Unequal Timeliness Protection Massive Access for Mission Critical Communications in S-IoT

Shiying Su, Jian Jiao, *Senior Member, IEEE*, Tao Yang, *Student Member, IEEE*, Liang Xu, Ye Wang, *Member, IEEE*, and Qinyu Zhang, *Senior Member, IEEE*

***Abstract*— In this paper, we propose three unequal timeliness (UT) protection massive access (UTMA) schemes in satellite-based Internet of Things (S-IoT) for mission critical communications (MCC) user equipments (UEs) with three types of timeliness requirements: independent successive UTMA (IS-UTMA), extended cognitive offloading UTMA (ECO-UTMA), and independent cognitive offloading UTMA (ICO-UTMA). First, MCC UEs are grouped according to their timeliness requirements, and a multi-dimensional codebook is introduced to resolve the UE collisions in massive access. Then, the IS-UTMA exclusively allocates time slots and pilots to different MCC UE groups to perform massive access, while the ECO- and ICO-UTMA allow timeliness critical group to share resources with timeliness tolerant group to improve the system timeliness. To capture the timeliness evaluation of each MCC UE group, we utilize age of information (AoI) to model the information freshness and derive closed-form expressions of average AoI (AAoI) by tracing the access failure probability (AFP) and instantaneous AoI. Furthermore, we establish the parameter optimization problems to minimize AAoI under desired AFP requirements. Extensive simulations validate the accurate of theoretical derivations, and demonstrate the effectiveness of the proposed UTMA scheme with joint optimized parameters, which can achieve minimum AAoI under desired AFP than the state-of-the-art schemes.**

***Index Terms*— Satellite-based Internet of Things, unequal timeliness protection, mission critical communications, age of information, grant free random access.**

I. INTRODUCTION

RECENTLY, the fifth-generation (5G) communications have been deployed in various fields such as industrial automation [1], smart agriculture, and remote healthcare [2], becoming an important cornerstone for providing timeliness massive access of status updates in these applications. However, due to rural environment and extreme weather conditions, the existing terrestrial networks cannot support these timeliness applications in a cost-effective manner [3]. Thanks to the extensive coverage of satellites, satellite-based Internet of Things (S-IoT) is one of the key directions for achieving ubiquitous intelligent connections in the next generation of mobile networks [4], and several mega constellations of low earth orbit (LEO) high-throughput satellites (HTS) are planned and begin to launch, such as Starlink and OneWeb [5], [6].

These upcoming S-IoT are expected to support mission critical communications (MCC) with an end-to-end delay can be as low as 30 ms [7], since the worst two-way propagation delay is expected to be 26 ms for LEO at 600 km [8]. However, massive MCC user equipments (UEs) in different applications exhibit various timeliness requirements for status updates. Recently, 3rd Generation Partnership Project (3GPP) Release 17 has discussed the access failure probability (AFP) and packet delay budget (PDB) requirements of the different MCC services for future non-terrestrial networks (NTN) [8]. Specifically, three types of timeliness requirements are considered in this paper: 1) timeliness stringent MCC UEs (TSUs) in the Intelligent Transport Systems (ITS), where PDB $\leq$ 30 ms and AFP $\leq 10^{-5}$ [8]; 2) timeliness critical MCC UEs (TCUs), such as the Mission Critical user plane Push To Talk voice [9], requires PDB of 60 ms and the AFP as low as $10^{-6} \sim 10^{-5}$ [8]; and 3) timeliness tolerant MCC UEs (TTUs), where the Mission Critical Data needs PDB 200 ms and the AFP 10^{-6} [8], [10]. Moreover, the exponential growth of MCC UEs due to the emerging timeliness applications leads to severe UE collisions in massive access. Hence, designing a unequal timeliness (UT) protection massive access (UTMA) scheme to simultaneously satisfy various timeliness requirement, and resolve UE collisions in S-IoT holds significant practical value.

In addition, maximizing the throughput or minimizing delay cannot fully guarantee the timeliness requirements, because the delay only captures the latency from the transmission of a status update to its successful decoding at the satellite [11]. Thus, a related indicator named age of information (AoI) is proposed to study the timeliness of massive access at satellite [12]. The AoI is defined as the elapsed time since the latest resolved status update generated by UE [13], and the satellite calculates the average AoI (AAoI) by fairly measuring the AoI of each type of MCC UEs to get the timeliness of them. Then, we can design different UTMA schemes for

Manuscript received 15 October 2023; revised 14 December 2023; accepted 16 January 2024. Date of publication 23 January 2024; date of current version 18 June 2024. This work was supported in part by the National Natural Sciences Foundation of China (NSFC) under Grant 62071141, Grant 62027802, and Grant 61831008, in part by the Shenzhen Science and Technology Program under Grant JSGG20220831110801003 and Grant GXWD20231127123203001, and in part by the Major Key Project of PCL Department of Broadband Communication. The associate editor coordinating the review of this article and approving it for publication was E. K. S. Au. *(Corresponding authors: Jian Jiao; Ye Wang.)*

Shiying Su, Jian Jiao, Tao Yang, Liang Xu, 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 Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: 21s152079@stu.hit.edu.cn; jiaojian@hit.edu.cn; yangtao_hit@foxmail.com; xuliang_hit@foxmail.com; zqy@hit.edu.cn).

Ye Wang is with the Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: wangy02@pcl.ac.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCOMM.2024.3357627.

Digital Object Identifier 10.1109/TCOMM.2024.3357627

each type of MCC UEs to satisfy their different timeliness requirements.

A. Related Works

Considering the hundreds of kilometers between HTS and UEs, the conventional grant-based (GB) random access will result in huge communication overhead due to the propagation delay, which cannot guarantee the timeliness [14]. The grant-free (GF) random access can reduce the propagation delay and signaling overhead caused by conventional GB random access schemes. Therefore, GF random access is an inevitable choice for massive access in S-IoT [15]. However, GF random access would lead to serious collisions, because the UEs randomly choose pilot sequences and perform access [16], if two or more UEs select the same pilot, the collision occurs.

The authors in [17] and [18] studied non-orthogonal multiple access (NOMA) for random access in satellite communications, where NOMA provides a degree of freedom in the power domain by transmitting superimposed signals in the same time-frequency resource block, which can resolve UE collisions. However, the power domain NOMA cannot extend to more than three UEs in each group under the desired AFP in practical, which limits the usage in massive access of S-IoT.

In addition, several works to resolve the UE collisions are proposed in [12], [19], [20], [21], and [22] from a time domain perspective. The authors in [19] propose a random access scheme with quality-of-service (QoS) guarantees, in which UEs are grouped based on their QoS requirements, and multi-slot access frames are divided into multiple time slots to resolve UE collisions and satisfy the QoS requirements. The authors in [20] propose a random access scheme based on the irregular repeat-slot ALOHA (IRSA) protocol, where different priority UE groups have independent transmission slots. However, the above two schemes do not consider performance optimization for low-priority UE groups. A multi-slot pilot allocation (MSPA) random access scheme with unequal access latency (UAL) protection is proposed in [21] and [22], where the UEs are divided into groups based on their UAL requirements under certain AFP. With a precise design to allocate the number of access slots and pilots, the MSPA scheme can guarantee the QoS of high priority UE groups while enhancing the performance of lower priority UE groups. Further, a grant free age-optimal (GFAO) random access protocol is proposed in [12] for S-IoT, where the GFAO protocol can achieve the lowest AoI under certain AFP by adjusting the number of access slots in each frame. However, the timeliness is sacrificed for reliability in the above mentioned time domain multiply slots random access schemes, and they are difficult to meet the extremely timeliness requirements in MCC scenarios.

Recently, Polyanskiy et al. propose a T-fold code domain random access scheme to resolve UE collisions in unsourced random access [23], [24], [25]. The T-fold scheme relies on the concatenated codes, where the outer code can recover the collided messages, and the inner code is used for error correction. The inner code enables the base station (BS) to decode the modulo-2 sum of all codewords transmitted in the same time slot. When the number of collided UEs $\leq T$, the outer code can recover each message from the modulo-2 sum with zero error probability. However, the complexity of codebook in T-fold scheme would significantly increase with T, which limits the implementation in massive access. The authors in [26] proposed a LT-collision resolution GF random access (LT-GFRA) scheme for massive access, where a pilot set containing L orthogonal pilots combines a T-order codebook to resolve up to LT UE collisions, which shows the potential to achieve stringent timeliness for massive access in S-IoT.

B. Contributions

In this paper, we design an UTMA scheme for MCC UEs in S-IoT, where the MCC UEs are grouped according to their different timeliness requirements. Then, we introduce a multi-dimensional codebook and orthogonal pilot set to resolve UE collisions, and design different UTMA schemes for each type of MCC UEs, which can guarantee the desired AFP and timeliness requirement simultaneously. The main contributions in this paper are concluded as follows.

- *Unequal Timeliness Protection Massive Access Scheme:* To the best of our knowledge, this is the first work on designing an UTMA scheme for MCC UEs in S-IoT. In our UTMA scheme, we group the TSUs, TCUs and TTUs into three priority groups according to their timeliness requirements, and each group is allocated different access stage and pilots, where each pilot corresponds to a T-order codebook. Specifically, we design three types of UTMA schemes: 1) Independent successive UTMA (IS-UTMA), where each group performs random access in successive time slot according to its priority; 2) Extended cognitive offloading UTMA (ECO-UTMA), where part of TTUs can offload in the TCUs access stage; and 3) Independent cognitive offloading UTMA (ICO-UTMA), where part of TTUs can offload in the TCUs access stage, and pilot set for both types of UEs are pre-allocated in advance.
- *Analysis for UTMA Scheme:* Through the analysis of UE collisions and decoding failure probability under the T-order codebook and L orthogonal pilots, we derive the closed-form expressions of AFP of each group under the shadowed-Rician fading channel for three UTMA schemes. Then, with the help of theoretical results of AFP, we model the timeliness of each group via AoI and track the instantaneous AoI evolution process to derive the closed-form expression of the AAoI. Monte Carlo simulation results validate the accuracy of the theoretical derivations.
- *Optimization and Validation of UTMA Scheme:* To meet the timeliness requirements of TCUs group and improve the AAoI of TTUs group under the desired AFP, we construct the joint parameter optimization problems of AAoI for the ECO- and ICO-UTMA schemes under diverse AFP constraints, number of pilots and number of UEs. To solve the non-convex joint parameter optimization problems, we decompose each optimization problem into two simplified sub-problems to obtain the optimal parameters for the ECO- and ICO-UTMA schemes. Simulation

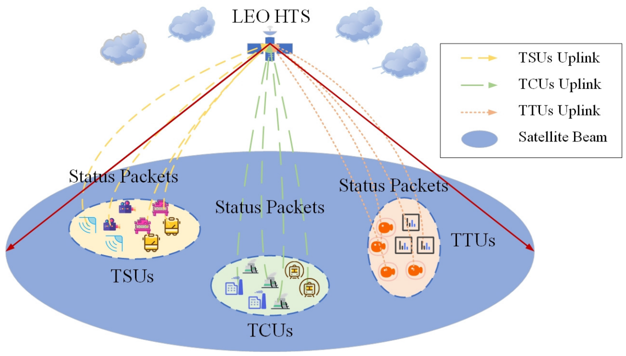


Fig. 1. The system model of unequal timeliness protection massive access for MCC UEs under a LEO HTS.

results show that the proposed UTMA schemes can provide UT protection for each group by resolve UE collisions in code domain, and exhibit significantly lower AAoI than existing scheme.

The remainder of this paper is organized as follows. Section II shows the system model and the principles of UTMA scheme. Section III describes the AoI and AFP analysis for the UTMA scheme. The performance optimization for UTMA scheme is depicted in section IV. Section V gives the simulation results and Section VI exhibits the conclusion.

II. System Model and UTMA Scheme

In this section, we first present the system model of the proposed UTMA scheme. Then, we concretely describe the three UTMA schemes, including the procedure of UTMA random access scheme.

A. System Model

As shown in Fig. 1, we consider a LEO S-IoT system with wide coverage of K MCC UEs, where the orbit type is circular orbiting around the Earth in a Walker Delta constellation [5], and a 0.5 m electronic phased array antenna can be utilized in Ka-band uplink communications for UEs [27]. Considering that the altitude of LEO HTS is set as 600 km with antenna gain no less than 50 dBi [27], the worst two-way propagation delay is expected to be 26 ms [8]. Without loss of generality, the covered MCC UEs are grouped into three priority UE groups according to their access timeliness requirements: TSUs, TCUs and TTUs, where the PDB of TSUs, TCUs and TTUs are set as 30 ms, 60 ms and 200 ms, respectively [8], [9], [10]. The access frame is divided into three access stages denoted as N_1, N_2, and N_3, and each access stage equals to one time slot. Assume that the length of time slot equals to the duration from an activated UE generating a status update packet to the HTS decoding the packet. Thus, due to the PDB requirement of each type MCC UEs, the length of one access frame and one time slot are set as 90 ms and 30 ms, respectively [8], [9], [10]. Moreover, the number of activated UEs in each group is denoted as K_{TSUs}, K_{TCUs} and K_{TTUs} correspondingly.

