \leq 2\left(\frac{\delta}{\max(\rho_0, \rho_1)}\right)^2.$$  \hspace{1cm} (14)

Secondly, we relax the discrete integer constraint $L$ into a connected subset of the real axis, i.e., $c_2': \bar{L}_\min \leq L \leq \bar{L}_\max$. Then, the optimal block-length $L^*$ can be computed by rounding $L$ down to $\lfloor L \rfloor$.

Thirdly, the average packet error probability can be approximated as [15]

$$P_{be} \approx \kappa \left(\frac{\gamma_b}{\theta} + \frac{1}{\theta} + \frac{\gamma_b \theta^2}{2\pi} - 1\right),$$

where $\theta = \exp\left(\frac{D}{L}\right) - 1$ and $\kappa = -\frac{L}{2\pi\exp\left(\frac{D}{L}\right) - 1}$. We consider the case with $D \ll L$, since the achievable covert rate is normally small. Then, we have $\theta \approx \frac{D}{L}$ and $\kappa \approx -\frac{L}{2\pi D}$.

In this work, we only focus on the case with $0 < P_{be} < 1$. By substituting (14) and (15) into (13), the optimization problem can be rewritten as

$$\min_{L} f(L) = \frac{2L \rho_1 W(1 - \delta)}{\sigma_w^2 \sqrt{\frac{1}{\pi D} + 1 - \frac{D}{\pi}} + \frac{L}{2W}}$$

subject to:

1) $c_1'$, $c_2'$,

2) $c_3 : \frac{1}{\gamma_b} (D - \sqrt{\pi D}) \leq L \leq \frac{1}{\gamma_b} (D + \sqrt{\pi D}),$  \hspace{1cm} (16)

where $c_3$ is based on $0 < P_{be} < 1$. With regard to $f(L)$, we define $f'(L)$ as its first-order derivative with respect to variable $L$. Then, we can derive two roots for $f'(L) = 0$

$$L_1 = -\frac{2}{\gamma_b \rho_1 W(1 - \delta)} \left(\frac{D}{\pi} - 1\right)^{1/2} + \frac{1}{\gamma_b} (D - \sqrt{\pi D})$$

and

$$L_2 = \frac{2}{\gamma_b \rho_1 W(1 - \delta)} \left(\frac{D}{\pi} - 1\right)^{1/2} + \frac{1}{\gamma_b} (D - \sqrt{\pi D}).$$

Based on
the constraints $c_1', c_2'$ and $c_3$, we note that the block-length $L$ is limited to a region, i.e., $L_{low} \leq L \leq L_{up}$ where $L_{up} = \min(L_{max}, \frac{1}{\gamma_0}(D + \sqrt{\pi D})/\max(\rho_0^2, \rho_1^2\pi^2))$ and $L_{low} = \max(L_{min}, \frac{1}{\gamma_0}(D - \sqrt{\pi D})$). Accordingly, the optimal block-length $L^*$ for the given transmit power $P$ and prior transmission probability $\rho_1$ can be derived as

$$L^* = \begin{cases} \lfloor L_2 \rfloor, & L_2 < L_{up}, \\ \lceil L_{up} \rceil, & L_2 \geq L_{up}. \end{cases} \tag{17}$$

For the effective covert rate $R_c$, we have $R'_c(L) \geq 0$, which indicates that $R_c$ increases as the block-length $L$ increases when $L_{low} \leq L \leq L_{up}$. Therefore, the optimal block-length that maximizes the effective covert rate may not be the optimal one that minimizes the average CAoI.

