A Joint Flow Scheduling Scheme for the Uplink of 5G-TSN in Industrial Internet of Things Systems

Jian Zhao, *Student Member, IEEE*, Tao Wang, *Student Member, IEEE*, Haonan Tong, *Member, IEEE*, Nuocheng Yang, *Student Member, IEEE*, Changchuan Yin, *Senior Member, IEEE*, and Dusit Niyato, *Fellow, IEEE*

***Abstract*—As a crucial enabler for the wireless upgrade of the industrial Internet of Things (IIoT), the integration of fifth-generation (5G) wireless network and the wired time-sensitive networking (TSN), i.e., 5G-TSN, is expected to support the deterministic transmission of massive data flows in IIoT systems. Therefore, we focus on the design of a joint flow scheduling scheme for the uplink of 5G-TSN. Specifically, in the 5G part, we propose a novel multiple configured grant (MCG) subframe structure to efficiently exploit the available spectrum resources. Meanwhile, in the TSN part, the age of synchronization (AoS) indicator, which measures the freshness of each flow arriving at the transparent bridge connecting 5G and TSN, determines which flow should be injected into the TSN. Then, to support massive device access, we formulate an optimization problem to maximize the number of successfully transmitted flows while satisfying the quality of service requirements. To solve the non-stationary and non-convex problem, we propose a cooperative multi-agent proximal policy optimization (CMA-PPO) based flow scheduling scheme to determine the MCG configuration with cognitive information of flows. The scheme can improve the number of successfully transmitted flows within the deadline constraints. Extensive numerical simulations demonstrate that, compared with benchmarks, the proposed scheme improves the number of successfully transmitted flows by up to 24.02% and about two times compared with the benchmarks in the 5G-TSN network.**

***Index Terms*—5G-TSN, multiple configure grants, age of synchronization, cooperative multi-agent proximal policy optimization.**

Received 16 December 2024; revised 13 May 2025; accepted 20 June 2025. Date of publication 25 June 2025; date of current version 29 December 2025. This work was supported by the National Natural Science Foundation of China under Grant 62471056, the National Key Research and Development Program of China under Grant 2024YFE0200300, Beijing Natural Science Foundation under Grant L223027, in part by the 111 Project under Grant B17007, in part supported by the National Research Foundation, Singapore and Infocomm Media Development Authority under its Future Communications Research & Development Programme (FCP-NTU-RG-2022-010 and FCP-ASTAR-TG-2022-003), Singapore Ministry of Education (MOE) Tier 1 (RG87/22 and RG24/24), the NTU Centre for Computational Technologies in Finance (NTU-CCTF), and the RIE2025 Industry Alignment Fund - Industry Collaboration Projects (IAF-ICP) (Award I2301E0026), administered by A*STAR, as well as in part supported by Alibaba Group and NTU Singapore through Alibaba-NTU Global e-Sustainability CorpLab (ANGEL). The associate editor coordinating the review of this article and approving it for publication was L. B. Le. *(Corresponding author: Changchuan Yin.)*

Jian Zhao, Tao Wang, Haonan Tong, Nuocheng Yang, and Changchuan Yin are with the Beijing Laboratory of Advanced Information Network, and the Beijing Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: zhaojian@bupt.edu.cn; taowang@bupt.edu.cn; hntong@bupt.edu.cn; yangnuocheng@bupt.edu.cn; ccyin@bupt.edu.cn).

Dusit Niyato is with the College of Computing and Data Science, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

Digital Object Identifier 10.1109/TCCN.2025.3583193

I. Introduction

THE INDUSTRIAL Internet of Things (IIoT) plays a crucial role in the fourth industrial revolution, characterized by the time-sensitive data transmission between numerous industrial sensors/devices and decision-making controler [1]. To enable the deterministic transmission through both wired and wireless networks [2], [3], the integrated fifth-generation (5G) and time-sensitive networking (TSN), i.e., 5G-TSN, has been proposed in the 3rd Generation Partnership Project (3GPP) Release 16 (R16) [4], [5]. Furthermore, 3GPP R17 defines the 5G system as a transparent bridge, which facilitates the flow control at TSN ingress/egress ports by introducing TSN translators (TTs) for the user and control planes [6]. The design of the transparent bridge simplifies the architecture of 5G-TSN and makes it feasible to implement. However, TSN utilizes time-domain access, while 5G performs multiple access in the time-frequency domain. This mismatch results in different scheduling resolutions when flows are injected from 5G into TSN [7]. Therefore, the deterministic transmission in 5G-TSN encounters substantial challenges, particularly in the joint design of 5G access and TSN flow forwarding, as well as the constrained transmission resources within massive access devices. To address these challenges, developing integrated 5G multiple-access and TSN flow forwarding protocols and an efficient resource management scheme is imperative.

Under the 5G-TSN framework, a suite of 3GPP technologies in [8], [9], [10], [11] have been instrumental in facilitating massive ultra-reliable and low-latency communications (mURLLC) for 5G. Specifically, the grant-free (GF) technology is proposed as an innovative alternative to grant-based (GB) access in Long Term Evolution (LTE) [8], which achieves low-latency wireless access by eliminating the need for scheduling requests. To further enhance transmission reliability, 3GPP has proposed the K-repetition GF access approach [9], which achieves reliable sequential transmission by sending a predefined number of consecutive replicas of the same data flow during successive time slots. In response to severe flow collisions caused by massive GF access, non-orthogonal multiple access (NOMA) was integrated into GF access, resulting in the GF-NOMA framework. GF-NOMA framework is designed to reduce collisions during the uplink massive access within 5G systems [10]. Based on GF-NOMA, the multiple configured-grants (MCG) transmission has been proposed in [11], adapting to varying scheduled resources' starting time slot offsets caused by the casual uplinks flows'

arrival time. MCG raises the MCG-GF-NOMA framework, which further reduces the latency of flows at each subframe. The MCG-GF-NOMA framework is advantageous in reducing collision events during the contention period for multiple active devices and decreasing latency by providing multiple start offset options. However, the MCG-GF-NOMA framework merely works on the 5G side network and does not guarantee latency in the TSN side network.

To ensure seamless time-sensitive data flow transmission across the entire 5G-TSN network, further enhancements of TSN are imperative, particularly in the realm of TSN scheduling. The dynamic nature of wireless channels introduces uncertainty of 5G side transmission latency, complicating the pre-configuration of the gate control list (GCL) in TSN switch equipped with the time-aware shaper (TAS) [12], [13]. To satisfy the requirements of real-time seamless flow injection from 5G to TSN without pre-configured information, TSN introduced the cyclic queuing and forwarding (CQF) technique, which implements ping-pong queues by cyclically alternating the sending and receiving status on switch ports [14]. Since the gate operations are executed repetitively, the injections of multiple flows into a single queue is feasible if the queue's capacity is not exceeded, making the temporal order of flows within the queue inconsequential. However, the lack of effective jointly scheduling scheme of 5G-TSN network leads to congestion in dynamic and high-load scenarios [15].

To solve the above issues, this paper aims to maximize the number of successfully transmitted time-sensitive flows by introducing a joint flow scheduling scheme in GB-Enhanced MCG-GF-NOMA based 5G-TSN uplink framework. Our contributions are summarized as follows:

1) *The 5G-TSN framework design:* We consider the deterministic transmission across 5G-TSN with massive device access. To efficiently exploit the available spectrum resource, we propose a novel MCG frame structure. Furthermore, the age of synchronization (AoS) indicator is utilized for the flow injection from the base station (BS) to the TSN. We then formulate a problem to maximize the number of successfully transmitted flows constrained by the QoS requirements.
2) *Multi-agent reinforcement learning with proximal policy optimization:* To address the formulated non-stationary and non-convex problem, we first propose a GB-enhanced MCG-GF-NOMA framework. Then, we utilize the cooperative multi-agent (CMA) reinforcement learning (RL) algorithm to optimize the MCG configuration of MCG-GF-NOMA framework based on cognitive information of flows. We adopt the proximal policy optimization (PPO) to learn the non-stationary environments, which enhances the accuracy of policy gradient calculations and improves efficient convergence.
3) *Simulation evaluations for the priority of the proposed framework:* We conduct extensive simulations to compare the proposed design with existing solutions. The simulation results show that the proposed scheme can improve the number of successfully transmitted flows by up to 24.02% and about two times, compared to benchmarks.

The rest of this paper is organized as follows. Section II overviews the related works and technical background of 5G-TSN. The system model is defined in Section III. Section IV presents the AoS-enhanced flow scheduling and problem formulation. Section V explains the proposed CMA-PPO approach. Section VI presents the simulation evaluations. And finally, we conclude the paper in Section VII.

II. Related Works and Technical Background of 5G-TSN

In this section, we provide the related works and technical background of 5G-TSN. First, we introduce the related open literature about 5G-TSN and summarize the novelty of this work from existing works. Then, we introduce the technical background of the considered system model.

A. Related Works

To guarantee seamless time-sensitive data flows' transmission over a 5G-TSN network, existing studies have focused on the design of scheduling schemes under the standards of the 3GPP and TSN [16], [17], [18], [19], [20]. In [16], the authors introduced a dynamic quality of service (QoS) mapping scheme and an adaptive semi-persistent scheduling (SPS) method for periodic data flows. This method's reliance on prior flow information of the periodic data flows limits its applicability across diverse scenarios. Especially, the SPS method cannot be deployed for event-triggered flows to guarantee deterministic transmissions. To accommodate the transmission needs of massive time-sensitive devices, a cooperative multi-agent double deep Q-network (CMA-DDQN) for MCG-GF-NOMA was proposed to support mURLLC [17], which only considered transmission within the 5G network. Further research on the deployment of MCG-GF-NOMA in the 5G-TSN network, as presented in [18], added the CQF as a paradigm to support deterministic transmission. CQF can realize deterministic flow scheduling within 5G-TSN networks without pre-configured information [19], [20]. The MCG-GF-NOMA framework inefficiently utilizes spectrum resources during the MCG configuration, leading to low deterministic transmission efficiency in the 5G network. Meanwhile, the CQF can only guarantee the deterministic transmission in TSN; the uncertain latency in the 5G side due to the dynamic nature of wireless channels may indirectly affect the deterministic transmission efficiency in TSN. An effective joint scheduling scheme over 5G-TSN networks is urgently needed.