Considering the multipath and shadow fading effects caused by the surrounding environment, such as buildings, trees, and terrain [28], we utilize the widely-used shadowed-Rician fading channel to model the satellite-to-ground link [4], [29], [30], [31], which aligns closely with the observed land mobile satellite (LMS) channel data and provides substantial analytical and numerical advantages for predicting system performance. Moreover, the terrestrial MCC UEs are quasi-static and the Doppler effects in our system can be modeled as a constant multiplicator factor [32], thus the LEO HTS can utilize a guard band that double than the Doppler shifts to relieve the influence on our system [14]. The probability density function (PDF) of channel power gain $r = |h|^2$ is as follows [29],

$$f(r) = \left(\frac{2bm}{2bm+\Omega}\right)^m \frac{1}{2b} \exp\left(-\frac{r}{2b}\right) \cdot {}_1F_1\left(m, 1, \frac{\Omega r}{2b(2bm+\Omega)}\right), \tag{1}$$

where b is the average power of multi-path components, m is Nakagami-m parameter, Ω is the average power of line of sight (LoS), and ${}_1F_1(a, b, c)$ is confluent hypergeometric function.

The Zadoff-Chu (ZC) sequence has excellent auto-correction and cross-correlation properties along with its low peak-to-average power ratio [33], which leads to an outstanding detection capability in NTN [34], [35]. Thus, we utilize ZC sequence as pilot sequence to estimate channel gain and synchronize. We denote a pilot set of L orthogonal pilot sequences as $\mathbf{D} = \{\xi_1, \xi_2, \xi_3, \ldots, \xi_L\}$ in our system, and the activated UE l can uniformly random select a pilot ξ_j from the allocated pilots. Then, the activated UE l encodes its status update by a T-order codebook. The design of T-order concatenated code is from [36], where the Bose-Chaudhuri-Hocquenghem (BCH) outer code can recover up to T UE collisions on any pilot, and the low density parity check (LDPC) inner code is used for channel error correction. Thus, the HTS can recover up to LT UEs in one time slot [26].

Therefore, assuming that $\mathbf{Z}_j$ is the set of UEs who select ξ_j in a time slot, we can divide the pilots into four cases as follows:

1) Singleton UE pilot: $|\mathbf{Z}_j| = 1$;

2) Fully decodable collision pilot: $2 \leq |\mathbf{Z}_j| \leq T$;

3) Partially decodable collision pilot: $|\mathbf{Z}_j| > T$;

4) Empty pilot: $|\mathbf{Z}_j| = 0$, which can be perfectly detected by the HTS and is ignored in the following.

Let X_l denote the status update of UE l selected ξ_j and encoded by the T-order codebook, Y_d denotes the status update packet received by HTS, and we have:

$$Y_d = \sum_{j=1}^{L} \sum_{l \in \mathbf{Z}_j} \sqrt{P_T} h_l \left(X_l \otimes \xi_j\right) + N, \tag{2}$$

where P_T is the transmission power, h_l is the channel gain, $\otimes$ denotes the Kronecker product and $N \sim \mathcal{CN}(0, \sigma^2)$ represents the additive white Gaussian noise (AWGN). Then, the HTS performs pilot detection on Y_d, and the status updates over pilot ξ_j is

$$Y_d^j = \sum_{l \in \mathbf{Z}_j} \sqrt{P_T} h_l X_l + N_d, \tag{3}$$

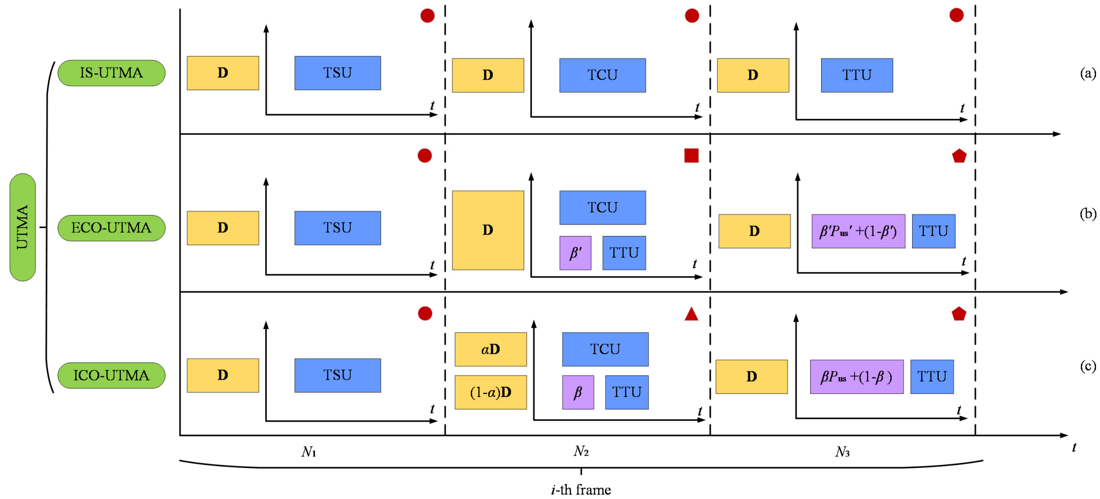


Fig. 2. Resource scheduling logic diagram of three UTMA schemes, where $P_{\text{us}}{}'$ and P_{us} are the AFP of offloaded TTUs in N_2 for ECO- and ICO-UTMA scheme, respectively.

and the channel estimation for ξ_j via the least-squares (LS) algorithm is

$$\hat{h}_{\xi_j} = \sum_{l \in \mathbf{Z}_j} h_l + \hat{N}. \tag{4}$$

Then, the HTS first performs successive cancellation (SC) to recover the encoded status update of the strongest UE in the Y_d^j. Hence, for $v < T$, in the v-th iteration of SC, the signal to interference plus noise ratio (SINR) is:

$$\text{SINR}_{v,\text{SC}} = \frac{P_T |h_v|^2}{\sum_{i=1}^{v-1} \delta P_T |h_i|^2 + \sum_{i=v+1}^{|\mathbf{Z}_j|} P_T |h_i|^2 + \sigma^2}, \tag{5}$$

where δ is the residual interference power coefficient. If the v-th SC is successful, the encoded status update of this UE will be subtracted from Y_d^j and the HTS will perform decoding to recover the corresponding packet. Otherwise, the HTS will perform joint decoding (JD) on the remaining signal, and the corresponding SINR is:

$$\text{SINR}_{v,\text{JD}} = \frac{\sum_{i=v}^{|\mathbf{Z}_j|} P_T |h_i|^2}{\sum_{i=1}^{v-1} \delta P_T |h_i|^2 + \sigma^2}. \tag{6}$$

Note that the encoding of the T-order codebook only consists of mapping a message to a primitive element α in $GF(2^m)$, and then computing its odd power $(\alpha, \alpha^3, \ldots, \alpha^{2T-1})$, which requires $\mathcal{O}(T^2)$ multiplications [25]. Further, the decoding of the received status updates is similar to standard Gorenstein Peterson Zierler (GPZ) decoding of the BCH code, and the decoding complexity of the SC then JD procedure is $\mathcal{O}(T + \rho k T \log^2 T \log\log T)$ [R35], where ρ denotes the normalized stopping set proportion of the SC decoder and k denotes the length of message. Therefore, the complexity of the proposed UTMA scheme can be $\mathcal{O}(T + T^2 + \rho k T \log^2 T \log\log T)$. In addition, the AFP of above SC then JD procedure for Y_d^j is analyzed in Section III-B.

B. The Proposed UTMA Schemes

As shown in Fig. 2, the activated MCC UEs are grouped into TSUs, TCUs and TTUs according to their access timeliness requirements, and perform access in three access stage N_1, N_2, N_3 in our UTMA scheme. At the beginning of each frame, the HTS broadcasts a control signal to activated MCC UEs for estimating the average channel gain and synchronization for the GF random access. In the independent successive (IS)-UTMA, each group performs random access in successive time slot according to its priority as shown in Fig. 2(a), and in the two cognitive offloading (CO)-UTMA schemes, part of TTUs can offload in the TCUs access stage as shown in Fig. 2(b) and (c), which can resolve the UE collisions in N_3 and improve the overall access timeliness of TTUs. Note that the pilot set $\mathbf{D}$ is pre-allocated for TCUs and TTUs in ICO-UTMA scheme.

1) In N_1: Only TSUs can perform access to ensure stringent access timeliness requirement and randomly select pilots from $\mathbf{D}$ in three types UTMA schemes. At the end of N_1, the HTS decodes the received status updates.

2) In N_2: There are three different cases in this access stage, and at the end of N_2, the HTS decodes the received status updates.

Case 1 In the IS-UTMA scheme, only TCUs perform access and can randomly select pilots from $\mathbf{D}$.

Case 2 In the ECO-UTMA scheme, all TCUs and K'_{31} TTUs perform access and randomly select pilots from $\mathbf{D}$, and the ratio of offloading TTUs is $\beta' = \frac{K'_{31}}{K_{\text{TTUs}}}$.

Case 3 In the ICO-UTMA scheme, $\mathbf{D}$ is divide into $\mathbf{D}_1 = \{\xi_1, \xi_2, \xi_3, \ldots, \xi_p\}$ and $\mathbf{D}_2 = \{\xi_{p+1}, \xi_{p+2}, \xi_{p+3}, \ldots, \xi_L\}$, where $\alpha = \frac{p}{L}$ is the pilot allocation ratio, and all TCUs and $\beta = \frac{K_{31}}{K_{\text{TTUs}}}$ of TTUs perform access and they randomly select pilots from $\mathbf{D}_1$ and $\mathbf{D}_2$, respectively.

3) In N_3: All the activated TTUs that failed to access and unscheduled in N_2 perform access and randomly select pilots from $\mathbf{D}$, and the HTS decodes the received status updates at the end of N_3.

C. Procedure of UTMA Random Access Scheme

Fig. 3 shows four types of 2-step GF random access procedure in different access stages corresponding to Fig. 2 with the same mark of cases, for example, the three access stages of IS-UTMA are marked the red solid circle as shown

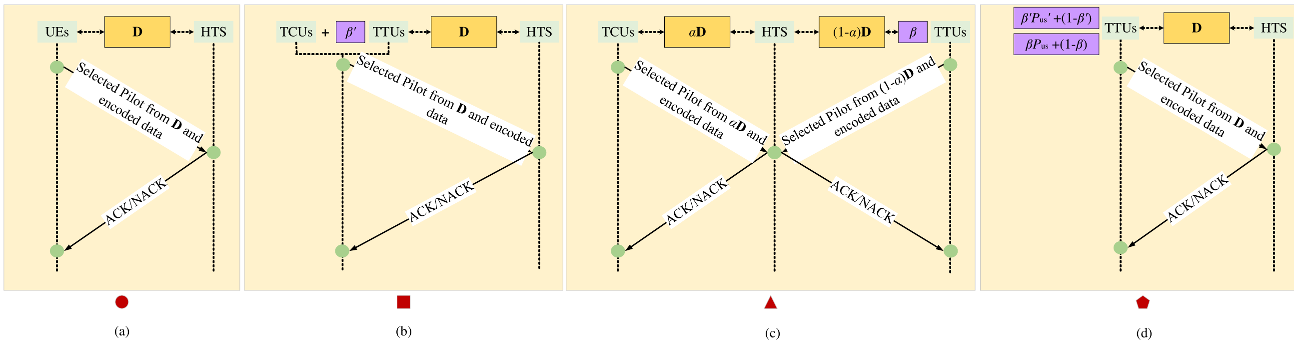


Fig. 3. Four types of 2-step GF random access procedure in the i-th frame in our UTMA scheme for N_1, N_2 and N_3, where the marks on different types are corresponding to Fig. 2 with the same mark of cases.

in Fig. 2(a), and their random access procedure in each access stage are the same as shown in Fig. 3 (a) marked with the red solid circle. Moreover, for the CO-UTMA scheme, the random access procedure in each access stage is introduced as follows:

Step 1 As shown in Fig. 2(a), each activated TSU uniformly random selects a pilot from $\mathbf{D}$ with equal probability $\frac{1}{L}$, and utilizes the T-order codebook to encoding its status update at the beginning of N_1. Then, the TSUs send the selected pilots and the encoded status updates to HTS.

Step 2 The HTS first performs pilot detection and channel estimation based on the received pilot signal. Then, the HTS decodes the received status updates from each pilot by the SC then JD decoder at the end of N_1, and recovers up to T TSUs collisions on each detected pilot. Once the HTS finishes the decoding, it broadcasts a feedback of the decoding states to all TSUs.

Step 3 Similar to Step 1, at the beginning of N_2, the activated TCUs and the offloading TTUs uniformly random select a pilot from $\mathbf{D}$ in ECO-UTMA scheme as shown in Fig. 2(b), or from $\mathbf{D}_1$ or $\mathbf{D}_2$ in ICO-UTMA scheme, respectively, as shown in Fig. 2(c). Then, the TCUs and offloading TTUs send the selected pilots and the encoded status updates to HTS.

Step 4 Similar to Step 2, the HTS decodes the received status updates at the end of N_2, and broadcasts a feedback of the decoding states to all TCUs and TTUs.

