This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

Automatic Repeat Request Design in AoI-Aware Broadcast with Heterogeneous Direct and Relay-Assisted Users

Xueer Wang, *Student Member, IEEE,* Haoyuan Pan, *Member, IEEE,* Tse-Tin Chan, *Member, IEEE,* and Jianqiang Li, *Senior Member, IEEE*

***Abstract*—This paper investigates the design of automatic repeat request (ARQ) protocols in age of information (AoI)-aware broadcast systems with heterogeneous users, including both direct and relay-assisted users. In this setup, a direct user receives status updates directly via a single-hop link, while a relay-assisted user can receive status updates through either a direct single-hop link or a two-hop relay-assisted link. While ARQ is commonly used to ensure reliable transmission in error-prone wireless networks, previous studies suggest that ARQ does not improve the average AoI in single-hop networks. However, its impact on AoI in systems with relays, particularly those involving heterogeneous users, remains unclear. We address this gap by analyzing the average AoI under different ARQ strategies, introducing a transmission limit $k \geq 0$ at the relay. Here, $k = 0$, $k = 1$, and $k > 1$ correspond to non-relay, non-ARQ-at-relay, and truncated-ARQ-at-relay strategies, respectively. Utilizing a unified Markov chain framework that models the transmission processes for each user type, we derive the theoretical average AoI. Our results show that the direct user benefits most from non-relay and non-ARQ strategies, similar to single-hop systems. In contrast, the relay-assisted user achieves optimal performance with truncated-ARQ-at-relay, leveraging both direct and relay-assisted links. For the overall system, applying a non-ARQ-at-relay approach strikes a balance in AoI between the direct and relay-assisted users, leading to a more stable and lower system-wide average AoI.**

***Index Terms*—Age of information (AoI), automatic repeat request (ARQ), heterogeneous systems, relay.**

This work was supported in part by the National Natural Science Funds for Distinguished Young Scholar under Grant 62325307, in part by Shenzhen Medical Research Fund under Grant C2404001, in part by the National Natural Science Foundation of China under Grants 62371302 and 62203134, in part by the Natural Science Foundation of Guangdong Province under Grants 2023B1515120038, in part by Shenzhen Science and Technology Innovation Commission under Grants 20220809141216003 and KJZD20230923113801004, in part by the Scientific Instrument Developing Project of Shenzhen University under Grant 2023YQ019, in part by Shenzhen Science and Technology Program under Grant ZDSYS20220527171400002, and in part by the Research Matching Grant Scheme from the Research Grants Council of Hong Kong. *(Corresponding author: Haoyuan Pan.)*

X. Wang and H. Pan are with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China (e-mails: wangxueer2023@email.szu.edu.cn, hypan@szu.edu.cn).

T.-T. Chan is with the Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China (e-mail: tsetinchan@eduhk.hk).

J. Li is with the National Engineering Laboratory for Big Data System Computing Technology, and the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China (e-mail: lijq@szu.edu.cn).

I. Introduction

With the rapid development of the Internet of Things (IoT) and Machine-Type Communication (MTC), ensuring timely status updates has become increasingly critical. In real-time monitoring applications, such as industrial automation [1], intelligent transportation systems [2], and smart health monitoring [3], acquiring the most current data is essential for maintaining system accuracy, reliability, and performance. Timely status updates enable informed decision-making based on accurate and up-to-date information, thereby mitigating the risks associated with data staleness [4], [5].

Traditional performance metrics like throughput and delay fall short of capturing the timeliness of status updates. To address this, the age of information (AoI) has been introduced as a key metric for measuring the freshness of information in timely status update systems. First proposed by [6], AoI is defined as the elapsed time since the generation of the latest successfully received information at the destination. Recent research has focused on optimizing the average AoI, the time-averaged measure of instantaneous AoI, across various network configurations, including single-hop point-to-point networks [7], random access networks [8], multi-hop relay networks [9], etc. In addition, studies have investigated the tradeoffs between AoI and other performance metrics to enhance real-time data delivery across diverse applications [10]–[13].

In error-prone wireless channels, managing corrupted packets is crucial for maintaining a low average AoI. Traditional automatic repeat request (ARQ) systems ensure reliable transmission by retransmitting corrupted packets until they are correctly received. However, applying the classical ARQ scheme directly to status update systems can lead to high average AoI due to potentially excessive retransmissions. In single-hop systems, prior works have shown that a non-ARQ scheme, where a new packet is immediately sent when the old packet is corrupted, can significantly outperform the classical ARQ scheme. This is because newer packets carry fresher information, whereas old packets become increasingly obsolete [14]. In contrast, two-hop relay systems present a different scenario that necessitates different ARQ strategies. Studies [15], [16] suggest that in the first hop, it is beneficial to generate and send new packets when current ones are corrupted, while in the second hop, retransmitting old packets is preferable to avoid the delays associated with relaying

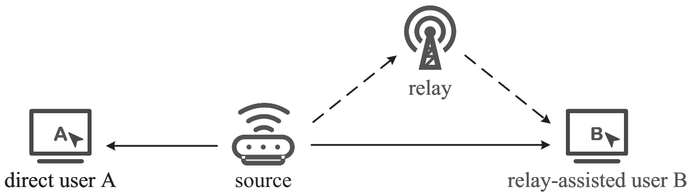


Fig. 1. A status update system with heterogeneous users: one direct user (User A) and one relay-assisted user (User B). User A receives update packets directly via a single-hop link. User B can receive update packets either directly through a single-hop link, or via a two-hop relay-assisted link.

new packets. Thus, unlike single-hop systems, two-hop relay systems benefit from a strategic combination of non-ARQ in the first hop and classical ARQ in the second hop.

While existing research has extensively explored the application of ARQ protocols in homogeneous system settings, such as non-relay single-hop and two-hop relay systems, this paper focuses on a more challenging yet practical scenario: heterogeneous broadcast systems with both direct and relay-assisted users. In such systems, the source broadcasts status updates to a direct user via a single-hop link and to a second user through both a direct and a relay-assisted link, as illustrated in Fig. 1. This hybrid structure combines elements of traditional single-hop and relay-based systems, but introduces new complexities in AoI behavior. In homogeneous systems, non-ARQ schemes are often suitable for single-hop communication, while conventional ARQ protocols are effective in dedicated relay systems. However, these existing strategies fall short in heterogeneous settings, where the two users exhibit fundamentally different AoI dynamics. More specifically, the direct user may need to wait for the successful update of the relay-assisted user before the next transmission, so its AoI evolution is directly affected by the relay-assisted user. Compared to the direct link, the relay introduces additional latency, which makes the relay-assisted user usually experience a higher instantaneous AoI when it successfully receives updates via the relay-assisted link. As a result, the presence of both direct and relay-assisted users in a heterogeneous setting complicates the AoI analysis, and identifying the optimal ARQ strategy for this hybrid system is essential for minimizing the average AoI.

In our heterogeneous system, interactions between direct and relay-assisted users significantly impact overall performance. Managing the number of packet forwarding transmissions by the relay is crucial in this context. We introduce a transmission limit $k \geq 0$ at the relay, where $k = 0$, $k = 1$, and $k > 1$ correspond to non-relay, non-ARQ-at-relay, and truncated-ARQ-at-relay [17] approaches, respectively. Consider a time-slotted system. When the source broadcasts an update packet, if $k = 0$, the system reduces to a non-relay configuration where both users rely solely on direct links for updates from the source. For $k \geq 1$, if the relay fails to receive an update packet, the source generates and sends a new packet in the next time slot, following the non-ARQ strategy in single-hop systems. If the relay successfully receives the packet but the relay-assisted user does not, the relay forwards the packet until it reaches the transmission limit k. Once this threshold is reached, the old packet is discarded, and a new one is generated and sent. Here, $k = 1$ represents a non-ARQ approach, while $k \to \infty$ aligns with the classical ARQ strategy at the relay. This study aims to investigate the optimal value of k to minimize the system's average AoI.

To derive the theoretical average AoI for both direct and relay-assisted users, we model the transmission process of each user type using a unified Markov chain framework that accommodates different approaches, such as non-relay, non-ARQ-at-relay, and truncated-ARQ-at-relay, by varying the parameter k. The states of these Markov chains are carefully designed to reflect both the successful updates of each user and the interactions between direct and relay-assisted users under the ARQ processes. In addition, given that IoT devices typically transmit small status update packets (e.g., tens of bytes), we estimate the packet error rate (PER) for each link in the heterogeneous broadcast system using short packet theory [18].

Our theoretical and simulation results suggest that different ARQ strategies should be tailored to individual users and the overall system. For the direct user, a non-relay approach ($k = 0$) yields the lowest average AoI, as the source can generate and broadcast a new packet in every time slot, minimizing the direct user's average AoI regardless of the success of previous transmissions to either user. However, when $k \geq 1$, the direct user must wait until the relay completes its forwarding process before receiving a new update, resulting in a higher average AoI. In contrast, the relay-assisted user achieves better AoI performance with truncated-ARQ-at-relay ($k > 1$) due to the combination of direct and relay-assisted links, with a smaller k typically minimizing its average AoI. For the overall system, employing a non-ARQ-at-relay approach ($k = 1$) effectively balances the average AoI between direct and relay-assisted users, leading to a more stable system-wide average AoI, even under varying signal-to-noise ratios (SNRs) across the heterogeneous broadcast system. Our findings provide valuable insights for optimizing ARQ design in AoI-aware broadcast systems with heterogeneous users.

In summary, key contributions of our work are

- We propose an AoI-aware ARQ strategy with relay transmission limits, designed specifically for heterogeneous broadcast systems where direct and relay-assisted users coexist. Unlike most existing studies that focus on homogeneous systems, our approach considers different user types and aims to achieve balanced average AoI performances of different users in a more complex network environment.
- We introduce a unified Markov chain framework to model the transmission processes of both direct and relay-assisted users with a tunable relay transmission limit k. This framework captures the dynamic interactions between users and provides a comprehensive analytical tool for evaluating AoI performance.
- We conduct comprehensive simulations to evaluate different strategies, namely the non-relay ($k = 0$), non-ARQ-at-relay ($k = 1$), and truncated-ARQ-at-relay ($k > 1$) strategies. Overall, an adjustable parameter k strikes a balance in average AoI between the direct and relay-

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

assisted users. In particular, applying a non-ARQ-at-relay approach results in a more stable and lower system-wide average AoI under different channel conditions.

II. Related Works

The AoI has been extensively studied in the literature. Initial research primarily focused on minimizing AoI within various queueing systems [7], [19]–[22], providing foundational insights into how constraints such as interference [10], throughput [11], and energy consumption [12] impact AoI optimization. Furthermore, handling packet loss through techniques like ARQ [14] and multiple-input multiple-output (MIMO) systems [23] has been critical for enhancing AoI performance under diverse communication conditions. However, these studies have predominantly addressed single-hop systems.

