Toward Goal-Oriented Semantic Communications: AoII Analysis of Coded Status Update System Under FBL Regime

Siqi Meng, Shaohua Wu, *Member, IEEE*, Aimin Li, *Student Member, IEEE*, and Qinyu Zhang, *Senior Member, IEEE*

***Abstract*—In the past decade, the emergence of beyond fifth generation (B5G) wireless networks has necessitated the timely updating of system states in Internet of Things (IoT) and cyber-physical systems, where Age of Information (AoI) has been a well-concentrated metric. However, the content-agnostic nature of AoI reflects its limitation of characterizing the significance of status update messages, which induces various variants for AoI including Age of Incorrect Information (AoII). AoII is a goal-oriented significance (etymological meaning of "semantics") metric that could overcome such shortcomings, and thus analyzing AoII performance can be a potential approach of realizing semantic communications. Nevertheless, AoII analysis of practical coded status update system under finite blocklength (FBL) regime is still in its nascent stages. To the best of our knowledge, our study represents the first analysis of AoII for FBL regime. We explicitly obtain the average AoII expressions for different transmission schemes including Automatic Repeat reQuest (ARQ), Hybrid ARQ (HARQ), and non-ARQ transmission schemes. Moreover, we theoretically prove that non-ARQ scheme outperforms ARQ schemes in terms of AoII, and numerically compare AoII performance between non-ARQ and HARQ schemes by formulating and solving the AoII-optimal block assignment problem. Extensive simulation results show the superiority of AoII-optimal transmission schemes.**

***Index Terms*—Age of Incorrect Information (AoII), finite blocklength (FBL) regime, goal-oriented communication, status update system.**

I. Introduction

OVER the past decade, with the unprecedented development of beyond fifth generation (B5G) wireless networks, timely updating the fresh system states has emerged as a consensus from both industry and academia in Internet of Things (IoT) and cyber-physical systems. To capture and reflect the freshness of data at the receiver, the age of information (AoI) metric is introduced to measure the timeliness of status update messages received from the remote transmitter [1]. Specifically, if we denote $U(t)$ as the timestamp that the latest acknowledged status update was generated, at time t, the AoI is defined as $\Delta_{\text{AoI}}(t) = t - U(t)$ which is the time duration of the freshest update from its generation to successful reception.

According to the definition of AoI, the original age metric increases linearly in time when the new status update is absent or the latest status update cannot be decoded successfully. Only if a status message is successfully received does the AoI drop to the age of the newly received message. Therefore, AoI can describe the freshness performance of *time-oriented* applications, including real-time status monitoring and emergency rescue. Nevertheless, the characteristics of continuously increasing penalty of AoI implies a limitation in describing the effectiveness yielded by status updates, for the AoI metric does not take into consideration various factors that affect effectiveness of status update system besides only timeliness, such as the change of source, the wrong estimation of the status updates, and different significance of status updates at different time. In other words, AoI adopts the penalty function where the costs of all status updates are considered as the same. AoI may not be an ideal indicator of performance for *goal-oriented* applications with wider scope where the ultimate effectiveness, rather than just timeliness, imposes substantial impact on terminals at the receiver. To surmount this inherent challenge, a series of variants for AoI have been proposed, which share a common feature that the cost or penalty on status staleness is customized so that these variants for AoI can better capture the ultimate effectiveness for goal-oriented application scenarios. Such variants for AoI include but are not limited to Query Age of Information (QAoI) [2], [3], Value of Information (VoI) [4], binary freshness metric [5], Version Age of Information (VAoI) [6], [7], and Age of Synchronization (AoS) [8].

Similar to VAoI and AoS, the Age of Incorrect Information (AoII), which is another new variant for AoI, is recently proposed in [9]. Specifically, denote the latest timestamp at which the estimated status $\hat{X}(t)$ at the receiver is the same as (i.e., synchronized with) real status $X(t)$ as $W(t)$, then the AoII

Manuscript received 26 February 2023; revised 16 August 2023 and 20 November 2023; accepted 15 December 2023. Date of publication 19 December 2023; date of current version 27 December 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB1806403; in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022B1515120002; and in part by the National Natural Science Foundation of China under Grant 62027802 and Grant 61871147. *(Corresponding author: Shaohua Wu.)*

Siqi Meng and Aimin Li are with the Department of Electronics and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China (e-mail: mengsiqi@stu.hit.edu.cn; liaimin@stu.hit.edu.cn).

Shaohua Wu and Qinyu Zhang are with the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China, and also with the Department of Broadband Communication, Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: hitwush@hit.edu.cn; zqy@hit.edu.cn).

Digital Object Identifier 10.1109/JSAIT.2023.3344586

is defined as $\Delta_{\mathrm{AoII}}(t) = f(t - W(t))$, which is a function of the time duration since the latest time stamp that the transceiver statuses are synchronized. Ideally, if the status updates are transmitted over an error-free channel, then the AoII is in fact equivalent as AoS. Moreover, if the status updates are different from each other, then each update can be treated as a version of the source status, and AoII is equivalent to VAoI.

AoII, as a goal-oriented metric, can capture the significance of different status updates, and AoII analysis can thus be a potential approach of realizing semantic (which can be interpreted as "significance" etymologically) communication. Compared to AoI and error indicator function which is defined as $\Delta_{\mathrm{error}}(t) = \mathbb{1}(\hat{X}(t) \neq X(t))$, AoII can reflect both status staleness and transceiver status asynchronization due to error-prone channels. Specifically, AoII imposes a zero penalty for synchronized timestamps and a penalty that linearly (or non-linearly) increases in asynchronization duration for asynchronized timestamps. However, AoI only captures the staleness of messages without considering the content of the messages (i.e., whether the transmitting messages really reflect the current source status), and error indicator function only describes whether the current timestamp is synchronized without considering the duration of (a)synchronized transceiver statuses. AoII reflect both the message content (synchronized status or not) and the (a)synchronization status duration, which can better capture the significance of messages as compared to AoI and error indicator function, since the messages that can make the transceiver statuses synchronized as rapid as possible are more significant.

In fact, the goal of status updating is to let the receiver acquire fresh and real-time statuses, and minimizing the AoII of the system just satisfies the demand of realizing this goal, and thus AoII can serve as a goal-oriented significance metric. As a major approach of semantic communication, goal-oriented communication aims at evaluating and refining the communication system such that the goal (or task) can be better achieved at the receiver. Hence, analyzing the AoII performance, which is a goal-oriented semantic metric to capture the significance of status update messages for status update system, is of vital significance for designing upcoming goal-oriented and further semantic communication systems [10], [11].

Studies are conducted in the aspects of AoII in the literature. For instance, in [12], the AoII of a remote estimation with the simplest binary Markov source is studied, and the optimal sampling policy is solved. Moreover, the AoII optimization under power constraint is discussed in [13]. In [14], the connection between AoII and semantic communications is established, and AoII is employed as a metric of semantic-empowered communication. The superiority of AoII-optimal policies over AoI-optimal and MSE-optimal policies is demonstrated. In [15], a more practical status update system model with random delays is introduced and the AoII performance of the system is analyzed. In the recent literature of [16], the average AoII expression of a system with two-state Markov source is derived, and the relationship between average AoII and average mean absolute error is theoretically analyzed. Another recent work [17] introduces limited retransmission with resource constraint to enhance the AoII performance by adopting hybrid automatic repeat request (HARQ) protocol. Similar Markov-based source model is also adopted in previous AoI analysis [18], [19], [20], [21], [22], and the above works do not consider specific transmission processes, such as encoding and decoding. Nevertheless, in a more general application scenario, coded status update system with finite block-length (FBL) regime, where status updates are encoded as finite-length code words, transmitted over an unreliable channel, and decoded by the receiver, should be further considered in order to measure the AoII performance for real goal-oriented communication systems.

In a practical system, FBL regime (or called short-packet communication) is usually adopted to meet the requirement of ultra low latency in B5G communications. In short-packet transmission, the asymptotic Shannon capacity is barely achievable. Instead, FBL analysis can measure the rate-distortion function by bounding the rate of channel codes under certain error probability [23]. In the literature, the FBL regime has been studied from a variety of perspective, such as analyzing the capacity performance of random access channels [24], minimizing the delay of IoT networks [25], and enhancing the throughput of security communication [26]. Recently, because of short-packet characteristics for status update messages, the time-oriented AoI performance of coded status update system under FBL regime has raised wide attention and been extensively researched [27], [28], [29], [30], [31], [32], [33], and these works all draw the conclusion that refining transmission schemes under FBL regime can reduce the average AoI under certain conditions. However, the performance analysis on goal-oriented semantic metrics under FBL regime is still in its initial stage, and how transmission schemes affect the AoII performance under FBL regime is still to be studied. Therefore, it is imperative to analyze the AoII performance of different transmission schemes under FBL regime.

In this paper, we study one of the goal-oriented semantic metrics, namely AoII of the coded status updated system under FBL regime for the first time. Different from the previous studies, we consider a more practical system model in the physical layer, where the status update messages are firstly encoded as finite-length code words, transmitted over an unreliable channel, and then decoded by the receiver. Moreover, due to limited bandwidth of IoT devices, some delay elements such as encoding/decoding delay and propagation delay are no longer negligible as compared to code word transmission delay. Hence, in the system model, similar to our previous works [30], [31], we further consider non-trivial typical delay elements in IoT scenario from all the practical encoding, propagation, decoding and feedback processes other than transmission delay. Specifically, we analyze the AoII performance of the system under different FBL transmission schemes including simple forward transmission, Automatic Repeat reQuest (ARQ), and HARQ, and we call the simple forward transmission scheme as non-ARQ transmission scheme. As the most general case, we firstly derive the average AoII expression for HARQ transmission scheme under FBL regime, which is a function of all the delay elements and the successful

TABLE I
THE CONTRIBUTIONS OF OUR WORK COMPARED TO THE LITERATURE

Contributions	This work	[1], [22]	[18], [19]	[20], [21]	[2]–[8]	[9], [12]–[14]	[15]	[16], [17]	[23], [24]	[32], [33]	[25]–[29]	[30], [31]
Goal-oriented Semantic Metric	✓				✓	✓	✓	✓				
Theoretical Expressions on Performance Metric	✓	✓			✓			✓	✓	✓	✓	✓
Non-trivial Delay Elements Besides Transmission Delay	✓						✓			✓		✓
FBL Regime	✓								✓	✓	✓	✓
Re-transmission Scheme	✓			✓				✓			✓	✓
Analytical Comparison of Different Regimes	✓											

decoding probabilities in every HARQ transmission round. Considering non-ARQ and ARQ as special cases of HARQ, we then mathematically express the average AoII for the other two schemes.

Furthermore, we compare the AoII performance among different FBL transmission schemes including non-ARQ, ARQ and HARQ both theoretically and numerically. Firstly, we provide a theoretical proof on the AoII performance comparison between non-ARQ and ARQ schemes, which implies that non-ARQ transmission scheme achieves lower average AoII than ARQ transmission schemes. Secondly, we formulate and solve a problem of AoII-optimal block assignment for FBL regime in order to compare AoII performance between non-ARQ and HARQ schemes. Since minimizing the message generation cost (e.g., encoding delay), minimizing the transmission cost (e.g., transmission and propagation delays), and maximizing the successful decoding probability contradict, the problem of AoII minimization is a non-trivial problem. The solutions show that AoII-optimal block assignments are equivalent to non-ARQ schemes at worse channel conditions, i.e., low signal-to-noise ratios (SNRs), and HARQ schemes at better channel conditions, i.e., high SNRs, which strikes the balance among message generation cost, transmission cost, and successful decoding probability. Also, compared to AoI-optimal transmission schemes, the number of code symbols in each round of AoII-optimal schemes is larger, since minimizing the goal-oriented AoII of the system requires for higher successful decoding probability of the very message representing the current status. Extensive numerical results are given to show the superiority of AoII-optimal transmission scheme as compared to AoI-optimal transmission scheme along with other baseline transmission schemes in terms of AoII.