B. Optimal $P$ to Minimize the Average CAoI

A larger transmit power $P$ can guarantee a lower packet error probability but also result in a higher covert communication outage probability. Similar to Section III-A, the optimization problem to minimize the average CAoI for the given $L$ and $\rho_1$ can be formulated as

$$\min_P f(P) = \frac{2L}{\rho_1 W(1-\delta)}\left(\frac{L P}{\sigma_b^2} \sqrt{\frac{1}{\pi D} + 1 - \sqrt{\frac{D}{\pi}}}\right) + \frac{L}{2W}$$

s.t. $c_1'$, $c_4 : \frac{\sigma_b^2}{L}(D - \sqrt{\pi D}) \leq P \leq \frac{\sigma_b^2}{L}(D + \sqrt{\pi D})$. \tag{18}

It is clear that the objective function $f(P)$ increases as the transmit power $P$ increases due to $f'(P) > 0$. Thus, the optimal transmit power $P^*$ is its maximum value that simultaneously satisfies the constraints $c_1'$ and $c_4$, which is derived as

$$P^* = \min\left(\frac{\sigma_b^2}{L}(D + \sqrt{\pi D}), \frac{\sigma_b^2}{L}(D - \sqrt{\pi D})\right)$$

We note that the optimal transmit power that maximizes the effective covert rate $R_c$ is the same as that minimizes the average CAoI since $R'_c(P) > 0$. In addition, the covert constraint $c_1'$ introduces a tradeoff between the transmit power and block-length in order to minimize the average CAoI, but the transmit power can be sufficient large to minimize the average AoI without the covertness constraint.

C. Optimal $\rho_1$ to Minimize the Average CAoI

A higher prior transmission probability $\rho_1$ indicates that Alice can timely transmit more data packets to Bob, but it also leads to that the covertness constraint becomes harder to be satisfied. Thus, there is an optimal prior transmission probability $\rho_1^*$ to minimize the average CAoI. Similar to Section III-A, the optimization problem for the given $P$ and $L$ is formulated as

$$\min_{\rho_1} f(\rho_1) = \frac{2L}{\rho_1 W(1-\delta)}\left(\frac{L P}{\sigma_b^2} \sqrt{\frac{1}{\pi D} + 1 - \sqrt{\frac{D}{\pi}}}\right) + \frac{L}{2W}$$

s.t. $c_1'$, $c_5 : 0 < \rho_1 < 1$. \tag{20}

It is obvious that $f'(\rho_1) < 0$, which indicates that the average CAoI decreases as the prior transmission probability $\rho_1$ increases. Hereby, we have $\rho_1 \geq \rho_0$ and the optimal prior transmission probability is its maximum value that simultaneously satisfies the constraints $c_1'$ and $c_5'$, where $c_5'$ is given as $0.5 \leq \rho_1 < 1$. Accordingly, the optimal prior transmission probability $\rho_1^*$ can be derived as

$$\rho_1^* = \delta \sqrt{2/D(\rho_0 || \rho_1^* || P_1)}.$$ \tag{21}

We note that the optimal prior transmission probability that maximizes the effective covert rate is same as that minimizes the average CAoI. Furthermore, Alice can keep transmitting to minimize the average AoI without the covertness constraint for each time slot, while it needs to opportunistically send the data packet to minimize the average CAoI.

IV. SIMULATION RESULTS

In this section, numerical results are presented to evaluate the impact of communication covertness and timeliness on the optimal parameters in the considered system. Some parameters are set as follows: $W = 1$ MHz, $\sigma_b^2 = -114$ dBm, $D = 10$ nats, $L_{max} = 1000$ and $L_{min} = 100$.

Fig. 2(a) conforms the correctness of the results. It also shows that the optimal block-length is balanced between communication covertness and timeliness rather than the largest one. Fig. 2(b) plots the optimal block-length versus the received SNR. We can observe that the optimal block-length to maximize the effective covert rate is the largest one when the SNR is smaller than a threshold, since the average AoI and CAoI are limited by the timeliness. In addition, we have the same optimal block-length for maximizing the effective rate and minimizing the average CAoI, which is lower than that for minimizing the average AoI, when SNR is larger than a

![Fig. 2. (a) Average CAoI versus the block-length with $\rho_1 = 0.5$. (b) Optimal block-length versus the received SNR with $\rho_1 = 0.5$ and $\delta = 0.15$.](image1)