With the present related works, the key points of the proposed joint scheduling scheme over 5G-TSN networks can be summarized as: **(1) GB-enhanced MCG-GF-NOMA framework:** The IIoT time-sensitive flows' deadlines were defined in the range from 1 ms to 5 ms [21], the deadline range can only allow part of the time-sensitive flows to retransmit with GB access scheme in the 5G-TSN networks, which can improve the reliability of transmission without collisions caused by random access in 5G network [9]. We introduce a GB-enhanced MCG-GF-NOMA framework, which can support GF-NOMA and GB-NOMA transmission with an effective MCG pre-configuration (including spectrum

TABLE I
RELATED WORKS COMPARISON NOTE: GREEN SQUARES ■ INDICATE THE LIMITATIONS OF THE WORKS AND BLUE STARS ★ HIGHLIGHT THE ADVANTAGES OF THE WORKS

Scenario	Ref.	5G Side Tech.	TSN Side Tech.	Limitations & Advantages
URLLC over 5G-TSN	[16]	SPS	TAS	■ Low efficiency in massive access ■ Relies on a priori information ★ Low mapping complexity ★ High resource efficiency
mURLLC over 5G	[17]	MCG-GF-NOMA	N/A	■ Only support low latency transmission over 5G ★ High massive access capacity
mURLLC over 5G-TSN	[18]	MCG-GF-NOMA	CQF	■ Without further design of the framework ★ Low energy consumption ★ Robust with environmental non-stationarity
	[22]	GF	CQF	■ Low efficiency in massive access ★ Feature aware flow injection
	This paper	GB enhanced MCG-GF-NOMA	AoS+CQF	★ Enhanced framework design ★ Joint flow scheduling scheme

allocation, transmission repetition value settings of GF-NOMA and GB-NOMA, and starting offset configurations). **(2) AoS-enhanced flow scheduling:** In the TSN, the authors in [22] utilized centralized training with a decentralized execution technique to manage flow injection from the BS's gateway to the TSN with the 5G transmission feature. The per-stream filtering and policing (PSFP) is considered a key technique to manage flow injection from the BS's gateway to the TSN, thus integrating the 5G network with TSN. Furthermore, to support broadband services, the authors in [23] proposed an age-of-information (AoI)-based flow scheduling scheme in TSN, which constructs an augmented logical topology according to the service constraints without considering the subframe structure design of supporting massive deterministic transmission. Age metrics (AoI, AoS, etc.) are defined to record the time elapsed since the generation time, which provides guidelines for future communication system design. Based on these works, we adopt an AoS-enhanced flow scheduling to manage flow injection with PSFP, thus enhancing the joint scheduling scheme of the flows over 5G-TSN. The comparison of the related works is shown in Table I.

B. Technical Background

1) 5G NR Frame Structure: The 5G NR frame structure is a highly adaptable and efficient design that accommodates the varied demands of 5G services, emphasizing low latency and the ability to scale with different operational requirements [17]. Specifically, the frame structure comprises a 10-millisecond radio frame, further divided into ten 1-millisecond subframes. Each subframe is further divided into 2^{μ} slots, where $\mu \in \{0, 1, 2, 3, 4\}$ denotes the numerology factor. The slot duration can then be calculated as $\frac{1}{2^{\mu}}$ ms. Each slot typically consists of 14 orthogonal frequency division multiplexing (OFDM) symbols, whose period is defined as a transmission time interval (TTI), with each symbol duration represented as $\frac{1}{2^{\mu}\times 14}$ ms. As a result of changes in the variable numerology factor μ, the number of TTIs (5G slots) may also vary. For instance, a larger numerology factor results in shorter symbol durations, thus shortening the TTI and enabling support for low-latency services. To further reduce the transmission latency, 5G NR introduced the concept of mini-slots, whose TTI spans just 2, 4, or 7 OFDM symbols,

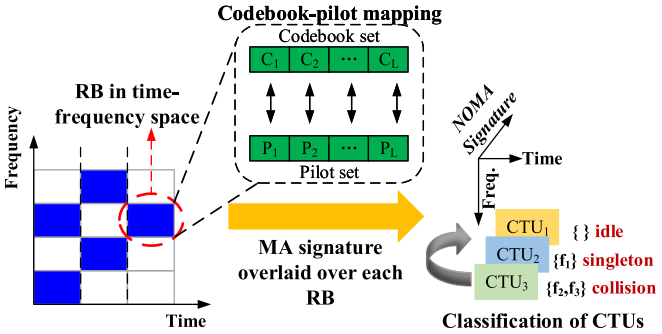


Fig. 1. An example of contention transmission unit (CTU).

instead of 14 symbols for a typical slot, thus reducing the TTI from $\frac{1}{2^{\mu}}$ ms to $\frac{N_{\text{symbol}}}{2^{\mu}\times 14}$ ms, where N_{symbol} denotes the number of symbols within a mini-slot [24]. For example, when $\mu = 2$ and $N_{\text{symbol}} = 7$, the TTI is equal to 125 μs. This feature is essential for meeting strict latency requirements of specific applications, such as autonomous driving and industrial automation.

2) MCG-GF-NOMA: GF technology has been introduced to address the latency challenges of time-sensitive 5G wireless communication services. This technology permits uplink transmissions from devices to the BS in an "arrive-and-go" fashion, bypassing the need for scheduling requests and resource grants and thereby significantly reducing transmission latency [10].

Considering the limited time-frequency resource units, the simultaneous access and transmission for massive devices is practically infeasible, thus prolonging the transmission latency. NOMA has been introduced to address the massive transmission of low latency by leveraging the non-orthogonal code domain resources to tackle this issue. As depicted in Fig. 1, in the three-dimensional time-frequency-code domain, the smallest transmission unit that devices compete for is termed a contention transmission unit (CTU), which comprises a multiple access (MA) physical resource and an MA signature [25]. Each MA physical resource spans a set of time-frequency resource blocks (RBs). The MA signatures represent a set of pilot sequences for channel estimation or device activity detection and a codebook designed for robust data transmission and interference mitigation. The mapping

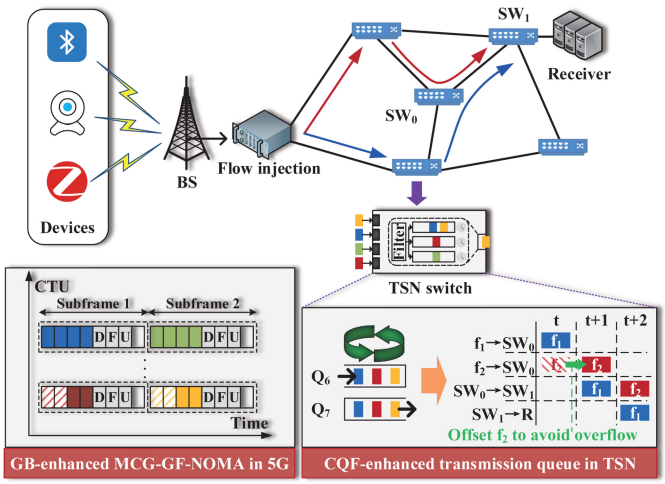


Fig. 2. Illustration of the integration of 5G and TSN, highlighting data flows from IIoT devices to the BS, then injected into the TSN.

between the pilot sequences and the codebook determines which CTUs each device can utilize for random access.

Each pilot sequence l is uniquely paired with a specific codebook, serving as the device's signature. NOMA enables multiple devices with distinct pilot sequences and codebooks to transmit over the same RB, thereby increasing the number of accessible devices without expanding time-frequency resources. We use $N_{\mathrm{CTU}} = F \times L$ to denote the number of CTUs in each subframe, where F and L denote the number of RBs in each subframe and the number of CTUs overlaid in each RB, respectively. Despite the increased CTU number brought by NOMA, under the GF-NOMA mechanism, access collisions occur when multiple non-collaborative devices select the same pilot sequence and the corresponding codebook, i.e., the same CTU. Thus, in the GF-NOMA transmission, we classify the CTUs into three categories: a) *idle*: CTUs not chosen by any device; b) *singleton*: CTUs chosen by only one device; and c) *collision*: CTUs chosen by two or more devices.

Besides, as defined in 3GPP Release 15, since subframe is the smallest scheduling unit, a device configured with a single CG must wait until the next subframe to transmit if data arrives during the transmission of a subframe, which can introduce additional latency for time-sensitive services [26]. To mitigate this issue, the MCG concept was proposed in 3GPP Release 16, enabling support for various start offsets in resource allocation so as to reduce the latency caused by waiting for the next subframe [11].

3) CQF-Based Forwarding and Per-Stream Filtering and Policing: CQF is an integral technology within the TSN framework, primarily operational within network switches. For each TSN gateway and switch, a total of eight queues are designated, with two specifically configured as dual ping-pong queues (denoted as Q6 and Q7 in Fig. 2). The switch meticulously manages these queues in a cyclical manner, orchestrating the en-queueing and de-queueing processes. This ensures that data flows enqueued in one time slot are promptly dequeued in the subsequent time slot and concurrently received by the downstream TSN switch. This cyclical mechanism is pivotal in maintaining transmission latency within a predefined threshold. Furthermore, to optimize queue utilization within the TSN network, especially for the seamless integration of 5G and TSN transmissions, the system incorporates the PSFP mechanism [27]. One key problem during the integration of 5G and TSN transmissions is that the latency in the 5G network is not deterministic due to the nature of the dynamic channel conditions. Before flow injection, the BS acts as a TSN logical gateway to re-encapsulate the flows from 5G devices [19]. The PSFP permits stream filtering, policing, and enqueue decisions through three components in order, i.e., the stream filters, stream gates, and flow meters. Stream filters classify flows into streams using a stream handle or the Priority Code Point (PCP) in the virtual LAN (VLAN) tag, supported by flow classification rules that combine fields such as MAC addresses, VLAN ID, and priority to map each stream to a dedicated stream gate. Stream gates regulate flow enqueuing via an open/closed state and an Internal Priority Value (IPV), where open gates permit flows to enter transmission queues aligned with the current IPV, while closed gates discard flows, with both states and IPVs dynamically updated via a cyclic time-based schedule. The flow meters enforce traffic specifications by marking flows as compliant or non-compliant and dropping excess traffic, while also enabling policed shaping through priority adjustments. At the end of the PSFP process, the flows are enqueued in the transmission queues according to the IPV of the traversed stream gate. The proposed AoS-enhanced flow scheduling manages the IPV of each flow, thus deciding which flow can be injected into the TSN and forwarded with CQF. PSFP serves as a preemptive measure to mitigate the risk of overflow in CQF-based forwarding. It operates by managing the data flows before they are injected into the TSN network, thereby enhancing the overall efficiency and reliability of the transmission process [28].

III. System Model

We consider a 5G-TSN network for the real-time transmission of an IIoT system as shown in Fig. 2, where event-triggered time-sensitive flows are generated by M devices, which are wirelessly transmitted to a BS, then injected into the TSN, and finally forwarded to the destined receiver through the wired TSN network. Meanwhile, we configure the TSN slot size to match 5G for proper slot alignment, as 5G URLLC features shorter mini-slots for effective TSN integration [18], i.e., when $\mu = 2$ and $N_{\mathrm{symbol}} = 7$, the 5G TTI is equal to 125 μs, we set TSN slot as 125 μs too. Moreover, the set of 5G-TSN networks is considered to meet the clock synchronization requirements of the BS, gateway, and TSN switches [29]. The flow generated by device m is denoted as f_m, which has a strict deadline T_m^{D}. The GB-enhanced MCG-GF-NOMA mechanism [11] is adopted for the multiple access and data transmission, the schematic frame structure of which is plotted on the bottom left of Fig. 2. Then the BS injects the flows into the TSN, where the flows are forwarded to the destined receiver based on the CQF scheduling. The main variables of the system are summarized in Table II. We represent the slot index as $t \in \{1, \ldots, T\}$. The