Contributions: The main contributions of this work are listed and compared to the literature in Table I. Specifically, we summarize the key contributions of this paper as below:

- We analyze the AoII performance of coded status update system under FBL regime for the first time. In this system, all the delay elements including encoding delay, transmission delay, propagation delay, decoding delay and feedback delay are considered, and the AoII evolution processes of the different FBL transmission schemes are analyzed.
- We derive the average AoII expression of the coded status update system under HARQ transmission scheme, which is related to the delay elements and the successful decoding probability. As special cases for HARQ, the average AoII expressions for non-ARQ and ARQ transmission scheme can also be derived given the code length and successful decoding probability of the original message.
- We theoretically prove that non-ARQ schemes achieve lower average AoII than ARQ schemes. Moreover, in order to compare the AoII performance between non-ARQ and HARQ schemes, we formulate and solve the problem of AoII-optimal block assignment for FBL regime. Extensive simulation results are given to show the superiority of AoII-optimal transmission scheme over various baseline schemes in terms of average AoII.

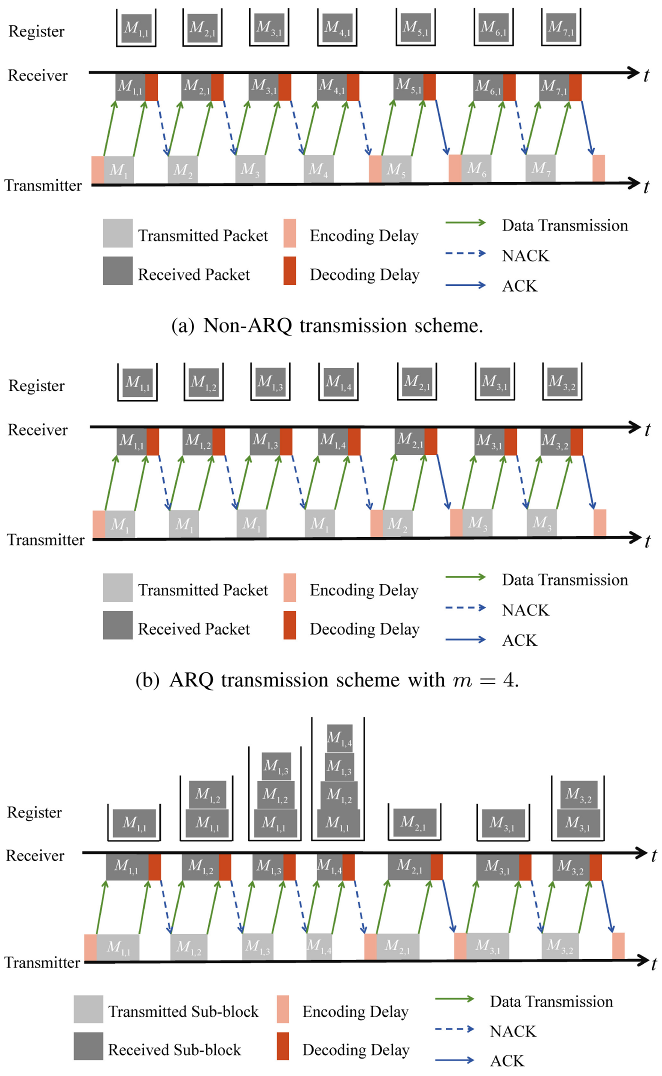
Organization: The remainder of this paper is organized as follows: Section II presents the system model of the HARQ- (or non-ARQ-) based coded status update system, and introduces the AoII metric. In section III, the average AoII expressions of coded status update system under various FBL transmission schemes are analyzed. Section IV focuses on the theoretical results and block assignment optimization problem in order to compare the AoII performance among different transmission schemes. Numerical results are demonstrated in Section V, followed by the conclusion in Section VI.

II. SYSTEM MODEL

A. Preliminaries

Before we introduce the system model including source model and transmission model, we firstly elaborate on some transmission protocols (also called transmission schemes in this paper) based on the FBL regime, including non-ARQ (also known as forward-error-correction-only), ARQ, and HARQ. The comparison of these transmission schemes is illustrated in Fig. 1.

The most simple transmission scheme is non-ARQ shown in Fig. 1(a), where the packets (also called messages in this paper) are generated, coded, and transmitted sequentially no matter whether the packets are decoded successfully or not. Different from non-ARQ, ARQ scheme shown in Fig. 1(b) introduces retransmission protocol in order to enhance the probability of successful decoding of the current packet. Specifically, after a failed decoding of a certain packet, a negative acknowledgment (NACK) feedback will be sent back to the transmitter. The transmitter will send the same



(a) Non-ARQ transmission scheme.

(b) ARQ transmission scheme with $m = 4$.

(c) HARQ transmission scheme with $m = 4$.

Fig. 1. Illustration and comparison of different transmission schemes. Here m represents the maximum number of total transmission rounds for one message.

packet again after receiving NACK feedback, and the receiver will continuously decode the newly received packet. This retransmission process will repeat until a packet is decoded successfully, and after this successful decoding, an acknowledgment (ACK) feedback will be sent back to the transmitter. If the transmitter receives ACK feedback, a new packet will be generated, coded, and transmitted.

The non-ARQ and ARQ transmission schemes both consider each packet as independent, i.e., the decoder at the receiver only utilizes the information of the latest received packet, which implies that the probability of successful decoding is constant if the channel condition does not change. Therefore, in ARQ scheme, the information of the previously received packet does not contribute to current packet decoding. To utilize the previous information, HARQ scheme shown in Fig. 1(c) is adopted more widely as a refinement of ARQ scheme. Specifically, as opposed to sending the same packet, the transmitter will send a new packet containing incremental information after receiving NACK feedback, and the receiver combines the newly received packet with all the previous packets stored in the register as an integrated packet for decoding. After a successful decoding, the receiver deletes the previous packets in the buffer for the arrival of new packets in the future. By such combination of previous information, the successful decoding probability will be significantly enhanced as the number of retransmissions increases.

Remark 1: The HARQ transmission scheme considered in this paper is a specific instance of the general HARQ scheme known as HARQ with incremental redundancy (HARQ-IR). In fact, there is also an implementation of HARQ called HARQ with chase combining (HARQ-CC), where the retransmission packet is also the original one and the receiver combines the received packets using maximum ratio combination (MRC) method. Since the packet length is the same in each retransmission and equal to the original packet length, HARQ-CC can be seen as a special case for HARQ-IR.

However, retransmission schemes, including ARQ and HARQ, will naturally introduce communication overhead, such as retransmission delay, extra decoding delay, and corresponding energy costs.[1] Considering that message generation will also cause communication overhead such as generation delay and energy cost, the problem of deciding whether or not to retransmit the message is a non-trivial problem. Therefore, analyzing and comparing the AoII performance of different transmission scheme with/without retransmission is of vital importance, which is a principal topic of this paper.

B. The Model of Coded Status Update System Under FBL Regime

1) Source Model: We assume that the source status information $X(t)$ is discrete and will change randomly as a discrete Markov chain, in which the status transition only depends on the latest status. We assume that the Markov source $X(t)$ is symmetric, that is, the transition probability matrix $\mathbf{P}$ is a symmetric $N \times N$ matrix with diagonal elements α and other elements μ, where N is the number of statuses, α is the probability that the status remains as the previous state, and μ is the probability of transition to one of the other statuses. Because of the Markov source, we have

$$(N-1)\mu + \alpha = 1. \tag{1}$$

Generally, a state in the real world has infinite number of statuses since it is usually continuous. However, in status update system under FBL regime, physical states should be firstly sampled and secondly quantified in order to generate bit sequences for coding. After sampling and quantification, the source states are discrete and possess finite number of statuses. Therefore, as a practical case, we assume that the number of statuses is finite, that is, source statuses are described as k-bit sequences which can represent 2^k number of different statuses. In order to precisely capture the subtle variation of source statuses, in this paper we assume that k is large enough

[1] As an initial work of AoII analysis under FBL regime, in this paper we only consider one type of communication overhead, namely the delay elements, and the energy analysis is beyond the scope of this paper. Delay-energy trade off in the AoII analysis under FBL regime will be an interesting topic in the future work.

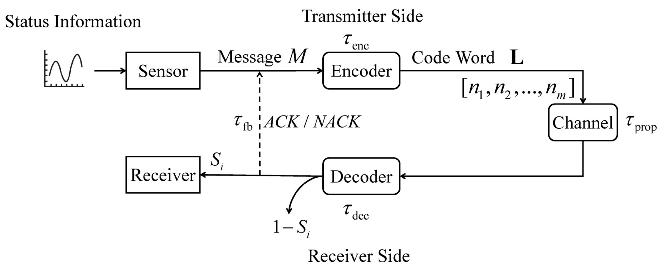


Fig. 2. The transmission model of coded status update system.

(i.e., $k \geq 100$).[2] In this case, the probability that a newly generated status is the same as one of the previously generated statuses is only 2^{-k}, which is negligible. In other words, the status will never (with extremely low probability) return to any one of the previous statuses in the future. Therefore, we have $N = 2^k$, $\alpha > 0$, and $\mu \to 0$. Also, we assume that the status information changes very slowly, i.e., $X(t)$ will be constant during the transmission process of the message involving information of $X(t)$ (or the status will not change after the generation of the same message), and will change once only after the successful decoding (or discard) of the current message. In this case, the generated status update messages will be different from all the previous messages (e.g., when the time stamp of the status is contained by messages, or when the statuses are represented by large enough number of bits).

These assumptions on the dynamics of source status are feasible for two principal reasons. Firstly, previous studies on AoII analysis [9], [14], [15], [16] also assume a Markov source model, where the status change occurs once within a time slot. Since the duration of a time slot is significantly longer than packet (i.e., message) transmission delay, the source can be assumed to be invariant during packet transmission. Thus, we assume that the source status changes slowly similarly. Secondly, for status update system deployed on terrestrial wireless networks, these assumptions also establish because of sufficiently large bandwidth. For instance, IEEE 802.11 suggests that the transmission rate of wireless communication be two megahertz [34], which implies that the bit duration is only 5×10^{-7} s. Assuming that the encoded status message has 200 bits, then the packet transmission delay is only 10^{-4} s, which is negligibly short as compared to status variation time.

2) Transmission Model: We consider the status update system as a transmitter-receiver pair shown in Fig. 2. In this model, the transmitter sends the status update messages that represent the status $X(t)$ over an unreliable channel under a certain FBL regime denoted by block assignment vector $\mathbf{N} = [n_1, n_2, \ldots, n_m]$ with $m \geq 1$, which includes non-ARQ, ARQ, or HARQ transmission scheme. The receiver recovers the messages to conduct estimation on the source status $X(t)$ as $\hat{X}(t)$ and gives feedback to the transmitter.

We introduce the transmission model taking HARQ transmission scheme as an example. Specifically, at the transmitter side, the sensor captures the varying statuses from the environment, with each represented by a k-bit status update message M, where k is large enough such that the generated messages are different from each other. Then, the encoder encodes M as n_m-bit code word $\mathbf{L}$, and then divides $\mathbf{L}$ into m child code words $l_i, i = 1, 2, \ldots, m$, with l_i containing $n_{i+1} - n_i$ bits. In the ith transmission round, the transmitter sends the child code word l_i to the receiver. At the receiver end, the decoder undertakes the decoding of the message based on the cumulatively received child code words $l_1, l_2, \ldots, l_i$. These code words have a total code length of n_i and are associated with a successful decoding probability of $S_i \in [0, 1]$. If the decoding process fails, an NACK will be sent to the transmitter. Upon receiving the NACK feedback, the transmitter will then proceed to send the next child code word l_{i+1} in the $(i+1)$th transmission round. This process will continue until the decoding succeeds, or all the m child code words have been transmitted. In these cases a new status update will be captured and repeat the above steps. In the non-ARQ transmission scheme, the transmission of data occurs only once without any retransmission process. Consequently, a new status update is always captured after the transmitter receives feedback. In contrast, in the ARQ transmission scheme, the transmitter continuously repeats the same packet, and the receiver decodes based on only the latest received packet, upon receiving an ACK from the receiver.