Step 5 In the CO-UTMA scheme, only the activated TTUs that failed to access and unscheduled in N_2 perform access in N_3. We assume that $P_{\text{us}}{}'$ and P_{us} are the AFP of the offloaded TTUs in N_2 for ECO- and ICO-UTMA schemes, respectively. Thus, $\beta' P_{\text{us}}{}' + (1-\beta')K_{\text{TTUs}}$ and $\beta P_{\text{us}} + (1-\beta)K_{\text{TTUs}}$ TTUs uniformly random select pilots from $\mathbf{D}$ and encodes their status updates, respectively. Then, the selected pilots and the encoded status updates are sent to the HTS.

Step 6 Similar to the Step 2, the HTS decodes the received status updates at the end of N_3 and broadcasts the decoding states to all TTUs.

III. Performance Analysis of UTMA Scheme

In this section, we present the analysis of the AAoI and AFP of each MCC UE group for three UTMA schemes. First, we utilize AoI to model the timeliness of each group. By analyzing the UE collisions and decoding failure probability, we derive the closed-form expressions of AFP for each group under the shadowed-Rician fading channel in three UTMA schemes. Further, we derive the closed-form expressions of AAoI by tracing the AFP and instantaneous AoI.

A. AoI Evolution in UTMA Scheme

Recall the instantaneous AoI observed at the receiver is defined as the difference between the current time and the generation time of the last recovered status update at the source. Therefore, we define $G_l(t)$ as the generation timestamp of the last status update from UE l that recovered by the HTS at time t, where $t = 0, 1, 2, \ldots$ is normalized to one frame. Then, the instantaneous AoI $\Lambda_l(t)$ of UE l observed at the HTS can be expressed as follows [38]:

$$\Lambda_l(t) = t - G_l(t). \tag{7}$$

In our UTMA scheme, an access frame is divided into three access stages, and each access stage equals to one time slot. Therefore, we assume that $\Lambda_l^q(f)$ represents the instantaneous AoI observed at the HTS for the UE l of group q in the f-th frame, where $q \in \{\text{TSUs, TCUs, TTUs}\}$. Fig. 4 shows the evolution of instantaneous AoI for UE l in different cases, and the analysis is as follows:

1) TSUs: The TSUs generate and send status updates at the beginning of each frame, and the HTS decodes the status updates at the end of N_1 to ensure the timeliness of TSUs. As shown in Fig. 4, a TSU l is failed to access in the first frame and Λ_l^{TSUs} increases with t, and it successfully accesses in N_1 of the second frame, and Λ_l^{TSUs} is set to 1 at $t = 4$. Define $D_l^q(f, N_j)$ as the state value of UE l of group q in the access stage N_j of the f-th frame. If the activated TSU l successfully accesses, $D_l^q = 0$, otherwise, $D_l^q = 1$. Therefore, the instantaneous AoI evolution of the TSU l in the f-th frame can be expressed as follows:

$$\Lambda_l^{\text{TSUs}}(f) = \begin{cases} 1, & \text{if } D_l^{\text{TSUs}}(f, N_1) = 0 \\ \Lambda_l^{\text{TSUs}}(f-1) + 3, & \text{if } D_l^{\text{TSUs}}(f, N_1) = 1. \end{cases} \tag{8}$$

2) TCUs: The activated TCUs generate status updates at the beginning of each frame and transmits at the beginning of N_2, and the HTS decodes the status update sent by the TCU at the end of N_2. As shown in Fig. 4, a TCU l successfully accesses in N_2 of the first and third frames and the corresponding Λ_l^{TCUs}

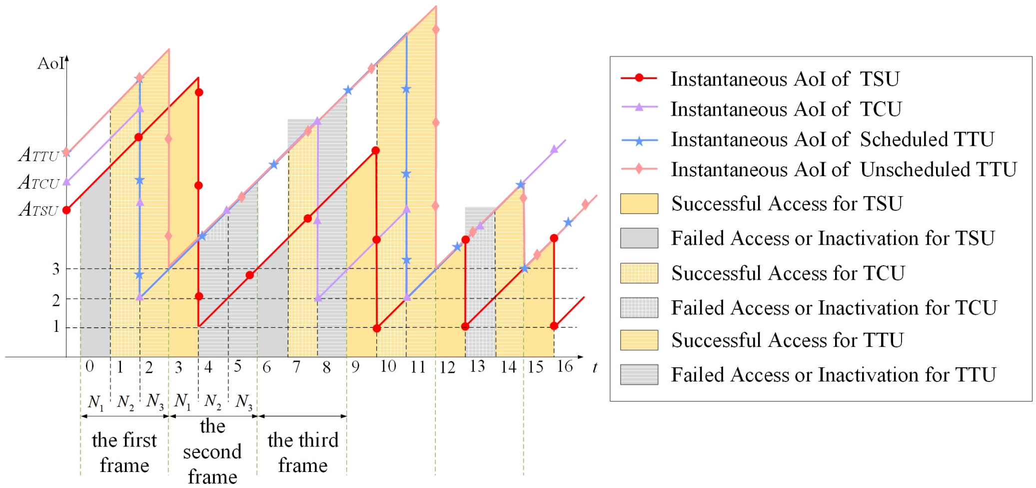


Fig. 4. The evolution of instantaneous AoI observed at the HTS for UE l in different cases.

is set to 2 at $t = 2$ and $t = 8$, respectively, and when TCU l is inactivated or fail to access in the second frame, Λ_l^{TSUs} increases with t. Therefore, if the TCU l successfully accesses in N_2 of the f-th frame, $\Lambda_l^{\text{TCUs}}(f)$ is set to 2 at the end of N_2. Otherwise, Λ_l^{TCUs} increases with t. The instantaneous AoI evolution of the TCU l in the f-th frame can be expressed as:

$$\Lambda_l^{\text{TCUs}}(f) = \begin{cases} 2, & \text{if } D_l^{\text{TCUs}}(f, N_2) = 0 \\ \Lambda_l^{\text{TCUs}}(f-1) + 3, & \text{if } D_l^{\text{TCUs}}(f, N_2) = 1. \end{cases} \tag{9}$$

3) TTUs: The instantaneous AoI evolution of TTU l can be divided into four different cases:

Case a The TTU l is offloading to send status update in N_2 of the f-th frame and successfully accesses, the corresponding $\Lambda_l^{\text{TTUs}}(f) = 2$ at the end of N_2, as the AoI of scheduled TTU $\Lambda_l^{\text{TTUs}}(1) = 2$ at $t = 2$ as shown in Fig. 4.

Case b The TTU l successfully accesses in N_3 of the f-th frame, and $\Lambda_l^{\text{TTUs}}(f) = 3$ at the end of N_3, as the AoI of unscheduled TTU $\Lambda_l^{\text{TTUs}}(1) = 3$ at $t = 3$ as shown in Fig. 4.

Case c The TTU l successfully accesses in N_2 of the $(f-1)$-th frame and fails to access in the f-th frame, $\Lambda_l^{\text{TTUs}}(f) = \Lambda_l^{\text{TTUs}}(f-1) + 4$, as the AoI of TTU $\Lambda_l^{\text{TTUs}}(2) = \Lambda_l^{\text{TTUs}}(1) + 4$ at $t = 6$ as shown in Fig. 4.

Case d The TTU l is unscheduled or failed to access in N_2 in the $(f-1)$-th frame, and also fails to access in the f-th frame, $\Lambda_l^{\text{TTUs}}(f) = \Lambda_l^{\text{TTUs}}(f-1) + 3$, as the AoI of TTU $\Lambda_l^{\text{TTUs}}(3) = \Lambda_l^{\text{TTUs}}(2) + 3$ at $t = 9$ as shown in Fig. 4.

Therefore, the instantaneous AoI evolution of the TTU l in the f-th frame can be expressed as:

$$\Lambda_l^{\text{TTUs}}(f) = \begin{cases} 2, & \text{if } D_l^{\text{TTUs}}(f, N_2) = 0 \\ 3, & \text{if } D_l^{\text{TTUs}}(f, N_3) = 0 \\ \Lambda_l^{\text{TTUs}}(f-1) + 4, & \text{if } D_l^{\text{TTUs}}(f-1, N_2) = 0 \\ & \text{and } D_l^{\text{TTUs}}(f) = 1 \\ \Lambda_l^{\text{TTUs}}(f-1) + 3, & \text{otherwise.} \end{cases} \tag{10}$$

B. Derivation of AFP in UTMA Scheme

As mentioned in Section II-A, the pilots can be divided into four cases: 1) Singleton UE pilot; 2) Fully decodable collision pilot; 3) Partially decodable collision pilot; 4) Empty pilot. For $T = 2$, denote γ as the decoding threshold, if the UE l selects ξ_j and fails to access, it may be caused by one of the following three cases.

Case I ξ_j is a singleton UE pilot, if $SINR_{1,\text{SC}} < \gamma$, decoding is failure and UE l fails to access.

Case II ξ_j is a fully decodable collision pilot, i.e., UE l and UE p both select ξ_j. If $|h_l|^2 > |h_p|^2$, the UE l fails to access due to $SINR_{1,\text{JD}} < \gamma$. Otherwise, $SINR_{2,\text{SC}} < \gamma$ for $|h_l|^2 < |h_p|^2$.

Case III ξ_j is partially decodable collision pilot, i.e., the UE l shares ξ_j with two or more UEs. Note that the HTS decodes the status updates on ξ_j in descending order according to their channel gain, and we can assume that the channel gain of UE l is ranked at v-th. If $|h_l|^2$ is not the lowest, then the UE l fails to access due to $SINR_{v,\text{JD}} < \gamma$. Otherwise, $SINR_{v,\text{SC}} < \gamma$ for UE l has the lowest $|h_l|^2$.

Therefore, the AFP of UE l can be derived based on the above three cases, i.e., $|\mathbf{Z}_j| = 1$, $|\mathbf{Z}_j| = 2$, and $|\mathbf{Z}_j| > 2$.

1) $|\mathbf{Z}_j| = 1$: The probability of **Case I** is as follows:

$$\Pr(|\mathbf{Z}_j| = 1) = \binom{K}{1}\left(\frac{1}{L}\right)\left(1 - \frac{1}{L}\right)^{K-1}. \tag{11}$$

In **Case I**, the decoding fails when $SINR_{1,\text{SC}} \leq \gamma$, where $\gamma = 2^R - 1$ and R is the inner code rate. To facilitate the subsequent analysis, the PDF of shadowed-Rician channel in Eq.(1) can be simplified as follows:

$$f(r) = \omega e^{-\eta r}, \tag{12}$$

where $\omega = \frac{P_T}{2b\sigma^2}(\frac{2bm}{2bm+\Omega})^m$, $\eta = \frac{mP_T}{(2bm+\Omega)\sigma^2}$. Therefore, by setting $b = 0.063$, $m = 1$, $\Omega = 0.000897$ [39], the decoding failure probability is:

$$\Pr\left(\frac{P_T|h_l|^2}{\sigma^2} < \gamma\right) = \Pr\left(|h_l|^2 < \epsilon\gamma\right) = 1 - e^{-w\epsilon\gamma}, \tag{13}$$

where $\epsilon = \frac{\sigma^2}{P_T}$. Thus, combining the UE collisions and decoding failure probability, the AFP of UE l under $|\mathbf{Z}_j| = 1$ is:

$$AFP_{l,|\mathbf{Z}_j|=1} = \Pr(|\mathbf{Z}_j| = 1)\Pr\left(\frac{P_T|h_l|^2}{\sigma^2} < \gamma\right) = \frac{K(L-1)^{K-1}\left[1-\exp(-\frac{w\gamma}{\epsilon})\right]}{L^K}. \tag{14}$$

2) $|\mathbf{Z}_j| = 2$: In **Case II**, the probability of both UEs choosing ξ_j is:

$$\Pr(|\mathbf{Z}_j| = 2) = \binom{K}{2}\left(\frac{1}{L}\right)^2\left(1-\frac{1}{L}\right)^{K-2}, \tag{15}$$

and the decoding failure probability is:

$$\Pr\left(\frac{P_T|h_1|^2}{P_T|h_2|^2+\sigma^2} < \gamma, \frac{\sum_{i=1}^{2} P_T|h_i|^2}{\sigma^2} < \gamma\right) + \Pr\left(\frac{P_T|h_2|^2}{\delta P_T|h_1|^2+\sigma^2} < \gamma\right) = \Pr\left(|h_1|^2+|h_2|^2 < \epsilon\gamma\right) + \Pr\left(|h_2|^2 - \delta\gamma|h_1|^2 < \epsilon\gamma\right), \tag{16}$$

where $|h_1|^2 + |h_2|^2 \sim Erlang(2, w)$. By deriving the PDF of $|h_2|^2 - \delta\gamma|h_1|^2$, the AFP can be obtained for $|\mathbf{Z}_j| = 2$:

$$AFP_{l,|\mathbf{Z}_j|=2} = \frac{\binom{K}{2}(L-1)^{K-2}}{L^K} \cdot \left[2 - \left(1 + \frac{w\gamma}{\epsilon} + \frac{1}{1+\delta\gamma}\right)\exp\left(-\frac{w\gamma}{\epsilon}\right)\right]. \tag{17}$$