TABLE II
SYSTEM NOTATION LIST

Notation	Description	Notation	Description
d_m	The distance between the m-th device and the BS	D_m	The total transmission latency of device m
D_m^{5G}	The 5G side transmission latency of device m	D_m^{TSN}	The TSN side transmission latency of device m
$\delta_m(\tau)$	Indicator of whether flow f_m successfully transmitted	$\mathcal{E}$	The set of connections between any two switches
$\mathcal{E}(f_m)$	The transmission path set of flow f_m	η	The path-loss attenuation factor
f_m	The flow of device m	$g_m(t)$	The AoS evolution of device m
$\mathcal{G}$	The TSN topology	$\gamma_{m,n,i}^{\mathrm{SINR}}(k)$	The SINR of device m
γ_{th}	The SINR threshold	$h_{m,n}$	The small-scale fading of device m at RB n
L_m	The large-scale path loss of device m	λ	The carrier wavelength
M	The number of devices	μ	The 5G NR numerology factor
$N_i^{\mathrm{CTU}}(\tau)$	The CTU number of CG_i in the τ-th subframe	$N_i^{\mathrm{s}}(\tau)$	The start offset of CG_i in the τ-th subframe
$N_i^{\mathrm{r}}(\tau)$	The repetition slot number of CG_i in the τ-th subframe	N_i^{GB}	The GB repetition slot number of CG_i in the τ-th subframe
N_{CG}	The number of CGs	$N_{m,n,i}^{\mathrm{sc}}(k)$	The flow number of $\mathcal{N}_{m,n,i}^{\mathrm{sc}}(k)$
N_{suc}^{τ}	Sum number of the successfully transmitted flows during the τ-th subframe	$\mathcal{N}_{n,i}^{\mathrm{cc}}(k)$	The set of devices that have chosen the collision CTUs at the RB n
$\mathcal{N}_{m,n,i}^{\mathrm{sc}}(k)$	The set of devices that have chosen the singleton CTUs for device m at the RB n	$P_{m,n,i}(k)$	The received power of m-th device at the k-th reception of CG_i at RB n
$o_{m,e}(t)$	Indicator of flow f_m occupying the forwarding resource	P	The transmitting power of each device
π_m	The flow injection indicator	Q_{size}	The maximum forwarding capacity of each switch
t	The time slot count	τ	The subframe count
t_{slot}	The time length of a slot t	T_m^{D}	The required deadline constraint of device m
T_m^{trans}	The GF repetition number when the m-th device decodes successfully	T_m^{retrans}	The GB repetition number when the m-th device decoded successfully
T_m^{RTT}	The RTT slot number of ACK/NACK feedback	T_m^{wait}	The wait time slot number of the flow of device m
T_m^{Pre}	Required TSN pre-configured transmission time	$\mathcal{V}$	The set of TSN switches

latency consumed by f_m can be represented as

$$D_m = D_m^{\mathrm{5G}} + D_m^{\mathrm{TSN}}, \tag{1}$$

where D_m^{5G} and D_m^{TSN} are the 5G side transmission latency and the TSN side transmission latency, respectively.

The AoS metric is applied in this paper to capture the unpredictable latency of wireless transmission [30], which measures the time elapsed since the generation time slot of a flow that has not arrived at the receiver. We use the index r_m to indicate whether f_m has been received, with $r_m = 1$ indicating that f_m is received, and $r_m = 0$ otherwise. We denote the generation slot index of f_m as t_m^{g}. Thus, the AoS (in the unit of a slot duration) evolution of f_m at slot t ($t > t_m^{\mathrm{g}}$) can be represented as

$$g_m(t) = \begin{cases} g_m(t-1) + 1, & r_m = 0, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $g_m(t_m^{\mathrm{g}}) = 0$.

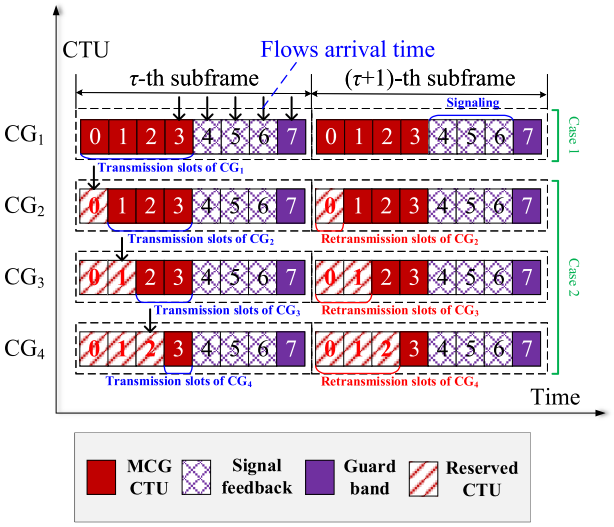


Fig. 3. Proposed subframe structure under the GB-enhanced MCG-GF-NOMA framework.

A. GB-Enhanced MCG-GF-NOMA Subframe Structure Design

As shown in Fig. 3, we consider $\mu = 2$, i.e., the time slot number of a subframe T_{frame} is 8. Each subframe consists of 4 transmission CTU slots, 3 signaling slots, and a guard band slot. Slots from 0 to 3 are defined as CTUs, and the following 3 slots are defined as signaling slots, which are used for the BS to transmit the ACK/NACK feedback and control information for the devices. At the end of each subframe, one slot is defined as a guard band to prevent the transmission flows and signaling flows from interfering with each other's transmission. The CTUs are defined as MCG CTU and reserved CTU, respectively. CGs have been classified into 4 different types. The four subframe structures are designed to transmit flows generated before different start offsets. In each CG, a start offset is defined thus that the flows generated at the device can be transmitted in MCG CTUs immediately in the next slot with "arrive-and-go" fashion. Meanwhile, the reserved CTUs can retransmit the flows that

are not successfully transmitted in the previous subframe, i.e., the reserved CTUs in the CG_2, CG_3, and CG_4 can be used to retransmit the flows which are under the deadline constraint in a GB manner. The flow transmission cases can be summarized as follows:

- **Case 1:** When the flows are generated at the end of the MCG CTUs, signaling slots, and guard band slot, the flows cannot be transmitted at the next CTU slot immediately. All the flows are postponed to the next subframe for transmission. The flows are transmitted from the start of the next subframe's MCG CTUs, i.e., the CG_1 is implemented to transmit the flows.
- **Case 2:** When flows are generated in the first three slots of the MCG CTU slots, the flows are transmitted by the CG_2, CG_3, and CG_4 according to the generation order.

The MCG-GF-NOMA framework can significantly decrease the signaling overhead. Still, the reliability cannot be guaranteed for each CG during GF access (e.g., CG_4 can only transmit one repetition, the transmission fails when a collision or decoding failure happens). Thus, we apply GB scheduling to improve transmission reliability. When the flow of device *m* allocated to CG_i is not successfully transmitted, and the deadline allows the retransmission, then the flow will be retransmitted in the reserved CTUs with an allocated CG_i. All the subframes in the system have the same number of slots all the time. Thus, for ease of presentation, we denote the *i*-th CG configuration information in the τ-th subframe by $\mathrm{CG}_i(\tau) = \{N_i^{\mathrm{CTU}}(\tau), N_i^{\mathrm{s}}(\tau), N_i^{\mathrm{r}}(\tau), N_i^{\mathrm{GB}}(\tau)\}$, where $N_i^{\mathrm{CTU}}(\tau)$ is the configured CTU number of *i*-th CG in the τ-th subframe, $N_i^{\mathrm{s}}(\tau)$ is the configured start offset of the *i*-th CG in the τ-th subframe, $N_i^{\mathrm{r}}(\tau)$ is the configured repetition slot number of *i*-th CG in the τ-th subframe, and N_i^{GB} is the configured GB repetition slot number of *i*-th CG in the τ-th subframe. For each flow, the transmission may have two stages: the GF-NOMA and GB-NOMA transmission stages. When a flow *m* is generated, the flow *m* is transmitted in the GF-NOMA manner. The BS randomly chooses one CTU to access the device *m* and transmit flow *m* in this subframe. However, a collision will occur when more than one device selects the same CTU. If the flow is not successfully transmitted, and the deadline allows the retransmission, i.e., we consider T_m^{pre} as the theoretical TSN side latency of the flow *m*, which can be calculated by the hop number between the BS and receiver. If $(g_m(t) + N_i^{\mathrm{GB}})t_{\mathrm{slot}} + T_m^{\mathrm{pre}} \leq T_m^{\mathrm{D}}$, then the flow *m* is transmitted in the GB-NOMA manner, otherwise, the flow *m* is not retransmitted. The BS allocates the reserved CTU via polling to each flow for transmission by using GB-NOMA. In particular, the GB-NOMA can avoid collisions occurred during random access. The detail of the transmission procedure with the GB-enhanced MCG-GF-NOMA framework is shown in Algorithm 1.

Algorithm 1 Transmission Procedure With the GB-Enhanced MCG-GF-NOMA Framework in 5G Network

```
for m do
    Allocate the device m to CG_i based on the generation time of flow m
    The BS randomly chooses one CTU of CG_i to transmit flow m
    if Flow m is successfully transmitted then
        Transmission procedure of flow m finishes
    else
        if (g_m(t) + N_i^GB) t_slot + T_m^pre ≤ T_m^D then
            The BS allocates the reserved CTU of CG_i via polling to transmit flow m
        else
            Flow m is not retransmitted, and the transmission procedure of flow m ends
        end if
    end if
end for
```

B. 5G Side Latency Model

For the 5G side transmission, each device has two ways to transmit the flow: (1) the device transmits the flow in the GF-NOMA manner with a certain CG; (2) if the transmission fails in the GF-NOMA manner and does not exceed the latency constraints, the flow is re-transmitted in the GB-NOMA manner at the next subframe. Then, we will introduce the details of the two ways.

1) Successful Transmission During GF Stage: When the flow of device *m* is transmitted successfully during the GF stage, the latency of 5G side transmission for device *m* to transmit the flow in the GF stage is given by

$$D_m^{5\mathrm{G}} = \left(T_m^{\mathrm{wait}} + T_m^{\mathrm{trans}}\right) t_{\mathrm{slot}}, \tag{3}$$

where T_m^{wait} is the time slot from the generation time to the transmission time of the flow of device *m*, and T_m^{trans} is the GF-based transmission repetition number that the flow of the *m*-th device is decoded successfully, which is determined by when the GF access repetition number satisfies flow decoding of the flow of *m*-th device. The waiting time slot for the flow of *m*-th device is determined by

$$T_m^{\mathrm{wait}} = \begin{cases} 1, & \left(t_m^{\mathrm{g}} \bmod T_{\mathrm{frame}}\right) < T_{\mathrm{CTU}}, \\ \left\lceil \frac{t_m^{\mathrm{g}}}{T_{\mathrm{frame}}} \right\rceil T_{\mathrm{frame}} - t_m^{\mathrm{g}}, & \text{otherwise}, \end{cases} \tag{4}$$

where T_{CTU} is the maximum CTU transmission slot number in a subframe.

2) Successful Transmission During GB Stage: When the flow of device *m* is transmitted successfully in the GB stage, the latency of 5G side transmission for device *m* to transmit the flow in the GB stage is given by

$$D_m^{5\mathrm{G}} = \left(T_m^{\mathrm{wait}} + T_m^{\mathrm{RTT}} + T_m^{\mathrm{retrans}}\right) t_{\mathrm{slot}}, \tag{5}$$

where T_m^{RTT} is the round trip time (RTT) slot number of ACK/NACK feedback, which is set as 4 in Fig. 3. T_m^{retrans} is the GB transmission repetition number that satisfies flow decoding of the flow of *m*-th device. The RTT time slot number for the flow of *m*-th device is given by

$$T_m^{\mathrm{RTT}} = T_{\mathrm{frame}} - (t_m^{\mathrm{g}} \bmod T_{\mathrm{frame}}). \tag{6}$$

C. Successive Interference Cancellation Decoding

The BS utilizes the successive interference cancellation (SIC) technique to decode the flows transmitted by the devices. In accordance with the principles of NOMA, during each iterative stage of SIC, the BS initially decodes the flow characterized by the strongest received power and subsequently subtracts the successfully decoded signal from the composite received signal. The large-scale path loss is modeled as the free space path-loss (FSPL) model [31], which is given by

$$L_m = \frac{\lambda^2}{16\pi^2 d_m^2}, \tag{7}$$

where d_m is the distance between the m-th device and the BS. λ denotes the carrier wavelength. Without loss of generality, we consider the small-scale fading of device m at RB n as $h_{m,n} \sim \mathcal{CN}(0,1)$.