In this paper, we consider the generate-at-will and zero-wait sampling model. In this model, a successfully decoded status update triggers the immediate generation of a new update upon receiving an ACK feedback. Also, if the decoding process still fails after m rounds of transmission, the transmitter will instantly generate a new status update after receiving the NACK feedback (for non-ARQ, the transmitter always instantly generates new status updates). In this paper, we employ the additive white Gaussian noise (AWGN) channel, where the code symbols will experience a random additive Gaussian noise, that is

$$Y' = \sqrt{\gamma} Y + N, \tag{2}$$

where Y is the unit-power code symbol, N is the random Gaussian noise with zero mean and unit variance, and Y' is the channel output. In this channel model, the SNR is γ. Note that we adopt the AWGN channel model as an example for AoII analysis. In fact, as will be shown later, the derived mathematical expression can evaluate the AoII performance under any FBL regime and any channel conditions.

We assume that time is slotted and normalized to the slot duration, and one slot duration is assumed to be one channel use (CU), which is equal to the time of a code symbol duration.[3] In our system model, we consider the inclusion of various types of delays as communication overhead in IoT scenarios. In addition to the transmission delay given by $n_{i+1} - n_i$ CUs in the ith round, we also account for the

[2]In the following, we also use FBL results derived in [23] to verify our numerical results based on this assumption.

[3]Here we consider symbol-level AoII analysis. Though this is the first work that considers symbol-level AoII, some recent works on AoI analysis which split time into slots to study symbol-level AoI can also be taken as reference [27], [28], [29], [30], [31]. Also, it is worth noting that the time slot duration in this paper is much shorter than the other works on AoII analysis [9], [14], [15], [16], where the duration is usually equal to or larger than message transmission delay. However, in this paper, the transmission delay is equal to several time slots (or CUs).

following delays. Firstly, before the first transmission round, the encoding process (also including the sensing process) of the code word $\mathbf{L}$ will incur τ_{enc} CUs encoding delay, which is closely related to encoder structure and code length. Secondly, in each round of transmission, the forward propagation of the electromagnetic wave through the free space will take up τ_{prop} CUs propagation delay, which only depends on the distance between the transceivers. Thirdly, the decoding process will cause τ_{dec} CUs decoding delay, which is highly dependent on decoder structure and code length.[4] Fourthly, the feedback process of each round, similar to the forward propagation process, will also experience τ_{fb} CUs feedback delay, which is usually one CU of feedback bit plus propagation delay.

C. Age of Incorrect Information

AoII is defined as the function of the time elapsed since the latest time slot that the estimated status $\hat{X}(t)$ is equal to the current source status $X(t)$, which can be mathematically expressed as

$$\Delta_{\text{AoII}}(t) = f(t - W(t)) \cdot g\Big(\hat{X}(t), X(t)\Big), \tag{3}$$

where $f(t)$ is an increasing function of t which serves as the time penalty, $g(\hat{X}(t), X(t))$ is the function of the difference between $X(t)$ and $\hat{X}(t)$ which serves as the information penalty, and $W(t)$ is the last time that the information penalty $g(\hat{X}(t), X(t)) = 0$. Also, a time slot is called to be *synchronized*, if the current estimation $\hat{X}(t)$ is equal to the real source status $X(t)$; conversely, a time slot is called to be *asynchronized* otherwise. Thus, AoII describes the time duration since the latest synchronized time slot. If we consider that synchronized time slots are more significant, AoII can serve as a classic significance metric. Because the word "semantics" can be interpreted as significance etymologically, AoII analysis can be an approach to realize goal-oriented semantic communication. The reason is that analyzing AoII performance under FBL regime can guide the codeword design that can send messages with more significant information in order to make the system synchronized as fast as possible, which is the ultimate goal of the status update system.

Generally, we can choose $f(t - W(t))$ as the linear time penalty $f(t) = t - W(t)$, the nonlinear time penalty $f(t) = \sum_{i=0}^{m} a_i(t - W(t))^i, m \in \mathbb{Z}$, or the threshold time penalty $f(t) = \mathbb{1}(t - W(t) > c), c > 0$. Also, there are various choices for the $g(\hat{X}(t), X(t))$, such as error indicator function $g(\hat{X}(t), X(t)) = \mathbb{1}(\hat{X}(t) \neq X(t))$, the mean square error (MSE) function $g(\hat{X}(t), X(t)) = (\hat{X}(t) - X(t))^2$, and the threshold function $g(\hat{X}(t), X(t)) = \mathbb{1}(|\hat{X}(t) - X(t)| \geq c), c > 0$. To simplify the analysis, we adopt the linear time penalty function and the error indicator information penalty function to formulate the AoII metric in this paper, that is

$$\Delta_{\text{AoII}}(t) = (t - W(t)) \cdot \mathbb{1}\Big(\hat{X}(t) \neq X(t)\Big). \tag{4}$$

[4]The decoding delay in each round may be slightly different due to variable code length, but for simplicity, we assume that the decoding delay is constant in each round.

D. AoII Evolution for Non-ARQ and HARQ Transmission Scheme

In this subsection, we will further study the AoII evolution of coded status update system under FBL regime including non-ARQ scheme $\mathbf{N}_{\text{non}} = [n_1]$ and HARQ scheme $\mathbf{N}_{\text{HARQ}} = [n_1, n_2, \ldots, n_m]$.[5] Note that ARQ scheme is a special case for HARQ when both the child code length and successful decoding probability are the same for each round.

Generally, for each time slot t, the AoII evolution depends on whether the statuses between the transceivers are synchronized: if they are synchronized, the AoII at current time slot will become to zero; otherwise, the AoII will increase by one as compared to previous time slot. Mathematically, the general AoII evolution can be expressed as

$$\Delta_{\text{AoII}}(t) = \begin{cases} 0, X(t) = \hat{X}(t) \\ \Delta_{\text{AoII}}(t-1) + 1, X(t) \neq \hat{X}(t) \end{cases}. \tag{5}$$

For status update system under FBL regime, the AoII evolution can be specified because of a series of assumptions. According to our assumption on source status information, the status will not change after a message representing this status is generated, and will change only once during the time interval of feedback delay. Since source status is constant after message generation, we have (a) when the status changes, the statuses between transceivers will instantly be asynchronized, and vice versa, and (b) the time slots in the life cycle of a message (i.e., the time interval from its generation to acknowledgment/discard) must be in the same state, namely all synchronized or all asynchronized. Thus the AoII evolution is monotone during the life cycle of the message: if the latest message makes the transceiver statuses synchronized, then AoII in the life cycle of current message will be all zeros (no matter whether the current message is decoded successfully or not); otherwise, AoII increases linearly during the life cycle of current message.

Moreover, after a successful decoding occurs at the receiver, since the source status does not change and the receiver really acquires the status involved in the message (i.e., $X(t) = \hat{X}(t)$), the statuses between transceivers are instantly synchronized with AoII updated as zero. However, if a failed decoding occurs after maximum rounds of transmission, the estimated status $\hat{X}(t)$ will remain as the status represented by the latest successfully decoded message as usual. Mathematically, the estimated status at the receiver can be expressed as follows.

$$\hat{X}(t) = \begin{cases} X(t), t \text{ is the time slot of successful decoding} \\ \hat{X}(t-1), \text{otherwise} \end{cases}. \tag{6}$$

Therefore, there are two situations for the AoII evolution in the future: if the latest message makes the transceiver statuses synchronized, this case is equivalent to a successful decoding; otherwise, AoII will continuously increase no matter whether the status changes or not, since we assume that the statuses will never return.

[5]For non-ARQ and HARQ transmission schemes, the variable n_1 may not be the same. Here the indices for each element only represent the round. In other words, the n_1 in $\mathbf{N}_{\text{non}}$ may not be equal to that in $\mathbf{N}_{\text{HARQ}}$ here.

According to the above analysis, the AoII evolution of status update system under FBL regime can be simplified, because the evolution within the life cycle of a message is identical.

(1) When the latest $(p-1)$th message makes the transceiver statuses synchronized (i.e., AoII at the end of life cycle of the latest message is zero), AoII evolution depends on whether the status changes within the feedback delay. Denote $\phi_{p,k}$ as the indicator of status change at the kth time slot of the feedback delay before the pth message is generated, $k \in 1, 2, \ldots, \tau_{\mathrm{fb}}$. $\phi_{p,k} = 1$ implies that the status changes at the slot k, and that $\phi_{p,j} = 0$ for $j > k$ since status changes only once during the feedback delay. If $\phi_{p,k} = 0$ for $1 \leq k \leq \tau_{\mathrm{fb}}$, the status does not change during the feedback delay. We assume that $\phi_{p,k}$ is a Markov process with transition probability β when $\phi_{p,k-1} = 0$, and $\phi_{p,k} \equiv 0$ when $\phi_{p,k-1} = 1$. Denote the life cycle of message p as $\mathcal{T}_p$. Thus, when $\phi_{p,k} = 0$ for $1 \leq k \leq \tau_{\mathrm{fb}}$, the AoII is zero during the feedback delay before the ith message is generated and the life cycle of message p. Conversely, when $\phi_{p,k} = 1$ with $1 \leq k \leq \tau_{\mathrm{fb}}$, since $X(t) = \hat{X}(t)$ when $t \leq k$ and $X(t) \neq \hat{X}(t)$ when $t > k$, the AoII evolution during the feedback delay before the pth message is generated and the life cycle of message p can be mathematically expressed as[6]

$$\Delta_{\mathrm{AoII}}(t) = \begin{cases} 0, 1 \leq t \leq k \\ \Delta_{\mathrm{AoII}}(t-1) + 1, k < t < \tau_{\mathrm{fb}} + \mathcal{T}_p \end{cases}. \quad (7)$$

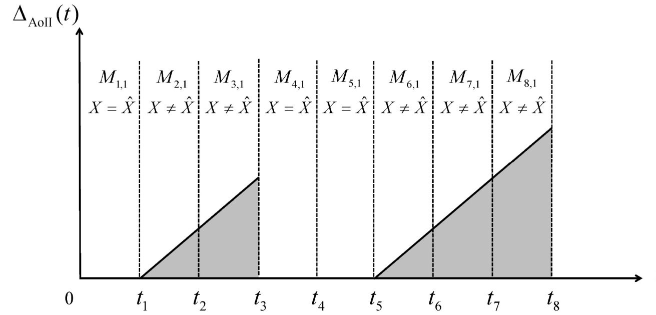
(2) When the latest $(p-1)$th message makes the transceiver statuses asynchronized (i.e., AoII at the end of life cycle of the latest message is larger than zero), AoII evolution becomes rather simple. Since newly generated pth message still makes the statuses between transceivers asynchronized, we have $X(t) \neq \hat{X}(t)$ for any t, and thus AoII will continuously increase during the feedback delay before the pth message is generated and the life cycle of message p, that is

$$\Delta_{\mathrm{AoII}}(t) = \Delta_{\mathrm{AoII}}(t-1) + 1, \text{for any } t. \quad (8)$$

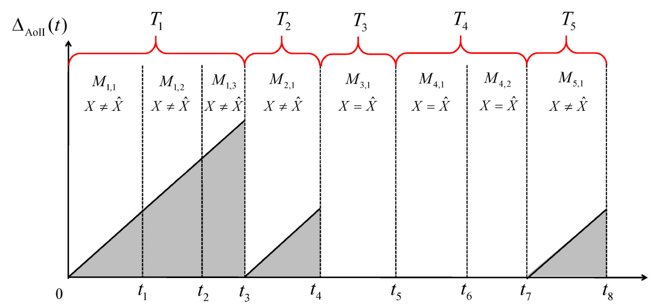
Based on equation (7) and (8), we can derive the AoII evolution under different transmission schemes. Specifically, an AoII evolution example under non-ARQ transmission scheme $\mathbf{N}_{\mathrm{non}} = [n_1]$ is shown in Fig. 3(a), and an example for HARQ transmission scheme $\mathbf{N}_{\mathrm{HARQ}} = [n_1, n_2, n_3]$ is shown in Fig. 3(b), respectively. In Fig. 3, M_j represents the jth status update and $M_{j,i}$ the ith round for the jth status update, and $T_p^{\mathrm{change}} = \{k||\phi_{p,k} = 1\}$ is a random variable representing the time interval from the successful decoding of $(p-1)$th message to the time slot in which source status changes. Note that if T_p^{change} does not exist, AoII will be continuously zero during the transmission of pth message. In this paper, we assume that the initial AoII of the system is $\Delta_{\mathrm{AoII}}(0) = 0$, that is, the initial state of the system is synchronized.