Subsequent research extended AoI analysis to multi-hop networks, where relays are employed to extend coverage, mitigate channel fading, and manage power constraints [9], [24]. Research in this domain often focuses on two-hop networks due to their practical relevance and complexity [25]–[30]. For example, [28] examined scenarios without a direct link between the source and destination, revealing that a smaller gap in signal-to-noise ratio (SNR) between the two hops results in better AoI performance. Conversely, [26], [27] studied relay networks with both relay-assisted and direct links, demonstrating that retransmissions via relays can significantly improve system timeliness. In addition, [25] explored time division multiple access (TDMA) and non-orthogonal multiple access (NOMA) schemes for relay-assisted communication, showing substantial AoI improvements over direct transmission under poor channel conditions. While these studies focused on homogeneous system settings with solely relay-assisted users, [31] considered the average AoI in a multiple access scenario involving both direct and relay-assisted users, albeit without ARQ. These works are primarily based on static relay settings. In parallel, recent studies have also explored the use of mobile relays to enhance network coverage and improve information freshness in more dynamic environments. For instance, unmanned aerial vehicles (UAV) and autonomous underwater vehicles have been employed as mobile relays to improve AoI performance in UAV-assisted wireless networks [38], [39], mobile ad hoc networks [40], and the Internet of Underwater Things [41]. In contrast, our research introduces a heterogeneous broadcast environment including heterogeneous users, emphasizing the management of relay transmissions and the integration of ARQ to minimize average AoI.

A significant body of work has also explored AoI performance under various ARQ protocols, including classical ARQ [14]–[16], truncated ARQ [32], [33], and hybrid ARQ (HARQ) [34], [35]. While traditional ARQ strategies primarily aim to ensure reliable communication [36], [37], the aforementioned works optimized AoI-related metrics, which is essential for real-time status updates in emerging applications such as IoT, autonomous systems, and industrial automation. For example, [14] analyzed the impact of ARQ on the average AoI in single-hop networks, demonstrating that a non-ARQ scheme significantly outperforms the classical ARQ approach. Reference [15] further showed that combining non-ARQ in the first hop with classical ARQ in the second hop yields superior AoI performance, using a Markov chain model to capture the transmission process. Additionally, [32] explored a truncated ARQ approach with distinct retransmission thresholds for each hop, indicating that widening the threshold gap can significantly decrease the average AoI. In the context of cognitive-radio-based Internet of Things (CR-IoT) with outdated channel state information, [33] found that truncated ARQ excels at high packet generation rates, whereas classical ARQ is preferable under intense interference. Furthermore, [35] illustrated that HARQ not only reduces the average AoI but also curtails energy consumption more effectively than classical ARQ.

Along this line, our work also focuses on AoI optimization rather than solely on reliability or throughput, which is relevant to real-time applications where timely information delivery is paramount. To the best of our knowledge, however, our work is the first to study the performance of truncated ARQ in AoI-aware heterogeneous systems. While truncated ARQ has been explored in the context of AoI optimization, its application in heterogeneous systems remains unexplored. This unique focus sets our work apart and addresses a critical gap in the literature.

III. Preliminaries

A. System Model and Age of Information (AoI)

We investigate a status update system with heterogeneous direct and relay-assisted users. The system consists of a source node, a half-duplex decode-and-forward (DF) relay node, and two user nodes, A and B, as depicted in Fig. 1. The source aims to broadcast status update packets to the direct user A, who is within the communication range of the source (i.e., a direct channel), and to the relay-assisted user B, who requires the assistance of a relay due to a weak channel condition between the source and the user node. In this setup, the relay-assisted user B can receive update packets from the source through either the direct or relay-assisted channel. Both users A and B aim to receive update packets from the source that are as fresh as possible.

In this study, the age of information (AoI) is employed to quantify the freshness of the update packets from the source at the destinations. At any time t, the instantaneous AoI of the source, as measured at the destination (i.e., user A or user B), is defined as $\Delta_i(t) = t - G_i(t)$, where $G_i(t)$ is the generation time of the most recently received update packet at user $i \in \{A, B\}$ from the source. A lower instantaneous AoI $\Delta_i(t)$ signifies higher information freshness for user i in obtaining the source status [5]. We consider a time-slotted system where time is divided into multiple slots, each corresponding to the duration required to send a status update packet. Consequently, $\Delta_i(t)$ can be measured in units of time slots.

Fig. 2 illustrates an example of the instantaneous AoI $\Delta_i(t)$. This work adopts a generate-at-will packet generation model [5], where the source can take measurements and generate an update packet whenever it has the opportunity to transmit. This approach ensures that the sampled status is as fresh as possible,

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

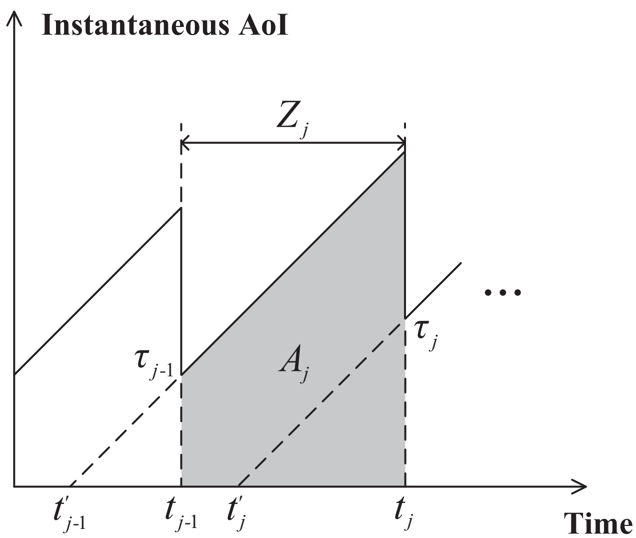


Fig. 2. An example of the instantaneous AoI $\Delta_i(t)$, where the $(j-1)$-th and j-th successful updates occur at times t_{j-1} and t_j, respectively.

like a sensor reading obtained just before the transmission opportunity. Each time slot is assumed to have unit length, which provides a normalized time basis for the subsequent AoI analysis. With respect to Fig. 2, the source generates and sends update packets at times t'_{j-1} and t'_j, which are received at times t_{j-1} and t_j, respectively, corresponding to the $(j-1)$-th and the j-th successful updates. Let τ denote the instantaneous AoI at the moment when the destination successfully receives an update packet. As depicted in Fig. 2, the instantaneous AoI $\Delta_i(t)$ drops to τ_{j-1} and τ_j at times t_{j-1} and t_j, respectively. Between these two consecutive successful updates, the instantaneous AoI $\Delta_i(t)$ increases linearly with time t.

The average AoI of a status update system is commonly evaluated in the literature [5]. The average AoI for user i, denoted as $\bar{\Delta}_i$, is defined as the time average of the instantaneous AoI, i.e.,

$$\bar{\Delta}_i = \lim_{T\to\infty} \frac{1}{T}\int_0^T \Delta_i(t)dt. \tag{1}$$

Hence, the average AoI of the system is simply $\bar{\Delta} = (\bar{\Delta}_A + \bar{\Delta}_B)/2$.

To compute the average AoI $\bar{\Delta}_i$, let Z denote the duration between two consecutive successful status updates. For example, in Fig. 2, Z_j represents the time between the $(j-1)$-th and j-th successful updates. The area A_j under the instantaneous AoI curve between the $(j-1)$-th and j-th successful updates is calculated as

$$A_j = \tau_{j-1} Z_j + \frac{1}{2}(Z_j)^2. \tag{2}$$

Then, the average AoI of user i, $\bar{\Delta}_i$ can be computed by

$$\begin{aligned}\bar{\Delta}_i &= \lim_{J\to\infty} \frac{\sum_{j=1}^{J} A_j}{\sum_{j=1}^{J} Z_j} = \frac{\mathbb{E}\left[\tau Z + \frac{1}{2}(Z)^2\right]}{\mathbb{E}[Z]} \\ &= \frac{\mathbb{E}\left[\tau Z\right]}{\mathbb{E}\left[Z\right]} + \frac{\mathbb{E}\left[Z^2\right]}{2\mathbb{E}\left[Z\right]},\end{aligned} \tag{3}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator.

In wireless communication systems, packet loss is inevitable due to wireless channel impairments. Reliable packet delivery over error-prone wireless channels typically employs packet retransmission managed through automatic repeat request (ARQ) protocols. While packet retransmission enhances the reliability of packet delivery, it also increases the time required to receive a packet, which can result in outdated packets by the time they are received. The source may have generated and sent a newer packet, thus achieving a lower AoI. Reference [14] studied the impact of ARQ on the average AoI in a point-to-point single-hop system, demonstrating that packet retransmission does not help reduce the average AoI under the generate-at-will model. This is because when the receiver fails to receive an older packet, the generate-at-will transmitter can generate and send a newer update packet that always contains fresher information, effectively adopting a non-ARQ strategy at the transmitter.

Expanding from single-hop systems to two-hop relay systems, [15] demonstrated that in a two-hop system, it is beneficial to generate and send new packets when packets are corrupted in the first hop, as the newly sent packets contain fresher information. Conversely, when packets are corrupted in the second hop, it is advantageous to retransmit old packets until they are successfully received. This is because reverting to the first hop incurs additional time for the relay to receive a packet from the source. Thus, a strategic approach combining non-ARQ in the first hop with classical ARQ in the second hop is necessary to optimize performance in two-hop relay systems.

B. ARQ with Heterogeneous Direct and Relay Channels

The heterogeneous direct and relay-assisted system considered in this work incorporates features of both single-hop and two-hop systems. The direct user A is single-hop from the source, favoring the non-ARQ scheme because new packets contain fresher information. In contrast, relay-assisted user B operates through both single-hop and two-hop links, necessitating a more nuanced ARQ strategy. Furthermore, the transmission processes of the two users are interdependent in such a heterogeneous setting. For example, the direct user must wait until the relay completes its forwarding process before receiving a new update. Thus, designing an effective ARQ protocol requires careful consideration, particularly when optimizing for the average AoI of the heterogeneous broadcast system.