The BS received power from m-th device at the k-th reception of CG i at RB n is given by

$$P_{m,n,i}(k) = P|h_{m,n}(k)|^2 L_m^{-\eta}, \tag{8}$$

where P denotes the transmitting power of each device and η is the path-loss attenuation factor.

We denote $\mathcal{N}^{\mathrm{sc}}_{m,n,i}(k)$ as the set of devices that choose the singleton CTUs of CG i on the n-th RB and decoded in the SIC iteration behind flow m. The flow number of $\mathcal{N}^{\mathrm{sc}}_{m,n,i}(k)$ is given by $N^{\mathrm{sc}}_{m,n,i}(k)$. Meanwhile, we denote $\mathcal{N}^{\mathrm{cc}}_{n,i}(k)$ as the set of devices that choose the collision CTUs of CG i on the n-th RB. We assume that the received power of each flow at k-th reception of CG i at RB n follows $P_{1,n,i}(k) \geq P_{2,n,i}(k) \geq \cdots \geq P_{N^{\mathrm{sc}}_{m,n,i}(k),n,i}(k)$. For each iteration of SIC decoding, the singleton CTU with the strongest received power is decoded by treating the received but not decoded powers of other CTUs over the same RB as the interference. Thus, the received signal-to-interference plus noise ratio (SINR) of the flow of device m, which is allocated to CG i at RB n, is given by

$$\gamma^{\mathrm{SINR}}_{m,n,i}(k) = \frac{P_{m,n,i}(k)}{B^{\mathrm{RB}} N_0 + \sum_{m' \in \mathcal{N}^{\mathrm{sc}}_{m,n,i}(k)} P_{m'}(k) + \sum_{m'' \in \mathcal{N}^{\mathrm{cc}}_{n,i}(k)} P_{m''}(k)}, \tag{9}$$

where B^{RB} and N_0 represent the bandwidth of an RB and the unilateral power spectral density of the white noise, respectively.

At each iterative stage, SIC decoding is successful when the SINR $\gamma^{\mathrm{SINR}}_{m,n,i}(k)$ at the repetition is larger than or equal to the SINR threshold γ_{th}; otherwise, SIC decoding fails. The detailed procedures of SIC decoding are outlined in Algorithm 2.

Algorithm 2 SIC Decoding Procedure

```
Input: γ_th.
Set m = 1
while m ≤ N^sc_{m,n,i}(k), do
    Based on (9), calculate the SINR γ^SINR_{m,n,i}(k) of flow m of CG i on the n-th RB at k-th repetition.
    if γ^SINR_{m,n,i}(k) ≥ γ_th then
        Flow of device m is decoded successfully, set m = m + 1
    else
        Flow m cannot be successfully decoded
        break
    end if
end while
```

D. TSN Side Latency Model

On the TSN side, to reduce the sojourn time, the flows will be immediately injected into the TSN network and delivered to its receiver once successfully received by the BS. In the multi-hop TSN network, latency is deterministic and foreseeable, relying only on the pre-configured time slot length and routing path. Meanwhile, we configure the TSN slot size to match 5G for proper slot alignment, as 5G URLLC features shorter mini-slots for effective TSN integration [18]. We consider the TSN topology as a directed graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ is the set of TSN switches and $\mathcal{E}$ represents the set of connections between any two switches. The transmission path of each flow f_m is described as a set of sequential links as $\mathcal{E}(f_m) = (e_1, \ldots, e_{\mathrm{hop}_m})$, where hop_m is the hop length of flow f_m. Each flow is injected between different switches periodically with the CQF protocol. For example, a flow generated in the first slot can be forwarded to the switch in the second slot. This transmission characteristic limits the offset time into a reasonable range, which guarantees the deterministic transmission latency. The TSN side latency of f_m from the BS to the receiver can be expressed based on the deduced forwarding time as follows:

$$D^{\mathrm{TSN}}_m = \sum_{e \in \mathcal{E}(f_m)} t_{\mathrm{slot}}, \quad \forall f_m \in \mathcal{F}, e \in \mathcal{E}(f_m), \tag{10}$$

where $\mathcal{E}(f_m)$ is the set of paths between switches that f_m passes through, and e is each element of $\mathcal{E}(f_m)$.

In each time slot, the length of each queue should not exceed the maximum forwarding capacity of switch Q_{size}. Thus, we define the indicator $o_{m,e}(t)$ to represent whether flow f_m occupies the queue forwarding resource at the e-th switch during time slot t. Then the queue forwarding resource constraint is written as

$$\sum_{m=1}^{M} o_{m,e}(t) \leq Q_{\mathrm{size}}, \quad \forall f_m \in \mathcal{F}, \forall e \in \mathcal{E}(f_m), \forall t. \tag{11}$$

IV. AoS-Enhanced Flow Scheduling and Problem Formulation

A. AoS-Enhanced Flow Scheduling

Due to the uncertainty of the arrival time of each event-triggered flow on the 5G side and the time-varying nature of the wireless channel, the flows cannot arrive at the BS within a deterministic 5G side latency. In the BS gateway connected with the TSN network, we implement PSFP [32] in the system to deploy filtering and control schemes for

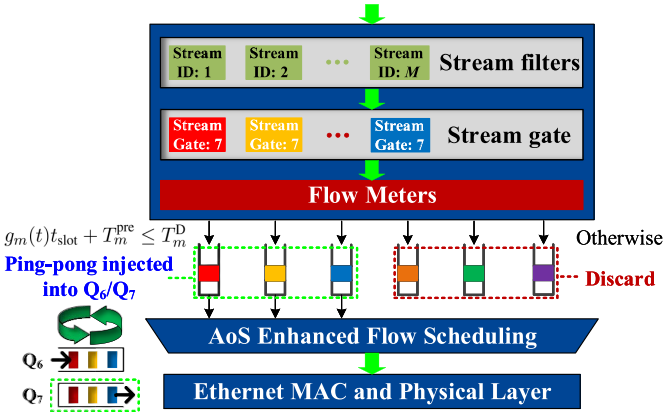


Fig. 4. Forwarding process in gateway combining the PSFP and the AoS-enhanced flow scheduling.

Algorithm 3 AoS-Enhanced Flow Scheduling

Initialize the flow injection policing
for t **do**
  Cognize the information of the arrival flows at t
  Calculate T_m^{pre} of each arrival flow f_m
  for m **do**
    if Flow f_m arrives the BS at t, and $g_m(t)t_{\text{slot}} + T_m^{\text{pre}} \le T_m^{\text{D}}$ **then**
      Inject flow f_m into the TSN network, and set $\pi_m = 1$
    else
      The flow f_m is discarded, and set $\pi_m = 0$
    end if
  end for
end for

each data flow to ensure that the input flow follows a new proposed flow scheduling. In the TSN network, each flow f_m once successfully pre-configured, the theoretical transmission latency on the TSN side can be calculated as

$$T_m^{\text{Pre}} = \sum_{e \in \mathcal{E}(f_m)} t_{\text{slot}}. \tag{12}$$

As shown in Fig. 4, we propose the AoS-enhanced flow scheduling in PSFP, which is deployed in the BS to manage the flow injection to the TSN. Since the CQF is implemented in the TSN, the stream filters in PSFP allow flows to be injected into a single stream gate in a ping-pong manner (injected into Q6 and Q7 in turn). The stream filter classifies each flow into a stream to a dedicated stream gate, and the stream gate state (i.e., open or closed) depends on the AoS of each flow. The stream filter also maps each flow to the flow meters, which are in charge of marking the flows, on the basis of defined stream specifications, and discarding them if the flow does not comply with these specifications. To render more flows transmitted under the deadline constraint in the TSN network, we propose the AoS-enhanced flow scheduling to record the real-time AoS metrics of each flow m and apply AoS into the flow meters to indicate which flow to inject into the TSN network. When a flow arrives at the gateway, the AoS $g_m(t)t_{\text{slot}}$ can be denoted as the 5G side latency of flow m, then the theoretically TSN side latency T_m^{pre}, can be calculated by the hop number between the BS and receiver. Thus, the 5G-TSN theoretical latency can be calculated as $g_m(t)t_{\text{slot}} + T_m^{\text{pre}}$. If $g_m(t)t_{\text{slot}} + T_m^{\text{pre}}$ is within the deadline of flow m, then the flow m complies with the specifications of the proposed AoS-enhanced flow scheduling and injected into the allocated TSN queue; otherwise, the flow m is discarded. We use π_m to indicate whether flow m is injected into the TSN network, which is defined as

$$\pi_m = \begin{cases} 1, \ g_m(t)t_{\text{slot}} + T_m^{\text{pre}} \le T_m^{\text{D}}, \\ 0, \ \text{otherwise}. \end{cases} \tag{13}$$

The detail of the AoS-enhanced flow scheduling is shown in Algorithm 3. It is worth noting that a successfully transmitted flow under the deadline constraint can be affected by the 5G side sampling decision. In the following, we describe the problem formulation of the 5G side sampling decision.

B. Problem Formulation

We aim to maximize the successfully transmitted flows. We use δ_m to indicate whether the flow of device m is successfully transmitted within the deadline constraint. The flow is generated at the device m during subframe τ in the 5G-TSN network. δ_m is given by

$$\delta_m(\tau) = \begin{cases} 1, \ \text{if } D_m \le T_m^{\text{D}}, \\ 0, \ \text{otherwise}. \end{cases} \tag{14}$$

Then the sum number of the successfully transmitted flows during the τ-th subframe is obtained as

$$N_{\text{suc}}^{\tau} = \sum_{m=1}^{M} \delta_m(\tau). \tag{15}$$

To jointly design the 5G network and TSN, we formulate an optimization problem whose objective is to maximize the number of long-term successfully transmitted time-sensitive flows, which is given by

$$\max_{\mathcal{A}_\tau} \quad \sum_{\tau=1}^{\infty} \mathbb{E}(N_{\text{suc}}^{\tau}) \tag{16}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_{\text{CG}}} N_i^{\text{CTU}}(\tau) = N_{\text{CTU}}, \forall \tau, \tag{16a}$$

$$\sum_{m=1}^{M} o_{m,e}(t) \le Q_{\text{size}}, \forall e, \forall t, \tag{16b}$$

where N_{CG} is the number of CGs, $\mathcal{A}_\tau = \{\text{CG}_1(\tau), \ldots, \text{CG}_{N_{\text{CG}}}(\tau)\}$ is the sampling indicator set of each CG, $\text{CG}_i(\tau) = \{N_i^{\text{CTU}}(\tau), N_i^{\text{s}}(\tau), N_i^{\text{r}}(\tau), N_i^{\text{GB}}(\tau)\}$ at each subframe τ, $N_i^{\text{CTU}}(\tau)$ is the configured MCG CTU number of i-th CG in the τ-th subframe, $N_i^{\text{s}}(\tau)$ is the configured starting offset of the i-th CG in the τ-th subframe, $N_i^{\text{r}}(\tau)$ is the configured repetition slot number of the i-th CG in the τ-th subframe, and N_i^{GB} is the configured reserved CTU

repetition slot number of the *i*-th CG in the τ-th subframe. Equation (16a) indicates that all the CGs are constrained by the spectrum resource, i.e., N_{CTU} is the total number of the configured MCG CTUs in each τ-th subframe. Equation (16b) is the flow forwarding resource constraint, which indicates that the size of the total flows forwarded within each time slot *t* should be no more than the maximum capacity per time slot of each switch.