![Fig. 3. (a) Average CAoI versus the received SNR with $L = 200$ and $\delta = 0.15$. (b) Probabilities versus the received SNR with $L = 200$ and $\delta = 0.15$.](image2)
threshold. This is due to the fact that the optimal block-length to maximize the effective covert rate and average CAoI are restricted by the covertness constraint. In Fig. 3(a), we observe that average CAoI achieved by the optimal \( \rho^*_1 \) is significantly lower than that achieved by \( \rho_1 = 0.5 \), which indicates the importance of optimizing \( \rho_1 \) to minimize the average CAoI. In Fig. 3(b), we observe that a higher received SNR results in a lower packet error probability, thus leads to a smaller average CAoI. We also observe that the optimal prior transmission probability decreases but is larger than 0.5 as the received SNR increases, which indicates Alice can transmit the data packet with a higher prior transmission probability to minimize the average CAoI when the transmit power is small.

\[
\text{APPENDIX}
\]

As shown in Fig. 1, assuming that there are \( K \) time slots between the \((i - 1)\)-th and the \(i\)-th valid data packets, the first-order moment of \( A_{i-1} \) is

\[
\mathbb{E}(A_{i}) = \sum_{k=1}^{\infty} kT \mathbb{P}(K = k) = \rho_1 (1 - \mathbb{P}_{\text{cop}})(1 - \mathbb{P}_{\text{be}}) T \sum_{k=1}^{\infty} k(1 - \rho_1 (1 - \mathbb{P}_{\text{cop}})(1 - \mathbb{P}_{\text{be}}))^{k-1} = \frac{L}{\rho_1 W (1 - \mathbb{P}_{\text{cop}})(1 - \mathbb{P}_{\text{be}})}. \tag{22}
\]

Then, the average arrival rate of valid data packet is

\[
\bar{\lambda}_b = \frac{1}{\mathbb{E}(A_{i})} = \frac{\rho_1}{L} W (1 - \mathbb{P}_{\text{cop}})(1 - \mathbb{P}_{\text{be}}). \tag{23}
\]

On the other hand, the first-order moment of interval time \( \Delta T \) for the \((j - 1)\)-th and the \(j\)-th data packets is given by

\[
\mathbb{E}(\Delta T) = \rho_1 \frac{L}{W} \sum_{k=1}^{\infty} (\rho_0)^{k-1} = \rho_1 \frac{L}{W} \sum_{k=0}^{\infty} (\rho_0)^{k} = \frac{L}{\rho_1 W}. \tag{24}
\]

Then, the second-order moment of interval time \( \Delta T \) is

\[
\mathbb{E}[(\Delta T)^2] = \frac{\rho_1 L^2}{W^2} \sum_{k=1}^{\infty} k^2(\rho_0)^{k-1} = \frac{L^2}{\rho_1 W^2}(\frac{\rho}{\rho_1} - 1). \tag{25}
\]

Based on the law of total probability, the second-order moment of \( A_j \) is derived in (26) on the top of the page, where \( \Delta \mathbf{T} \) is due to that \((\Delta T)_m \) and \((\Delta T)_l \) are independent. As shown in Fig. 1, \( S_i \) denotes the rectangular trapezoids area. Thus, we have \( \mathbb{E}(S_i) = 0.5 E(A^2_j) + T E(A_i) \), which can be further derived as

\[
\mathbb{E}(S_i) = \frac{2L^2 + \rho_1 L^2 (1 - \mathbb{P}_{\text{cop}})(1 - \mathbb{P}_{\text{be}})}{2(\rho_1 W (1 - \mathbb{P}_{\text{cop}})(1 - \mathbb{P}_{\text{be}}))}. \tag{27}
\]

Finally, by substituting (23) and (27) into (11), the closed-form expression for the average CAoI is obtained in (12).

\[
\text{REFERENCES}
\]