3) $|\mathbf{Z}_j| > 2$: In **Case III**, there are $|\mathbf{Z}_j| = k$ ($k = 3, 4, \ldots, K$) UEs selecting ξ_j, and UE l is ranked at ($v = k-i+1$)-th ($i = 1, 2, \ldots, k$). In this case, there are $(k-i)$ UEs successfully access, and the UE l and i UEs with worese channel gains fails to access, and the AFP for UE l is:

$$\sum_{k=3}^{K} \Pr(|\mathbf{Z}_j| = k) \sum_{i=1}^{k} \Pr(i \text{ UEs failed to decode}), \tag{18}$$

where the probability of k UEs choosing the same pilot ξ_j is:

$$\Pr(|\mathbf{Z}_j| = k) = \binom{K}{k}\left(\frac{1}{L}\right)^k\left(1-\frac{1}{L}\right)^{K-k}. \tag{19}$$

When k UEs choose ξ_j, the decoding failure probability of $i (i \neq k)$ UEs is

$$\Pr\left(\frac{P_T|h_{k-i+1}|^2}{\sum_{j=1}^{k-i}\delta P_T|h_j|^2 + \sum_{j=k-i+2}^{k} P_T|h_j|^2 + \sigma^2} < \gamma, \frac{\sum_{j=k-i+1}^{k} P_T|h_j|^2}{\sum_{j=1}^{k-i}\delta P_T|h_j|^2 + \sigma^2} < \gamma\right) = \Pr\Big(-\delta\gamma|h_1|^2 - \cdots - \delta\gamma|h_{k-i}|^2 + |h_{k-i+1}|^2 + \cdots + |h_k|^2 < \epsilon\gamma\Big), \tag{20}$$

where $X_t = |h_t|^2 (t = k-i+1, \ldots, k) \sim \exp(\frac{1}{w})$ and $Y_u = -\delta\gamma|h_u|^2 (t = 1, \ldots, k-i) \sim \exp(\frac{\delta\gamma}{w})$. Therefore, we can derive $H_i = \sum_{i=k-i+1}^{k} X_i \sim Erlang(i, w)$, $G_{k-i} = \sum_{i=1}^{k-i} Y_i \sim Erlang(k-i, \frac{w}{\delta\gamma})$.

Let $Z = H_i + G_{k-i}$, which can be derived by taking the Laplace transform of the PDF of H_i and G_{k-i} due to they are independent of each other. Then, we can obtain $f(z)$ via the Laplace inverse transform [40]. Thus, we can derive the closed-form of decoding failure probability in Eq.(20) as follows,

$$\Pr(i \text{ UEs failed to decode}) = \sum_{j=1}^{k-i} \frac{\binom{k-j-i}{k-j-1}(\delta\gamma)^i}{(1+\delta\gamma)^{k-j}} + \sum_{j=1}^{i} \frac{\binom{k-j-i}{i-j}(\delta\gamma)^{i-j}}{(1+\delta\gamma)^{k-j}} \cdot \left[1 - \exp(-\frac{w\gamma}{\epsilon})\sum_{m=0}^{j-1}\frac{(\frac{w\gamma}{\epsilon})^m}{\Gamma(m)}\right], \tag{21}$$

where $\Gamma(\cdot)$ is the incomplete Gamma function. Similarly, we can derive the decoding failure probability when $i = k$ as:

$$\Pr\left(\frac{P_T|h_1|^2}{\sum_{j=2}^{k} P_T|h_j|^2 + \sigma^2} < \gamma, \frac{\sum_{j=1}^{k} P_T|h_j|^2}{\sigma^2} < \gamma\right) = 1 - \exp\left(-\frac{w\gamma}{\epsilon}\right)\sum_{m=0}^{k-1}\frac{\left(\frac{w\gamma}{\epsilon}\right)^m}{\Gamma(m)}. \tag{22}$$

Thus, the AFP of **Case III** is:

$$AFP_{l,|\mathbf{Z}_j|>2} = \sum_{i=3}^{K} \frac{\binom{K}{k}(L-1)^{K-k}}{L^K} \cdot \left\{\left\{\sum_{i=1}^{k-1}\left\{\sum_{j=1}^{k-i}\frac{\binom{k-j-i}{k-j-1}(\delta\gamma)^i}{(1+\delta\gamma)^{k-j}} + \sum_{j=1}^{i}\frac{\binom{k-j-i}{i-j}(\delta\gamma)^{i-j}}{(1+\delta\gamma)^{k-j}} \cdot \left[1-\exp(-\frac{w\gamma}{\epsilon})\sum_{m=0}^{j-1}\frac{(\frac{w\gamma}{\epsilon})^m}{\Gamma(m)}\right]\right\}\right\} + \left[1-\exp\left(-\frac{w\gamma}{\epsilon}\right)\sum_{m=0}^{k-1}\frac{(\frac{w\gamma}{\epsilon})^m}{\Gamma(m)}\right]\right\}. \tag{23}$$

Further, by combining Eq. (14), Eq. (17) and Eq. (23), we can derive the closed-form expression of AFP for UE l in one time slot as Eq. (24), shown at the bottom of the next page. Then, as shown in Fig. 2, by substituting the number of activated UEs and available pilots in different access stages into Eq. (24), the AFP of UE l for each group in three UTMA schemes can be derived as Eq. (25), Eq. (26) and Eq. (27), shown at the bottom of the next page., respectively, where $L^{(q)}$ denotes the number of available pilots for group q in their access stages.

C. Derivation of AAoI in UTMA Scheme

Assuming F is the number of access frames and H_q is the number of UEs in group q, thus the AAoI of group q can be expressed as:

$$\Delta_q = \frac{\sum_{j=1}^{H_q}\sum_{f=1}^{F}\Lambda_l^q(f)}{H_q F}. \tag{28}$$

Utilizing the theoretical results of AFP in each group p for three UTMA schemes, we can derive the corresponding AAoI. Let $P_{s,l}^q$ denote the successful access probability of UE l, and $P_{us,l}^q$ is the probability of inactivated or failure to access of UE l. Then, we have $P_{s,l}^q$ for UE l of group q for each UTMA scheme as:

$$P_{s,l}^q = 1 - P_{us,l}^q = \lambda_q(1 - AFP_l^q), \tag{29}$$

where $\lambda_q = \frac{K_q}{H_q}$ is the active probability of group q. Therefore, the AoI of TSU and TCU in the f-th frame can be expressed as follows:

$$\Lambda_l^{\text{TSUs}}(f) = \sum_{x=1}^{f}(3x-2)P_{s,l}^{\text{TSUs}}\left(P_{us,l}^{\text{TSUs}}\right)^{x-1}, \tag{30a}$$

and

$$\Lambda_l^{\text{TCUs}}(f) = \sum_{x=1}^{f}(3x-1)P_{s,l}^{\text{TCUs}}\left(P_{us,l}^{\text{TCUs}}\right)^{x-1}. \tag{30b}$$

Then, consider the AoI of TTU L, let $P_{ss2,l}^{\text{TTUs}} = \lambda_{\text{TTUs}}\beta(1 - AFP_{l,2}^{\text{TTUs}})$ denote the probability of TTU l successfully accessing in N_2, $P_{ss3,l}^{\text{TTUs}} = \lambda_{\text{TTUs}}\beta AFP_{l,2}^{\text{TTUs}}(1 - AFP_{l,3}^{\text{TTUs}})$ is the probability of TTU l failed to access in N_2 but successfully accessing in N_3, and $P_{us3,l}^{\text{TTUs}} = \lambda_{\text{TTUs}}(1-\beta)(1 - AFP_{l,3}^{\text{TTUs}})$ is the probability of TTU l unscheduled in N_2 but successfully accessing in N_3. In addition, we assume that $P_{s3,l}^{\text{TTUs}} = P_{ss3,l}^{\text{TTUs}} + P_{us3,j}^{\text{TTUs}}$ is the probability of TTU l successfully access in N_3. Thus, the AoI of TTU l in three UTMA schemes are as follows:

*1) **IS-UTMA:***

$$\Lambda_l^{\text{TTUs}}(f) = \sum_{x=1}^{f} 3xP_{s,l}^{\text{TTUs}}\left(P_{us,l}^{\text{TTUs}}\right)^{x-1}, \tag{30c}$$

*2) **ECO- and ICO-UTMA:***

$$\Lambda_l^{\text{TTUs}}(f) = \begin{cases} 2P_{ss2,l}^{\text{TTUs}} + 3(1 - P_{ss2,l}^{\text{TTUs}}), & f = 1, \\ 2P_{ss2,l}^{\text{TTUs}} + 3P_{s3,l}^{\text{TTUs}} + 6(P_{us,l}^{\text{TTUs}})^2, & f = 2, \\ 2P_{ss2,l}^{\text{TTUs}} + 3P_{s3,l}^{\text{TTUs}} + 3f(P_{us,l}^{\text{TTUs}})^{f-1} & \\ +[\sum_{x=2}^{f-1} 3x(P_{ss2,l} + P_{s3,l}^{\text{TTUs}})(P_{us,l}^{\text{TTUs}})^{x-1}], & f \geq 3. \end{cases} \tag{30d}$$

By substituting Eq. (30) into Eq. (28), we have the closed-form expressions of AAoI for each group in three UTMA schemes as shown in Eq. (31), Eq. (32) and Eq. (33), at the bottom of the next page.

IV. Optimization of UTMA Scheme

In this section, we first construct the joint parameter optimization problems of AAoI for ECO- and ICO-UTMA

$$AFP_l = \frac{K(L-1)^{K-1}\left[1-\exp(-\frac{w\gamma}{\epsilon})\right] + \binom{K}{2}(L-1)^{K-2}\cdot\left[2-\left(1+\frac{w\gamma}{\epsilon}+\frac{1}{1+\delta\gamma}\right)\exp\left(-\frac{w\gamma}{\epsilon}\right)\right]}{L^K}$$
$$+\sum_{i=3}^{K}\frac{\binom{K}{k}(L-1)^{K-k}}{L^K}\cdot\left\{\left\{\sum_{i=1}^{k-1}\left\{\sum_{j=1}^{k-i}\frac{\binom{k-j-i}{k-j-1}(\delta\gamma)^i}{(1+\delta\gamma)^{k-j}}+\sum_{j=1}^{i}\frac{\binom{k-j-i}{i-j}(\delta\gamma)^{i-j}}{(1+\delta\gamma)^{k-j}}\right.\right.\right.$$
$$\left.\left.\cdot\left[1-\exp(-\frac{w\gamma}{\epsilon})\sum_{m=0}^{j-1}\frac{(\frac{w\gamma}{\epsilon})^m}{\Gamma(m)}\right]\right\}\right\}+\left[1-\exp\left(-\frac{w\gamma}{\epsilon}\right)\sum_{m=0}^{k-1}\frac{(\frac{w\gamma}{\epsilon})^m}{\Gamma(m)}\right]\Bigg\}. \tag{24}$$

*1) **IS-UTMA:***

$$AFP_l^q = AFP_l|_{K=K_q,L=L^{(q)}}\left(q\in\{\text{TSUs, TCUs, TTUs}\}\right). \tag{25}$$

*2) **ECO-UTMA:***

$$\begin{cases} AFP_l^{\text{TSUs}} = AFP_l|_{K=K_{\text{TSUs}}}, \\ AFP_l^{\text{TCUs}} = AFP_l|_{K=K_{\text{TCUs}}+\beta K_{\text{TTUs}}}, \\ AFP_l^{\text{TTUs}} = (1+\beta AFP_l|_{K=K_{\text{TCUs}}+\beta K_{\text{TTUs}}}-\beta)\cdot AFP_l|_{K=\beta K_{\text{TTUs}}\cdot AFP_l|_{K=K_{\text{TCUs}}+\beta K_{\text{TTUs}}}+(1-\beta)K_{\text{TTUs}}}. \end{cases} \tag{26}$$

*3) **ICO-UTMA:***

$$\begin{cases} AFP_l^q = AFP_l|_{K=K_q,L=L^{(q)}}\left(q\in\{\text{TSUs, TCUs}\}\right), \\ AFP_l^{\text{TTUs}} = \left(1+\beta AFP_l|_{K=\beta K_{\text{TTUs}},L=(1-\alpha)L}-\beta\right)\cdot AFP_l|_{K=\beta K_{\text{TTUs}}\cdot AFP_l|_{K=\beta K_{\text{TTUs}},L=(1-\alpha)L}+(1-\beta)K_{\text{TTUs}}}. \end{cases} \tag{27}$$

schemes under diverse AFP constraints, number of pilots and MCC UEs. Then, we decompose each optimization problem into two simplified sub-problems to obtain the optimal parameters for the ECO- and ICO-UTMA schemes. In practical, the related data is collected by the HTS and the optimization of UTMA scheme can be performed in the ground data center, and the HTS broadcasts the optimized parameters to the covered MCC UEs.