Equation (9) indicates that the counts of idle, singleton, and collision CTUs determine the SINR for each device *m*, thereby affecting the number of successfully transmitted flows per subframe τ. In consequence, the total successfully transmitted flow number in (16) depends on the sampling indicator of each kind of CG. Moreover, (16a) and (16b) show that, to transmit time-sensitive flow over the 5G-TSN network within the deadline constraint successfully, the flow must satisfy a series of requirements.

Since the BS does not fully know all the information of the communication scenario, scheduling policies $\mathcal{A}_\tau$ must be optimized based on partial observation. As a result, a flow's state transition is uncertain due to other flows and the varying environment, which is hard to solve within polynomial time using traditional optimization methods [33], [34], [35]. Moreover, conventional gradient-based optimization techniques are impractical because the search space expands as the number of parameters increases. One promising solution is the RL-based algorithm that can adapt to the dynamic environment. However, since each CG in the BS has partial observation from the environment, the CMA algorithm should be used to address complex cooperative control problems in the non-stationary environment [36]. Compared to traditional RL-based CMA algorithms, i.e., CMA-DDQN, CMA-PPO can reduce the complexity of hyperparameter tuning while ensuring training stability [37].

V. Proposed Cooperative Multi-Agent Proximal Policy Optimization Approach

In this section, we introduce a CMA-PPO based approach to tackle the problem (16) by optimizing $\mathcal{A}_\tau$, which decomposes the selections in high-dimensional action space into multiple parallel sub-tasks. Each CG in the BS can only partially transmit flows and observe flow information at each subframe τ, leading to suboptimal rewards. To solve this problem, we propose a CMA approach where each CG acts as an agent at the BS. These agents select appropriate actions $a \in \mathcal{A}_\tau$ based on the partial observation $s \in \mathcal{S}$. The goal is to manage the deployment of each CG effectively, thereby maximizing the successfully transmitted time-sensitive data flows. At the beginning of each subframe τ, each agent of the CGs first observes the different CG's state which contains a set of previous state observations $\mathcal{O}_\tau^{\mathrm{old}} = \{s_1, \ldots, s_{\tau-1}\}$, i.e., the state information from the first subframe to the $\tau-1$-th subframe. Based on the observation s_τ, each agent chooses $a_\tau \in \mathcal{A}_\tau$ and the environment transits to a state $s_{\tau+1}$ and obtains reward $r_{\tau+1}$ as the feedback. Then, the PPO agent updates the policy function according to the reward to achieve the optimization goal. Next, the proposed CMA-PPO-based algorithm's action, state, and reward design are specified. Finally, we detail the proposed CMA-PPO approach and analyze its complexity and adaptability.

A. States Design

The state is a representation of the current environmental situation. The state of each subframe τ is given by $s_\tau = \{N_\tau^{\mathrm{cc}}, N_\tau^{\mathrm{sc}}, N_\tau^{\mathrm{ic}}, N_\tau^{\mathrm{suc}}, N_\tau^{\mathrm{fail}}\}$. The five parameters are given by: the number of the collision CTUs, N_τ^{cc}, the number of the singleton CTUs, N_τ^{sc}, the number of the idle CTUs, N_τ^{ic}, the number of flows that have been successfully decoded at the BS and transmitted to the receiver under the deadline constraint, N_τ^{suc}, and the number of flows that have not been successfully decoded at the BS or not transmitted to the receiver under the deadline constraint, N_τ^{fail}.

B. Actions Design

CGs of the MCG-GF-NOMA framework are configured at each subframe τ to access massive devices with random flows. To improve the long-term number of successfully transmitted time-sensitive flows, the BS cognizes the CTU observation $\mathcal{O}_\tau^{\mathrm{old}}$ at each subframe τ. At each step τ, the action of each *i*-th CG is expressed as $a_\tau^i = \{N_i^{\mathrm{CTU}}(\tau), N_i^{\mathrm{s}}(\tau), N_i^{\mathrm{r}}(\tau), N_i^{\mathrm{GB}}(\tau)\}$, where $N_i^{\mathrm{CTU}}(\tau)$ is the number of CTUs chosen as the *i*-th CG at subframe τ, $N_i^{\mathrm{s}}(\tau)$ is the starting time slot in a subframe of *i*-th CG at subframe τ, $N_i^{\mathrm{r}}(\tau)$ is the set number of the transmission repetitions of the *i*-th CG at subframe τ, $N_i^{\mathrm{GB}}(\tau)$ is the GB repetitions number of the *i*-th CG at subframe τ. The training configured parameters can be revised by considering the constraint in (16a). As shown in (16a), the CTU number of each *i*-th CG is constrained by the spectrum resources, i.e., $\sum_{i=1}^{N_{\mathrm{CG}}} N_i^{\mathrm{CTU}}(\tau) = N_{\mathrm{CTU}}$. In addition, each flow is constrained by T_m^{D}. Since extra GB retransmission only happens when a flow exceeds T_m^{D}, the additional $N_i^{\mathrm{GB}}(\tau)$ number will influence GB retransmission efficiency. However, considering multiple CGs results in the increment of observation space, exponentially increasing state space size, i.e., the number of available actions corresponds to the possible combinations of configurations $|\mathcal{A}_\tau| = \prod_{i=1}^{N_{\mathrm{CG}}}(|\mathcal{N}_i^{\mathrm{CTU}}(\tau)| \times |\mathcal{N}_i^{\mathrm{s}}(\tau)| \times |\mathcal{N}_i^{\mathrm{r}}(\tau)| \times |\mathcal{N}_i^{\mathrm{GB}}(\tau)|)$.

C. Rewards Design

To maximize the number of successfully transmitted flows under the constraints, we denote the reward $r_{\tau+1}$ as

$$r_{\tau+1} = N_\tau^{\mathrm{suc}}, \tag{17}$$

where N_τ^{suc} is the total number of flows that have been successfully decoded and transmitted to the receiver under the deadline constraint at subframe τ.

D. CMA-PPO Approach

The numerous actions and states will inevitably result in considerable computational latency, greatly affecting the

performance of RL algorithms. To address this issue, a centralized value function and a distributed policy function combined framework is proposed, which can achieve distributed execution while obtaining the maximum reward.

The Actor network selects actions based on a policy $\tilde{\pi}$; meanwhile, another Critic network parameter θ evaluates the rewards of the chosen actions. The global rewards are summed as

$$r_\tau(\theta) = \sum_{i=1}^{N_{\mathrm{CG}}} \sum_{\tau=1}^{\tau_{\mathrm{frame}}} r(\omega_i)\tilde{\pi}_\theta\Big(a_\tau^i \mid s_\tau^i\Big), \tag{18}$$

where ω_i represents the sequence of experiences for the i-th CG agent engaging with the environment, which is retained within the experience buffer. $\sum_{\tau=1}^{\tau_{\mathrm{frame}}} \tilde{\pi}_\theta(a_\tau^i \mid s_\tau^i)$ represents the probability of a sequence ω_i occurring. Within the continuous engagement between the agent and the environment, the most recent reward holds greater significance compared to past rewards. Consequently, a discounted reward $r(\omega) = \sum_{\tau=1}^{\tau_{\mathrm{frame}}} \xi^\tau r$ is employed, where ξ denotes the discount factor. To optimize the agent's cumulative reward, the policy network parameter θ is iteratively adjusted via gradient descent, which can be formulated as

$$\nabla \hat{r}_\theta = \mathbb{E}_{\omega\sim\tilde{\pi}_\theta(\omega)}\Big[\nabla \log \tilde{\pi}_\theta\Big(a_\tau^i \mid s_\tau^i\Big) A_\theta(s_\tau, a_\tau)\Big], \tag{19}$$

where A_θ refers to the advantage function, which is utilized as a substitute for the reward. The advantage function could guarantee a_τ as the best action in state s_τ when selecting by $\tilde{\pi}_\theta(a_\tau \mid s_\tau)$. We consider $V_\phi(s_\tau)$ to be the value function with state s_τ, then, the advantage function is given by

$$A_\theta(s_\tau, a_\tau) = \tilde{\pi}_\theta(a_\tau|s_\tau) - V_\phi(s_\tau). \tag{20}$$

Due to PPO's on-policy approach, each agent we train is similar to the agent cognizing the data samples. The sampling method is crucial because it ensures that the adopted sampling data can be reproduced consistently. We consider the probability ratio between old policies $\tilde{\pi}_{\theta_{\mathrm{old}}}$ and new policies $\tilde{\pi}_\theta$ is shown as

$$u_\tau(\theta) = \left[\frac{\tilde{\pi}_\theta(a_\tau \mid s_\tau)}{\tilde{\pi}_{\theta_{\mathrm{old}}}(a_\tau \mid s_\tau)}\right]. \tag{21}$$

Based on (21), (18) and (19) are rewritten as

$$\nabla \hat{r}_\theta = \mathbb{E}_{\omega\sim\tilde{\pi}_{\theta_{\mathrm{old}}}(\omega)}\left[\frac{\nabla\tilde{\pi}_\theta\big(a_\tau^i \mid s_\tau^i\big)}{\tilde{\pi}_{\theta_{\mathrm{old}}}\big(a_\tau^i \mid s_\tau^i\big)} A_{\theta_{\mathrm{old}}}(s_\tau, a_\tau)\right], \tag{22}$$

and

$$J_{\theta_{\mathrm{old}}}(\theta) = \mathbb{E}_{\omega\sim\tilde{\pi}_{\theta_{\mathrm{old}}}(\omega)}\left[\frac{\nabla\tilde{\pi}_\theta(a_\tau \mid s_\tau)}{\tilde{\pi}_{\theta_{\mathrm{old}}}(a_\tau \mid s_\tau)} A_{\theta_{\mathrm{old}}}(s_\tau, a_\tau)\right]. \tag{23}$$

The high variance between $\tilde{\pi}_\theta$ and $\tilde{\pi}_{\theta_{\mathrm{old}}}$ can substantially reduce sample efficiency and disrupt the training process. Thus, the Kullback-Leibler (KL) divergence employed to measure the variability between two probability distributions is written as

$$D_{\mathrm{KL}}\big(\tilde{\pi}_\theta \| \tilde{\pi}_{\theta_{\mathrm{old}}}\big) = \int_{-\infty}^{+\infty} \tilde{\pi}_\theta(a_\tau|s_\tau) \log \frac{\tilde{\pi}_\theta(a_\tau|s_\tau)}{\tilde{\pi}_{\theta_{\mathrm{old}}}(a_\tau|s_\tau)} \mathrm{d}s_\tau \tag{24}$$

Meanwhile, to constrain the update of θ within an acceptable range, the objective function of PPO can be clipped by clipping the parameter ϵ, which is given by

$$L(\theta) = \hat{\mathbb{E}}\big[\min\big(u_\tau(\theta)A_{\theta_{\mathrm{old}}}(s_\tau, a_\tau), \mathrm{clip}(u_\tau(\theta), 1-\epsilon, 1+\epsilon)A_{\theta_{\mathrm{old}}}(s_\tau, a_\tau)\big)\big], \tag{25}$$

where the clip function is designed to prevent huge policy updates, thereby alleviating the issue of catastrophic performance loss. This clip fuction aims to eliminate incentives for the new policy to deviate substantially from the old policy. Consequently, if $A_{\theta_{\mathrm{old}}}(s_\tau, a_\tau) > 0$, $u_\tau(\theta)$ is constrained to $1+\epsilon$. Otherwise, $u_\tau(\theta)$ is clipped as $1-\epsilon$.