The AoII evolution under non-ARQ scheme is rather simple: it will only increase when the generated (or currently transmitting) message is different from the latest successfully received message, and be updated as zero when a message is successfully decoded, as demonstrated in Fig. 3(a), where M_2, M_3, M_6, M_7, and M_8 make the transceiver statuses asynchronized.

[6]Here, we assume that $t = 1$ represents the first time slot of feedback delay.



(a) Non-ARQ transmission scheme.



(b) HARQ transmission scheme with $m = 3$.

Fig. 3. The AoII evolution under different transmission schemes. Note that T_p^{change} is not illustrated in the two subfigures, since here feedback delay elements are assumed to be small enough that can be neglected as compared to the transmission delay. Though, T_p^{change} is a principal factor that affect the average AoII performance, as will be elaborated in the next section.

AoII evolution under HARQ transmission scheme is more complex due to retransmission. Specifically, if a message fails in decoding, the transmitter will retransmit some incremental information for the next decoding until the maximum mth round. Only if a message is decoded successfully or a message fails in decoding after m rounds does the source status start changing. It is illustrated in Fig. 3(b) that the message M_1 is decoded successfully within three rounds with AoII updated as zero at t_3. Then, status changes after T_2^{change} time slots with AoII increasing from $t_3 + T_2^{\mathrm{change}}$ to t_4. After that, M_3 gets successfully decoded in one round with AoII updated as zero at t_4. The status does not change until t_7, where M_5 represents a different status from the latest source status and the AoII starts increasing from $t_7 + T_5^{\mathrm{change}}$.

To show how T_p^{change} affects the AoII evolution, we present a discrete point graph that illustrates the AoII evolution under HARQ transmission scheme using a toy example in Fig. 4, where the block assignment vector $\mathbf{N} = [3, 5, 6]$ is taken into account, along with all relevant delay elements. Here $M_{j,i}$ similarly represents the ith round of the message M_j, and we assume that the AoII will be updated to zero at the time slot of the last channel use of successful decoding. Specifically, after a successful decoding at $t = 0$ (which results from the assumption of $\Delta_{\mathrm{AoII}}(0) = 0$), the status changes at the second slot of feedback delay, i.e., $T_1^{\mathrm{change}} = 2$. Therefore, AoII starts increasing from $t = 2$, and at $t = 22$ the message M_1 gets successfully decoded after three rounds with AoII updated as 0 instantly. After that, the status changes at the first time slot

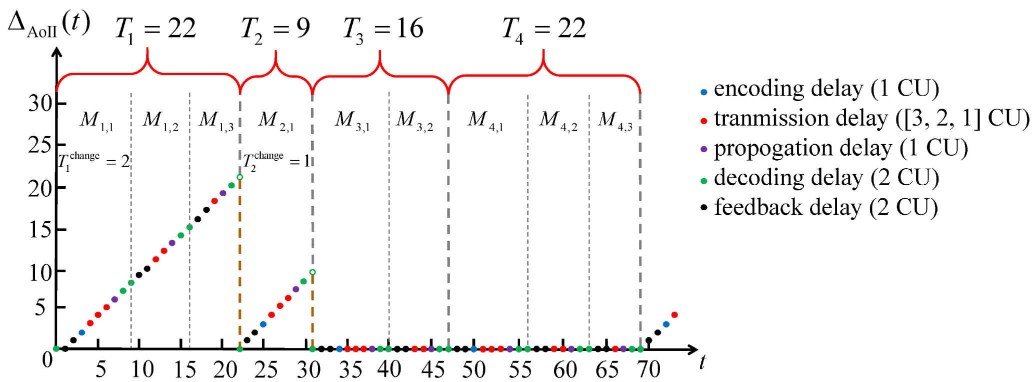


Fig. 4. The AoII evolution process of a status update system under the transmission scheme of $\mathbf{N} = [3, 5, 6]$, with encoding delay $\tau_{\text{enc}} = 1$, decoding delay $\tau_{\text{dec}} = 2$, propagation delay $\tau_{\text{prop}} = 1$, and feedback delay $\tau_{\text{fb}} = 1 + \tau_{\text{prop}} = 2$.

of feedback delay, i.e., $T_2^{\text{change}} = 1$, and thus AoII starts increasing at $t = 23$. Similar to M_1 and M_2, the AoII evolution for the rest messages can also be analyzed.

III. Analysis of the Average AoII of HARQ, Non-ARQ, and ARQ Transmission Schemes

In this section, we aim to derive the average AoII expression for the HARQ transmission scheme. Note that non-ARQ scheme can be considered as a special case of HARQ when the number of rounds is one, and thus the average AoII for non-ARQ transmission scheme can be derived.

Firstly, we give the average AoII expression under HARQ scheme in the following theorem.

Theorem 1: Given the HARQ transmission scheme $\mathbf{N} = [n_1, n_2, \ldots, n_m]$ and the successful decoding probability vector $[S_1, S_2, \ldots, S_m]$, where S_i represents the probability of successful decoding at round i. Let the encoding delay be τ_{enc}, propagation delay τ_{prop}, decoding delay τ_{dec}, and feedback delay τ_{fb}, and the transition probability of source status β. Then, if τ_{fb} is considerably large, the average AoII can be approximately estimated as

$$\bar{A} \approx \frac{E(C_{1,j}) + E(C_{2,j})}{E(T_j)}, \tag{9}$$

where

$$\begin{aligned} E(C_{1,j}) = {} & \frac{\beta}{1-\beta} \cdot \Bigg[\frac{1}{2}\Big(E\Big(T_j^2\Big) + E(T_j)\Big) \cdot \sum_{n=1}^{\tau_{\text{fb}}} (1-\beta)^n \\ & - \frac{1}{2}\big(2E(T_j) + 1\big) \cdot \sum_{n=1}^{\tau_{\text{fb}}} n(1-\beta)^n + \frac{1}{2}\sum_{n=1}^{\tau_{\text{fb}}} n^2 (1-\beta)^n \Bigg], \end{aligned} \tag{10}$$

$$\begin{aligned} E(C_{2,j}) = {} & \frac{\beta(1-S_m)}{(1-\beta)^{\tau_{\text{fb}}-1-\tau_m}} \cdot \Bigg[\frac{1}{2}\Big(E\Big(T_j^2\Big) + E(T_j)\Big) \cdot \sum_{n=1+\tau_m}^{\tau_{\text{fb}}+\tau_m} (1-\beta)^n \\ & - \frac{1}{2}\big(2E(T_j) + 1\big) \cdot \sum_{n=1+\tau_m}^{\tau_{\text{fb}}+\tau_m} n(1-\beta)^n + \frac{1}{2} \sum_{n=1+\tau_m}^{\tau_{\text{fb}}+\tau_m} n^2(1-\beta)^n \Bigg] \\ = {} & \frac{\beta(1-S_m)}{(1-\beta)^{\tau_{\text{fb}}-1}} \Bigg[\frac{1}{2}\Big(E\Big(T_j^2\Big) + (1-2\tau_m)E(T_j) + \tau_m^2 - \tau_m\Big) \cdot \sum_{n=1}^{\tau_{\text{fb}}} (1-\beta)^n \\ & - \frac{1}{2}\big(2E(T_j) + 1 - 2\tau_m\big) \cdot \sum_{n=1}^{\tau_{\text{fb}}} n(1-\beta)^n + \frac{1}{2}\sum_{n=1}^{\tau_{\text{fb}}} n^2(1-\beta)^n \Bigg], \end{aligned} \tag{11}$$

where $\tau_i = n_i + \tau_{\text{enc}} + i(\tau_{\text{prop}} + \tau_{\text{dec}} + \tau_{\text{fb}}), i = 1, 2, \ldots, m$ represents the equivalent number of CUs after the ith round, and the two moments related to T_j, which represents the time interval of two successful decoding processes, can be calculated as

$$\mathbb{E}\big(T_j\big) = \frac{\tau_m}{S_m} - \sum_{i=1}^{m-1} (\tau_{i+1} - \tau_i)\frac{S_i}{S_m}, \tag{12}$$

$$\begin{aligned} \mathbb{E}\Big(T_j^2\Big) = {} & \frac{\tau_m^2(2-S_m)}{S_m^2} - \sum_{i=1}^{m-1} \Bigg\{ \frac{S_i}{S_m} \\ & \cdot (\tau_{i+1} - \tau_i)\Bigg[\tau_{i+1} + \tau_i + 2\tau_m \frac{1-S_m}{S_m}\Bigg]\Bigg\}. \end{aligned} \tag{13}$$

Before we prove the above theorem, we firstly show the main idea of the calculation of average AoII. Firstly, the AoII at each time slot can be seen as an ergodic random process, and we can use the time average (i.e., long-term average AoII) to evaluate the statistic average of AoII. A typical choice of long-term average AoII is the average AoII during the time interval between two successful decoding processes, since this evaluation on average AoII is also ergodic. Secondly, to derive the average AoII during a certain time interval, we should acquire the AoII evolution (i.e., the AoII value at each time slot) by exploring the synchronized and asynchronized time slots. Thirdly, we sum up the AoII values at each time slot and divide by the average time interval to derive the time average (also statistic average) AoII.

Proof:

(1) *Choosing a time interval*: Considering that AoII must be updated as zero after a successful decoding, and does not decrease until the next successful decoding, we choose the time interval between two successful decoding processes to analyze the average AoII. Specifically, we denote this time interval as T_j, where j represents the number of messages that have been successfully decoded. In order to derive the mathematical expression about the random process T_j, we first give some notations.

We denote $\xi_{j,i}$ as the feedback indicator at the ith round of the jth message M_j under the HARQ transmission scheme $\mathbf{N} = [n_1, n_2, \ldots, n_m]$. Denote an ACK feedback as $\xi_{j,i} = 1$, and an NACK feedback as $\xi_{j,i} = 0$. If $\xi_{j,i} = 0$ and $i < m$, the transmitter will send the $(i+1)$th child code word for the ith message. If either $\xi_{j,i} = 1$ or $i = m$, then the transmitter will stop sending the current message and instantly sensing and generating another message M_{i+1}. Note that because of serial transmission of HARQ scheme, if one message is decoded at the ith round, then the feedback after the ith round is also ACK; if one message is wrongly decoded at ith round, then it must be wrongly decoded before the ith round. Therefore, we have $\{\xi_{j,i} = 1\} \subseteq \{\xi_{j,m} = 1\}$ for $i < m$, and $\{\xi_{j,i} = 0\} \subseteq \{\xi_{j,k} = 0\}$ for $k < i$. We denote Z_j as the number of the messages generated until the jth successful decoding, i.e., $Z_j = \inf\{k > Z_{j-1}:\xi_{k,m} = 1\}$, and let $Z_0 = 0$. Denote $R_j = Z_j - Z_{j-1} - 1$ as the number of incorrectly (i.e., failed) decoded messages between two correctly decoded messages. Let V_j be the round at which the message Z_j gets successfully decoded. For the probability distribution of R_j and V_j, we have the following lemma.

Lemma 1: The probability distribution of R_j and V_j are

$$\mathbb{P}\big(R_j = r\big) = S_m(1 - S_m)^r, \tag{14}$$

$$\mathbb{P}\big(V_j = v\big) = \frac{S_v - S_{v-1}}{S_m}. \tag{15}$$

Proof: Please refer to the Appendix. ■

Let $\tau_i = n_i + \tau_{\text{enc}} + i(\tau_{\text{prop}} + \tau_{\text{dec}} + \tau_{\text{fb}})$ be the equivalent number of CUs after the ith round, and then we can calculate the moments of the random variable R_j and τ_{V_j} as the following lemma.

Lemma 2: The first moment and the second moment of R_j is

$$\mathbb{E}\big(R_j\big) = \frac{1 - S_m}{S_m}, \tag{16}$$

$$\mathbb{E}\Big(R_j^2\Big) = \frac{(1 - S_m)^2 + 1 - S_m}{S_m^2}. \tag{17}$$

The pth moment of τ_{V_j} is

$$\mathbb{E}\Big(\tau_{V_j}^p\Big) = \tau_m^p - \sum_{i=1}^{m-1}\big(\tau_{i+1}^p - \tau_i^p\big)\frac{S_i}{S_m}. \tag{18}$$

Proof: According to $\mathbb{E}(R_j) = \sum_{i=0}^m i\mathbb{P}(R_j = i)$, $\mathbb{E}(R_j^2) = \sum_{i=0}^m i^2\mathbb{P}(R_j = i)$, and equation (14), we can easily derive equation (16) and (17), respectively.

