Deterministic Flow Scheduling in 5G-TSN for IIoT: A Learning-Based Adaptive MCG-GF-NOMA Approach

Jian Zhao*, Tao Wang*, Zhaohui Yang†, Nuocheng Yang*, Haonan Tong‡, and Changchuan Yin*
* Beijing Key Laboratory of Network System Architecture and Convergence,
Beijing University of Posts and Telecommunications, Beijing, China 100876.
† College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, Zhejiang, China 310027.
‡ Key Laboratory of Target Cognition and Application Technology (TCAT),
Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China.
Emails: {zhaojian, taowang, yangnuocheng, ccyin}@bupt.edu.cn, yang_zhaohui@zju.edu.cn, tonghn@aircas.ac.cn.

2025 IEEE/CIC International Conference on Communications in China (ICCC) | 979-8-3315-4444-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICCC65529.2025.11148637

***Abstract*—The Industrial Internet of Things (IIoT) demands ultra-reliable low-latency communication (URLLC) to support massive time-sensitive devices with stringent requirements on deterministic transmission, ultra-dense connectivity, and bounded latency. To address this challenge, the convergence of 5G and Time-Sensitive Networking (TSN), termed 5G-TSN, has emerged as a key enabler for deterministic industrial networks. In this paper, we propose a joint 5G-TSN flow scheduling framework that guarantees deterministic transmission for massive time-sensitive data flows. Our approach integrates Grant-Free Non-Orthogonal Multiple Access (GF-NOMA) in 5G with Cyclic Queuing and Forwarding (CQF) in TSN. Specifically, we introduce a learning-based adaptive Multi-Configured Grant (MCG) subframe structure to optimize spectral efficiency in GF-NOMA. We then formulate a deadline-aware flow scheduling problem to maximize the number of successfully transmitted flows under strict latency constraints. To address this non-stationary and non-convex problem, we develop a Proximal Policy Optimization (PPO)-based algorithm that dynamically adjusts MCG configurations based on real-time ACK feedback and deadline requirements. Extensive simulations demonstrate that our solution outperforms existing benchmarks, achieving up to 23.81% improvement in the number of successfully transmitted flows under heavy flow IIoT scenarios compared with the MCG subframe structure.**

***Index Terms*—Industrial Internet of Things, 5G-TSN, multiple configured grant, grant-free, non-orthogonal multiple access.**

I. Introduction

The evolution of 3GPP standards has catalyzed significant advancements in massive ultra-reliable and low-latency communications (mURLLC) for the integration of the fifth-generation wireless networks and time-sensitive networking (5G-TSN). A pivotal innovation in this context is grant-free (GF) access, first introduced in 3GPP Release 15 as a paradigm-shifting alternative to conventional grant-based (GB) mechanisms in long-term evolution (LTE) [1]. By enabling devices to transmit data flows to the base station (BS) in an arrive-and-go mode, GF transmission eliminates the need for scheduling request (SR) procedures, thereby substantially reducing access latency. To address reliability challenges, the K-repetition GF mechanism [2], [3] enhances robustness through consecutive replica transmissions by leveraging time diversity. This method is pivotal in bolstering the robustness of data transmission. The integration of GF and non-orthogonal multiple access (NOMA) yields GF-NOMA systems [4], [5], effectively mitigating flow collisions in scenarios requiring massive device connectivity. Building upon these advancements, the multiple configured grants (MCG) structure has been proposed, catering to varying starting offsets of resources in relation to the arrival time of uplink flows [6]. This innovation has given rise to the MCG-GF-NOMA system, which is instrumental in reducing latency at each subframe. The MCG subframe structure is advantageous in reducing collision events during the contention period for multiple active devices and decreasing latency by providing many starting offset options.