A. Optimization of ECO-UTMA Scheme

Fig. 5(a) validates the derivation accuracy of Eq. (26) and Eq. (32) for AFP^{TCUs} and Δ_{TCUs}, respectively, and we can observe that AFP^{TCUs} and Δ_{TCUs} are monotonically non-decreasing functions of β'. Further, Fig. 5(b) shows that the simulated AFP^{TTUs} and Δ_{TTUs} agree well with the theoretical performance of Eq. (26) and Eq. (33b), respectively, and AFP^{TTUs} and Δ_{TTUs} are also monotonically non-increasing functions of β' under different K_{TTUs}.

In the ECO-UTMA scheme, K_{TCUs} TCUs and $\beta' K_{\text{TTUs}}$ TTUs share L pilots in $\mathbf{D}$ at N_2, and the optimization for ECO-UTMA scheme is to find the optimal β', i.e., β'^{*}, to minimize Δ_{TCUs} and Δ_{TTUs} under their desired AFP. Therefore, the optimization problem for ECO-UTMA scheme can be constructed as follows:

$$\min_{\beta'} \quad \Delta_q \left(q \in \{\text{TCUs}, \text{TTUs}\}\right), \quad s.t. \quad \begin{cases} AFP^q \leq AFP^{q*}, \\ 0 \leq \beta' \leq 1, \end{cases} \tag{34}$$

where the two constraints are the desired AFP and the value range of offloading ratio, respectively.

Thus, according to Fig. 5(a), we can achieve the lowest Δ_{TCUs} by set $\beta' = 0$, and the ECO-UTMA scheme degenerates to IS-UTMA scheme. To improve Δ_{TTUs}, we can find the upper bound of β' as $\beta_1'^{*}$ to meet the desired $AFP^{\text{TCU}*}$, and decompose the optimization problem Eq.(34) into two sub-problems as follows.

Sub-Problem 1: The upper bound $\beta_1'^{*}$ of β' must satisfy the desired $AFP^{\text{TCUs}*}$, then, for $\beta' = \beta_1'^{*}$, we can achieve the lowest Δ_{TTUs}, and we have:

$$\min_{\beta'} \quad \Delta_{\text{TTUs}}, s.t. \quad \begin{cases} AFP^{\text{TCUs}}(\beta') \leq AFP^{\text{TCUs}*}, \\ 0 \leq \beta' \leq 1. \end{cases} \tag{35}$$

Hence, from Fig. 5(a), $\beta_1'^{*}$ can be express as follows:

$$\beta_1'^{*} = \left\{\beta_1'^{*} | AFP^{\text{TCUs}}(\beta_1'^{*}) = AFP^{\text{TCUs}*}\right\}. \tag{36}$$

Sub-Problem 2: Sub-problem 1 has found a upper bound for β' to satisfy the desired $AFP^{\text{TCUs}*}$ and achieve the lowest Δ_{TTUs}. Further, we can also find a lower bound for β' as $\beta_0'^{*} = \left\{\beta_0'^{*} | AFP^{\text{TTUs}}(\beta_0'^{*}) = AFP^{\text{TTUs}*}\right\}$ to satisfy the desired $AFP^{\text{TTUs}*}$, and we have:

$$\min_{\beta'} \quad \Delta_{\text{TTUs}},$$

---

*1) **TSUs:***

$$\Delta_{\text{TSUs}} = \frac{\sum_{l=1}^{H_{\text{TSUs}}} \sum_{f=1}^{F} \sum_{x=1}^{f} (3x-2) \frac{K_{\text{TSUs}}\left(1-AFP_l^{\text{TSUs}}\right)}{H_{\text{TSUs}}} \left[1 - \frac{K_{\text{TSUs}}\left(1-AFP_l^{\text{TSUs}}\right)}{H_{\text{TSUs}}}\right]^{x-1}}{H_{\text{TSUs}} F}, \tag{31}$$

*2) **TCUs:***

$$\Delta_{\text{TCUs}} = \frac{\sum_{l=1}^{H_{\text{TCUs}}} \sum_{f=1}^{F} \sum_{x=1}^{f} (3x-1) \frac{K_{\text{TCUs}}\left(1-AFP_l^{\text{TCUs}}\right)}{H_{\text{TCUs}}} \left[1 - \frac{K_{\text{TCUs}}\left(1-AFP_l^{\text{TCUs}}\right)}{H_{\text{TCUs}}}\right]^{x-1}}{H_{\text{TCUs}} F}, \tag{32}$$

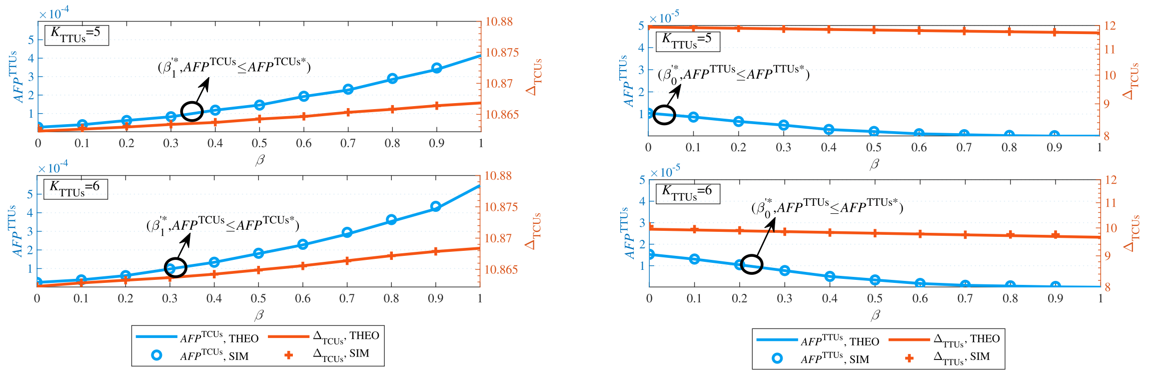
*3) **TTUs:***

*3.1) **IS-UTMA:***

$$\Delta_{\text{TTUs}} = \frac{\sum_{l=1}^{H_{\text{TTUs}}} \sum_{f=1}^{F} \sum_{x=1}^{f} 3x \frac{K_{\text{TTUs}}\left(1-AFP_l^{\text{TTUs}}\right)}{H_{\text{TTUs}}} \left[1 - \frac{K_{\text{TTUs}}\left(1-AFP_l^{\text{TTUs}}\right)}{H_{\text{TTUs}}}\right]^{x-1}}{H_{\text{TTUs}} F}, \tag{33a}$$

*3.2) **ECO- and ICO-UTMA:***

$$\Delta_{\text{TTUs}} = \frac{1}{H_{\text{TTUs}} F} \cdot \sum_{l=1}^{H_{\text{TTUs}}} \Bigg\{ \left\{4P_{ss2,l}^{\text{TTUs}} + 3\left(1 - P_{ss2,l}^{\text{TTUs}}\right) + 3P_{s3,l}^{\text{TTUs}} + 6P_{us,l}^{\text{TTUs}}\right\} + \sum_{f=3}^{F} \left\{2P_{ss2,l}^{\text{TTUs}} + 3P_{s3,l}^{\text{TTUs}} + \left\{\sum_{x=2}^{f-1} 3x \left(P_{ss2,l}^{\text{TTUs}} + P_{s3,l}^{\text{TTUs}}\right)\left(P_{us,l}^{\text{TTUs}}\right)^{x-1}\right\} + 3f\left(P_{us,l}^{\text{TTUs}}\right)^{f-1}\right\} \Bigg\}. \tag{33b}$$



(a) Δ_{TCUs} and AFP^{TCUs} versus β, where $K_{\text{TCUs}} = 2$ and $H_{\text{TCUs}} = 8$.

(b) Δ_{TTUs} and AFP^{TTUs} versus β, where $K_{\text{TCUs}} = 2$ and $H_{\text{TTUs}} = 20$.

Fig. 5. The AAoI and AFP of TCUs and TTUs versus β, where $AFP^{\text{TCUs}*} = 10^{-5}$, $AFP^{\text{TTUs}*} = 10^{-6}$ and $L = 169$.

$$s.t. \quad \begin{cases} AFP^{\text{TTUs}}(\beta') \leq AFP^{\text{TTUs}*}, \\ \beta_0'^* \leq \beta' \leq \beta_1'^*. \end{cases} \tag{37}$$

Note that if $\beta_0'^* \leq \beta_1'^*$, we can choose β'^* as $\beta_1'^*$ to achieve the lowest Δ_{TTUs} under the desired $AFP^{\text{TCUs}*}$ and $AFP^{\text{TTUs}*}$. Otherwise, $AFP^{\text{TCUs}*}$ and $AFP^{\text{TTUs}*}$ can not be simultaneously satisfied, and we select $\beta'^* = \beta_1'^*$ to satisfy the desired AFP requirement for the higher priority group, and then consider minimizing the AAoI of the lower priority group. Recall that AFP^{TCUs} is a monotonically non-decreasing function of β' for ECO-UTMA scheme, then β'^* can be easily obtained in practical as $\beta'^* K_{\text{TTUs}}$ should be an integer.

B. Optimization of ICO-UTMA Scheme

Fig. 6(a) validates the accuracy of Eq. (27) and Eq. (32) for AFP^{TCUs} and Δ_{TCUs}, respectively, where both AFP^{TCUs} and Δ_{TCUs} are monotonically non-increasing functions of α. In addition, the simulated AFP^{TTUs} and Δ_{TTUs} agree well with the theoretical performance of Eq. (27) and Eq. (33b) as shown in Fig. 6(b), respectively, and we can observe that AFP^{TTUs} is a monotonically non-increasing functions of β when g_{TTUs} is low.

The optimization goal of ICO-UTMA scheme is to find the optimal α and β to minimize Δ_{TCUs} and Δ_{TTUs} under their desired AFP, where αL and $(1-\alpha)L$ pilots are pre-allocated for K_{TCUs} TCUs and βK_{TTUs} TTUs in ICO-UTMA scheme, respectively. Thus, the optimization problem for ICO-UTMA scheme can be expressed as follows:

$$\min_{\alpha,\beta} \quad \Delta_q \left(q \in \{\text{TCUs, TTUs}\}\right)$$
$$s.t. \quad \begin{cases} AFP^q \leq AFP^{q*}, \\ 0 \leq \alpha \leq 1, \\ 0 \leq \beta \leq 1, \end{cases} \tag{38}$$

where the three constraints are the desired AFPs, the value ranges of pilot allocation ratio and offloading ratio, respectively.

The objective function Eq. (38) is not jointly convex versus all variables, thus simplification is considered. In ICO-UTMA scheme, only α affects Δ_{TCUs}, while both α and β affect Δ_{TTUs}. Thus, since the ICO-UTMA scheme avoids UE collisions between different groups, we can adopt a two-step optimization strategy to optimize (38). Combining the curve trend in Fig. 6(a), we can achieve the lowest Δ_{TCUs} by set $\alpha = 1$ and $\beta \neq 0$, and the ICO-UTMA scheme degenerates to ECO-UTMA scheme. To avoid collisions from TTUs and improve Δ_{TTUs}, we can find the minimum α as $\alpha_{\min}$ to meet the desired AFP^{TCUs}, and use it to analyze the optimal β as β^* to minimize Δ_{TTUs} under the desired $AFP^{\text{TTUs}*}$. Consequently, we decompose Eq. (38) into two simplified sub-problems as follows.

Sub-Problem 1: For desired $AFP^{\text{TCUs}*}$, $\alpha_{\min}$ can be obtained as follows:

$$\min_{\alpha} \quad \alpha,$$
$$s.t. \quad \begin{cases} AFP^{\text{TCUs}}(\alpha) \leq AFP^{\text{TCUs}*}, \\ 0 \leq \alpha \leq 1, \end{cases} \tag{39}$$

and we have:

$$\alpha_{\min} = \left\{\alpha_{\min} | AFP^{\text{TCUs}}(\alpha_{\min}) = AFP^{\text{TCUs}*}\right\}. \tag{40}$$

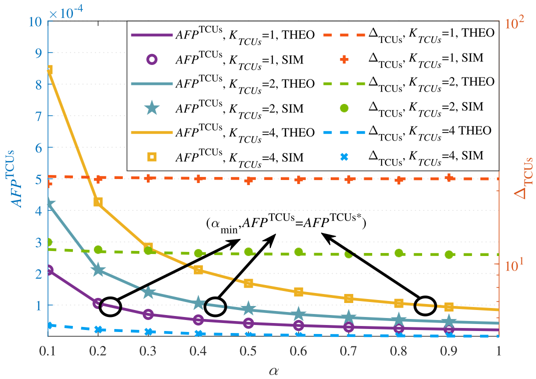
Note that for the ICO-UTMA scheme in practical scenarios, AFP^{TCUs} is a monotonically non-increasing function of α. Thus, $\alpha_{\min}$ can be also easily found due to $\alpha_{\min} L$ is an integer.