When the source broadcasts an update packet to the direct user A, the relay-assisted user B, and the relay, if the relay fails to receive that packet, the source should generate and send a new update packet in the next time slot. This aligns with the non-ARQ scheme used in single-hop systems and the first hop of two-hop systems. For packet forwarding, if the relay successfully receives the update packet from the source but the relay-assisted user fails to receive it via the direct link, the relay should forward the update packet to the relay-assisted user, irrespective of the reception result at the direct user. This is necessary because the channel condition of the

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

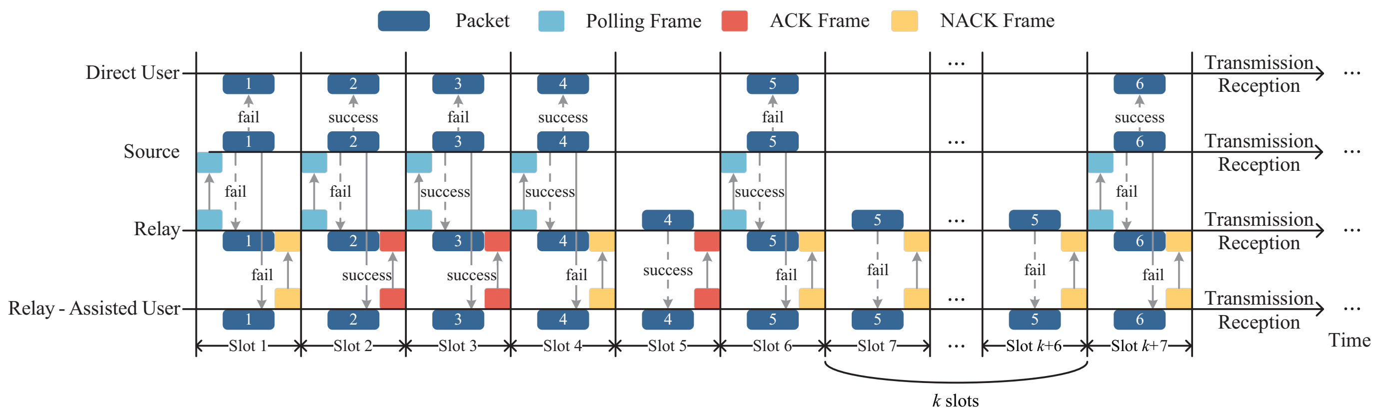


Fig. 3. An example of the medium access control (MAC) protocol for the heterogeneous broadcast system with a transmission limit k at the relay. The time slots from slot 7 to slot $k+6$ correspond to the k transmissions.

direct link for the relay-assisted user is typically weak, leading to a higher instantaneous AoI compared to the direct user, which receives most of its update packets directly. From a system-wide perspective, the relay should not delay packet forwarding to wait for a successful update at the direct user, as this would prioritize the direct user and potentially increase the average AoI for the relay-assisted user.

To effectively manage the timeliness of information delivery in a heterogeneous broadcast system, this study investigates the design of an ARQ protocol at the relay node. Specifically, we consider a truncated ARQ scheme in which the number of transmission attempts at the relay is limited by a parameter $k \geq 0$. The case of $k = 0$ corresponds to a non-relay scenario, where the relay does not participate in forwarding packets. When $k = 1$, the relay uses a non-ARQ protocol, forwarding each packet only once. For $k > 1$, the relay performs a truncated ARQ protocol, attempting retransmissions up to k times. If the relay fails to successfully deliver a packet to the relay-assisted user within k attempts, it discards the outdated packet and triggers the source to generate and transmit a fresh status update in the next time slot. This mechanism prevents the system from wasting resources on stale information and introduces a tunable design parameter k that directly influences the freshness of information received by both users.

The primary objective of this work is to determine the optimal value of k that minimizes the average AoI of the entire system, taking into account the competing needs of direct and relay-assisted users. By modeling the impact of limited relay transmissions on AoI dynamics, we aim to provide a theoretical framework for optimizing ARQ design in heterogeneous broadcast settings. The next section presents the detailed protocol and the corresponding analytical derivation of average AoI under this setting.

IV. AoI Analysis of The Heterogeneous Broadcast System

In this section, we analyze the average AoI in a heterogeneous broadcast system using a unified Markov chain framework tailored for each user type. This framework accommodates various transmission approaches, such as non-relay, non-ARQ-at-relay, and truncated-ARQ-at-relay, by adjusting the parameter k. We begin by describing the medium access control (MAC) protocol of the system. Next, we derive the average AoI for both the direct and relay-assisted users. Finally, we calculate the system-wide average AoI, integrating the individual user analyses.

A. Protocol Details of the Heterogeneous Broadcast System

Fig. 3 illustrates an example of the MAC protocol for the heterogeneous broadcast system. The relay coordinates the transmission of different nodes in the system. Each time slot consists of three mini-slots: the first mini-slot is reserved for polling frames sent by the relay; the second mini-slot is used for update packet transmissions sent by the source or the relay; the third mini-slot is reserved for acknowledgment (ACK) or negative acknowledgment (NACK) frames sent by the relay-assisted user. For simplicity, we assume the control frames, including the polling, ACK, and NACK frames, are error-free. It is worth noting that our scheme is an extension of the conventional ARQ protocol and does not introduce any additional computational overhead.

The source generates and sends a new update packet only upon receiving a polling frame from the relay (i.e., the generate-at-will model). As shown in slot 1 of Fig. 3, the relay sends a polling frame to the source at the beginning of the time slot. Upon receiving the polling frame, the source broadcasts an update packet to the direct user, the relay-assisted user, and the relay. The direct user simply tries to decode the packet transmitted from the source. For the direct link associated with the relay-assisted user, if the relay-assisted user fails to decode the packet, it sends a NACK frame to the relay at the end of the slot (e.g., see slots 1 and 4); otherwise, it sends an ACK frame to the relay (e.g., see slots 2 and 3).

Regarding the relay-assisted link, if the relay fails to receive the update packet from the source (e.g., see slots 1 and 2), or if both the relay and the relay-assisted user successfully receive the update packet from the direct link (e.g., see slot 3), the relay sends a new polling frame to the source in the next time slot to request a new generated update packet. Suppose only the relay successfully receives the update packet (e.g., see slot 4). In that case, it will not send a polling frame in the next slot (as the source does not need to send a new update packet) and instead starts forwarding the received update packet to the relay-assisted user in the next time slot, regardless of the

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

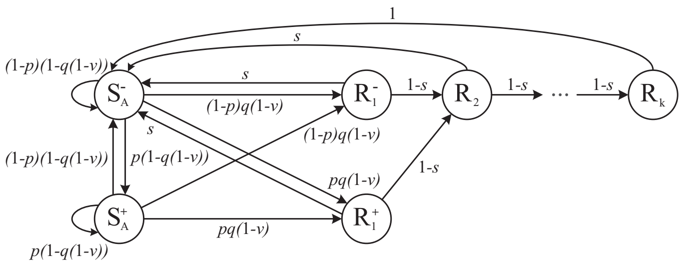


Fig. 4. The Markov chain for the direct user A in the heterogeneous broadcast system with a transmission limit k at the relay. The state transition probabilities are defined by p, q, s, and v, where p is the success probability for user A's single-hop reception, v is the success probability for user B's single-hop reception, and q and s represent the success probabilities for the first and second hops in the relay-assisted scenario for user B.

reception result of the direct user. The relay is allowed to forward the update packet up to k times to the relay-assisted user. Suppose that in slot 5 of Fig. 3, the relay-assisted user receives the update packet forwarded by the relay after the first transmission attempt. Then it sends an ACK frame to the relay at the end of slot 5. Subsequently, the relay polls a new update packet in slot 6.

However, if the relay-assisted user fails to receive the update packet forwarded by the relay, it sends a NACK frame to the relay (e.g., see slot 7). After k attempts, e.g., from slot 7 to slot $k+6$ in Fig. 3, the relay discards the old packet and sends a new polling frame in the next time slot (i.e., see slot $k+7$) to inform the source to generate and send a new update packet. Note that the direct user remains silent during the truncated ARQ process at the relay.

B. Average AoI Derivation

We now model the transmission process of each user using a Markov chain to derive the average AoI for each user. We assume that the duration of control frames (i.e., polling frames and ACK/NACK frames) is negligible compared to update packets.

1) The Average AoI of the Direct User: Let us first focus on the average AoI of the direct user A. Suppose the source broadcasts a new update packet in the current time slot. The direct user A may either successfully receive the update packet (i.e., a successful update) or fail to receive it (i.e., a failed update). Depending on the reception results at the relay and the direct link of the relay-assisted user, in the next time slot, the source broadcasts a new update packet, or the relay starts forwarding the update packet. We denote the transmission states as follows: S_A indicates the source broadcasts a new update packet (note: we add a subscript of A to differentiate the state of the relay-assisted user B presented later), and R_w indicates the relay forwards the update packet for the w-th time, where $w \in \{1, 2, \ldots, k\}$.

We model the transmission process of the direct user A using a Markov chain, depicted in Fig. 4. The states of the Markov chain are defined as follows:

- State S_A^+: The source broadcasts a new update packet in the current time slot, and the direct user A successfully receives it. In the next time slot, the transmission state moves to S_A.
- State S_A^-: The source broadcasts a new update packet in the current time slot, and the direct user A fails to receive it. In the next time slot, the transmission state moves to S_A.
- State R_1^+: Both the direct user A and the relay successfully receive the update packet broadcast by the source in the current time slot, and the transmission state moves to R_1 in the next time slot.
- State R_1^-: The direct user A fails to receive the update packet broadcast by the source, but the relay successfully receives it in the current time slot, and the transmission state moves to R_1 in the next time slot.
- State R_w (forwarding attempt w by the relay): The relay forwards the update packet for the w-th time, where $w \in \{2, \ldots, k\}$.

Notice that state R_1 is split into R_1^+ and R_1^-. This is because when the relay starts to forward the update packet to the relay-assisted user B (i.e., the transmission state R_1 is entered), whether the direct user A has a successful update in the current time slot affects its AoI evolution during the waiting period when the relay forwards the update packet to user B.

Let p, q, s, and v denote the successful reception probabilities for the links from the source to the direct user A, from the source to the relay, from the relay to the relay-assisted user B, and from the source to the relay-assisted user B, respectively. Together with the states defined above, let us explain the Markov chain depicted in Fig. 4.

Suppose the Markov chain starts in either state S_A^+ or S_A^-, depending on whether the direct user A successfully receives an update packet in the current time slot. In the subsequent time slot, if in transmission state S_A, the source broadcasts an update packet to the direct user A, the relay-assisted user B, and the relay. The Markov chain can transit to one of the states S_A^+, S_A^-, R_1^+, or R_1^-, depending on the reception outcomes at these nodes.

For example, if the direct user A successfully receives the update packet, except for the case where the relay-assisted user B fails to receive the packet from the direct link but the relay successfully receives it, the Markov chain transits to state S_A^+ with probability $p(1-q(1-v))$. That is, the direct user receives an update in the current time slot, and the next time slot is in a transmission state S_A where the source broadcasts a new update packet. Suppose the relay-assisted user fails to receive the packet from the direct link, but the relay successfully receives it. If the direct user successfully receives the update packet (with probability $pq(1-v)$), the Markov chain transits to state R_1^+. This state signifies that the relay will start forwarding the packet to the relay-assisted user in the next time slot. However, if the direct user fails to receive the update packet (with probability $(1-p)q(1-v)$), the Markov chain transits to state R_1^-. This state indicates that the relay will begin forwarding the packet, irrespective of whether the direct user received it, provided the relay successfully received the packet from the source and the relay-assisted user failed to receive it from the direct link.