To address the dual challenges of queue overflow and deterministic transmission scheduling in 5G-TSN networks, recent studies have developed injection time planning strategies based on incremental heuristics and adaptive scheduling policies [7], [8]. These solutions are critical for industrial IoT (IIoT) deployments where dense arrays of sensors and/or devices require strict latency guarantees for time-sensitive services [9]. While the MCG-GF-NOMA framework [10] has emerged as a pivotal enabler of mURLLC by supporting grant-free access and non-orthogonal resource sharing. Further research, as presented in [11], has implemented MCG-GF-NOMA and cyclic queuing and forwarding (CQF) in tandem within the 5G-TSN networks. Nevertheless, the conventional MCG-GF-NOMA design induces spectrum resource wastage due to suboptimal grant configurations, while persistent scheduling resolution mismatches between 5G and TSN domains demand urgent attention. The delayed acquisition of device status exacerbates these challenges, as updates at the BS prevent the dynamic optimization of configured grant (CG) pre-configuration parameters (including spectrum allocation, repetition values, and starting offsets), resulting in degraded transmission efficiency for time-sensitive flows. Furthermore, the time-varying characteristics of wireless channels introduce unpredictable transmission latencies, significantly reducing the success rate of deadline-constrained flow delivery.

With the above backgrounds, this paper aims to maximize the number of successfully transmitted time-sensitive flows over GF-NOMA-based 5G-TSN network for IIoT by introducing a learning-based adaptive MCG configuration for flow scheduling. Our key contributions are:

1) This paper considers the deterministic transmission in

This work was supported by Beijing Natural Science Foundation under Grant L223027, the National Natural Science Foundation of China under Grant 62471056, in part by the 111 Project under Grant B17007, in part supported by National Key R&D Program of China (Grant No. 2023YFB2904804), Young Elite Scientists Sponsorship Program by CAST 2023QNRC001, Zhejiang Key R&D Program under Grant 2023C01021, the Fundamental Research Funds for the Central Universities, K2023QA0AL02.

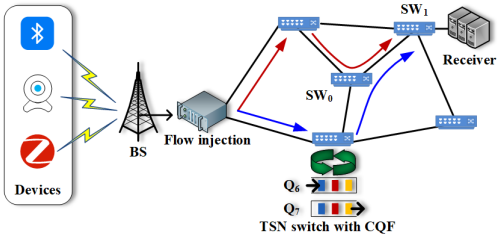


Fig. 1. 5G-TSN network system model.

the uplink of GF-NOMA empowered 5G-TSN with massive device access. We propose a novel adaptive MCG subframe structure to exploit the available spectrum resource fully. Further, the CQF is applied in the TSN to forward the time-sensitive flows, where a non-stationary and non-convex problem is formulated to maximize the number of successfully transmitted flows within the deadline.

2) To solve the formulated problem, we first propose a strategy to determine the adaptive MCG configuration based on the deadline constraints and the ACK reply. Then, we utilize cooperative multi-agent proximal policy optimization (CMA-PPO) to optimize the parameters of adaptive MCG configuration with non-stationary environments, which enhances the performance significantly.

The rest of this paper is organized as follows. The 5G-TSN system model is introduced in Section II. Section III formulates the problem. Section IV elucidates the proposed cooperative multi-agent PPO method. Section V presents the simulation results. And finally, we conclude the paper in Section VI.

II. 5G-TSN System Model

We consider the real-time transmission in the uplink of a GF-NOMA-based 5G-TSN network for IIoT as shown in Fig. 1, where event-triggered time-sensitive flows are generated by M devices, which are wirelessly transmitted to a BS, and then forwarded to the destined receiver through the TSN network. The flow generated by device m is denoted by f_m, which has a strict deadline T_m^{D}. The novel adaptive MCG subframe structure is adopted for multiple access and data transmission in the GF-NOMA-based 5G system. Then the BS injects the flows into the TSN network, which is forwarded to the destined receiver based on the CQF policy. We index the slot index as $t \in \{1, \ldots, T\}$. The slot number consumed by f_m can be represented as $D_m = D_m^{\mathrm{5G}} + D_m^{\mathrm{TSN}}$, where D_m^{5G} and D_m^{TSN} are the 5G side transmission slot number and the TSN side transmission slot number, respectively.