According to $\mathbb{E}(\tau_{V_j}^p) = \sum_{i=0}^m \tau_{V_j}^p\mathbb{P}(V_j = i)$ and equation (15), the equation (18) can be derived. ■

Let t_j be the time slot of the successful decoding of the jth successfully decoded message. Under HARQ transmission scheme $\mathbf{N} = [n_1, n_2, \ldots, n_m]$, upon a message is correctly acknowledged or wrongly decoded after m rounds, a new message will instantly be generated and transmitted, so we can recursively express t_j by R_j and τ_{V_j} as follows:

$$t_j = t_{j-1} + R_j\tau_m + \tau_{V_j}. \tag{19}$$

We denote T_j as the time interval between the jth and $(j-1)$th successfully decoded messages, that is $T_j = t_j - t_{j-1}$. Based on equation (19) we have

$$T_j = R_j\tau_m + \tau_{V_j}. \tag{20}$$

It is easy to know that the random process T_j is ergodic, since R_j and V_j can be any value in their value set. For easy reference, we have marked this random variable in the AoII evolution illustrated in Fig. 3(b) and Fig. 4. Then we can calculate the first-order and second-order moments of T_j, which is the basis of the proof, as shown in the lemma below.

Lemma 3: The first moment and the second moment of T_j can be calculated be (12) and (13), respectively.

Proof: By calculating $\mathbb{E}(T_j) = \mathbb{E}(R_j\tau_m + \tau_{V_j}) = \tau_m\mathbb{E}(R_j) + \mathbb{E}(\tau_{V_j})$, and substituting (16) and (18) to the above equation, the equation (12) can be derived.

By calculating $\mathbb{E}(T_j^2) = \mathbb{E}(R_j\tau_m + \tau_{V_j})^2 = \tau_m^2\mathbb{E}(R_j^2) + \mathbb{E}(\tau_{V_j}^2) + 2\tau_m\mathbb{E}(R_j\tau_{V_j})$, and considering that R_j and τ_{V_j} are independent (since they represents different status messages captured independently), we have $\mathbb{E}(R_j\tau_{V_j}) = E(R_j)E(\tau_{V_j})$, and then by substituting (17) and (18) to the above equation, we can prove (13). ■

(2) *AoII evolution analysis during the time interval T_j:* According to the analysis in Section III, AoII evolution during T_j is approximately monotone: (a) if status changes within the first τ_{fb} time slots of T_j, AoII will start increasing at the time slot of change, and continuously increases until the last time slot of T_j, and (b) if status does not change, AoII will continuously be zero until the last time slot of T_j. In order to calculate average AoII of these two cases, we also introduce a random variable q_j which denotes the time slot at which the source status changes, that is

$$q_j = \begin{cases} k, \phi_{1,k} = 1 \\ \tau_m + k, \phi_{2,k} = 1 \\ 2\tau_m + k, \phi_{3,k} = 1 \\ \cdots \end{cases}, \tag{21}$$

where $\phi_{p,k} = 1$ represents that a status change occurs at the kth time slot of the feedback delay before the pth message is generated in the time interval T_j. Thus, q_j is a random variable ranging from one to infinity, which makes it difficult to derive the average AoII considering all the cases of q_j. Hopefully, given the status transition probability in each time slot β, we can calculate the probability distribution of q_j as follows:

$$P\big(q_j = q\big) = \begin{cases} \beta \cdot (1 - \beta)^{q-1}, \phi_{1,k} = 1 \\ \beta \cdot (1 - \beta)^{\tau_{\text{fb}} - \tau_m + q - 1}, \phi_{2,k} = 1 \\ \beta \cdot (1 - \beta)^{2(\tau_{\text{fb}} - \tau_m) + q - 1}, \phi_{3,k} = 1 \\ \cdots \end{cases}. \tag{22}$$

From equation (22), according to the assumption on the feedback delay, the probability of the event that status changes after several messages have been transmitted is negligibly small when the feedback delay is large. For instance, when $\beta = 0.1$ and $\tau_{\text{fb}} = 21$, the probability of $\phi_{3,k} = 1$ is approximately only one tenth that of $\phi_{2,k} = 1$, which can be usually considered as negligible. Moreover, the average AoII under the condition that $\phi_{3,k} = 1$ is significantly smaller than that when $\phi_{2,k} = 1$ because of longer zero-value duration, which can further illustrate that the average AoII in the cases where status remains during several message transmissions can be neglected. Therefore, we next estimate the average AoII by only considering the cases where $\phi_{1,k} = 1$ or $\phi_{2,k} = 1$, and the other cases are neglected due to small probability and negligible values.

According to equation (7), AoII will increase by one if the current time slot is asynchronized, and be zero otherwise. According to the source change model, the status will change only once during a feedback delay at q_j. Since the first time slot of T_j must be synchronized because of a successful decoding just before T_j, and the status changes at q_j, the time slots within time interval $[t_{j-1}, q_j]$ are all synchronized while the rest of time slots are all asynchronized. Therefore, for any time slot $\mathcal{T}$ in $[t_{j-1}, t_j - 1]$, the instantaneous AoII can be mathematically expressed as

$$\Delta_{\text{AoII}}(\mathcal{T}) = \begin{cases} 0, t_{j-1} + 1 \le \mathcal{T} < t_{j-1} + q_j \\ \mathcal{T} - q_j + 1, q_j \le \mathcal{T} \le t_j \end{cases}. \tag{23}$$

(3) *Calculation of the cumulative and average AoII:* Firstly, we derive the cumulative AoII during T_j for the case where status changes before the first message is generated, denoted by $C_{1,j}$. Since q_j is a random variable, $C_{1,j}$ is also a random variable with probability distribution as follows:

$$P\left(C_{1,j} = \sum_{n=k}^{\tau_{\text{fb}}} \left(T_j - n + 1\right)\left(T_j - n\right)\right) = \beta(1-\beta)^{k-1}. \tag{24}$$

Similarly, we derive the cumulative AoII during T_j for the case where the status changes after one message has been transmitted, denoted by $C_{2,j}$, and its probability distribution is expressed as

$$P\left(C_{2,j} = \sum_{n=k}^{\tau_{\text{fb}}} \left(T_j - n + 1 - \tau_m\right)\left(T_j - n - \tau_m\right)\right)$$
$$= (1 - S_m)\beta(1-\beta)^{k-1+\tau_{\text{fb}}}, \tag{25}$$

since the message transmitted before status change must be failed after m rounds of transmission.

Thus, the average AoII can be estimated as

$$\bar{A} \approx E\left(\frac{C_{1,j} + C_{2,j}}{T_j}\right) = \frac{E\left(C_{1,j}\right) + E\left(C_{2,j}\right)}{E\left(T_j\right)}, \tag{26}$$

where $E(C_{1,j})$, $E(C_{2,j})$, and $E(T_j)$ can be calculated by (10), (11), and (12), respectively. Note that we can calculate the sum terms in (10), (11) by the following three formulas:

$$\sum_{n=1}^{N} a^n = \frac{a\left(1 - a^N\right)}{1 - a}, \tag{27}$$

$$\sum_{n=1}^{N} n a^n = \frac{a\left(1 - a^N\right)}{(1 - a)^2} - \frac{N a^{N+1}}{1 - a}, \tag{28}$$

$$\sum_{n=1}^{N} n^2 a^n = \frac{a}{(1 - a)^2} + \frac{2a^2\left(1 - a^{N-1}\right)}{(1 - a)^3} - \frac{(2N - 1)a^{N+1}}{(1 - a)^2} - \frac{N^2 a^{N+1}}{1 - a}. \tag{29}$$

■

The equation (26) or (9) is a unified explicit expression for average AoII of coded status update system. For example, under non-ARQ transmission scheme $\mathbf{N} = [n_m]$ where $m = 1$ with successful decoding probability S_m, the expressions of $E(T_j)$ and $E(T_j^2)$ reduce as follows.

$$\mathbb{E}\left(T_j\right) = \frac{\tau_m}{S_m}, \tag{30}$$

$$\mathbb{E}\left(T_j^2\right) = \frac{\tau_m^2 (2 - S_m)}{S_m^2}. \tag{31}$$

By substituting (30) and (31) into (10) and (11), we can obtain the estimation on average AoII expression under non-ARQ transmission scheme.

As a unified explicit expression, (26) can evaluate the AoII performance of coded status update system for any FBL regimes and under any channel conditions, providing that the delay elements, the transmission scheme vector, and the successful decoding probabilities are known. For instance, in order to utilize this expression as a goal-oriented measure for practical IoT scenarios, we use the FBL results [23] to evaluate S_i of a code word with length n_i over the AWGN channel with SNR γ, which can be expressed as follows:

$$S_i \approx 1 - Q\left(\frac{C(\gamma) - k/n_i - \frac{1}{2} n_i \log_2(n_i)}{\sqrt{V(\gamma)/n_i}}\right), \tag{32}$$

where k is the message length, $C(\gamma)$ is the channel capacity with $C(\gamma) = \frac{1}{2}\log_2(1+\gamma)$, $V(\gamma)$ is the channel dispersion with $V(\gamma) = (1 - \frac{1}{(1+\gamma)^2})\log_2^2 e$. By substituting (32) into the average AoII expressions (26), we can evaluate the average AoII performance of different FBL regimes over the AWGN channel. For instance, in HARQ transmission scheme, we substitute the elements in block assignment vector $[n_1, n_2, \ldots, n_m]$ into (32) and obtain the successful decoding probability vector $[S_1, S_2, \ldots, S_m]$. Since we have $n_i < n_{i+1}$ for any $i = 1, 2, \ldots, m-1$ because of incremental transmission, it is easy to obtain $S_i < S_{i+1}$, that is, the successful decoding probability increases as the number of retransmission increases. Although S_i estimated by (32) in HARQ scheme is still some way from the real value due to the additional information in previous rounds $1, 2, \ldots, i-1$, we can still neglect this difference when the SNR is sufficiently large that S_i is close to one.[7] In ARQ transmission scheme $[n_1, 2n_1, \ldots, mn_1]$, since the successful decoding probability only depends on the original message length, the successful decoding probability vector can be obtained by $[S_1, S_1, \ldots, S_1]$, where S_1 can be calculated by substituting n_1 into (32).

In order to evaluate the AoII performance over fading channel, we have two principal methods: one is to acquire the probability distribution of fading parameters and derive the successful decoding probabilities under all the fading parameters; the other is to directly evaluate the average successful decoding probability over a certain fading channel, e.g., Rayleigh fading channel. The second method is usually sufficient for evaluating the average AoII. For instance, given the block assignment vector and successful decoding probability vector, which can be obtained by the block error rate of Spinal codes over Rayleigh fading channel [35], we can evaluate the average AoII of Spinal coded status update system under FBL regime.[8]

[7]Equation (32) gives only an estimation of S_i for an independent transmission of n_i symbols at the ith round. Since the decoder combines the previous information with current information at the ith round under HARQ transmission protocol, the n_i symbols at the ith round cannot be seen as independent transmission, and thus (32) cannot represent the real value of S_i.

[8]Although Spinal codes are rateless codes, we can also use Spinal codes as fixed-rate codes by restricting the maximum allowable code length. Thus, FBL results, including the block error rate, can be derived and adopted to evaluate AoII performance.

IV. AoII-Optimal Transmission Scheme: Non-ARQ Versus ARQ/HARQ

In this section, we compare the AoII performance among different transmission schemes both theoretically and numerically. Firstly, since simple ARQ scheme does not provide any coding gain in retransmission processes, it can be intuitively inferred that ARQ scheme will decrease AoII performance as compared to non-ARQ scheme, and theoretical proof will further validate this inference. Secondly, as has been mentioned in Section II, retransmission scheme optimization under the constraint of message generation delay and code length is a non-trivial problem. Thus, for more complex HARQ schemes, we numerically compare the AoII performance with non-ARQ scheme through formulation and solution of block assignment optimization problem.