In state R_1^+ or state R_1^-, the relay begins forwarding the

$$\Omega_A = \begin{pmatrix} \omega_{S_A^+S_A^+} & \omega_{S_A^+S_A^-} & \omega_{S_A^+R_1^+} & \omega_{S_A^+R_1^-} & \omega_{S_A^+R_2} & \omega_{S_A^+R_3} & \cdots & \omega_{S_A^+R_k} \\ \omega_{S_A^-S_A^+} & \omega_{S_A^-S_A^-} & \omega_{S_A^-R_1^+} & \omega_{S_A^-R_1^-} & \omega_{S_A^-R_2} & \omega_{S_A^-R_3} & \cdots & \omega_{S_A^-R_k} \\ \omega_{R_1^+S_A^+} & \omega_{R_1^+S_A^-} & \omega_{R_1^+R_1^+} & \omega_{R_1^+R_1^-} & \omega_{R_1^+R_2} & \omega_{R_1^+R_3} & \cdots & \omega_{R_1^+R_k} \\ \omega_{R_1^-S_A^+} & \omega_{R_1^-S_A^-} & \omega_{R_1^-R_1^+} & \omega_{R_1^-R_1^-} & \omega_{R_1^-R_2} & \omega_{R_1^-R_3} & \cdots & \omega_{R_1^-R_k} \\ \omega_{R_2S_A^+} & \omega_{R_2S_A^-} & \omega_{R_2R_1^+} & \omega_{R_2R_1^-} & \omega_{R_2R_2} & \omega_{R_2R_3} & \cdots & \omega_{R_2R_k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_{R_kS_A^+} & \omega_{R_kS_A^-} & \omega_{R_kR_1^+} & \omega_{R_kR_1^-} & \omega_{R_kR_2} & \omega_{R_kR_3} & \cdots & \omega_{R_kR_k} \end{pmatrix}$$
$$= \begin{pmatrix} p(1-q(1-v)) & (1-p)(1-q(1-v)) & pq(1-v) & q(1-p)(1-v) & 0 & 0 & \cdots & 0 \\ p(1-q(1-v)) & (1-p)(1-q(1-v)) & pq(1-v) & q(1-p)(1-v) & 0 & 0 & \cdots & 0 \\ 0 & s & 0 & 0 & 1-s & 0 & \cdots & 0 \\ 0 & s & 0 & 0 & 1-s & 0 & \cdots & 0 \\ 0 & s & 0 & 0 & 0 & 1-s & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \quad (4)$$

update packet to the relay-assisted user B. If the relay-assisted user successfully receives the update packet (with probability s), the Markov chain moves to state S_A^-, indicating that the source should send a new update packet in the next time slot. On the other hand, if the relay-assisted user does not successfully receive the update packet (with probability $1-s$), the Markov chain transits to state R_2, where the relay attempts to forward the update packet to the relay-assisted user for the second time. This state transition pattern continues for states $R_w, w \in \{2, \ldots, k-1\}$, either transiting to state R_{w+1} or S_A^-. When the Markov chain reaches state R_k, after the relay has attempted to forward the update packet k times, it transits back to state S_A^- with probability 1, regardless of the reception result at the relay-assisted user.

With the Markov chain modeling, we can derive the average AoI of the direct user A. Let Q_A represent the state space of the Markov chain in Fig. 4, i.e., $Q_A = \{S_A^+, S_A^-, R_1^+, R_1^-, R_2, \ldots, R_k\}$. Let $V_A = \{S_A^+, R_1^+\}$ represent the set of states where the direct user A has a successful update. We use Ω_A to represent the state transition matrix, which can be written as (4), where ω_{xy} is the state transition probability from state $J = x$ to state $J = y$, where $x, y \in Q_A$.

To compute the average AoI of the direct user A, we need to compute $\mathbb{E}[Z]$, $\mathbb{E}\left[Z^2\right]$, and $\mathbb{E}[\tau Z]$. We first compute $\mathbb{E}[Z]$ and $\mathbb{E}\left[Z^2\right]$. Let $M_{\alpha V_A}$ denote the expected time required to traverse from state $J_0 = \alpha$ to state $J_Z = V_A$ for the first time through a series of intermediate states $J_1, J_2, \ldots, J_{Z-1}$, where Z is the duration between two consecutive successful updates. Based on the properties of the Markov chain, $M_{\alpha V_A}$ can be expressed as

$$\begin{aligned} M_{\alpha V_A} &= \mathbb{E}\left[T_{V_A} | J_0 = \alpha\right] \\ &= 1 \cdot \Pr\left(J_1 = V_A | J_0 = \alpha\right) \\ &\quad + \sum_{\beta \neq V_A} \mathbb{E}\left[1 + T_{V_A} | J_1 = \beta\right] \Pr\left(J_1 = \beta | J_0 = \alpha\right). \end{aligned} \quad (5)$$

where T_{V_A} is a random variable representing the time to reach state $J = V_A$ from state $J_0 = \alpha$ for the first time. Similarly, we use $N_{\alpha V_A}$ to represent the expectation of the second moment of the expected time required to traverse from state $J_0 = \alpha$ to state $J_Z = V_A$. Hence, $N_{\alpha V_A}$ can be expressed as

$$\begin{aligned} &N_{\alpha V_A} \\ &= \mathbb{E}\left[\left(T_{V_A}\right)^2 | J_0 = \alpha\right] \\ &= 1 \cdot \Pr\left(J_1 = V_A | J_0 = \alpha\right) \\ &\quad + \sum_{\beta \neq V_A} \mathbb{E}\left[\left(1 + T_{V_A}\right)^2 | J_1 = \beta\right] \Pr\left(J_1 = \beta | J_0 = \alpha\right). \end{aligned} \quad (6)$$

The duration between two consecutive successful updates, Z, is equal to the time required to start from state J_0 to state J_Z, where both J_0 and J_Z belong to V_A. Since V_A contains two states, i.e., $V_A = \{S_A^+, R_1^+\}$, the expectation of the first moment of Z, $\mathbb{E}[Z]$, is the summation of two terms

$$\mathbb{E}[Z] = \frac{\pi_{S_A^+}}{\pi_{S_A^+} + \pi_{R_1^+}} M_{S_A^+ V_A} + \frac{\pi_{R_1^+}}{\pi_{S_A^+} + \pi_{R_1^+}} M_{R_1^+ V_A}, \quad (7)$$

where each term is weighted by the probability of the initial state, i.e., whether the initial state is S_A^+ or R_1^+, reflecting their contributions to the final result of $\mathbb{E}[Z]$. $\pi = (\pi_{S_A^+}, \pi_{S_A^-}, \pi_{R_1^+}, \pi_{R_1^-}, \pi_{R_2}, \ldots, \pi_{R_k})$ is the stationary distribution of the Markov chain depicted in Fig. 4, which is used to compute the weights in (7). Similarly, $\mathbb{E}\left[Z^2\right]$ can be computed by

$$\mathbb{E}\left[Z^2\right] = \frac{\pi_{S_A^+}}{\pi_{S_A^+} + \pi_{R_1^+}} N_{S_A^+ V_A} + \frac{\pi_{R_1^+}}{\pi_{S_A^+} + \pi_{R_1^+}} N_{R_1^+ V_A}, \quad (8)$$

According to the property of a Markov chain, the stationary distribution π can be obtained by solving $\pi \Omega_A = \pi$, from which we have

$$\frac{\pi_{S_A^+}}{\pi_{S_A^+} + \pi_{R_1^+}} = 1 - q(1-v), \qquad \frac{\pi_{R_1^+}}{\pi_{S_A^+} + \pi_{R_1^+}} = q(1-v). \quad (9)$$

To compute $\mathbb{E}[Z]$ and $\mathbb{E}[Z^2]$, we now need to compute $M_{\alpha V_A}$ and $N_{\alpha V_A}$ for $\alpha = \{S_A^+, R_1^+\}$. According to (5), we have different equations based on different states $J_0 = \alpha \in$

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

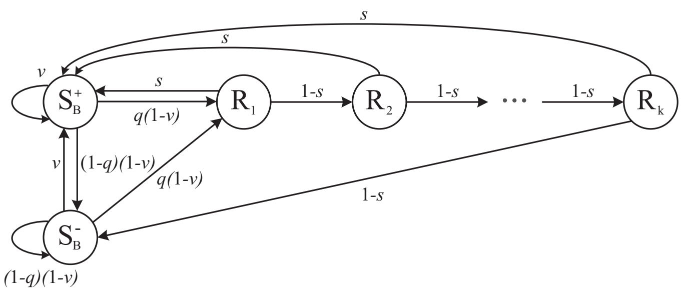


Fig. 5. The Markov chain for the relay-assisted user B in the heterogeneous broadcast system with a transmission limit k at the relay.

Q_A (see Appendix A for the details). Solving these equations (see Appendix B for the detailed mathematical computation), we obtain $M_{S_A^+ V_A}$ and $M_{R_1^+ V_A}$. Substituting them into (7), we have

$$\mathbb{E}\left[Z\right] = \frac{q(1-v)\phi + s}{ps}, \tag{10}$$

where $\phi = 1 - (1-s)^k$. Similarly, we can list different equations of (6) based on different states $J_0 = \alpha \in Q_A$ (also see Appendix A for the details). After solving the equations, we obtain $N_{S_A^+ V_A}$ and $N_{R_1^+ V_A}$, which are substituted into (8)

$$\begin{aligned} \mathbb{E}\left[Z^2\right] = & \frac{2q^2(1-v)^2(1-p)\phi^2}{p^2 s^2} \\ & + \frac{q(1-v)(ps(2k-3) + 2(p+2s))\phi}{p^2 s^2} \\ & + \frac{-2kpq(1-v) + s(2-p)}{p^2 s}. \end{aligned} \tag{11}$$

It is easy to observe that $\mathbb{E}[\tau Z] = \mathbb{E}[Z]$ because for the direct user A, the instantaneous AoI upon a successful update τ is always one time slot. Finally, we compute the average AoI of the direct user A, $\bar{\Delta}_A$ (12), by substituting (10) and (11) into (3).

2) The Average AoI of the Relay-Assisted User: We now focus on the average AoI of the relay-assisted user B. Similar to the case of the direct user A, we denote the transmission states as follows: S_B indicates the source broadcasts a new update packet (note: we add a subscript of B since we now focus on the relay-assisted user B), and R_w indicates the relay forwards the update packet for the w-th time, $w \in \{1, 2, \ldots, k\}$.

We model the transmission of the relay-assisted user B using the Markov chain shown in Fig. 5. The states of the Markov chain are defined as follows:

- State S_B^+: The source broadcasts a new update packet in the current time slot, and the relay-assisted user B successfully receives it. In the next time slot, the transmission state moves to S_B.
- State S_B^-: The source broadcasts a new update packet in the current time slot, and the relay-assisted user B fails to receive it. In the next time slot, the transmission state moves to S_B.
- State R_w (forwarding attempt w by the relay): The relay forwards the update packet for the w-th time, where $w \in \{1, \ldots, k\}$.

To explain the Markov chain depicted in Fig. 5 for the relay-assisted user B, we start by considering the initial state of the Markov chain, which can be either S_B^+ or S_B^-. This initial state depends on whether the relay-assisted user successfully received an update packet in the current time slot. If the relay-assisted user successfully receives the update packet through the direct link, the Markov chain moves to state S_B^+ with probability v. This indicates that user B has received the packet directly and will be in state S_B^+ in the next time slot, i.e., the source sends a new update packet.

If the relay-assisted user B fails to receive the update packet directly from the source, but the relay successfully receives it, the Markov chain moves to state R_1 with probability $q(1-v)$. This state signifies that the relay has received the packet and will start forwarding it to the relay-assisted user. In contrast, if neither the relay-assisted user nor the relay successfully receives the update packet, the Markov chain transits to state S_B^- with probability $(1-q)(1-v)$. In this scenario, the source will broadcast a new update packet in the next time slot.