Sub-Problem 2: After $\alpha_{\min}$ is determined, we can find β^* to meet the desired $AFP^{\text{TTUs}*}$ and minimize Δ_{TTUs} under H_q and the system load $g_q = \frac{K_q}{L}$, and we have:

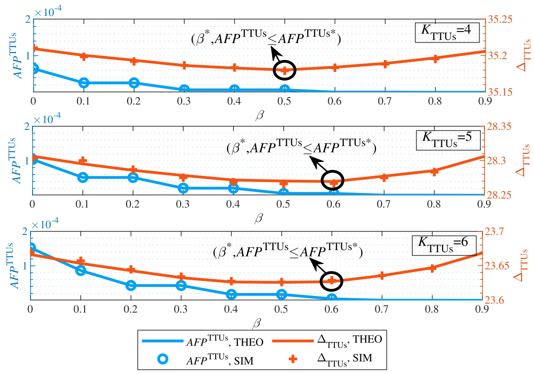
$$\min_{\beta} \quad \Delta_{\text{TTUs}},$$
$$s.t. \quad \begin{cases} AFP^{\text{TTUs}}(\alpha_{\min}, \beta) \leq AFP^{\text{TTUs}*}, \\ 0 \leq \beta \leq 1. \end{cases} \tag{41}$$

Therefore, according to Fig. 6(b), we can find a lower bound of β as $\beta_{\min}$ to meet the desired $AFP^{\text{TTUs}*}$. In addition, Fig. 6(b) shows that Δ_{TTUs} is a convex function of β, and there is an extreme point denoted as β_e to achieve a minimum Δ_{TTUs}. Thus, the minimum Δ_{TTUs} can be achieved if and only if $\beta = \beta_{\min}$ or $\beta = \beta_e$ under the desired $AFP^{\text{TTUs}*}$, and the selection rule for β^* is as follows:

$$\beta^* = \begin{cases} \beta_{\min}, & \text{if } \beta_{\min} > \beta_e, \\ \beta_e, & \text{if } \beta_{\min} \leq \beta_e. \end{cases} \tag{42}$$



(a) Δ_{TCUs} and AFP^{TCUs} versus α, where $H_{\text{TCUs}} = 8$.



(b) Δ_{TTUs} and AFP^{TTUs} versus β, where $H_{\text{TTUs}} = 20$ and $\alpha_{\min} = 0.85$.

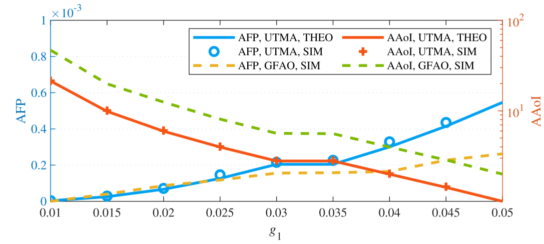
Fig. 6. The AAoI and AFP of TCUs and TTUs versus α and β, where $AFP^{\text{TCUs}*} = 10^{-5}$, $AFP^{\text{TTUs}*} = 10^{-6}$ and $L = 169$.

Considering that AFP_{TTUs} is a monotonically non-increasing function of β, and Δ_{TTUs} is a convex function of β, we can easily obtain β^* according to Eq. (42) for the ICO-UTMA scheme because $\beta^* K_{\text{TTUs}}$ is an integer.

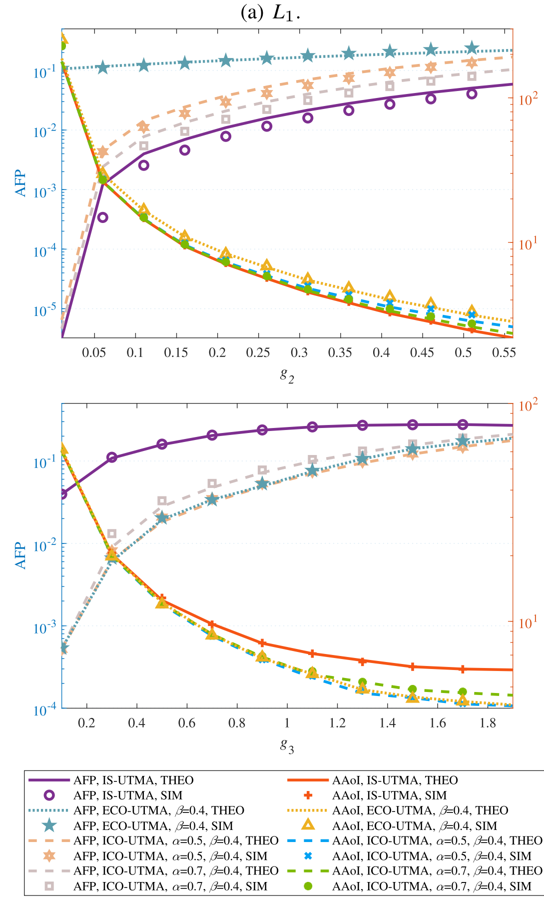
In addition, our scheme can be verified to provide an AFP as low as 10^{-6} within one time slot from Fig. 5 and Fig. 6. Therefore, setting each access stage equals to one time slot in our system model is sufficient to meet the AFP requirements of MCC UEs. In practical scenarios, we can set the length of one time slot according to the worst two-way propagation delay corresponding to the altitude of HTS. Moreover, we can set the number of time slots for each type of MCC UEs according to their AFP requirements, allowing multiple access opportunities to ensure the AFP requirements under the specified PDB requirements.

V. Simulation Results and Discussions

Fig. 5 and Fig. 6 in Section IV validate that the UTMA scheme can provide UT protection for three types of MCC UEs under the desired AFP. In this section, we deploy the UTMA scheme to massive access scenario, where we still divide UEs into three groups with different access priorities, i.e., L_1, L_2 and L_3, and the UEs in L_1 and L_3 have the highest and lowest access priority, respectively. First, we provide extensive Monte Carlo simulations to validate the accuracy of our theoretical analysis for AFP and AAoI with different values of α and β. Then, we discuss the AAoI performance versus g_i and T for ICO-UTMA scheme under different parameters,



(a) L_1.



(b) L_2 and L_3.

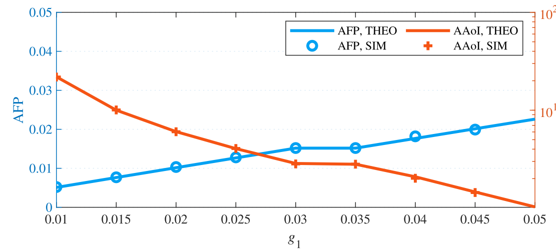
Fig. 7. The variation of AAoI and AFP derived in Section III versus g_i in UTMA scheme, where SNR = 20dB.

TABLE I
Simulation Parameters

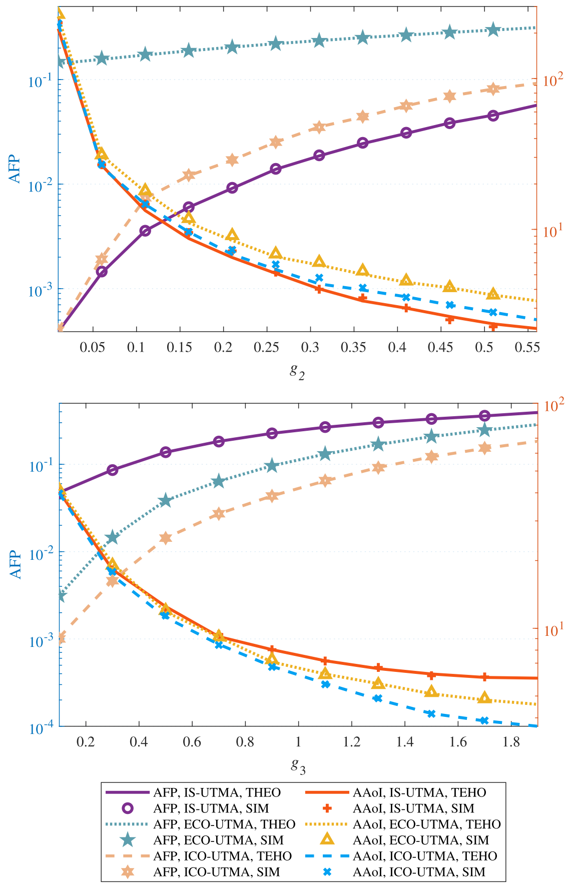
Description	Values
Number of access frames F	1000000
Number of UEs in group i	$\{H_1 = 8,\ H_2 \in [94, 338],$ $H_3 \in [338, 507]\}$
Number of orthogonal pilots L	169
Maximum collision resolution order T	2
SNR(dB)	[-5, 20]
Pilot allocation ratio α	{0.5, 0.7}
Offloading ratio β	0.4

and analyze the value range of g_2 and g_3 to find the optimal AAoI. Further, we compare the UTMA scheme with existing schemes [14].

Table I summarizes the Monte Carlo simulation parameters, where we set $F = 1000000$ to obtain the more precise AAoI. Then, H_i is set to be 8, $[94, 338]$, and $[338, 507]$, respectively,

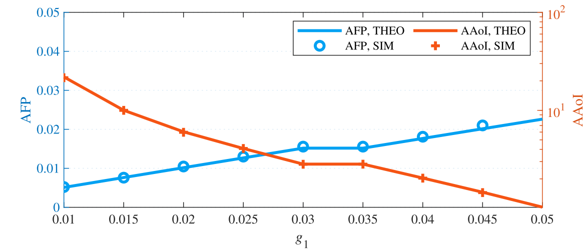


(a) L_1.

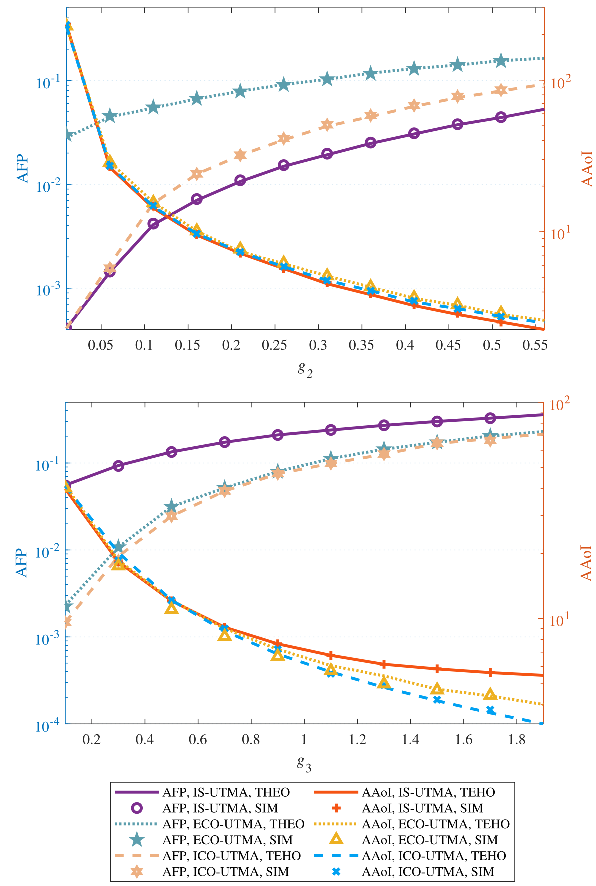


(b) L_2 and L_3.

Fig. 8. The variation of AAoI and AFP derived in Section III versus g_i in UTMA scheme, where SNR = 5dB.



(a) L_1.

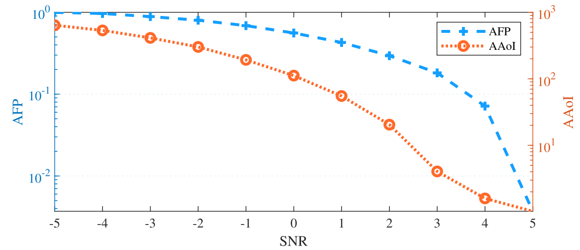


(b) L_2 and L_3.

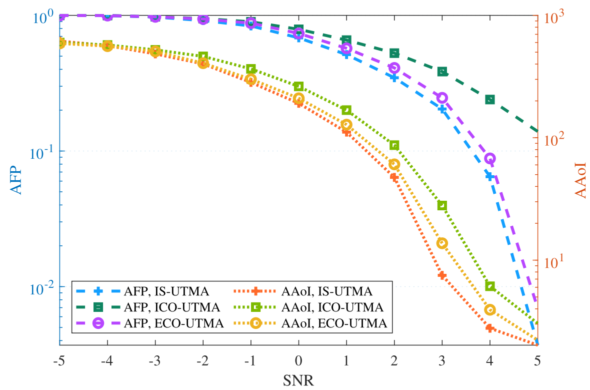
Fig. 9. The variation of AAoI and AFP derived in Section III versus g_i in UTMA scheme, where SNR = 10dB.

where 8 is the density requirement for MCC UEs in future NTN [8], and we simulate large H_2 and H_3 to validate the AFP performance of our UTMA schemes under a generic crowded random access scenario [41]. To obtain the largest number of orthogonal pilot sequences, we set $L = 169$. In addition, considering a tradeoff between the performance of AFP and the complexity, we set $T = 2$ [26]. Further, we set the Signal-to-Noise Ratio (SNR) $\in [-5, 20]$ dB, and α and β are the optimized parameters via the optimization in Section IV.