In the decentralized reward of CMA-PPO, each agent may compete CTU resources with the others, adopting selfish behaviors to obtain the highest reward, which can detrimentally affect the overall performance. To transform such selfish behaviors into cooperative ones, assigning global rewards to the value networks of all agents can be effective. The optimization objective of the value network with parameter ϕ is calculated as

$$V_\phi^i(\phi) = \min \mathbb{E}\sum_{\mathcal{B}_i}\Big(V_\phi^i(s_\tau) - \hat{r}_\tau\Big)^2, \tag{26}$$

where $\mathcal{B}_i$ represent the global experience buffer storing experience of each CG i.

Additionally, the optimization objectives for the policy network are defined by

$$J_{\theta_{\mathrm{old}}}(\theta) = \mathbb{E}_{\omega_i\sim\tilde{\omega}_{\theta_{\mathrm{old}}}(\omega_i)}\left[\frac{\tilde{\pi}_\theta\big(a_\tau^i \mid o_\tau^i\big)}{\tilde{\pi}_{\theta_{\mathrm{old}}}\big(a_\tau^i \mid o_\tau^i\big)} A_{\theta_{\mathrm{old}}}\Big(o_\tau^i, a_\tau^i\Big)\right], \tag{27}$$

where o_τ^i is the local observation to interact with the environment of the i-th CG. ω_i represents the sequence of o_τ^i, and the parameters θ of each agent is updated based on ω_i.

The details of the proposed CMA-PPO algorithm are shown in Algorithm 4.

E. Complexity and Adaptability Analysis

Compared with the online training approach, the PPO pre-training can be done offline in the BS, where most computational tasks, such as policy estimation, are executed with powerful computing ability. Thus, the complexity of the PPO selection action is determined by the number of layers, the sizes of input and output, and the size of the action space $|\mathcal{A}_\tau| = \prod_{i=1}^{N_{\mathrm{CG}}}(|\mathcal{N}_i^{\mathrm{CTU}}(\tau)| \times |\mathcal{N}_i^{\mathrm{s}}(\tau)| \times |\mathcal{N}_i^{\mathrm{r}}(\tau)| \times |\mathcal{N}_i^{\mathrm{GB}}(\tau)|)$. We define N_{hid} as the number of hidden layer dimensions in CMA-PPO, and the complexity of action selecting in CMA-PPO is calculated as $O(|\mathcal{A}_\tau| \times N_{\mathrm{hid}})$, which is of linear level. Meanwhile, the CMA-PPO performs fast and stable convergence in dynamic and unpredictable environments.

VI. Simulation Evaluations

In this section, we show the effectiveness of the proposed joint flow scheduling scheme under different settings over 5G-TSN networks.

Algorithm 4 CMA-PPO Aided Uplink MCG Configuration

Input: Episode time T_{episode}, action space $\mathcal{A}$, policy parameters θ, value function parameters ϕ, and the hyper-parameter ϵ.
Initialize the policy parameters θ, value function parameters ϕ, hyperparameter ϵ, and the experience buffer $\mathcal{B}$
for episode $= 1$ to T_{episode} **do**
  Cognize the initial state *S*, and set reward *r* as 0
  Initialize the actions of each CG *i*
  for each step τ **do**
    Each agent *i* selects the action a_τ^i based on its local state observation o_τ^i
    Execute each action a_τ^i, get new local state observation $o_{\tau+1}^i$, record $(o_\tau^i, a_\tau^i, o_{\tau+1}^i)$ in sequence of ω_i, and store ω_i in $\mathcal{B}_i$
    for CG *i* **do**
      Cognize the set of actions in the environment and calculate the reward $\hat{r}_\tau(\theta)$
      Estimate advantage function A_θ^i based on the current value function V_ϕ^i
      Calculate policy update θ^i based on the clip function
      Store $(s_\tau^i, s_{\tau+1}^i, a_\tau^i, \hat{r}_\tau^i)$ in replay buffer
      Fit value function by taking $V_\phi^i(\phi)$ via gradient descent algorithm.
    end for
  end for
end for

A. Simulation Setups

For our simulation, we consider a 5G-TSN uplink network with a circular area having a radius $d_m = 1$ km, a BS located at the center, and *M* devices randomly located within the cell. All performance results are evaluated by averaging over 1000 subframes. The detailed list of additional parameters used can be found in Table III. We display our simulation results of MCG configuration in the proposed GB-enhanced MCG-GF-NOMA framework compared with the MCG-GF-NOMA framework on the 5G side network. Moreover, we compare the proposed AoS-enhanced CQF scheduling with the CQF scheduling.

1) 5G-TSN Network Setups: In the 5G side, the mini-slots of $N_{\text{sys}} = 7$ OFDM symbols are utilized for transmissions with 60 kHz ($\mu = 2$) sub-carrier space, which follows the main guidelines for 5G NR performance evaluations presented in [38]. We set the RB number as 4, and the CTU number as 64. During GF and GB stages, devices access the BS with random access and GB manner, respectively. The transmitting power of each device is set as 23 dBm. And in each stage, the BS decodes with SIC decoding. For the successful SIC decoding, the SINR threshold is set as -3 dB.

In the TSN side, a TSN network with directed graph topology is established, which contains 7 switches. Meanwhile, a TSN gateway is set in the BS. The forwarding queue length of each switch is set as 3, and we assume there is no cache in each switch.

TABLE III
SIMULATION PARAMETERS

Parameter	Value
The number of devices M	10000,50000
Wireless bandwidth of subcarrier	60 kHz
The number of RBs	4
The number of CTUs	64
The transmitting power of each device P	23 dBm
Noise power N_0	-132 dBm
Path-loss parameter η	4
Distance to the BS d	0-1000 m
Numerology factor μ	2
Slot number in a subframe	8
Subframe duration	1 ms
SINR decoding threshold γ_{th}	-3 dB
The set of start offset	{1, 2, 3, 4}
Learning rate	0.001
Discount factor ξ	0.9
PPO clip parameter ϵ	0.1
Switch flow number capacity Q_{size}	3
Subframe length in a episode T_{sub}	1000
Deadline of emergency stop flow	1 ms
Deadline of closed-loop control flow	2 ms
Deadline of level alarm flow	5 ms

2) Flow Setups: In the IIoT scenarios, massive flows are generated according to aperiodic inter-arrival processes over the TTIs, which have Markov properties. To simulate the bursty flow process, we adopt the Beta and the Poisson distribution arrival processes to model the flow arrival intensity according to the flow model defined in 3GPP [21], [39].

When the flows follow a Beta distribution, the active devices could only transmit with the CG with the closest starting offset, then the flows could be divided into different CGs to transmit, and the period of each CG *i* is defined as the period from τ_{i-1} to τ_i. The flows generated during (τ_{i-1}, τ_i) are configured to transmit in CG *i*. The flow instantaneous rate in subframe τ is given by

$$\rho(\tau) = \frac{\tau^{\alpha-1}(T_{\text{sub}} - \tau)^{\beta-1}}{T_{\text{sub}}^{\alpha+\beta-1}\text{Beta}(\alpha, \beta)}, \tag{28}$$

where T_{sub} is the total subframe number of the bursty flows and $\text{Beta}(\alpha, \beta) = \int_0^1 \tau^{\alpha-1}(1-\tau)^{\beta-1}\text{d}\tau$ is the Beta functions with the constant parameters α and β. Then, the flow's arrival rate in the *i*-th CG period is given by

$$B_i = \int_{\tau_{i-1}}^{\tau_i} \rho(\tau)\text{d}\tau, \tag{29}$$

where (τ_{i-1}, τ_i) represents the duration during which the flows configured in CG *i* are generated.

When the flow follows the Poisson distribution [40], we consider the actual number of active devices in subframe τ

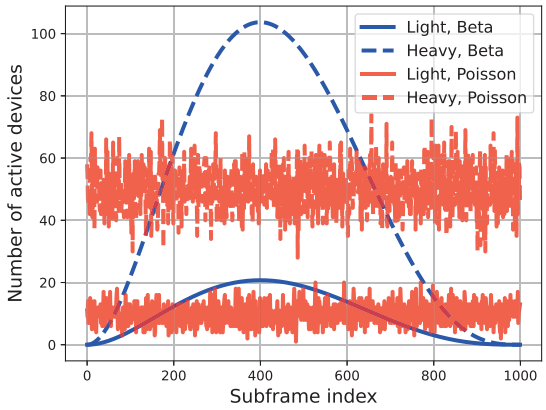


Fig. 5. Arrival flows with different distributions.

as N_τ. The probability density function that j devices are active in subframe τ is given by

$$\Pr(N_\tau = j) = \frac{w^j}{j!} e^{-w}, \tag{30}$$

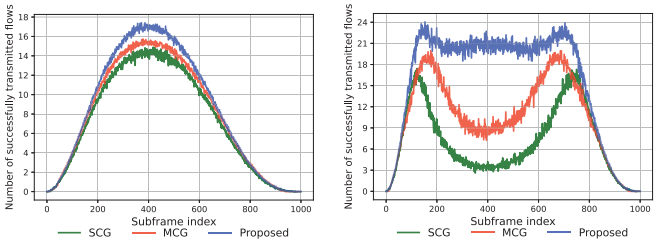
where w is the mean value of each subframe τ, which can be calculated as $w = \sum_{j=0}^{\infty} j \Pr(N_\tau = j)$, i.e., when 10000 devices attempt to access the BS in 1000 subframes with the Poisson distribution, the w is set as 10 at each subframe τ. The e is the natural constant.

The arrival flow distributions are shown in Fig. 5. We set the Beta distribution parameters as $\alpha = 3$ and $\beta = 4$. Meanwhile, we configure the device numbers for both the light and heavy access scenarios to 10000 and 50000 devices. To better evaluate the performance in the IIoT scenarios, we simulate three types of time-sensitive aperiodic flows, as in [41], including close-loop control, emergency stop, and level alarm. The three types of aperiodic flows have different latency requirements specified in Table III. Specifically, close-loop control is the main aperiodic flow transmitted for network control, while emergency stop have a higher latency requirement than level alarm.

B. Simulation Results

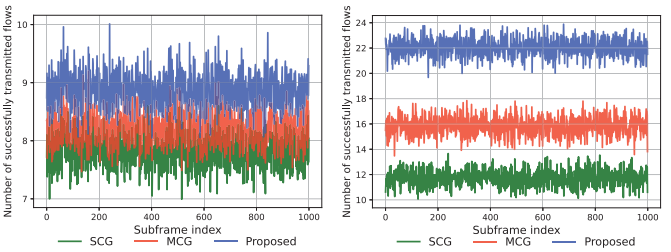
In this part, we first evaluate the priority of the proposed GB-enhanced MCG-GF-NOMA framework in the 5G side and then verify the number of successfully transmitted flows in the 5G-TSN network. To better evaluate the proposed AoS-enhanced CQF scheduling with different numbers of injected flows in TSN, we assess the TSN side transmission with varying numbers of time-sensitive flows. Finally, we simulate the coverage performance of the proposed MCG-GF-NOMA framework aided by the CMA-PPO in light and heavy access scenarios.

In Fig. 6, we show how the number of successfully transmitted flows in the 5G side changes as the subframe index varies with Beta distribution in light and heavy access scenarios. From Fig. 6(a), we observe that the GB-enhanced MCG-GF-NOMA framework can improve the number of successfully transmitted flows by 13.71% and 18.29% at the flow peak, compared to the MCG-GF-NOMA and SCG-GF-NOMA frameworks in the light access scenario, respectively.



(a) Light access flows with Beta distribution. (b) Heavy access flows with Beta distribution.