At all states $R_w, w \in \{1, \ldots, k-1\}$, the state transitions are identical: transitioning to state R_{w+1} if the relay-assisted user fails to receive the update packet, or to state S_B^+ if the update packet is successfully received. Upon reaching state R_k, after the relay's last attempt to forward the packet, the Markov chain proceeds to state S_B^+ with probability s, indicating successful reception by user B. Alternatively, it moves to state S_B^- with probability $1-s$, indicating that user B failed to receive the update packet after the ARQ process, leading the relay to discard the old packet.

Notice that for the direct user A, we introduce separate states R_1^+ and R_1^- because the successful or failed reception of an update packet by user A significantly impacts its AoI evolution during the waiting period when the relay forwards the update to user B. In both cases, the system transitions to the same forwarding state R_1 in the next slot, where the relay begins transmission to user B. However, user As AoI evolves differently depending on whether it received the update, which makes it essential to distinguish R_1^+ and R_1^- in the Markov model. For the relay-assisted user B, only state R_1 is required. This is because the relay forwards an update only if user B fails to receive it directly from the source, and the relay has successfully received it. If both the relay and user B receive the update in the broadcast phase, the source generates a new update in the next slot instead of initiating relay forwarding. Thus, splitting the state as in user As case is unnecessary.

We derive the average AoI of the relay-assisted user B using the Markov chain in Fig. 5. Similar to the case of the direct user A, let Q_B represent the state space, i.e., $Q_B = \{S_B^+, S_B^-, R_1, R_2, \ldots, R_k\}$. Denote by $V_B = \{S_B^+\}$ the set of states that the relay-assisted user can have a successful update. Let Ω_B represent the state transition matrix, as written in (13).

As in the case of the direct user, we can calculate $\mathbb{E}\left[Z\right]$ and $\mathbb{E}\left[Z^2\right]$ by $\mathbb{E}\left[Z\right] = M_{S_B^+ S_B^+}$ and $\mathbb{E}\left[Z^2\right] = N_{S_B^+ S_B^+}$, which respectively represent the expectation of the first and second moments of the time required to traverse from state $J_0 = S_B^+$ to state $J_Z = S_B^+$ for the first time through a series of states $J_1, J_2, \ldots, J_{Z-1} \notin S_B^+$. Here, Z is the duration

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

$$\bar{\Delta}_A = \frac{q^2(1-v)^2(1-p)\phi^2}{ps(q(1-v)\phi+s)} + \frac{q(1-v)(ps(2k-1)+2(p+2s))\phi}{2ps(q(1-v)\phi+s)} + \frac{-2kpq(1-v)+s(p+2)}{2p(q(1-v)\phi+s)}, \text{ where } \phi = 1-(1-s)^k. \tag{12}$$

$$\Omega_B = \begin{pmatrix} \omega_{S_B^+S_B^+} & \omega_{S_B^+S_B^-} & \omega_{S_B^+R_1} & \omega_{S_B^+R_2} & \cdots & \omega_{S_B^+R_k} \\ \omega_{S_B^-S_B^+} & \omega_{S_B^-S_B^-} & \omega_{S_B^-R_1} & \omega_{S_B^-R_2} & \cdots & \omega_{S_B^-R_k} \\ \omega_{R_1S_B^+} & \omega_{R_1S_B^-} & \omega_{R_1R_1} & \omega_{R_1R_2} & \cdots & \omega_{R_1R_k} \\ \omega_{R_2S_B^+} & \omega_{R_2S_B^-} & \omega_{R_2R_1} & \omega_{R_2R_2} & \cdots & \omega_{R_2R_k} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega_{R_kS_B^+} & \omega_{R_kS_B^-} & \omega_{R_kR_1} & \omega_{R_kR_2} & \cdots & \omega_{R_kR_k} \end{pmatrix} = \begin{pmatrix} v & (1-q)(1-v) & q(1-v) & 0 & \cdots & 0 \\ v & (1-q)(1-v) & q(1-v) & 0 & \cdots & 0 \\ s & 0 & 1-s & 0 & \cdots & 0 \\ s & 0 & 0 & 1-s & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ s & 1-s & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{13}$$

between two consecutive successful updates. Following the same computation method as in deriving the average AoI of the direct user, $\mathbb{E}[Z]$ and $\mathbb{E}[Z^2]$ are given by

$$\mathbb{E}[Z] = \frac{q(1-v)\phi+s}{s(q(1-v)\phi+v)}, \tag{14}$$

$$\begin{aligned}\mathbb{E}[Z^2] = &\frac{q^2(1-v)^2(2-s)\phi^2}{s^2(q(1-v)\phi+v)^2} \\ &+ \frac{q(1-v)(2ks(-s+v)+(s+v)(2-s))\phi}{s^2(q(1-v)\phi+v)^2} \\ &+ \frac{-2kqs(1-v)(-s+v)+s^2(2-v)}{s^2(q(1-v)\phi+v)^2},\end{aligned} \tag{15}$$

where $\phi = 1-(1-s)^k$.

Next, we compute $\mathbb{E}[\tau Z]$. If user B receives the update packet through the direct link, the instantaneous AoI drops to $\tau = 1$, with a probability of v. If user B receives the update packet from the relay after w times of packet forwarding, the instantaneous AoI drops to $\tau = 1+w$, with a probability of $q(1-s)^{w-1}s$. Furthermore, the probability that the relay-assisted user B can successfully receive the update packet from the relay using truncated ARQ is $\sum_{w=1}^{k} q(1-s)^{w-1}s$. Therefore,

$$\Pr(\tau = 1) = \frac{v}{v+\sum_{\omega=1}^{k} q(1-v)(1-s)^{w-1}s}, \tag{16}$$

$$\Pr(\tau = 1+\omega) = \frac{q(1-v)(1-s)^{w-1}s}{v+\sum_{\omega=1}^{k} q(1-v)(1-s)^{w-1}s}. \tag{17}$$

As a result, $\mathbb{E}[\tau Z]$ can be computed by

$$\begin{aligned}&\mathbb{E}[\tau Z] \\ &= (\Pr(\tau=1) + \textstyle\sum_{\omega=1}^{k}(1+\omega)\Pr(\tau=1+\omega))M_{S_B^+S_B^+} \\ &= \frac{q(1-v)((ks+s+1)\phi-ks)+sv}{s(q(1-v)\phi+v)} M_{S_B^+S_B^+}.\end{aligned} \tag{18}$$

Finally, we substitute the components into the average AoI formula (3) to obtain the average AoI of the relay-assisted user B, $\bar{\Delta}_B$, expressed in (19). After computing the average AoI of the direct user A and the relay-assisted user B, we can obtain the average AoI of the heterogeneous status update system $\bar{\Delta} = \frac{1}{2}(\bar{\Delta}_A + \bar{\Delta}_B)$.

V. Performance Evaluation

In this section, we evaluate the average AoI of the heterogeneous broadcast system under different scenarios. We first present the packet error rate (PER) for short packets to obtain successful transmission probabilities p, q, s and v. Then, we present the average AoI of the two users and compare the system average AoI.

A. Packet Error Rate (PER) for Short-Packet Communication

In status update systems, update packets are typically short. Information theory states that the PER cannot be zero for a finite block length. To estimate the PER of short packets, we apply the Polyanskiy-Poor-Verd bound [18], and the PER of short update packets can be approximated by

$$\varepsilon \approx Q\left(\frac{(\ln 2)\sqrt{n}\left(\log_2(1+\gamma)-\frac{b}{n}\right)}{\sqrt{1-(1+\gamma)^{-2}}}\right) \triangleq Q(b,n,\gamma), \tag{20}$$

where b and n represent the number of source and coded bits of the update packet, respectively, and $\frac{b}{n}$ is the channel coding rate. $Q(\cdot)$ denotes the Q-function, i.e., $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$.

In a general block-fading channel, the average PER of short update packets, $\bar{\varepsilon}$, can be derived as

$$\bar{\varepsilon} \approx \int_0^\infty f_\gamma(x)\, Q(b,n,x)\, dx, \tag{21}$$

where $f_\gamma(x)$ denotes the probability density function (PDF) of the SNR γ. To derive a closed-form expression for $\bar{\varepsilon}$, we can use a linear approximation. The Q-function in (20) can be further expressed as

$$Q(b,n,\gamma) \approx \begin{cases} 1, & \gamma \le \delta - \frac{1}{2\beta}, \\ \frac{1}{2} - \beta(\gamma-\delta), & \delta - \frac{1}{2\beta} \le \gamma \le \delta + \frac{1}{2\beta}, \\ 0, & \gamma \ge \delta + \frac{1}{2\beta}, \end{cases} \tag{22}$$

where $\delta = 2^{b/n}-1$ and $\beta = \sqrt{n/\left(2\pi\left(2^{2b/n}-1\right)\right)}$. Substituting (22) into (21), the average PER $\bar{\varepsilon}$ can also be expressed as

$$\bar{\varepsilon} \approx \beta \int_{\delta-1/(2\beta)}^{\delta+1/(2\beta)} F_\gamma(x)\, dx, \tag{23}$$

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

$$\bar{\Delta}_B = \frac{q^2(1-v)^2(2ks+s+4)\phi^2}{2s(q(1-v)\phi+v)(q(1-v)\phi+s)} + \frac{q(1-v)(2ks(qv-q+v)+s(s+v+4)+2v)\phi}{2s(q(1-v)\phi+v)(q(1-v)\phi+s)} + \frac{-2kqsv(1-v)+s^2(v+2)}{2s(q(1-v)\phi+v)(q(1-v)\phi+s)}, \text{ where } \phi = 1-(1-s)^k. \quad (19)$$

where $F_\gamma(x)$ is the cumulative distribution function (CDF) of the SNR γ.

Under the Rayleigh fading channel, the CDF of the SNR γ can be expressed as

$$F_\gamma(x) = \begin{cases} 1-e^{-x/\bar{\gamma}}, & x \geq 0, \\ 0, & x < 0, \end{cases} \quad (24)$$

where $\bar{\gamma}$ is the average SNR. Substituting (24) into (23), we can finally express the average PER of short update packets, $\bar{\varepsilon}$, as

$$\begin{aligned} \bar{\varepsilon} &\approx \beta \int_{\delta-1/(2\beta)}^{\delta+1/(2\beta)} 1-e^{-x/\bar{\gamma}} dx \\ &= 1-\beta\bar{\gamma}\left(\mathrm{e}^{-\frac{1}{\bar{\gamma}}\left(\delta-\frac{1}{2\beta}\right)} - \mathrm{e}^{-\frac{1}{\bar{\gamma}}\left(\delta+\frac{1}{2\beta}\right)}\right). \end{aligned} \quad (25)$$

Utilizing the PER expression (25), we can estimate successful transmission probabilities p, q, s and v (i.e., one minus the corresponding PER) for different $\bar{\gamma}$ values. Then, we can substitute these into the corresponding average AoI formulas. Furthermore, the number of source bits per update packet in our simulations is 128 bits. We optimize the code rate (i.e., the number of coded bits per update packet) to obtain the optimal system average AoI under different SNRs $\bar{\gamma}$ and transmission limits k. Notice that the unit of the average AoI per user (or the system average AoI) is the number of channel uses, which is simply the multiplication of the number of coded bits n per update packet and the average AoI (or the system average AoI) in the number of time slots.