A. Theoretical Proof of AoII Performance Comparison Between Non-ARQ and ARQ

For average AoII performance under non-ARQ scheme $[n]$ and ARQ scheme $[n, 2n, 3n, \dots,]$ with successful decoding probability S in each round, we have the following theorem, which shows that non-ARQ scheme has better AoII performance than ARQ schemes.

Theorem 2: We denote the average AoII of non-ARQ scheme $[n]$ with successful decoding probability vector $[S]$ as $\bar{A}_{\text{non-ARQ}}$, and denote the average AoII of ARQ scheme $[n, 2n, 3n, \dots,]$ with successful decoding probability vector $[S, S, S, \dots,]$ as $\bar{A}_{\text{ARQ}}$. Since $S < 1$ for any erroneous channel, we always have $\bar{A}_{\text{non-ARQ}} < \bar{A}_{\text{ARQ}}$.

Proof: We can directly do the difference of the two average AoII expression as follows:

$$\begin{aligned}&\bar{A}_{\text{non-ARQ}} - \bar{A}_{\text{ARQ}}\\ &= \frac{E\left(C_{1,j}^{\text{non-ARQ}}\right) + E\left(C_{2,j}^{\text{non-ARQ}}\right)}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{E\left(C_{1,j}^{\text{ARQ}}\right) + E\left(C_{2,j}^{\text{ARQ}}\right)}{E\left(T_j^{\text{ARQ}}\right)}\\ &= \left[\frac{E\left(C_{1,j}^{\text{non-ARQ}}\right)}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{E\left(C_{1,j}^{\text{ARQ}}\right)}{E\left(T_j^{\text{ARQ}}\right)}\right] + \left[\frac{E\left(C_{2,j}^{\text{non-ARQ}}\right)}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{E\left(C_{2,j}^{\text{ARQ}}\right)}{E\left(T_j^{\text{ARQ}}\right)}\right].\end{aligned} \tag{33}$$

Since $E(C_{1,j})$ and $E(C_{2,j})$ have similar forms, we only prove that $\frac{E(C_{1,j}^{\text{non-ARQ}})}{E(T_j^{\text{non-ARQ}})} - \frac{E(C_{1,j}^{\text{ARQ}})}{E(T_j^{\text{ARQ}})} < 0$ as an example, and the second term can be similarly proved to be negative. Thus, we can rewrite the first term or (33) as

$$\begin{aligned}&\frac{E\left(C_{1,j}^{\text{non-ARQ}}\right)}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{E\left(C_{1,j}^{\text{ARQ}}\right)}{E\left(T_j^{\text{ARQ}}\right)}\\ &= \gamma\left[\frac{E\left(\left(T_j^{\text{non-ARQ}}\right)^2\right)}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{E\left(\left(T_j^{\text{ARQ}}\right)^2\right)}{E\left(T_j^{\text{ARQ}}\right)}\right]\\ &\quad + \delta\left[\frac{1}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{1}{E\left(T_j^{\text{ARQ}}\right)}\right],\end{aligned} \tag{34}$$

where γ and δ are both constants with $\gamma = \frac{\beta}{2(1-\beta)} \cdot \sum_{n=1}^{\tau_{\text{fb}}} (1-\beta)^n$ and $\delta = -\frac{\beta}{2(1-\beta)} \sum_{n=1}^{\tau_{\text{fb}}} n(1-\beta)^n$. Since $\gamma > 0$ and $\delta < 0$, if we can prove that $\frac{E((T_j^{\text{non-ARQ}})^2)}{E(T_j^{\text{non-ARQ}})} - \frac{E((T_j^{\text{ARQ}})^2)}{E(T_j^{\text{ARQ}})} < 0$ and $\frac{1}{E(T_j^{\text{non-ARQ}})} - \frac{1}{E(T_j^{\text{ARQ}})} > 0$ respectively, the theorem can be proved.

Next, we prove by (a) considering a two-round ARQ scheme, (b) demonstrating that the above two inequalities establish, and (c) elaborating that ARQ schemes with infinite rounds can be sequentially divided into a series of two-round ARQ scheme.

Firstly, under non-ARQ scheme $[n]$ and two-round ARQ scheme $[n, 2n]$, according to (12) and (13), we have

$$E\left(T_j^{\text{non-ARQ}}\right) = \frac{\tau}{S}, E\left(\left(T_j^{\text{non-ARQ}}\right)^2\right) = \frac{\tau^2(2-S)}{S^2}, \tag{35}$$

$$E\left(T_j^{\text{ARQ}}\right) = \frac{2\tau}{S} - \tau, E\left(\left(T_j^{\text{ARQ}}\right)^2\right) = \frac{8\tau^2(1-S) + S^2\tau^2}{S^2}, \tag{36}$$

where $\tau = n + \tau_{\text{enc}} + \tau_{\text{prop}} + \tau_{\text{dec}} + \tau_{\text{fb}}$. Thus, we have

$$\begin{aligned}&\frac{E\left(\left(T_j^{\text{non-ARQ}}\right)^2\right)}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{E\left(\left(T_j^{\text{ARQ}}\right)^2\right)}{E\left(T_j^{\text{ARQ}}\right)}\\ &= \frac{(2-S)^2 - 8(1-S) - S^2}{S^2(2-S)} = \frac{4(S-1)}{S^2(2-S)} < 0,\end{aligned} \tag{37}$$

$$\frac{1}{E\left(T_j^{\text{non-ARQ}}\right)} - \frac{1}{E\left(T_j^{\text{ARQ}}\right)} = \frac{\tau(2-S-1)}{S} > 0. \tag{38}$$

Therefore, we can obtain that the left hand side of (34) is lower than zero, and thus the theorem establishes when the number of ARQ round is two.

Lastly, we qualitatively show that the theorem also establishes when the number of ARQ rounds is infinite (also including finite ARQ rounds). In fact, the ARQ scheme $[n, 2n, \dots, mn]$ can be separated into several two-round ARQ schemes, that is $[n, 2n]$, $[2n, 3n]$,..., and $[(m-1)n, mn]$. For each two-round ARQ scheme we have $\bar{A}_{[(i+1)n]} < \bar{A}_{[in,(i+1)n]}$ for any $i \in [1, m-1]$, where $\bar{A}_{\mathbf{N}}$ represents the average AoII of scheme $\mathbf{N}$. Then, it is easy to integrate those above inequalities and conclude that $\bar{A}_{[mn]} < \bar{A}_{[n,2n,\dots,mn]}$ because of serial transmission of message packets. In other words, each retransmitted packet will increase the average AoII, and thus non-ARQ scheme must achieves the minimum AoII as compared to ARQ schemes. ■

The phenomenon behind this theorem is easy to explain: the retransmitted messages do not contain any new information that can enhance the successful decoding probability and thus reduce AoII, while introducing longer transmission delay along with propagation delay. Therefore, simple retransmission schemes (i.e., ARQ schemes) achieve higher AoII than non-ARQ scheme. The conclusion drawn by this theorem is also consistent with previous work [16], where two transmission schemes P1 (non-ARQ transmission scheme) and P2 (truncated ARQ transmission scheme) are compared in terms of AoII, and simulation shows that P1 performs better. That is, the transmission of newly captured status information instead of retransmitting old information yields better AoII performance.

B. AoII-Optimal Block Assignment: Non-ARQ vs. HARQ

It can be implied from equation (9) that the average AoII of transmission scheme $\mathbf{N} = [n_1, n_2, \ldots, n_m]$ is related to the length of child code words and the successful decoding probability in each round. However, the AoII optimization of these two parameters contradicts. Specifically, under the same SNR, if the code length is larger, the successful decoding probability will grow and the average AoII can get decreased; conversely, the long code length will inherently result in longer transmission delay and larger AoII. Therefore, optimizing the HARQ block assignment (including non-ARQ when $m = 1$) of FBL regime to minimize the AoII is a non-trivial problem due to the natural trade-off between the transmission delay and the successful decoding probability.

1) Problem Formulation: The problem of solving the AoII-optimal block assignment can be described as follows:

(a) Objective function: the average AoII $\bar{A}$ which we have derived in (9).

(b) Decision variable: the transmission scheme $\mathbf{N}_{\text{optimal}} = [n_1, n_2, \ldots, n_m]$. Note that m is also an invisible decision variable, which is related to the following constraints on code length.

(c) Constraints: In the HARQ transmission scheme, the code length is increasing with the number of rounds, that is $n_{i+1} > n_i, i = 1, 2, \ldots, m-1$. In order to restrict the size of solution space, the maximum code length should be no larger than a certain integer $n_{\max}$, i.e., $n_m \leq n_{\max}$. Also, in (32), the condition that the left-hand-side is approximately equal to the right-hand-side is $k \geq 100$ and $n_i \geq k$. Therefore, the minimum code length should be no smaller than a certain integer $n_{\min} \geq 100$, i.e., $n_1 \geq n_{\min}$. According to code length constraint, the variable m should satisfy $1 \leq m \leq n_{\max} - n_{\min} + 1$.

Problem 1: AoII-optimal block assignment of FBL regime:

$$\begin{aligned} \min \quad & \bar{A} \\ s.t. \quad & c_1 : n_{\min} \leq n_1 < n_2 < \cdots < n_m \leq n_{\max}, \\ & c_2 : 1 \leq m \leq n_{\max} - n_{\min} + 1, \\ & c_3 : m, n_1, n_2, \ldots, n_m \in \mathbb{Z}^+. \end{aligned} \tag{39}$$

2) Solutions and Discussions: Problem 1 is a non-linear integer programming problem. The optimal algorithm that can solve the problem is an exhaustive algorithm, which lists all the candidate transmission schemes and selects the scheme with minimum average AoII as the solution, which is elaborated in Algorithm 1. Note that the complexity of this algorithm is $O(2^{n_{\max}-n_{min}+1})$, and thus the algorithm is only practical for short incremental code length. Since the scope of this paper is principally the theoretical results on average AoII performance, we adopt the optimal while complex exhaustive search as an initial example of optimization. Also, some heuristic algorithms for longer incremental length, which are elaborated in Algorithm 2 in our previous work [31], can be utilized here to solve this problem.

Algorithm 1: The Algorithm for Solving Problem 1

Input: The signal-to-noise ratio γ; The message length k; The minimum code length $n_{\min}$; The maximum code length $n_{\max}$; The delay elements τ_{enc}, τ_{prop}, τ_{dec} and τ_{fb};
Output: The optimal block assignment vector $\mathbf{N}_{\text{optimal}}$;

```
Initialization: Ā_min = ∞;
for b in {0,1}^(n_max − n_min + 1) do
    Map the vector b to the transmission scheme vector N;
    According to the obtained N, calculate the average age Ā by equation (9) ;
    if Ā < Ā_min then
        Update Ā_min = Ā;
        Update N_optimal = N;
return N_optimal
```

TABLE II
THE OPTIMAL TRANSMISSION SCHEME SOLVED BY ALGORITHM 1, WHERE $\tau_{\text{PROP}} = 20$

SNR (dB)	AoI-optimal scheme [31]	AoII-optimal scheme
-1∼0.3	[120]	[120]
0.4	[119]	[120]
0.5	[117]	[120]
0.6	[112, 120]	[120]
0.7	[110, 120]	[111, 120]
0.8	[108, 120]	[110, 120]
0.9	[107, 120]	[108, 120]
1	[105, 120]	[107, 120]
1.1	[104, 120]	[105, 120]
1.2	[102, 118]	[104, 120]
1.3	[101, 116]	[103, 120]
1.4	[100, 115]	[101, 120]
1.5∼1.6	[100, 114]	[100, 114, 120]
1.7∼1.8	[100, 113, 120]	[100, 113, 120]
1.9∼2	[100, 112, 120]	[100, 112, 120]
...	...	...

In Algorithm 1, in order to exhaustively search all the candidate block assignment, we introduce a vector $\mathbf{b} = [b_1, b_2, \ldots, b_{n_{\max}-n_{\min}+1}] \in \{0,1\}^{n_{\max}-n_{\min}+1}$ and construct a one-to-one map from $\mathbf{b}$ to the transmission scheme $\mathbf{N}$. Similar to our previous work [31], the map is constructed as follows:

(a) Find out all the one-value positions in the vector $\mathbf{b}$ and store them in the set $\mathcal{S}$.

(b) Sort the elements s_i in $\mathcal{S}$ so that it satisfies that $s_1 < s_2 < \cdots < s_{|\mathcal{S}|}$.