Fig. 6. Number of successfully transmitted flows with Beta distribution in 5G side.



(a) Light access flows with Poisson distribution. (b) Heavy access flows with Poisson distribution.

Fig. 7. Number of successfully transmitted flows with Poisson distribution in 5G side.

Meanwhile, Fig. 6(b) shows that the GB-enhanced MCG-GF-NOMA framework can improve the number of successfully transmitted flows by more than about one and six times at the flow peak, compared to the MCG-GF-NOMA and SCG-GF-NOMA frameworks in the heavy access scenario, respectively. This is because the proposed MCG-GF-NOMA framework can retransmit the flows that failed to be transmitted in the previous subframe without exceeding the deadline, and the GB manner during retransmission can avoid collision, thus improving the number of successfully transmitted flows, especially when the number of flows increases.

In Fig. 7, we show how the number of successfully transmitted flows in the 5G side changes as the subframe index varies with the Poisson distribution in light and heavy access scenarios. From Fig. 7(a), we can see that the proposed MCG-GF-NOMA framework can improve the overall number of successfully transmitted flows up to 8.68% and 12.97% compared to the MCG-GF-NOMA and SCG-GF-NOMA frameworks for all the subframes in the light access scenario, respectively. On the other hand, Fig. 7(b) shows that the proposed MCG-GF-NOMA framework can improve the overall number of successfully transmitted flows significantly up to 40.13% and 88.06% compared with the MCG-GF-NOMA and SCG-GF-NOMA frameworks for all the subframes in the heavy access scenario, respectively.

In Fig. 8, we show how the number of successfully transmitted flows varies over the 5G-TSN network with Beta and Poisson distributions in light and heavy access scenarios. From Fig. 8, we can see that compared with the GB-enhanced MCG-GF-NOMA and MCG-GF-NOMA frameworks, the SCG-GF-NOMA framework has a low efficiency in transmitting time-sensitive flows over the 5G-TSN network

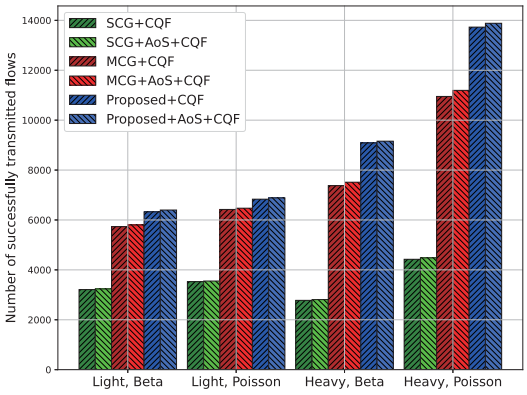
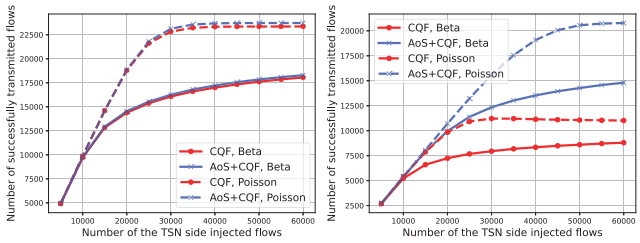


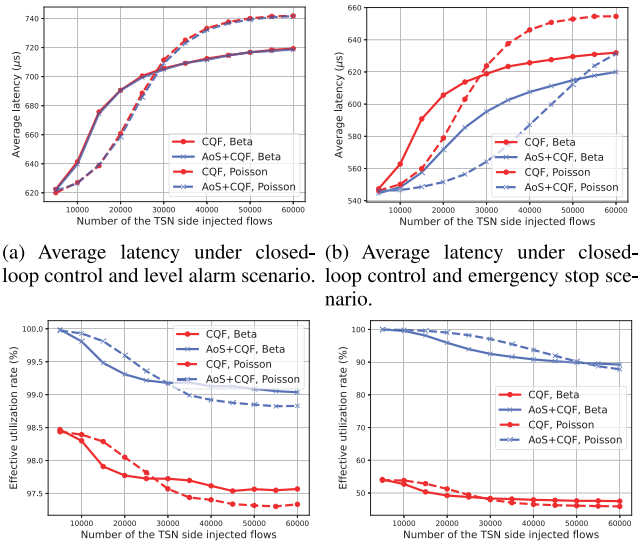
Fig. 8. Number of successfully transmitted flows over 5G-TSN network.



(a) Number of successfully transmitted flows under closed-loop control and level alarm scenario.

(b) Number of successfully transmitted flows under closed-loop control and emergency stop scenario.

Fig. 9. Number of successfully transmitted flows with different numbers of injected flows.



(a) Average latency under closed-loop control and level alarm scenario.

(b) Average latency under closed-loop control and emergency stop scenario.

(c) Effective utilization rate under closed-loop control and level alarm scenario.

(d) Effective utilization rate under closed-loop control and emergency stop scenario.

Fig. 10. Average latency and effective utilization rate under different flow scenarios.

in both light and heavy access scenarios. This is because transmitting flows with the SCG-GF-NOMA framework can lead to larger latency, i.e., the flows have to start transmitting in another subframe instead of transmitting in another CG of the same subframe with a different start offset. Meanwhile, we can see that the GB-enhanced MCG-GF-NOMA framework increases the number of successfully transmitted flows up to 24.02% compared with the MCG-GF-NOMA framework, especially with the Poisson distribution scenario. This is due to the fact that the proposed MCG-GF-NOMA framework can retransmit the flows with collision, which are more common in the Poisson distribution scenario. However, the proposed AoS-enhanced CQF scheduling exhibits substantial advantages compared to the CQF scheduling. The reason is that the limitations of the current 5G side spectrum transmission ability, the number of flows injected into the TSN has not reached the switch's forwarding capacity to prove the strength of the AoS-enhanced CQF scheduling. This result only proves the superiority of the proposed MCG-GF-NOMA framework. Hence, we will vary the number of the TSN-side injected flows to show the advantage of the AoS-enhanced CQF scheduling.

In Fig. 9, we demonstrate how the number of successfully transmitted flows changes as the number of injected flows varies. Fig. 9(a) demonstrates the number of successfully transmitted flows in closed-loop control and level alarm scenario. From Fig. 9(a), we can see that the AoS-enhanced CQF and CQF schedulings transmit nearly the same flow numbers in both scenarios. At the same time, in emergency stop and closed-loop control scenario, Fig. 9(b) shows that the AoS-enhanced CQF scheduling can significantly improve the number of successfully transmitted flows with Beta distribution and Poisson distribution by 82.22% and 70.29%, respectively, compared with the CQF scheduling as injected flow number increases. This is due to that the AoS-enhanced CQF scheduling can discard flows exceeding the deadline constraints, guaranteeing that the flows injected into the TSN network can be transmitted to the receiver within the deadline constraints.

In Fig. 10, we demonstrate how the average latency and effective utilization rate vary with different numbers of injected flows. Fig. 10(a) compares the average latency in closed-loop control and level alarm scenario as the number of injected flows varies. The average latencies of the AoS-enhanced CQF and CQF schedulings are nearly identical in both Beta and Poisson distributions. However, Fig. 10(b) shows that, compared with the CQF scheduling, the AoS-enhanced CQF scheduling can decrease the transmission latency. Fig. 10(c) illustrates that the effective utilization rates of the AoS-enhanced CQF and CQF schedulings for forwarding resources, at the TSN side, with various numbers of injected flows, are almost the same in the closed-loop control and level alarm scenario. Fig. 10(d) shows that the AoS-enhanced CQF scheduling can improve the TSN side effective utilization rate in emergency stop and closed-loop control scenario services by up to 50% compared with the CQF scheduling. This stems from the fact that the proposed AoS-enhanced CQF scheduling can discard flows that exceed deadline constraints, thus reducing the average latency and improving the effective utilization rate.

In Fig. 11 and Fig. 12, we show the convergence performance of the CMA-PPO algorithm aided GB-enhanced MCG-GF-NOMA framework over the 5G-TSN network. We

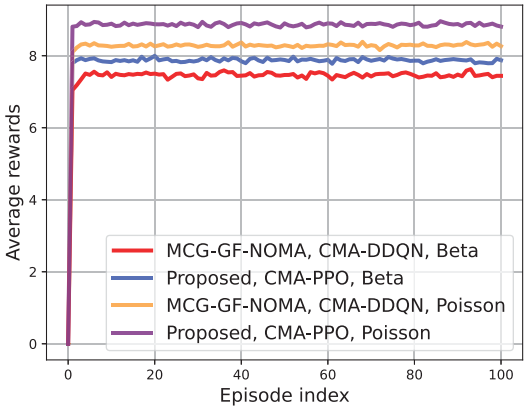


Fig. 11. Rewards of different algorithms in light access scenario.

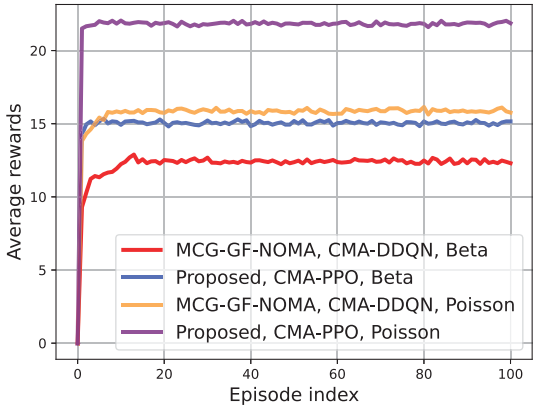


Fig. 12. Rewards of different algorithms in heavy access scenario.

compare it with the CMA-DDQN algorithm aided MCG-GF-NOMA framework in both light and heavy access scenarios with Beta distribution and Poisson distribution. In the TSN network, we deploy AoS-enhanced CQF policy to forward flows. Fig. 11 and Fig. 12 show that the CMA-PPO algorithm aided GB-enhanced MCG-GF-NOMA framework can achieve more successfully transmitted flows over the 5G-TSN network, Meanwhile, the manifestation is even more pronounced in the heavy access scenario. From both Fig. 11 and Fig. 12, we can observe that the proposed algorithm converges faster than the CMA-DDQN aided MCG-GF-NOMA algorithm. This is because the proposed CMA-PPO algorithm can adjust the learning rate at each iteration to speed up the convergence. Moreover, the CMA-PPO algorithm can achieve a greater reward than that of the CMA-DDQN algorithm when the algorithm converges. This is due to the fact that the proposed algorithm can improve the total reward by decreasing the learning rate when the updated policy approximates the local optimal solution.

VII. Conclusion

In this paper, we have proposed a novel joint flow scheduling scheme for the 5G-TSN uplink network to improve the number of successfully transmitted flows. In the 5G side network, we have designed a novel MCG-GF-NOMA framework. Meanwhile, we introduced an AoS-enhanced CQF scheduling in the TSN side network. Then, we have formulated an optimization problem based on the 5G-TSN system, aiming to maximize the number of successfully transmitted flows under different IIoT scenarios and deadline constraints. Next, we have proposed the CMA-PPO algorithm to solve the optimization problem by controlling the MCG configuration with cognitive information of flows. The simulation results have evidently shown that the proposed joint flow scheduling scheme aided by the CMA-PPO algorithm significantly improves the number of successfully transmitted flows compared with the MCG-GF-NOMA aided by the CMA-DDQN algorithm and the SCG-GF-NOMA. Future work could explore the integration of this framework with emerging 6G technologies and investigate its applicability in other time-sensitive transmission scenarios.