B. Simulation Results

1) SNR-Balanced Scenarios: Let $\bar{\gamma}_{SA}$, $\bar{\gamma}_{SB}$, $\bar{\gamma}_{SR}$, and $\bar{\gamma}_{RB}$ denote the average SNR of the links from the source to the direct user A, from the source to the relay-assisted user B, from the source to the relay, and from the relay to the relay-assisted user B, respectively. We first consider SNR-balanced scenarios, where $\bar{\gamma}_{SA} = \bar{\gamma}_{SR} = \bar{\gamma}_{RB} = \bar{\gamma}$, varying from $-2dB$ to $4dB$, and $\bar{\gamma}_{SB}$ is fixed at $-3dB$. Figs. 6(a) and (b) plot the optimal average AoI versus the transmission limit k for the direct user A and relay-assisted user B, respectively, across different average SNRs. Both theoretical and simulation results are presented in these figures. In the simulations, packet decoding outcomes in each time slot are captured to compute instantaneous AoI, which is then used to derive the average AoI. It is important to note that $k = 0$ corresponds to a non-relay configuration, where both users rely solely on direct links for updates from the source.

The results in Fig. 6 confirm that simulation results align with theoretical analysis, thus validating our theoretical framework detailed in Section IV. As the transmission limit k increases, Fig. 6(a) shows that the optimal average AoI for the direct user A increases due to the waiting delay introduced by the relay's ARQ transmissions. This finding aligns with previous studies on pure single-hop systems [14]. Specifically, a non-relay approach ($k = 0$) yields the lowest average AoI since the source can generate and broadcast a new packet in every time slot. However, as k increases, the direct user must wait for the relay to complete its forwarding process before receiving a new update, leading to a higher average AoI.

Fig. 6(b) shows that for the relay-assisted user B, the optimal average AoI increases with k under low SNRs (e.g., $\bar{\gamma} = -2dB$). This is because with poor channel conditions, allowing too many retransmissions at the relay increases the instantaneous AoI upon successful reception, ultimately leading to high average AoI as well. This suggests that when the channel condition of the relay-assisted link is not significantly better than the direct link, the relay-assisted user B prefers a small k, even a non-relay setup ($k = 0$) (which helps lower the instantaneous AoI upon successful reception). However, as $\bar{\gamma}$ increases, the relay-assisted user tends to benefit more from the truncated-ARQ-at-relay approach, particularly when the SNR of the relay-assisted link is significantly better than that of the direct link. In such cases, a larger k becomes preferable, as it increases the likelihood of successful packet reception during the truncated ARQ process.

The above observations under higher SNR conditions underscore the importance of the relay-assisted link's channel quality in determining the optimal transmission limit k of the relay-assisted user. In scenarios where the relay's link quality is better, additional retransmissions facilitate a more reliable communication than the direct link, thereby improving the information freshness of the relay-assisted user. In addition, this phenomenon aligns with the findings from previous studies on two-hop systems, where ARQ should be adopted in the second hop [15]. However, a large k results in a prolonged waiting time for the direct user A before the next update. Hence, this leads to different average AoI performances between the two users, as indicated in Figs. 6(a) and (b). Furthermore, we notice from Fig. 6(b) that with an increased $\bar{\gamma}$, a larger k marginally reduces the average AoI of the relay-assisted user. This is due to the low probability of successfully receiving a packet after a higher number of retransmissions. Usually, the destination can finally receive the packet from the relay after a small number of retransmissions. As a result, excessive retransmissions may not provide substantial advantages for the relay-assisted user, but affects the overall system performance, as presented below.

Fig. 6(c) presents the average AoI of the entire system. When the SNRs of the links are as low as $-2dB$ or $0dB$, the average AoI of the system increases monotonically with k, which is effectively the average result of Figs. 6(a) and (b). However, as the SNR improves, the average AoI of the system

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

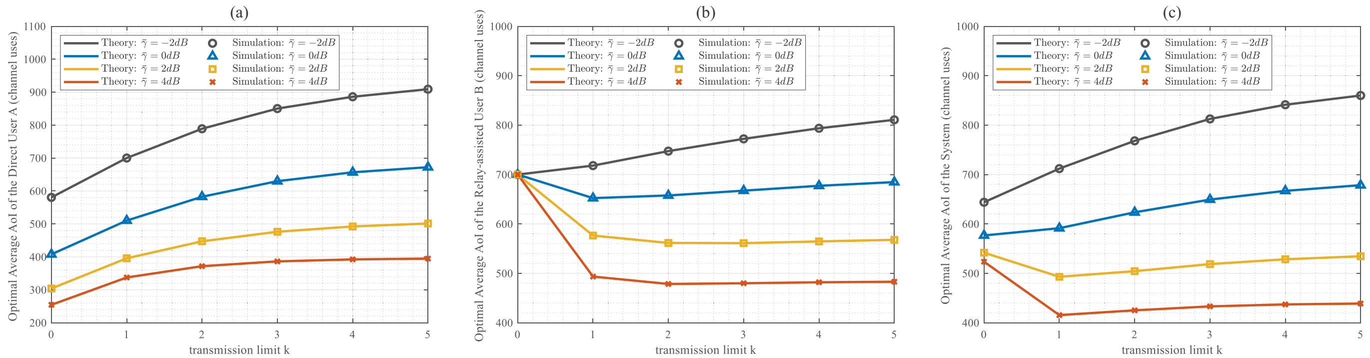


Fig. 6. The optimal average AoI versus $\bar{\gamma}$ under different transmission limit k, when $\bar{\gamma}_{SA} = \bar{\gamma}_{SR} = \bar{\gamma}_{RB} = \bar{\gamma}$, ranging from $-2dB$ to $4dB$, and $\bar{\gamma}_{SB} = -3dB$: (a) the direct user A, (b) the relay-assisted user B, and (c) the overall heterogeneous system.

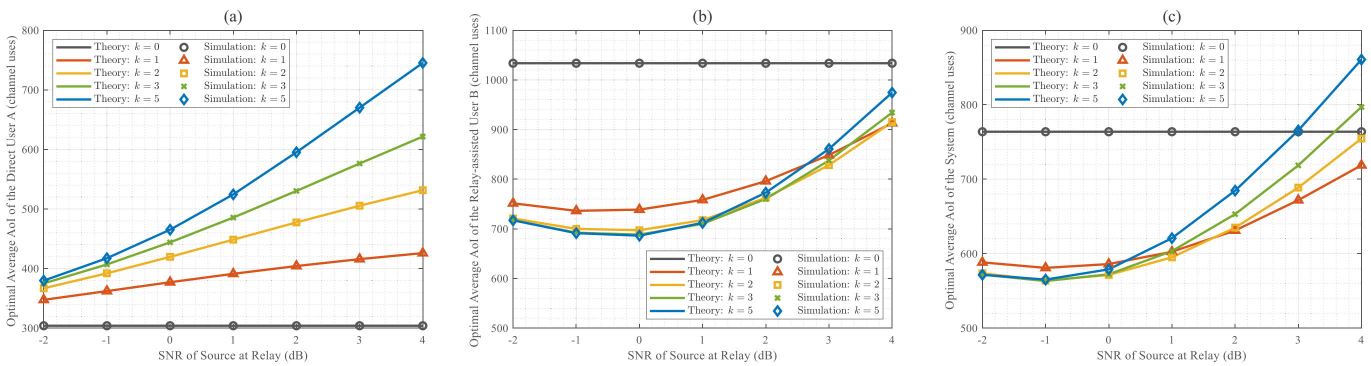


Fig. 7. The optimal average AoI versus $\bar{\gamma}_{SR}$ under different transmission limit k, when $\bar{\gamma}_{SA} = 2dB$, and $\bar{\gamma}_{SB} = -5dB$ and $\bar{\gamma}_{SR} + \bar{\gamma}_{RB} = 2dB$: (a) the direct user A, (b) the relay-assisted user B, and (c) the overall heterogeneous system.

first decreases and then increases, achieving the lowest value when $k = 1$. This indicates that $k = 1$, i.e., a non-ARQ-at-relay scheme, offers a balanced performance for the average AoI between the direct and relay-assisted users, resulting in a low and stable system-wide average AoI. Therefore, selecting a small transmission limit becomes vital to achieve better information freshness performance across heterogeneous users. This highlights the difference between heterogeneous and homogeneous systems in terms of information freshness, as the performance trade-off between direct and relay-assisted users must be carefully managed in the heterogeneous broadcasting scenarios.

2) SNR-Imbalanced Scenarios: Next, we explore SNR-imbalanced scenarios, as illustrated in Fig. 7, where $\bar{\gamma}_{SA}$, $\bar{\gamma}_{SB}$, $\bar{\gamma}_{SR}$, and $\bar{\gamma}_{RB}$ may differ. Specifically, $\bar{\gamma}_{SA}$ is set to $2dB$, and $\bar{\gamma}_{SB}$ is set to $-5dB$. The sum of $\bar{\gamma}_{SR} + \bar{\gamma}_{RB}$ is fixed at $2dB$, with $\bar{\gamma}_{SR}$ ranging from $-2dB$ to $4dB$, and $\bar{\gamma}_{RB}$ consequently ranging from $4dB$ to $-2dB$. These different SNR pairs simulate varying relay positions. Figs. 7(a), (b), and (c) plot the optimal average AoI versus $\bar{\gamma}_{SR}$ for the direct user A, the relay-assisted user B, and the overall system, respectively, under different transmission limits k. Again, the simulation results are consistent with our theoretical analysis.

For individual users, the phenomenon observed in the SNR-imbalanced scenario exhibits a pattern consistent with that in the SNR-balanced scenario shown in Fig. 6. Specifically, the direct user A prefers a non-relay and non-ARQ scheme, as indicated by the monotonic increase in its average AoI with k. For the relay-assisted user B, when γ_{SR} increases from $-2dB$ to $1dB$ (and γ_{RB} decreases from $4dB$ to $1dB$), a larger k leads to a lower average AoI due to the higher SNR at the second hop relative to the first hop — a scenario where retransmissions are advantageous, e.g., better channel conditions at the second hop increase the likelihood of successful packet delivery with retransmission. Conversely, when the first hop has a better SNR than the second hop, fewer retransmissions, such as $k = 1$, typically achieve better average AoI performance, as retransmissions in this case may introduce unnecessary delays so that the AoI performance benefits from returning back to the first hop. Despite these variations, the non-relay approach yields the highest average AoI for the relay-assisted user.