(c) The element in the transmission scheme $\mathbf{N}$ can be calculated by $n_i = n_{\min} + s_i - 1$.

Table II shows some examples of the solved AoII-optimal transmission scheme, where $n_{\min} = 100$, $n_{\max} = 120$, $\tau_{\text{enc}} = 20$, $\tau_{\text{prop}} = 20$, $\tau_{\text{dec}} = 30$, $\tau_{\text{fb}} = 21$, $\beta = 0.1$, and the

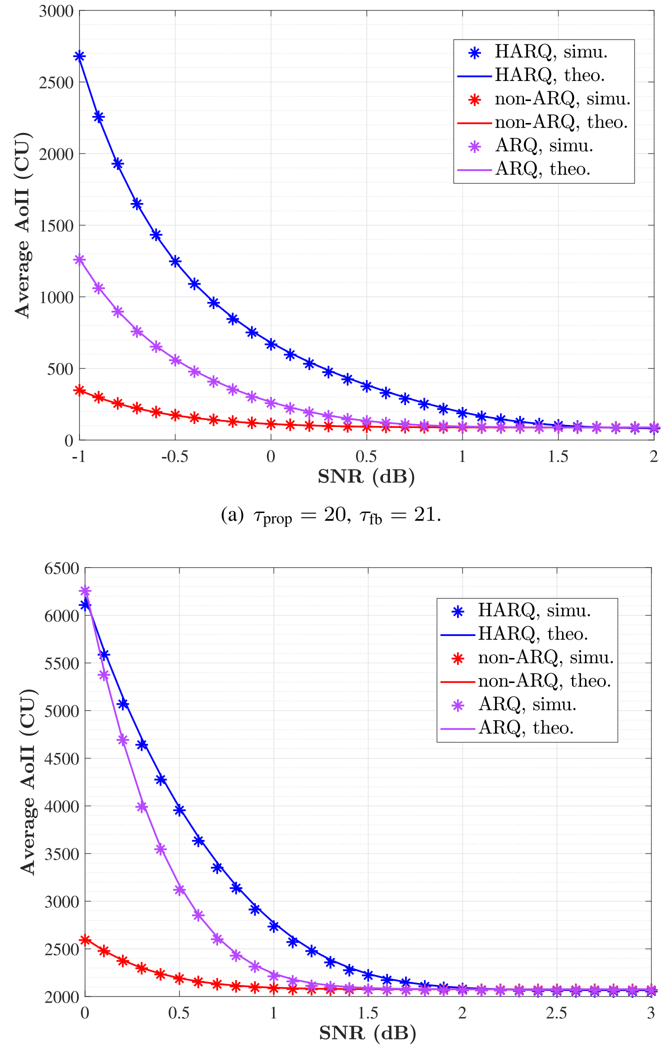
TABLE III
THE OPTIMAL TRANSMISSION SCHEME SOLVED BY ALGORITHM 1, WHERE $\tau_{\text{PROP}} = 2000$

SNR (dB)	**AoI-optimal scheme** [31]	**AoII-optimal scheme**
-1∼1.2	[120]	[120]
1.3	[118]	[120]
1.4	[116]	[120]
1.5	[114]	[119]
1.6	[113]	[117]
1.7	[111]	[115]
1.8	[109]	[113]
1.9	[107]	[111]
2	[106]	[110]
2.1	[104]	[108, 120]
2.2	[102]	[106, 120]
2.3	[101]	[104, 120]
2.4	[100]	[103, 120]
2.5	[100]	[101, 120]
2.6	[100]	[100, 120]
2.7∼2.8	[100]	[100, 119]
2.9∼3	[100]	[100, 118]
...	...	...

SNR step length is 0.1 dB.[9] In this case, the propagation (feedback) delay and transmission delay are of the same order of magnitude. It is demonstrated in Table II that AoII-optimal block assignment vectors are non-ARQ when SNR is lower than 0.6 dB, and HARQ when SNR is higher than 0.6dB. Moreover, for AoII-optimal block assignment vector, the number of symbols in the first round tends to decline while the number of maximum transmission rounds tends to increase as SNR grows. The reason why HARQ achieves better AoII performance at high SNR is that the successful decoding probability at the first round is sufficiently large despite short code length, which makes AoII more likely to be updated as zero.

Table III demonstrates the solved AoII-optimal transmission scheme, where $n_{\min} = 100$, $n_{\max} = 120$, $\tau_{\text{enc}} = 20$, $\tau_{\text{prop}} = 2000$, $\tau_{\text{dec}} = 30$, $\tau_{\text{fb}} = 2001$, $\beta = 0.1$, and the SNR step length is 0.1 dB. In this case, the propagation (feedback) delay is far greater than transmission delay. Compared to the solutions in Table II, the AoII-optimal scheme under large propagation (feedback) delay has the following characteristics: firstly, the SNR threshold of non-ARQ/HARQ is significantly higher, i.e., when SNR is larger than 2.1 dB, HARQ schemes outperform non-ARQ ones in terms of AoII; secondly, under the same SNR, the number of maximum rounds is lower by approximately one. We can explain the above characteristics as follows. Since the propagation and feedback processes occur after every round of retransmission and long feedback delay causes high probability of status change, AoII tends to grow rapidly due to long transmission process. Therefore, in order to strike a balance between propagation delay and successful decoding probability, schemes with as few maximum rounds as possible are adopted as AoII-optimal schemes.

[9]By our simulation, the transition probability of source status does not affect the solutions. The reason behind is that the transition probability does not affect the AoII evolution during transmission.



(a) $\tau_{\text{prop}} = 20$, $\tau_{\text{fb}} = 21$.

(b) $\tau_{\text{prop}} = 2000$, $\tau_{\text{fb}} = 2001$.

Fig. 5. The average AoII simulation and theoretical results of non-ARQ, ARQ and HARQ transmission schemes with different propagation delay, where $k = 100$, $\tau_{\text{enc}} = 20$, $\tau_{\text{dec}} = 30$, $\beta = 0.1$.

Furthermore, let us compare the solved AoII-optimal schemes with the AoI-optimal transmission schemes [31] which are also listed in Table II and Table III. Overall, at the same SNR, the number of symbols in each round for AoII-optimal schemes is larger than that for AoI-optimal schemes, and the number of maximum rounds for AoII-optimal schemes also tends to be larger. The reason behind is that the goal-oriented AoII metric sufficiently considers the content-aware information, namely the source status change, while AoI is content-agnostic. Specifically, since AoI is a content-agnostic metric, either a newly generated message or retransmission of old message may reduce AoI, because a successful decoding of any message regardless of message content will make AoI updated. Therefore, AoI-optimal schemes, especially when the propagation delay is long, tend to have fewer code length and fewer transmission rounds. However, after a status change,

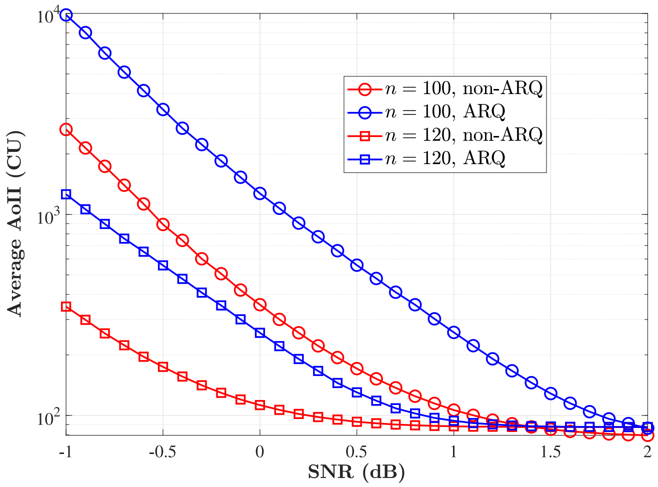


Fig. 6. The average AoII comparison of non-ARQ and ARQ transmission schemes, where $k = 100$, $\tau_{\text{enc}} = 20$, $\tau_{\text{prop}} = 20$, $\tau_{\text{dec}} = 30$, $\tau_{\text{fb}} = 21$, $\beta = 0.1$.

the only way to reduce AoII is to decode the very message that represents the changed status as accurately and rapidly as possible. Thus, retransmission schemes and non-ARQ schemes with longer code length are adopted to contribute to successful decoding on the current message.

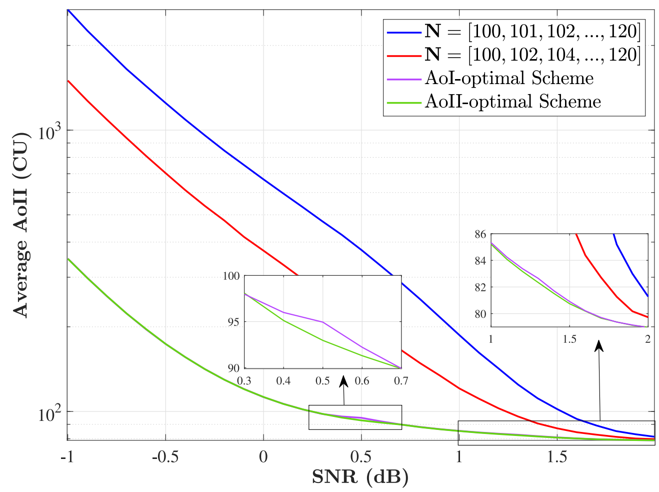
V. Numerical Results

A. Validation on Theoretical Results of Average AoII for Different Transmission Schemes

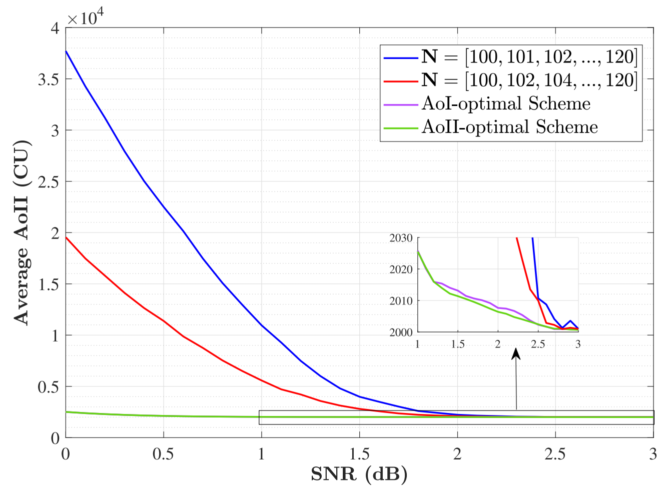
Firstly, we aim to validate our theoretical results derived in Section III by comparing the simulation results (represented by "simu.") and theoretical values of average AoII (represented by "theo.") under different transmission schemes. Fig. 5 compares the numerical and theoretical results on average AoII of non-ARQ, ARQ and HARQ transmission schemes. Specifically, we choose the non-ARQ transmission scheme as $\mathbf{N}_{\text{non}} = [120]$, the ARQ transmission scheme (in fact truncated ARQ scheme) $\mathbf{N}_{\text{ARQ}} = [120, 240, 360, 480]$ with $m = 4$, and the HARQ scheme $\mathbf{N}_{\text{non}} = [100, 110, 120]$. It can be observed from Fig. 5 that the numerical results fit close with theoretical results for all the three transmission schemes under a wide range of SNR and different propagation delays, which verifies that the derived average AoII expressions (9) can evaluate the AoII performance of coded status update systems under various FBL regimes.

B. Average AoII Performance: Non-ARQ Scheme Versus ARQ Schemes

Fig. 6 compares the average AoII of non-ARQ and ARQ transmission scheme, and for the two ARQ transmission schemes, the maximum number of rounds is four. Fig. 6 demonstrates that non-ARQ transmission schemes achieve lower AoII than ARQ transmission schemes for the same code length and other parameters. In other words, retransmitting the old messages instead of capturing and transmitting new messages will lead to decline in AoII performance, which further verifies the validity of Theorem 2.



(a) $\tau_{\text{prop}} = 20$, $\tau_{\text{fb}} = 21$.



(b) $\tau_{\text{prop}} = 2000$, $\tau_{\text{fb}} = 2001$.