References

[1] A. Nasrallah et al., "Ultra-low latency (ULL) networks: The IEEE TSN and IETF DetNet standards and related 5G ULL research," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 88–145, 1st Quart., 2018.

[2] K.-C. Chen, S.-C. Lin, J.-H. Hsiao, C.-H. Liu, A. F. Molisch, and G. P. Fettweis, "Wireless networked multirobot systems in smart factories," *Proc. IEEE*, vol. 109, no. 4, pp. 468–494, Apr. 2021.

[3] F. Lin, C. Chen, N. Zhang, X. Guan, and X. Shen, "Autonomous channel switching: Towards efficient spectrum sharing for industrial wireless sensor networks," *IEEE Internet Things J.*, vol. 3, no. 2, pp. 231–243, Apr. 2016.

[4] A. Mahmood et al., "Industrial IoT in 5G-and-beyond networks: Vision, architecture, and design trends," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 4122–4137, Jun. 2022.

[5] J. Farkas, L. L. Bello, and C. Gunther, "Time-sensitive networking standards," *IEEE Commun. Stand. Mag.*, vol. 2, no. 2, pp. 20–21, Jun. 2018.

[6] S. Senk, H. K. Nazari, H.-H. Liu, G. T. Nguyen, and F. H. P. Fitzek, "Open-source testbeds for integrating time-sensitive networking with 5G and beyond," in *Proc. IEEE Consum. Commun. Netw. Conf.*, Jan. 2023, pp. 1–7.

[7] Z. Satka, M. Ashjaei, H. Fotouhi, M. Daneshtalab, M. Sjödin, and S. Mubeen, "A comprehensive systematic review of integration of time sensitive networking and 5G communication," *J. Syst. Archit.*, vol. 138, May 2023, Art. no. 102852.

[8] "5G NR physical layer procedures for data," 3GPP, Sophia Antipolis, France, Rep. TS 38.214, Mar. 2020.

[9] Y. Liu, Y. Deng, M. Elkashlan, A. Nallanathan, and G. K. Karagiannidis, "Analyzing grant-free access for URLLC service," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 741–755, Aug. 2020.

[10] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, "Grant-free non-orthogonal multiple access for IoT: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1805–1838, 3rd Quart., 2020.

[11] T. K. Le, U. Salim, and F. Kaltenberger, "An overview of physical layer design for ultra-reliable low-latency communications in 3GPP releases 15, 16, and 17," *IEEE Access*, vol. 9, pp. 433–444, 2020.

[12] *IEEE Standard for Local and Metropolitan Area Networks–Bridges and Bridged Networks-Amendment 25: Enhancements for Scheduled Traffic*, IEEE Standard 802.1Qbv-2015, 2016.

[13] M. Kim, J. Min, D. Hyeon, and J. Paek, "TAS scheduling for real-time forwarding of emergency event traffic in TSN," in *Proc. Int. Conf. Inf. Commun. Technol. Converg.*, Oct. 2020, pp. 1111–1113.

[14] Z. Cheng, D. Yang, W. Zhang, J. Ren, H. Wang, and H. Zhang, "DeepCQF: Making CQF scheduling more intelligent and practicable," in *Proc. IEEE Int. Conf. Commun.*, May 2022, pp. 1–6.

[15] W. Tan and B. Wu, "Long-distance deterministic transmission among TSN networks: Converging CQF and DIP," in *Proc. IEEE Int. Conf. Netw. Protocols*, Nov. 2021, pp. 1–6.

[16] Y. Cai, X. Zhang, S. Hu, and X. Wei, "Dynamic QoS mapping and adaptive semi-persistent scheduling in 5G-TSN integrated networks," *China Commun.*, vol. 20, no. 4, pp. 340–355, Apr. 2023.

[17] Y. Liu, Y. Deng, M. Elkashlan, A. Nallanathan, and G. K. Karagiannidis, "Optimization of grant-free NOMA with multiple configured-grants for mURLLC," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1222–1236, Apr. 2022.

[18] M. Li, S. Guo, C. Chen, C. Chen, X. Liao, and X. Guan, "DecAge: Decentralized flow scheduling for industrial 5G and TSN integrated networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 1, pp. 543–555, Aug. 2024.

[19] Y. Huang, S. Wang, X. Zhang, T. Huang, and Y. Liu, "Flexible cyclic queuing and forwarding for time-sensitive software-defined networks," *IEEE Trans. Netw. Service Manag.*, vol. 20, no. 1, pp. 533–546, Aug. 2023.

[20] Y. Chen et al., "CFS: A CQF-based fast scheduling scheme for time-sensitive networking," in *Proc. Chin. Control Conf.*, Jul. 2023, pp. 6221–6226.

[21] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, and J. M. Lopez-Soler, "A survey on 5G usage scenarios and traffic models," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 905–929, 2nd Quart., 2020.

[22] H. Liu, M. Li, F. Gu, Q. Li, W. Zhang, and S. Guo, "End-to-end flow scheduling optimization for industrial 5G and TSN integrated networks," in *Proc. IEEE Global Commun. Conf.*, Dec. 2024, pp. 1797–1802.

[23] Y. Zhang, Q. Xu, C. Chen, M. Li, X. Guan, and T. Q. S. Quek, "Seamless scheduling for NFV-enabled 5G-TSN network: A full-path AoI based method," *IEEE Trans. Ind. Informat.*, vol. 20, no. 12, pp. 13513–13525, Dec. 2024.

[24] R. Kumar, D. Sinwar, and V. Singh, "QoS aware resource allocation for coexistence mechanisms between eMBB and URLLC: Issues, challenges, and future directions in 5G," *Comput. Commun.*, vol. 213, pp. 208–235, Jan. 2024.

[25] N. Ye, H. Han, L. Zhao, and A.-H. Wang, "Uplink nonorthogonal multiple access technologies toward 5G: A survey," *Wireless Commun. Mobile Comput.*, no. 1, pp. 1–26, Jun. 2018.

[26] M. Gerami and B. Singh, "Configured grant for ultra-reliable and low-latency communications: Standardization and beyond," *IEEE Commun. Stand. Mag.*, vol. 6, no. 4, pp. 40–47, Dec. 2022.

[27] F. Ihle, S. Lindner, and M. Menth, "P4-PSFP: P4-based per-stream filtering and policing for time-sensitive networking," *IEEE Trans. Netw. Service Manag.*, vol. 21, no. 5, pp. 5273–5290, Oct. 2024.

[28] G. Patti, L. L. Bello, and L. Leonardi, "Deadline-aware online scheduling of TSN flows for automotive applications," *IEEE Trans. Ind. Informat.*, vol. 19, no. 4, pp. 5774–5784, Apr. 2023.

[29] J. Yan, W. Quan, X. Jiang, and Z. Sun, "Injection time planning: Making CQF practical in time-sensitive networking," in *Proc. IEEE Conf. Comput. Commun.*, 2020, pp. 616–625.

[30] H. Tang, J. Wang, Z. Tang, and J. Song, "Scheduling to Minimize age of Synchronization in wireless broadcast networks with random updates," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4023–4037, Jun. 2020.

[31] S. Sun et al., "Propagation path loss models for 5G urban micro- and macro-cellular scenarios," in *Proc. IEEE Veh. Technol. Conf.*, May 2016, pp. 1–6.

[32] *IEEE Standard for Local and Metropolitan Area Networks–Bridges and Bridged Networks-Amendment 28: Per-Stream Filtering and Policing*, IEEE Standard 802.1Qci-2017, 2017.

[33] H. V. Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. Nat. Conf. Artif. Intell.*, Mar. 2016, pp. 2094–2100.

[34] R. Zhang et al., "Interactive AI with retrieval-augmented generation for next generation networking," *IEEE Netw.*, vol. 38, no. 6, pp. 414–424, Nov. 2024.

[35] Y. Xu et al., "Decentralization of generative AI via mixture of experts for wireless networks: A comprehensive survey," 2025, *arXiv:2504.19660*.

[36] K. Liu et al., "Deadline-constrained multi-agent collaborative transmission for delay-sensitive applications," *IEEE Trans. Cogn. Commun. Netw.*, vol. 9, no. 5, pp. 1370–1384, Oct. 2023.

[37] C. Yu et al., "The surprising effectiveness of PPO in cooperative multi-agent games," in *Proc. 36th Int. Conf. Neural Inf. Process. Syst.*, Dec. 2022, pp. 24611–24624.

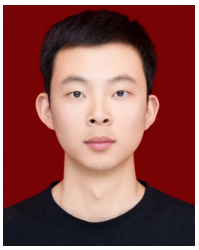
[38] S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, "NR: The new 5G radio access technology," *IEEE Commun. Stand. Mag.*, vol. 1, no. 4, pp. 24–30, Dec. 2017.

[39] S. K. Sharma and X. Wang, "Toward massive machine type communications in ultra-dense cellular IoT networks: Current issues and machine learning-assisted solutions," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 426–471, 1st Quart., 2020.

[40] T. N. Weerasinghe, V. Casares-Giner, I. A. M. Balapuwaduge, F. Y. Li, and M.-C. Vochin, "A pseudo-Bayesian subframe based framework for grant-free massive random access in 5G NR networks," in *Proc. Int. Conf. Comput. Commun. Netw.*, Jul. 2022, pp. 1–7.

[41] J. García-Morales, M. C. Lucas-Estañ, and J. Gozalvez, "Latency-sensitive 5G RAN slicing for industry 4.0," *IEEE Access*, vol. 7, pp. 143139–143159, 2019.

Jian Zhao (Student Member, IEEE) received the B.S. degree from the School of Computer Science and Technology and the M.E. degree from the School of Software Engineering from Shandong University, Jinan, Shandong, China. He is currently pursuing the Doctor's degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include wireless communication, industrial Internet of Things, and time sensitive networking.



Tao Wang (Student Member, IEEE) is currently pursuing the Doctor's degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. His research interests include wireless communication, extremely large scale antenna array, and time sensitive networking.

Haonan Tong (Member, IEEE) received the B.S. degree from the School of Electrical and Information Engineering, Tianjin University, Tianjin, China, and the Ph.D. degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. From 2023 to 2024, he was a Visiting Researcher with the College of Computing and Data Science, Nanyang Technological University, Singapore. His research interests include machine learning in wireless communication, digital twin, and semantic communications.

Nuocheng Yang (Student Member, IEEE) received the B.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2021, where he is currently pursuing the Ph.D. degree with the Information and Communication Engineering Department. His research interests include reinforcement learning, computing power networks, and machine learning in wireless networks.

Changchuan Yin (Senior Member, IEEE) received the Ph.D. degree in signal and information processing from the Beijing University of Posts and Telecommunications, Beijing, China, in 1998, where he is currently serving as a Professor with the School of Information and Communication Engineering. In 2004, he worked as a Visiting Scholar with the Faculty of Science, the University of Sydney, Sydney, NSW, Australia. From 2007 to 2008, he held a visiting position with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. His research interests include wireless networks and statistical signal processing. He was a co-recipient of the IEEE Marconi Prize Paper Award in 2023 and the IEEE International Conference on Wireless Communications and Signal Processing Best Paper Award in 2009. He has contributed as a Symposium Co-Chair and Technical Program Committee Member for numerous IEEE conferences.

Dusit Niyato (Fellow, IEEE) received the B.Eng. degree from the King Mongkuts Institute of Technology Ladkrabang, Thailand, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Canada. He is a Professor with the College of Computing and Data Science, Nanyang Technological University, Singapore. His research interests are in the areas of mobile generative AI, edge intelligence, quantum computing and networking, and incentive mechanism design.