It is important to note that even though $\bar{\gamma}_{SA}$ is fixed at $2dB$ in our simulations, the average AoI of the direct user A is influenced by the decoding outcomes at the relay and the relay-assisted user B. A larger k or a smaller SNR $\bar{\gamma}_{RB}$ increases the number of retransmissions required by relay-assisted user B to successfully receive updates from the relay, which in turn increases the direct user's wait time and thus its average AoI. In addition, it is interesting to see that for the relay-assisted user B, the lowest average AoI is achieved when neither hop's SNR (either $\bar{\gamma}_{SR}$ or $\bar{\gamma}_{RB}$) is excessively low or high. In other words, the average AoI of the relay-assisted user B suffers when either hop has a poor SNR.

Given that in most cases, the relay-assisted user B typically

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

has a higher AoI due to the two-hop update process, the system average AoI, depicted in Fig. 7(c), is significantly influenced by the AoI of the relay-assisted user B. Fig. 7(c) shows that for low $\bar{\gamma}_{SR}$ (or high $\bar{\gamma}_{RB}$), $k=5$ yields the lowest system average AoI, though with marginal reduction compared to $k=1$. Conversely, for high $\bar{\gamma}_{SR}$ (or low $\bar{\gamma}_{RB}$), the lowest system average AoI is achieved with $k=1$, providing a significant improvement in system AoI performance compared to $k=5$. This difference occurs because $k=5$ significantly increases the average AoI for the direct user A, as indicated in Fig. 7(a). Overall, $k=1$ still maintains a low and stable system average AoI in the SNR-imbalanced scenarios, serving as a better strategy for heterogeneous update systems with both direct and relay-assisted users.

VI. Conclusions

In this paper, we have investigated the design of ARQ schemes for AoI-aware heterogeneous broadcast systems involving both direct and relay-assisted users. In this heterogeneous setup, a direct user receives update packets directly via a single-hop link, and a relay-assisted user receives update packets through either a single-hop direct link or a two-hop relay-assisted link. Determining the optimal ARQ strategy for this hybrid system, both for individual users and the overall system, is essential for minimizing the average AoI from various perspectives.

In our ARQ design for the heterogeneous system, managing the number of packet forwarding transmissions by the relay is crucial. We consider a general transmission limit $k \geq 0$ at the relay, where $k=0$, $k=1$, and $k>1$ correspond to non-relay, non-ARQ-at-relay, and truncated-ARQ-at-relay approaches, respectively. To account for different ARQ approaches, we model the transmission process of each user type using a unified Markov chain framework that accommodates different k, with Markov states designed to capture successful updates and the interactions of the ARQ process between the two users.

Theoretical and simulation results reveal that tailored ARQ strategies are vital for optimizing AoI performance for individual users and the overall system. Specifically, while non-relay and non-ARQ strategies are generally effective for the direct user, the truncated-ARQ-at-relay approach is advantageous for the relay-assisted user. Overall, our findings indicate that a non-ARQ-at-relay strategy tends to maintain a stable and low average AoI across the system, even under diverse channel conditions in the heterogeneous setup. These insights are significant for designing ARQ schemes in AoI-aware heterogeneous broadcast systems, providing a foundation for optimizing performance in complex environments with both direct and relay-assisted communication.

Appendix A
Detailed Equations of $M_{\alpha V_A}$ and $N_{\alpha V_A}$ in Deriving the Average AoI of the Direct User

Equations for $M_{\alpha V_A}$: According to (5), the equations of $M_{\alpha V_A}$ with different $\alpha \in Q_A$ can be expressed by

$$\begin{cases} M_{S_A^+ V_A} = 1 + \omega_{S_A^+ S_A^-} M_{S_A^- V_A} + \omega_{S_A^+ R_1^-} M_{R_1^- V_A} \\ M_{S_A^- V_A} = 1 + \omega_{S_A^- S_A^-} M_{S_A^- V_A} + \omega_{S_A^- R_1^-} M_{R_1^- V_A} \\ M_{R_1^+ V_A} = 1 + \omega_{R_1^+ S_A^-} M_{S_A^- V_A} + \omega_{R_1^+ R_2} M_{R_2 V_A} \\ M_{R_1^- V_A} = 1 + \omega_{R_1^- S_A^-} M_{S_A^- V_A} + \omega_{R_1^- R_2} M_{R_2 V_A} \\ M_{R_2 V_A} = 1 + \omega_{R_2 S_A^-} M_{S_A^- V_A} + \omega_{R_2 R_3} M_{R_3 V_A} \\ \quad\vdots \\ M_{R_{k-1} V_A} = 1 + \omega_{R_{k-1} S_A^-} M_{S_A^- V_A} + \omega_{R_{k-1} R_k} M_{R_k V_A} \\ M_{R_k V_A} = 1 + \omega_{R_k S_A^-} M_{S_A^- V_A} \end{cases}. \tag{26}$$

Equations for $N_{\alpha V_A}$: According to (6), the equations of $N_{\alpha V_A}$ with different $\alpha \in Q_A$ can be expressed by

$$\begin{cases} N_{S_A^+ V_A} &= 1 + \omega_{S_A^+ S_A^-}(N_{S_A^- V_A} + 2M_{S_A^- V_A}) \\ &\quad + \omega_{S_A^+ R_1^-}(N_{R_1^- V_A} + 2M_{R_1^- V_A}) \\ N_{S_A^- V_A} &= 1 + \omega_{S_A^- S_A^-}(N_{S_A^- V_A} + 2M_{S_A^- V_A}) \\ &\quad + \omega_{S_A^- R_1^-}(N_{R_1^- V_A} + 2M_{R_1^- V_A}) \\ N_{R_1^+ V_A} &= 1 + \omega_{R_1^+ S_A^-}(N_{S_A^- V_A} + 2M_{S_A^- V_A}) \\ &\quad + \omega_{R_1^+ R_2}(N_{R_2 V_A} + 2M_{R_2 V_A}) \\ N_{R_1^- V_A} &= 1 + \omega_{R_1^- S_A^-}(N_{S_A^- V_A} + 2M_{S_A^- V_A}) \\ &\quad + \omega_{R_1^- R_2}(N_{R_2 V_A} + 2M_{R_2 V_A}) \\ N_{R_2 V_A} &= 1 + \omega_{R_2 S_A^-}(N_{S_A^- V_A} + 2M_{S_A^- V_A}) \\ &\quad + \omega_{R_2 R_3}(N_{R_3 V_A} + 2M_{R_3 V_A}) \\ &\quad\vdots \\ N_{R_k V_A} &= 1 + \omega_{R_k S_A^-}(N_{S_A^- V_A} + 2M_{S_A^- V_A}) \end{cases}. \tag{27}$$

Appendix B
Detailed Computation of $M_{\alpha V_A}$ in Deriving the Average AoI of the Direct User

This appendix explains the detailed computation of $M_{\alpha V_A}$ in deriving the average AoI of the direct user (specifically, the $\mathbb{E}[Z]$ term), where $V_A = \{S_A^+, R_1^+\}$. We obtain $M_{S_A^+ V_A}$ and $M_{R_1^+ V_A}$ by solving the equations of $M_{\alpha V_A}$ (26). Note that the computation of $N_{\alpha V_A}$ (when computing the $\mathbb{E}[Z^2]$ term) follows the same the method by solving the equations of $N_{\alpha V_A}$ with different $\alpha \in Q_A$. Hence, we omit the detailed computation of $N_{S_A^+ V_A}$ and $N_{R_1^+ V_A}$ for clarity.

In the following, we use $M_{S_A^+ V_A}$ as an example. We first focus on two special cases with $k=0$ and $k=1$, followed by the general case of $k>1$.

- Case: $k=0$ (the non-relay case)
  When $k=0$, both users have only the direct links. Hence, (26) is simplifies to

$$\begin{cases} M_{S_A^+ V_A} = 1 + \omega_{S_A^+ S_A^-} M_{S_A^- V_A} \\ M_{S_A^- V_A} = 1 + \omega_{S_A^- S_A^-} M_{S_A^- V_A} \end{cases}. \tag{28}$$

When $k = 0$, we have $\omega_{S_A^+ S_A^-} = \omega_{S_A^- S_A^-} = 1 - p$. Thus, we can compute $M_{S_A^+ V_A}$ by

$$M_{S_A^+ V_A} = \frac{1}{p}. \tag{29}$$

- Case: $k = 1$ (the non-ARQ-at-relay case)
When $k = 1$, the relay forwards the packet to the destination once without retransmission. (26) is simplifies to

$$\begin{cases} M_{S_A^+ V_A} = 1 + \omega_{S_A^+ S_A^-} M_{S_A^- V_A} + \omega_{S_A^+ R_1^-} M_{R_1^- V_A} \\ M_{S_A^- V_A} = 1 + \omega_{S_A^- S_A^-} M_{S_A^- V_A} + \omega_{S_A^- R_1^-} M_{R_1^- V_A} \\ M_{R_1^+ V_A} = 1 + \omega_{R_1^+ S_A^-} M_{S_A^- V_A} \\ M_{R_1^- V_A} = 1 + \omega_{R_1^- S_A^-} M_{S_A^- V_A} \end{cases}. \tag{30}$$

When $k = 1$, the state transition matrix for user A is given by (31). Solving the equations and substituting the corresponding state transition probabilities, we obtain

$$M_{S_A^+ V_A} = \frac{1 + q(1-p)(1-v)}{p}. \tag{34}$$

- Case: $k > 1$ (the truncated-ARQ-at-relay case)
First, in the first two equations of (26), since $\omega_{S_A^+ S_A^-} = \omega_{S_A^- S_A^-}$ and $\omega_{S_A^+ R_1^-} = \omega_{S_A^- R_1^-}$, we observe that $M_{S_A^+ V_A} = M_{S_A^- V_A}$. As a result, $M_{S_A^+ V_A}$ depends solely on $M_{R_1^- V_A}$. Then, by substituting $M_{R_k V_A}$ (the last equation of (26)) into $M_{R_{k-1} V_A}$ (the second to last equation of (26)), we obtain (32). Next, we substitute $M_{R_{k-1} V_A}$ into $M_{R_{k-2} V_A}$, $M_{R_{k-2} V_A}$ into $M_{R_{k-3} V_A}$, ..., $M_{R_2 V_A}$ into $M_{R_1^- V_A}$, $M_{R_1^- V_A}$ into $M_{S_A^+ V_A}$. By doing so, we progressively eliminate all relay states $M_{R_k V_A}, M_{R_{k-1} V_A}, \ldots, M_{R_2 V_A}, M_{R_1^- V_A}$, and ultimately obtain a closed-form solution for $M_{S_A^+ V_A}$, i.e., see (33). Finally, by substituting the corresponding state transition probabilities into (33), we obtain the final result of $M_{S_A^+ V_A}$

$$M_{S_A^+ V_A} = \frac{1 + \sum_{n=1}^{k} q(1-p)(1-v)(1-s)^{n-1}}{p}. \tag{35}$$

Moreover, it is worth noting that the above formula also holds for the special cases of $k = 0$ and $k = 1$.

References

[1] Y. Wang, S. Wu, C. Lei, J. Jiao, and Q. Zhang, "A review on wireless networked control system: The communication perspective," *IEEE Internet Things J.*, vol. 11, no. 5, pp. 7499-7524, Mar. 2024.