Fig. 7. The average AoII comparison of AoII-optimal, AoI-optimal and baseline transmission schemes, where $k = 100$, $\tau_{\text{enc}} = 20$, $\tau_{\text{dec}} = 30$, $\beta = 0.1$.

C. Average AoII Performance: AoII-Optimal Schemes Versus HARQ Schemes

In Fig. 7, we compare the average AoII performance of the proved AoII-optimal transmission scheme with baseline transmission schemes including AoI-optimal scheme and two HARQ transmission schemes. Specifically, the total code length n_m is the same for four transmission schemes, but the number of rounds m is different. It can be concluded from Fig. 7 that the average AoII is increasing as the number of rounds increases. In other words, finer block assignments, which have larger number of transmission rounds with smaller retransmission bit lengths, will cause larger average AoII. This conclusion is consistent with the solutions of optimization problem shown in Table II. That is, a long transmission delay caused by longer code length will not affect the AoII performance, while a high error probability caused by shorter code length will significantly increase AoII.

Moreover, as compared to Fig. 7(a), Fig. 7(b) shows that the AoII-optimal schemes outperform significantly the baseline schemes under a long propagation delay. A qualitative explanation is that fewer rounds of transmission occur under the AoII-optimal schemes as compared to baseline schemes, causing shorter propagation delay and lower AoII. Besides, it is worth noting that although the AoII performance of AoI-optimal transmission scheme is similar to that of AoII-optimal scheme, the performance of the latter is still slightly better, as shown in the partial enlarged view of the two subfigures.

VI. Conclusion

In this paper, the AoII performance of coded status updated system under FBL regime is analyzed. The average AoII expressions of the system under non-ARQ, ARQ, and HARQ transmission schemes are respectively derived. Furthermore, both theoretical and numerical results show that non-ARQ schemes have better AoII performance than ARQ schemes, and solutions of AoII-optimal block assignment problem demonstrate that HARQ schemes outperform non-ARQ schemes in terms of AoII when the channel conditions are better. This paper shows an example of analyzing goal-oriented semantic metrics for practical FBL-regime-based communications systems, and will serve as an approach of realizing highly effective semantic communication.

Appendix
Proof of Lemma 1

According to the definition, the random variable R_j is geometrically distributed with successful decoding probability of S_m, so the probability distribution can be expressed as (14).

The probability of event $\{V_j = v\}$ is equivalent to the probability of event that the decoding succeeds at round v and fails at all the previous rounds, that is

$$\mathbb{P}\big(V_j = v\big) = \frac{\mathbb{P}\Big(\big\{\xi_{j,v} = 1\big\} \cap_{i=1}^{v-1} \big\{\xi_{j,i} = 0\big\}\Big)}{\mathbb{P}\big(\big\{\xi_{j,m} = 1\big\}\big)}. \tag{40}$$

Because of serial transmission of HARQ scheme, it holds that $\{\xi_{j,k} = 0\} \subseteq \cap_{i=1}^{k-1}\{\xi_{j,i} = 0\}$. Therefore, the event $\{V_j = v\}$ can be simplified as the difference of event $\{\xi_{j,v} = 1\}$ and $\{\xi_{j,v-1} = 1\}$, and the equation (40) can be simplified as

$$\mathbb{P}\big(V_j = v\big) = \frac{\mathbb{P}\big(\big\{\xi_{j,v} = 1\big\}/\big\{\xi_{j,v-1} = 1\big\}\big)}{\mathbb{P}\big(\big\{\xi_{j,m} = 1\big\}\big)}, \tag{41}$$

which is equivalent to (15).

References

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

[2] F. Chiariotti et al., "Query age of information: Freshness in pull-based communication," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 1606–1622, Mar. 2022.

[3] M. E. Ildiz, S. Avşar, and E. Uysal, "An inequality for query age of information and age of information," in *Proc. 30th Signal Process. Commun. Appl. Conf. (SIU)*, May 2022, pp. 1–4.

[4] Z. Wang, M.-A. Badiu, and J. P. Coon, "A framework for characterizing the value of information in hidden Markov models," *IEEE Trans. Inf. Theory*, vol. 68, no. 8, pp. 5203–5216, Apr. 2022.

[5] M. Bastopcu, B. Buyukates, and S. Ulukus, "Gossiping with binary freshness metric," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2021, pp. 1–6.

[6] B. Buyukates, M. Bastopcu, and S. Ulukus, "Version age of information in clustered gossip networks," *IEEE J. Sel. Areas Inf. Theory*, vol. 3, no. 1, pp. 85–97, Mar. 2022.

[7] R. D. Yates, "The age of gossip in networks," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2021, pp. 2984–2989.

[8] J. Zhong, R. D. Yates, and E. Soljanin, "Two freshness metrics for local cache refresh," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 1924–1928.

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

[10] E. Uysal et al., "Semantic communications in networked systems: A data significance perspective," *IEEE Netw.*, vol. 36, no. 4, pp. 233–240, Jul./Aug. 2022.

[11] D. Gündüz et al., "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, Jan. 2023.

[12] C. Kam, S. Kompella, and A. Ephremides, "Age of incorrect information for remote estimation of a binary Markov source," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Jul. 2020, pp. 1–6.

[13] Y. Chen and A. Ephremides, "Minimizing age of incorrect information for unreliable channel with power constraint," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

[14] A. Maatouk, M. Assaad, and A. Ephremides, "Semantics-empowered communications through the age of incorrect information," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 3995–4000.

[15] Y. Chen and A. Ephremides, "Minimizing age of incorrect information in the presence of timeout," 2022, *arXiv:2207.02926*.

[16] S. Saha, H. S. Makkar, V. B. Sukumaran, and C. R. Murthy, "On the relationship between mean absolute error and age of incorrect information in the estimation of a piecewise linear signal over noisy channels," *IEEE Commun. Lett.*, vol. 26, no. 11, pp. 2576–2580, Nov. 2022.

[17] K. Bountrogiannis, A. Ephremides, P. Tsakalides, and G. Tzagkarakis, "Age of incorrect information with hybrid ARQ under a resource constraint for N-ary symmetric Markov sources," 2023, *arXiv:2303.18128*.

[18] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 534–556, Jan. 2020.

[19] S. Leng and A. Yener, "Learning to transmit fresh information in energy harvesting networks," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 4, pp. 2032–2042, Dec. 2022.

[20] S. Wu, Z. Deng, A. Li, J. Jiao, N. Zhang, and Q. Zhang, "Minimizing age-of-information in HARQ-CC aided NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1072–1086, Feb. 2023.

[21] E. T. Ceran, D. Gündüz, and A. György, "A reinforcement learning approach to age of information in multi-user networks with HARQ," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1412–1426, Mar. 2021.

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

[23] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[24] R. C. Yavas, V. Kostina, and M. Effros, "Gaussian multiple and random access channels: Finite-blocklength analysis," *IEEE Trans. Inf. Theory*, vol. 67, no. 11, pp. 6983–7009, Nov. 2021.

[25] Q. He, Y. Zhu, P. Zheng, Y. Hu, and A. Schmeink, "Multi-device low-latency IoT networks with blind retransmissions in the finite blocklength regime," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12782–12795, Dec. 2021.

[26] I. Zabir, A. Swami, and Y. Hua, "Secrecy throughput enhancement with ANECE and multi-antenna beamforming in finite blocklength," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 2169–2174.

[27] M. Xie, Q. Wang, J. Gong, and X. Ma, "Age and energy analysis for LDPC coded status update with and without ARQ," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10388–10400, Oct. 2020.

[28] J. You, S. Wu, Y. Deng, J. Jiao, and Q. Zhang, "An age optimized hybrid ARQ scheme for polar codes via Gaussian approximation," *IEEE Wireless Commun. Lett.*, vol. 10, no. 10, pp. 2235–2239, Oct. 2021.

[29] S. C. Bobbili, P. Parag, and J.-F. Chamberland, "Real-time status updates with perfect feedback over erasure channels," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5363–5374, Sep. 2020.

[30] S. Meng, S. Wu, A. Li, J. Jiao, N. Zhang, and Q. Zhang, "Analysis and optimization of the HARQ-based spinal coded timely status update system," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6425–6440, Oct. 2022.

[31] A. Li, S. Wu, J. Jiao, N. Zhang, and Q. Zhang, "Age of information with hybrid-ARQ: A unified explicit result," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 7899–7914, Dec. 2022.

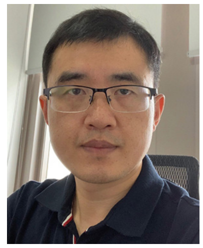
[32] J. Cao et al., "Toward industrial metaverse: Age of information, latency and reliability of short-packet transmission in 6G," *IEEE Wireless Commun.*, vol. 30, no. 2, pp. 40–47, Apr. 2023.

[33] J. Cao, X. Zhu, Y. Jiang, Z. Wei, and S. Sun, "Information age-delay correlation and optimization with finite block length," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7236–7250, Nov. 2021.

[34] *IEEE Standard for Information Technology–Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks–Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications–Amendment 4: Enhancements for Very High Throughput for Operation in Bands Below 6 GHz*, IEEE Standard 802.11ac(TM)-2013 (Amendment IEEE Standard 802.11-2012, IEEE Standard 802.11ae-2012, IEEE Standard 802.11aa-2012, and IEEE Standard 802.11ad-2012), 2013.

[35] A. Li, S. Wu, J. Jiao, N. Zhang, and Q. Zhang, "Spinal codes over fading channel: Error probability analysis and encoding structure improvement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 8288–8300, Dec. 2021.

Siqi Meng received the B.E. degree in electronic and information engineering from the Harbin Institute of Technology (Shenzhen) in 2021, where he is currently pursuing the Ph.D. degree with the Department of Electronic Engineering. His research interests include wireless communications, advanced channel coding techniques, and semantic communications.



Shaohua Wu (Member, IEEE) received the Ph.D. degree in communication engineering from the Harbin Institute of Technology in 2009. From 2009 to 2011, he held a postdoctoral position with the Department of Electronics and Information Engineering, Shenzhen Graduate School, Harbin Institute of Technology, where he has been with since 2012. From 2014 to 2015, he was a Visiting Researcher with BBCR, University of Waterloo, Canada. He is currently a Full Professor with the Harbin Institute of Technology (Shenzhen), China. He is also a Professor with Peng Cheng Laboratory, Shenzhen, China. His research interests include satellite and space communications, advanced channel coding techniques, space-air-ground-sea integrated networks, and B5G/6G wireless transmission technologies. He has authored or coauthored over 100 papers in these fields and holds over 40 Chinese patents.

Aimin Li (Student Member, IEEE) received the B.E. degree from the Harbin Institute of Technology (Shenzhen) in 2020, where he was awarded the highest honor of Undergraduate Thesis, where he is currently pursuing the Ph.D. degree. His current research interests include age of information and goal-oriented semantic communications. He has served as a Reviewer for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, and the IEEE COMMUNICATIONS LETTER.

Qinyu Zhang (Senior Member, IEEE) received the bachelor's degree in communication engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 1994, and the Ph.D. degree in biomedical and electrical engineering from the University of Tokushima, Tokushima, Japan, in 2003, where he was an Assistant Professor from 1999 to 2003. From 2003 to 2005, he was an Associate Professor with the Shenzhen Graduate School, HIT, and was the Founding Director of the Communication Engineering Research Center, School of Electronic and Information Engineering. Since 2005, he has been a Full Professor and the Dean of EIE School, HIT. He is also a Professor with Peng Cheng Laboratory, Shenzhen, China. His research interests include aerospace communications and networks, wireless communications and networks, cognitive radios, signal processing, and biomedical engineering. He has been a TPC Member for the Infocom, IEEE ICC, IEEE Globecom, IEEE Wireless Communications and Networking Conference, and other flagship conferences in communications. He was an Associate Chair of Finance of the International Conference on Materials and Manufacturing Technologies 2012, the TPC Co-Chair of the IEEE/CIC ICCC 2015, and the Symposium Co-Chair of the CHINACOM 2011, and the IEEE Vehicular Technology Conference 2016 (Spring). He was the Founding Chair of the IEEE Communications Society Shenzhen Chapter. He is on the editorial board of some academic journals, such as *Journal of Communication*, *KSII Transactions on Internet and Information Systems*, and *Science China Information Sciences*.