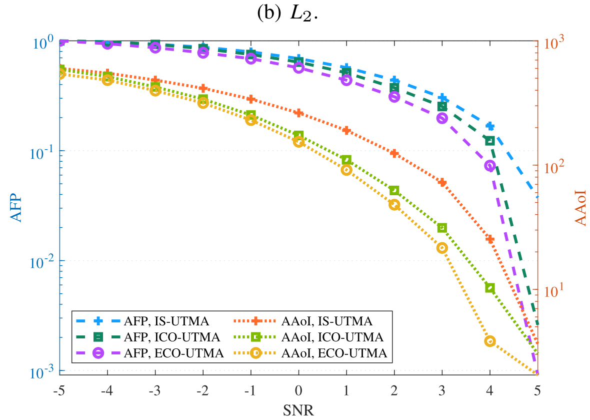
Fig. 7 illustrates the AAoI performance of each group versus $g_i (i \in \{1, 2, 3\})$ for our three UTMA schemes under SNR $=$ 20 dB, and the theoretical derivations agree well with the simulation results under different parameters. Fig. 7(a) compares the AAoI and AFP performance of L_1 under the UTMA and GFAO schemes [14]. Since the UTMA scheme introduces a multi-dimensional codebook to resolve UE collisions, which can achieve a lower AFP in one time slot, thus the UTMA scheme has good timeliness performance. On contrast, the UEs attempt to access in multiply slots in GFAO scheme, sacrificing timeliness for reliability and the AFP of GFAO outperforms that of the UTMA scheme. Fig. 7(b) shows that as g_i increases, all AAoI curves for each group are monotonically non-increasing functions under different α and β, where the AAoI is higher in the low load region and then decreases significantly. Because when g_i increases from 0, the number of activated UEs increases and the large AAoI caused by inactivated UEs will decrease rapidly, and the AAoI may decrease when $g_i \leq T$. Moreover, the IS-UTMA outperforms ECO- and ICO-UTMA in L_2 due to it does not provide offloading, and the ICO-UTMA scheme outperforms the ECO-UTMA scheme in terms of AAoI because it avoids UE collisions between different groups. As shown in Fig. 7(b), the ECO-UTMA scheme can provide AFP $10^{-2} \sim 10^{-1}$ for all g_2, while the ICO-UTMA scheme can provide AFP $\leq 10^{-4}$ if $g_i \leq 0.025$, and $10^{-2} \sim 10^{-1}$ if g_i is large. Moreover, due



(a) L_1.



(b) L_2.

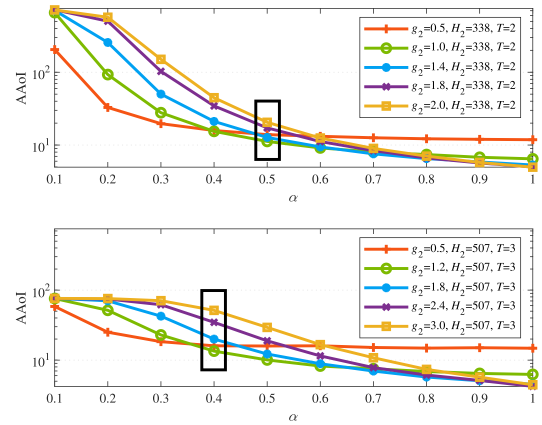


(c) L_3.

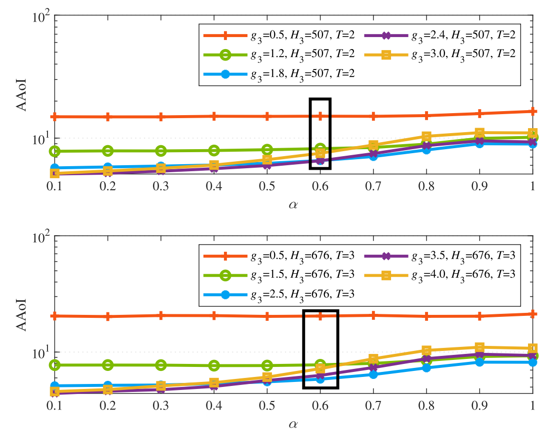
Fig. 10. The variation of AAoI and AFP performance of each group versus SNR in three UTMA schemes, where $g_1 = 0.01$, $g_2 = 0.02$, $g_3 = 0.1$, $L = 169$, $T = 2$, $\alpha = 0.7$ and $\beta = 0.4$.

to offloading, where the UE collisions in N_3 is resolved, the ECO- and ICO-UTMA schemes are beneficial to L_3, providing AFP as low as 10^{-4}, while the IS-UTMA scheme can only provide AFP $10^{-2} \sim 10^{-1}$.

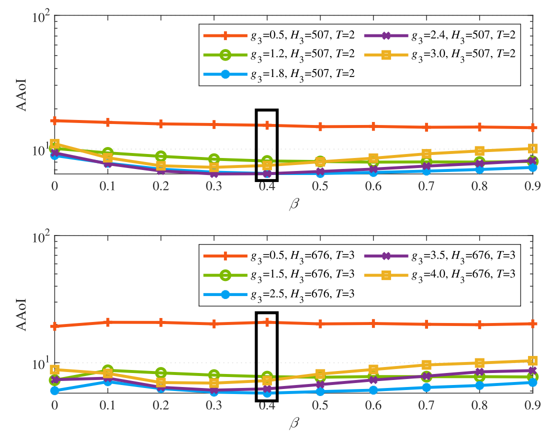
We utilize SNR=20 dB to validate that our UTMA scheme can achieve the $10^{-6} \sim 10^{-5}$ AFP requirements of MCC as shown in Fig. 7, and also simulate the performance of 5 and 10 dB in Fig. 8 and Fig. 9, respectively. It can be observed that the lowest achievable AFP can reach 10^{-3} at SNR=5 dB and 10^{-4} at SNR=10 dB in Fig. 8 and Fig. 9, respectively, and the theoretical derivations agree well with the simulation results. Further, Fig. 10 shows the AAoI and AFP performance of each group versus SNR $\in [-5, 5]$ dB for three UTMA schemes, and as the SNR increases, all AAoI and AFP curves for each group are monotonically decreasing. We can observed that the proposed UTMA schemes can achieve AFP 10^{-2} in SNR $\in [-5, 0]$ dB and 10^{-3} in SNR $\in [0, 5]$ dB, respectively. Thus, our UTMA schemes can guarantee the timely and reliable status updates in S-IoT.



(a) L_2, where $H_2 = \{338, 507\}$ and $T = \{2, 3\}$.



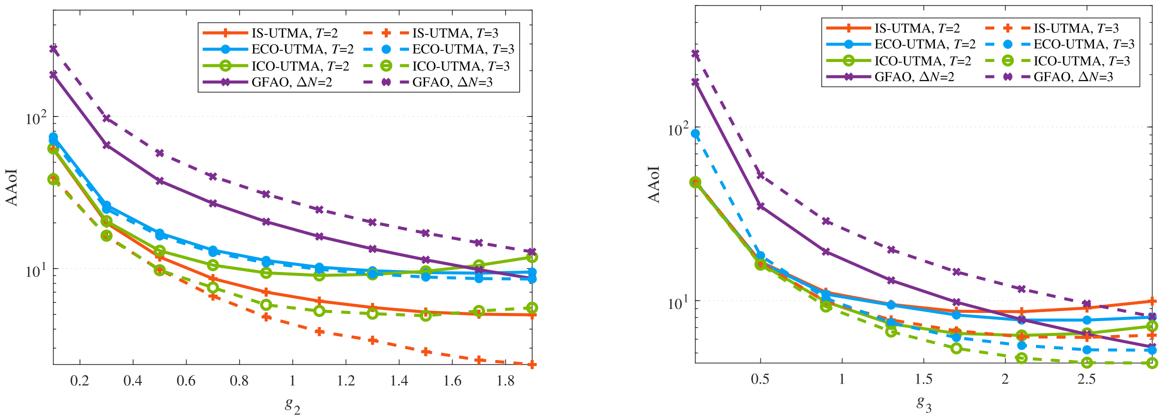
(b) L_3, where $\beta = 0.4$, $H_3 = \{507, 676\}$ and $T = \{2, 3\}$.



(c) L_3, where $\alpha = 0.6$, $H_3 = \{507, 676\}$ and $T = \{2, 3\}$.

Fig. 11. The AAoI of L_2 and L_3 versus α and β under different g_i, H_i and T in ICO-UTMA.

Fig. 11 shows the AAoI performance of L_2 and L_3 versus α and β under different g_i, H_i and T in ICO-UTMA scheme, which can resolve up to LT UE collisions in a time slot. Ideally, the ICO-UTMA scheme allows αLT UEs in L_2 and $(1-\alpha)LT$ UEs in L_3 to access successfully in N_2. Therefore, the AAoI performance of L_2 deteriorates if $K_2 > \alpha LT$, i.e., $g_2 T > \alpha LT$. Two black boxes are illustrated in Fig. 11(a), where the AAoI begins to increase when $g_2 > \alpha T$ with $\alpha = 0.5$ and $T = 2$ or $\alpha = 0.4$ and $T = 3$. Thus, we can derive that g_2 should satisfy $g_2 \leq \alpha T$ to guarantee the AAoI performance



(a) L_2, where $\alpha = 0.6$, $\beta = 0.4$ and $K_3 = 507$.

(b) L_3, where $\beta = 0.6$ and $K_2 = 338$.

Fig. 12. The AAoI of L_2 and L_3 versus load for proposed IS-, ECO-, ICO-UTMA scheme and GFAO scheme, where $H_2 = 338$, $H_3 = 507$, $T = \{2, 3\}$ and $\Delta N = \{2, 3\}$.

TABLE II
THE COMPARISON OF THE MINIMUM ACHIEVABLE AAOI FOR IS-, ECO-, ICO-UTMA SCHEMES AND GFAO SCHEME UNDER DIFFERENT T AND ΔN

Order T	**Number of Time Slots ΔN**	**Load g_i**	**Comparison of AAoI Δ_i**
2	2	$0 \leq g_1 \leq 0.05$	UTMA<GFAO
2	2	$0 \leq g_2 \leq 1.5$	IS-UTMA<ICO-UTMA<ECO-UTMA<GFAO
2	2	$1.5 < g_2 \leq 1.7$	IS-UTMA<ECO-UTMA<GFAO<ICO-UTMA
2	2	$1.7 < g_2 \leq 2$	IS-UTMA<GFAO<ECO-UTMA<ICO-UTMA
3	3	$0 \leq g_2 \leq 2$	IS-UTMA<ICO-UTMA<ECO-UTMA<GFAO
2	2	$0 \leq g_3 \leq 1.9$	ICO-UTMA<ECO-UTMA<IS-UTMA<GFAO
2	2	$1.9 < g_3 \leq 2.1$	ICO-UTMA<ECO-UTMA<GFAO<IS-UTMA
2	2	$2.1 < g_3 \leq 2.5$	ICO-UTMA<GFAO<ECO-UTMA<IS-UTMA
2	2	$2.5 < g_3 \leq 3$	GFAO<ICO-UTMA<ECO-UTMA<IS-UTMA
3	3	$0 \leq g_3 \leq 3$	ICO-UTMA<ECO-UTMA<IS-UTMA<GFAO

of L_2. Similarly, according to the relationship between K_i and $L^{(i)}$, g_3 should satisfy $\beta g_3 L \leq (1-\alpha)LT$ in N_2 and $(1-\beta)g_3 L \leq LT$ in N_3 to guarantee the AAoI performance of L_3. Hence, $g_3 \leq \min\left\{\frac{(1-\alpha)T}{\beta}, \frac{T}{1-\beta}\right\}$. Similarly, four black boxes are illustrated in Fig. 11(b) and Fig. 11(c), where the AAoI begins to increase when $g_3 > \frac{(1-\alpha)T}{\beta}$ with $\beta = 0.4$ and $T = \{2, 3\}$ or $\alpha = 0.6$ and $T = \{2, 3\}$.

Fig. 11 shows the AAoI performance of L_2 and L_3 versus load g_2 and g_3 for the proposed IS-, ECO-, ICO-UTMA schemes and the GFAO scheme under different T and ΔN, where ΔN is denoted as the number of time slots in an access stage for the GFAO scheme. In one time slot, our UTMA scheme can resolve T UE collisions over one pilot through the T-order codebook. Therefore, to be fairly comparison, we assume that ΔN is equal to T, i.e., $\Delta N = T \in \{2, 3\}$. In addition, Table II summarizes the comparison of the minimum achievable AAoI for the IS-, ECO-, ICO-UTMA and the GFAO schemes under T and ΔN in Fig. 11.