[2] M. Xie, B. An, X. Jia, M. Zhou, and J. Lu, "Simultaneous update of sensing and control data using free-ride codes in vehicular networks: An age and energy perspective," *Comput. Netw.*, vol. 252, Jun. 2024.

[3] Y. Liu, X. Wang, G. Zheng, X. Wan, and Z. Ning, "An AoI-aware data transmission algorithm in blockchain-based intelligent healthcare Systems," *IEEE Trans. Consum. Electr.*, vol. 70, no. 1, pp. 1180-1190, Feb. 2024.

[4] H. Wang, Y. Wang, X. Xie, and M. Li, "A scheduling scheme for minimizing age under delay tolerance in IoT systems with heterogeneous traffic," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 16902-16914, May 2024.

[5] Y. Sun, I. Kadota, R. Talak, and E. Modiano, Age of Information: A New Metric for Information Freshness. San Rafael, CA, USA: Morgan & Claypool, 2019.

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

[7] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Proc. Conf. Inf. Sci. Syst. (CISS)*, pp. 1-6, Mar. 2012.

[8] Y. Liu, L. X. Cai, Q. Chen, H. Zhang, F. Hou, and T. H. Luan, "Minimizing age of information in nonorthogonal random access networks," *IEEE Internet Things J.*, vol. 11, no. 14, pp. 24886-24902, Jul. 2024.

[9] J. Lou, X. Yuan, S. Kompella, and N.-F. Tzeng, "AoI and throughput tradeoffs in routing-aware multi-hop wireless networks," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, 2020, pp. 476-4851.

[10] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," in *Proc. MobiHoc*, 2018, pp. 6170.

[11] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. INFOCOM*, 2018, pp. 18441852.

[12] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. IEEE ISIT*, 2015, pp. 30083012.

[13] N. Akar and E. Karasan, "Is proportional fair scheduling suitable for age-sensitive traffic?," *Comput. Netw.*, vol. 226, 2023.

[14] 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. 1038810400, Oct. 2020.

[15] J. Feng, H. Pan, T.-T. Chan, and J. Liang, "Timely status update: Should ARQ be used in two-hop networks?," in *Proc. IEEE ICC*, 2022, pp. 35143519.

[16] H. Pan, T.-T. Chan, V. C. M. Leung, and J. Li, "Age of information in physical-layer network coding enabled two-way relay networks," *IEEE Trans. Mob. Comput.*, vol. 22, no. 8, pp. 44854499, Aug. 2023.

[17] H. Wang, X. Xie, X. Li, and J. Yang, "Scheduling schemes for age optimization in IoT systems with limited retransmission times," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21458-21468, Nov. 2022.

[18] 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. 23072359, May 2010.

[19] C. Kam, J. P. Molnar, and S. Kompella, "Age of information for queues in tandem," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, pp. 1-6, Oct. 2018.

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

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

[22] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *Proc. IEEE ISIT*, pp. 66-70, Jul. 2013.

[23] B. Yu and Y. Cai, "Age of information in grant-free random access with massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 10, no. 7, pp. 1429-1433, Jul. 2021.

[24] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *Proc. 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, 2017, pp. 486-493.

[25] H. Pan, J. Feng, T.-T. Chan, V. C. M. Leung, and J. Li, "Timely status update in relay-assisted cooperative communications," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 1574515761, Jul. 2023.

[26] J. Zhang and Y. Xu, "Age of information in relay-assisted status updating systems," *arXiv preprint arXiv:2107.01833*, 2021.

[27] M. Moradian and A. Dadlani, "Age of information in scheduled wireless relay networks," *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, pp. 1-6, May 2020.

[28] M. Xie, J. Gong, X. Jia, and X. Ma, "Age and energy tradeoff for multicast networks with short packet transmissions," *IEEE Trans. Commun.*, vol. 69, no. 9, pp. 6106-6119, Sep. 2021.

[29] A. Arafa and S. Ulukus, "Timely updates in energy harvesting two-hop networks: Offline and online policies," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4017-4030, Aug. 2019.

[30] J. P. Champati, H. Al-Zubaidy, and J. Gross, "Statistical guarantee optimization for AoI in single-hop and two-hop FCFS systems with periodic arrivals," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 365-381, Jan. 2021.

[31] T.-T. Chan and H. Pan, "Enhancing information freshness for coordinated direct and relay transmission," *Journal of Communications and Networks*, vol. 25, no. 3, pp. 312-321, Jun. 2023.

[32] M. Xie, J. Gong, S. Cai, and X. Ma, "Age-energy tradeoff for two-hop status update systems with heterogeneous truncated ARQ," *IEEE Wireless Commun. Lett.*, vol. 10, no. 7, pp. 1488-1492, Jul. 2021.

This article has been accepted for publication in IEEE Internet of Things Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2025.3568170

$$\Omega_A = \begin{pmatrix} \omega_{S_A^+S_A^+} & \omega_{S_A^+S_A^-} & \omega_{S_A^+R_1^+} & \omega_{S_A^+R_1^-} \\ \omega_{S_A^-S_A^+} & \omega_{S_A^-S_A^-} & \omega_{S_A^-R_1^+} & \omega_{S_A^-R_1^-} \\ \omega_{R_1^+S_A^+} & \omega_{R_1^+S_A^-} & \omega_{R_1^+R_1^+} & \omega_{R_1^+R_1^-} \\ \omega_{R_1^-S_A^+} & \omega_{R_1^-S_A^-} & \omega_{R_1^-R_1^+} & \omega_{R_1^-R_1^-} \end{pmatrix} = \begin{pmatrix} p(1-q(1-v)) & (1-p)(1-q(1-v)) & pq(1-v) & q(1-p)(1-v) \\ p(1-q(1-v)) & (1-p)(1-q(1-v)) & pq(1-v) & q(1-p)(1-v) \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{31}$$

$$\begin{aligned} M_{R_{k-1}V_A} &= 1 + \omega_{R_{k-1}S_A^-} M_{S_A^-V_A} + \omega_{R_{k-1}R_k} M_{R_kV_A} \\ &= 1 + \omega_{R_{k-1}S_A^-} M_{S_A^-V_A} + \omega_{R_{k-1}R_k}(1 + \omega_{R_kS_A^-} M_{S_A^-V_A}) \\ &= 1 + \omega_{R_{k-1}R_k} + (\omega_{R_{k-1}S_A^-} + \omega_{R_{k-1}R_k}\omega_{R_kS_A^-})M_{S_A^-V_A}. \end{aligned} \tag{32}$$

$$M_{S_A^+V_A} = \frac{1 + \omega_{S_A^+R_1^-} + \omega_{S_A^+R_1^-}\omega_{R_1^-R_2}\left(1 + \sum_{i=2}^{k-1}\prod_{j=2}^{i}\omega_{R_jR_{j+1}}\right)}{1 - \omega_{S_A^+S_A^-} - \omega_{S_A^+R_1^-}\omega_{R_1^-S_A^-} - \omega_{S_A^+R_1^-}\omega_{R_1^-R_2}\left(\sum_{i=2}^{k}\prod_{j=2}^{i-1}\omega_{R_jR_{j+1}}\omega_{R_iS_A^-}\right)}. \tag{33}$$

[33] Y. Chen, Y. Cai, G. Ding, B. Yu, and C. Xu, "Age of information for short-packet relay communications in cognitive-radio-based Internet of Things with outdated channel state information," *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 3, pp. 722-737, Jun. 2023.

[34] 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.

[35] J. Huang, J. Gong, and W. Zhan, "Energy-age tradeoff in status update communication systems based on HARQ," in *Proc. 13th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, pp. 1-5, Oct. 2021.

[36] S. W. H. Shah, M. M. U. Rahman, A. N. Mian, O. A. Dobre and J. Crowcroft, "Effective capacity analysis of HARQ-enabled D2D communication in multi-tier cellular networks", *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 9144-9159, Sep. 2021.

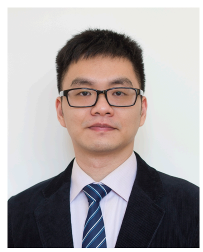
[37] S. W. H. Shah, A. N. Mian, S. Mumtaz, A. Al-Dulaimi, I. Chih-Lin and J. Crowcroft, "Statistical QoS analysis of reconfigurable intelligent surface-assisted D2D communication", *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7343-7358, Jul. 2022.

[38] M. Samir, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Online altitude control and scheduling policy for minimizing AoI in UAV-assisted IoT wireless networks," *IEEE Trans. Mobile Comput.*, vol. 21, no. 7, pp. 24932505, Jul. 2022.

[39] C. M. W. Basnayaka, D. N. K. Jayakody, and Z. Chang, "Ageof-information-based URLLC-enabled UAV wireless communications system," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 1021210223, Jun. 2022.

[40] Y. Inoue and T. Kimura, "Age-effective information updating over intermittently connected MANETs," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 14271440, May 2021.

[41] Z. Fang, J. Wang, C. Jiang, Q. Zhang, and Y. Ren, "AoI-inspired collaborative information collection for AUV-assisted Internet of Underwater Things," *IEEE Internet Things J.*, vol. 8, no. 19, pp. 1455914571, Oct. 2021.



Haoyuan Pan (Member, IEEE) received the B.E. and Ph.D. degrees in Information Engineering from The Chinese University of Hong Kong (CUHK), Hong Kong, in 2014 and 2018, respectively. He was a Post-Doctoral Fellow with the Department of Information Engineering, CUHK, from 2018 to 2020. Since 2020, he has been with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, where he is currently an Associate Professor. He received the Best Paper Award at IEEE Wireless Communications and Networking Conference (WCNC) 2024. His research interests include wireless communications and networks, Internet of Things (IoT), semantic communications, and age of information (AoI).

Tse-Tin Chan (Member, IEEE) received his B.Eng. (First Class Hons.) and Ph.D. degrees in Information Engineering from The Chinese University of Hong Kong (CUHK), Hong Kong SAR, China, in 2014 and 2020, respectively. He is currently an Assistant Professor with the Department of Mathematics and Information Technology, The Education University of Hong Kong (EdUHK), Hong Kong SAR, China. Prior to this, he was an Assistant Professor with the Department of Computer Science, The Hang Seng University of Hong Kong (HSUHK), Hong Kong SAR, China, from 2020 to 2022. His research interests include wireless communications and networking, Internet of Things (IoT), age of information (AoI), and AI in wireless communications.

Xueer Wang (Student Member, IEEE) received the B.S. degree from Fujian Normal University, Fuzhou, China, in 2023. She is currently working toward the M.E. degree in Shenzhen University, Shenzhen, China. Her research interests include age of information (AoI), and Internet of Things (IoT).

Jianqiang Li (Senior Member, IEEE) received the B.S. and Ph.D. degrees from the South China University of Technology, Guangzhou, China, in 2003 and 2008, respectively. He is a professor with the College of Computer and Software Engineering, Shenzhen University, Shenzhen, China. He is the recipient of the National Science Fund for Distinguished Young Scholars. He led five projects of the National Natural Science Foundation and three projects of the Natural Science Foundation of Guangdong Province, China. His major research interests include robotics, hybrid systems, Internet of Things, and embedded systems.