Fig. 11 and Table II show that the IS-UTMA scheme can significantly decrease the Δ_2 than that of the GFAO scheme. The numerical results confirm that the IS-UTMA scheme can achieve a minimum Δ_2 approximately $\frac{1}{2}$ and $\frac{1}{6}$ of the GFAO under $T = 2$ and $T = 3$. For example, when $1.5 < g_2 \leq 1.7$ under $T = 2$, the IS-UTMA scheme can achieve $\Delta_2 = 5.0$, but the GFAO scheme can only achieve $\Delta_2 = 10.5$. In addition, when $0 \leq g_2 \leq 2$ under $T = 3$, the IS-UTMA scheme can achieve $\Delta_2 = 2.3$, where the GFAO scheme has $\Delta_2 = 13.0$. In general, larger T allows to resolve more UE collisions in one time slot, and thus $T = 3$ has a lower Δ than $T = 2$. However, when $g_2 \leq 1.5$, in the ECO-UTMA scheme, Δ of $T = 3$ only slightly decreases than that of $T = 2$ for L_2, because only $g_2 L + \beta K_3$ UEs in N_2 to perform access, which leads to the resolvable UE collisions LT is redundant. In addition, in ICO-UTMA scheme, the resolvable UE collisions αLT for UEs in L_2 is redundant, and thus the ICO-UTMA scheme outperforms the ECO-UTMA scheme when $g_2 \leq 1.5$. On the other hand, considering that the UE collisions become serious when $g_2 > 1.5$, the ECO-UTMA scheme outperforms the ICO-UTMA scheme as the ECO-UTMA scheme offer L_2 a degree of freedom in selecting pilots.

Further, to avoid collisions from L_3 and improve Δ_3, the ECO- and ICO-UTMA schemes allow L_3 to offload in L_2 access stage. From Fig. 11 and Table II, we can observe that the ECO- and ICO-UTMA schemes can provide UT protection for L_3 than the IS-UTMA scheme, and significantly decrease Δ_3 than that of the IS-UTMA and GFAO schemes when $g_3 \leq T$. For example, when $0 \leq g_3 \leq 3$ under $T = 3$, the ICO- and ECO-UTMA schemes can achieve $\Delta_3 = 4.4$ and $\Delta_3 = 5.3$, respectively, but the IS-UTMA and GFAO schemes can only achieve $\Delta_3 = 6.1$ and $\Delta_3 = 8.1$, respectively.

Moreover, due to the pre-allocated $\mathbf{D}$ for L_2 and L_3, the ICO-UTMA scheme can relieve the UE collisions between L_2 and L_3. As shown in Table II, the ICO-UTMA scheme outperforms the ECO-UTMA scheme in Δ_3 for providing UT protection under different T.

VI. Conclusion

In this paper, we proposed three UTMA schemes to guarantee the various timeliness requirements for three type MCC UEs in S-IoT, and a multi-dimensional codebook was introduced to resolve UE collisions. Specially, the IS-UTMA scheme allowed each group to perform random access in successive time slot according to its priority, while the ECO- and ICO-UTMA scheme allowed timeliness tolerant group to offload in timeliness critical group access stage to improve the timeliness. To capture the timeliness evaluation of each MCC UE group, we derived the closed-form expression of the AAoI of each UE group for three UTMA schemes by tracing AFP and instantaneous AoI. Moreover, we constructed the joint parameter optimization problems for the ECO- and ICO-UTMA schemes to minimize AAoI under desired AFP requirements. Extensive simulations validated the accuracy of the theoretical derivations, and showed that the proposed UTMA scheme can exhibit significantly lower AAoI than state-of-the-art schemes by resolving UE collisions in code domain.

References

[1] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Mar. 2021.

[2] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare Internet of Things: A survey of emerging technologies," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1121–1167, 2nd Quart., 2020.

[3] Q. Hu, J. Jiao, Y. Wang, S. Wu, R. Lu, and Q. Zhang, "Multitype services coexistence in uplink NOMA for dual-layer LEO satellite constellation," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2693–2707, Feb. 2023.

[4] N. Saeed, A. Elzanaty, H. Almorad, H. Dahrouj, T. Y. Al-Naffouri, and M. S. Alouini, "CubeSat communications: Recent advances and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1839–1862, 3rd Quart., 2020.

[5] *Solutions for NR to support non-terrestrial networks (NTN)*, Sophia Antipolis, France, 3GPP Standard TR 38.821 V16.1.0, May 2021.

[6] *SpaceX Non-Geostationary Satellite System (Attachment A)*, document SAT-MOD-20181108-00083, FCC, 2018. [Online]. Available: https://www.fcc.report/IBFS/SAT-MOD-20181108-00083/1569860.pdf

[7] *Service Requirements for the 5G System*, 3GPP Standard TS 22.261 V19.1.0, Dec. 2022.

[8] *System Architecture for the 5G System (5GS)*, 3GPP Standard TS 23.501 V18.0.0, Dec. 2022.

[9] *Mission Critical Push to Talk (MCPTT); Stage 1*, 3GPP Standard TS 22.179 V17.0.0, Dec. 2019.

[10] *Mission Critical (MC) Data*, 3GPP Standard TS 22.282 V16.4.0, Dec. 2018.

[11] H. Chen et al., "Age-of-information dependent random access for massive IoT networks," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Jul. 2020, pp. 930–935.

[12] T. Yang, J. Jiao, S. Wu, R. Lu, and Q. Zhang, "Grant free age-optimal random access protocol for satellite-based Internet of Things," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3947–3961, Jun. 2022.

[13] J. Jiao et al., "Age-optimal downlink NOMA resource allocation for satellite-based IoT network," *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 11575–11589, Sep. 2023.

[14] H. Hong et al., "Age of incorrect information minimization for semantic-empowered NOMA system in S-IoT," *IEEE Trans. Wireless Commun.*, early access, 2023.

[15] J. Jiao, S. Wu, R. Lu, and Q. Zhang, "Massive access in space-based Internet of Things: Challenges, opportunities, and future directions," *IEEE Wireless Commun.*, vol. 28, no. 5, pp. 118–125, Oct. 2021.

[16] Y. Huang et al., "Age of information minimization for frameless ALOHA in grant-free massive access," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9778–9792, Dec. 2023.

[17] J. Yang et al., "Simplified random access design for satellite Internet of Things with NOMA," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun./Jul. 2021, pp. 128–132.

[18] Y. Liu, W. Yi, Z. Ding, X. Liu, O. A. Dobre, and N. Al-Dhahir, "Developing NOMA to next generation multiple access: Future vision and research opportunities," *IEEE Wireless Commun.*, vol. 29, no. 6, pp. 120–127, Dec. 2022.

[19] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Random access for M2M communications with QoS guarantees," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2889–2903, Jul. 2017.

[20] L. Toni and P. Frossard, "Prioritized random MAC optimization via graph-based analysis," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 5002–5013, Dec. 2015.

[21] J. Jiao, L. Xu, S. Wu, R. Lu, and Q. Zhang, "MSPA: Multislot pilot allocation random access protocol for mMTC-enabled IoT system," *IEEE Internet Things J.*, vol. 8, no. 24, pp. 17403–17416, Dec. 2021.

[22] J. Jiao, L. Xu, S. Wu, Y. Wang, R. Lu, and Q. Zhang, "Unequal access latency random access protocol for massive machine-type communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5924–5937, Sep. 2020.

[23] S. S. Kowshik and Y. Polyanskiy, "Fundamental limits of many-user MAC with finite payloads and fading," *IEEE Trans. Inf. Theory*, vol. 67, no. 9, pp. 5853–5884, Sep. 2021.

[24] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient coded random access for the wireless uplink," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4694–4708, Aug. 2020.

[25] O. Ordentlich and Y. Polyanskiy, "Low complexity schemes for the random access Gaussian channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2528–2532.

[26] Z. Rao et al., "Code-domain collision resolution grant-free random access for massive access in IoT," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4611–4624, Jul. 2023.

[27] *Application for Approval for Orbital Deployment and Operating Authority for the SpaceX GEN2 NGSO Satellite System*, document SAT-LOA-20200526-00055, FCC, 2020. [Online]. Available: https://fcc.report/IBFS/SAT-LOA-20200526-00055/2378669

[28] B. Vucetic, "Propagation," in *Satellite Communications-Mobile and Fixed Services*, M. J. Miller, B. Vucetic, and L. Berry, Eds. Boston, MA, USA: Kluwer, 1993, pp. 57–101.

[29] A. Abdi, W. C. Lau, M. Alouini, and M. Kaveh, "A new simple model for land mobile satellite channels: First- and second-order statistics," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 519–528, May 2003.

[30] A. Abdi et al., "A simple alternative to the lognormal model of shadow fading in terrestrial and satellite channels," in *Proc. IEEE 54th Veh. Technol. Conf. (VTC Fall)*, Atlantic City, NJ, USA, Oct. 2001, pp. 2058–2062.

[31] G. L. Stuber, *Principles of Mobile Communication* Boston, MA, USA: Kluwer, 1996.

[32] L. You, K.-X. Li, J. Wang, X. Gao, X.-G. Xia, and B. Ottersten, "Massive MIMO transmission for LEO satellite communications," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1851–1865, Aug. 2020.

[33] H. S. Jang, S. M. Kim, H.-S. Park, and D. K. Sung, "An early preamble collision detection scheme based on tagged preambles for cellular M2M random access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5974–5984, Jul. 2017.

[34] D. Ansorregui, B. Mouhouche, Y. Oh, and H. Lee, "Joint detection and channel estimation for broadcast OFDM systems using hybrid zadoff Chu pilots," in *Proc. Int. Conf. Commun., Signal Process., Appl. (ICCSPA)*, Feb. 2015, pp. 1–5.

[35] L. Zhen, H. Kong, Y. Zhang, W. Wang, and K. Yu, "Efficient collision detection based on zadoff-chu sequences for satellite-enabled M2M random access," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[36] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2523–2527.

[37] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient random access for the quasi-static fading MAC," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 2768–2772.

[38] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Orlando, FL, USA, Mar. 2012, pp. 2731–2735.

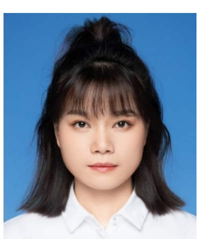
[39] J. Jiao, Z. Ni, S. Wu, Y. Wang, and Q. Zhang, "Energy efficient network coding HARQ transmission scheme for S-IoT," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 1, pp. 308–321, Mar. 2021.

[40] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Product*, 7th ed. New York, NY, USA: Academic, 2007.

[41] J. Sedin et al., "5G massive machine type communication performance in non-terrestrial networks with LEO satellites," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

Shiying Su received the B.S. degree in communication engineering from the Harbin Institute of Technology (Shenzhen), Shenzhen, China, in 2021, where she is currently pursuing the M.S. degree in communication engineering. Her research interests include satellite communications, the age of information, and massive access.

Jian Jiao (Senior Member, IEEE) received the M.S. and Ph.D. degrees in communication engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2007 and 2011, respectively. From 2011 to 2015, he was a Post-Doctoral Research Fellow with the Communication Engineering Research Centre, HIT Shenzhen Graduate School, Shenzhen, China. From 2016 to 2017, he was a China Scholarship Council Visiting Scholar with the School of Electrical and Information Engineering, The University of Sydney, Sydney, Australia. Since 2017, he has been with HIT (Shenzhen), where he has been a Professor with the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology, since 2022. He is currently an Associate Professor with the Peng Cheng Laboratory, Shenzhen. He is also serving as an Editor for *Science China Information Sciences*. His current research interests include semantic communications, satellite communications and networking, and coding techniques. He is also serving as an Editor for *Science China Information Sciences*.



Tao Yang (Student Member, IEEE) received the bachelor's degree in communication engineering from Yanshan University, Qinhuangdao, in 2019, and the master's degree in electronics and communication engineering from the Harbin Institute of Technology (Shenzhen), Shenzhen, China, in 2022, where she is currently pursuing the Ph.D. degree in information and communication engineering with the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology. Her research interests include space communication protocols and semantic communications.

Liang Xu received the B.S. degree in communication engineering from Nanchang University, Nanchang, China, in 2018, and the M.S. degree in information and communication engineering from the Harbin Institute of Technology (Shenzhen), Shenzhen, China, in 2021, where he is currently pursuing the Ph.D. degree in information and communication engineering with the Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology, HIT (Shenzhen). His research interests include space communication protocols, massive access, the age of information, and semantic communications.

Ye Wang (Member, IEEE) received the M.S. and Ph.D. degrees in information and communication engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 2009 and 2013, respectively. From 2013 to 2014, he was a Post-Doctoral Research Fellow with the University of Ontario Institute of Technology, Canada. From 2015 to 2021, he was an Assistant Professor with HIT (Shenzhen). Since 2022, he has been an Associate Professor with the Peng Cheng Laboratory, Shenzhen. His research interests include satellite communications, resource allocation, and mobile internet.

Qinyu Zhang (Senior Member, IEEE) received the bachelor's degree in communication engineering from the Harbin Institute of Technology (HIT) in 1994 and the Ph.D. degree in biomedical and electrical engineering from the University of Tokushima, Japan, in 2003. From 1999 to 2003, he was an Assistant Professor with the University of Tokushima. He has been with HIT (Shenzhen) since 2003, he is currently a Full Professor and the Vice President. He has been awarded the National Science Fund for Distinguished Young Scholars, Young and Middle-Aged Leading Scientist of China, the Chinese New Century Excellent Talents in University, and obtained three scientific and technological awards from governments. His research interests include aerospace communications and networks, wireless communications, and networks.