A. Adaptive MCG Subframe Structure

As shown in Fig. 2, we consider the time slot number of a subframe T_{frame} is 8, where each subframe consists of 4 contention transmission units (CTUs) slots and 4 signaling slots. Slots from 0 to 3 are defined as CTUs, and the following 4 slots are defined as signaling slots, which are used for the BS to transmit the ACK/NACK feedback and control information

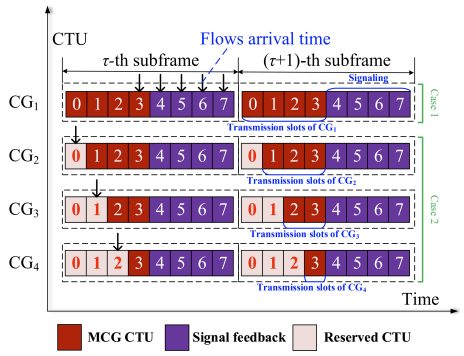


Fig. 2. Proposed adaptive MCG subframe structure.

for the devices. CGs have been classified into 4 different types. The four-frame structures are designed to transmit different flows generated at different times. The flow transmission cases can be summarized as follows:

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

The MCG-GF-NOMA system can significantly decrease the signaling overhead. Still, the reliability can not be guaranteed for each CG during GF transmission (e.g., CG_4 can only transmit one repetition, and the transmission has a high probability of failing due to a collision or decode failure happening). Thus, we propose an adaptive MCG-enhanced GF-NOMA system that applies GB scheduling to improve transmission reliability. The first half of the transmission CTUs of CG_2, CG_3, and CG_4 are set as reserved CTUs to retransmit the flows that are not successfully transmitted in the previous subframe. 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 CG_i. All the subframes in the system have the same number of slots all the time. Thus, for ease of presentation, we denote i-th type 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 starting slot of 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.

B. 5G Side Latency Model

For the 5G side transmission, each device has two ways to transmit the flow: firstly, the device transmits the flow in GF-NOMA manner in a certain CG; if it does not succeed, and

does not exceed the latency constraints, it can be re-transmitted with GB-NOMA manner in the next subframe, which are:

1) Successful Transmission during the GF Stage: The number of 5G side transmission slots fo device m to transmit the flow in the GF stage is given by $D_m^{\text{5G}} = T_m^{\text{wait}} + T_m^{\text{trans}}$, 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 at which the flow of the m-th device is decoded successfully, which is determined by the point at which the GF transmission repetition number meets the decoding requirement for the flow of m-th device. The waiting time slot number for the flow of m-th device is determined by:

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

where T_{CTU} is the maximum CTU slot number in a subframe, and the generation slot index of f_m is denoted as t_m^{g}.

2) Successful Transmission during the GB Stage: The 5G side transmission slot number required for successfully transmitting the data flow of device m during the GB stage is given by $D_m^{\text{5G}} = T_m^{\text{wait}} + T_m^{\text{RTT}} + T_m^{\text{retrans}}$, where T_m^{RTT} is the round-trip time (RTT) of the ACK/NACK feedback time slot number, which is set as 4 in Fig. 2. T_m^{retrans} is the GB transmission repetition number that satisfies flow decoding of the flow of the m-th device. The RTT time slot number of the flow of m-th device is calculated as $T_m^{\text{RTT}} = T_{\text{frame}} - (t_m^{\text{g}} \bmod T_{\text{frame}})$.

C. Successive Interference Cancellation Decoding

The BS employs successive interference cancellation (SIC) to decode the transmitted data flows. With the NOMA principles, during each SIC iteration, the BS initially decodes the data flow that has the highest received power. Subsequently, it removes the successfully decoded signal from the overall received signal. The large-scale path loss is modeled using the free space path-loss (FSPL) model [12], which is given by:

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

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

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

$$P_{m,n,i}(k) = P|h_{m,n}(k)|^2 L_m^{-\eta}, \tag{3}$$

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

We denote $\mathcal{N}_{m,n,i}^{\text{sc}}(k)$ as the set of devices that choose the singleton CTUs of CG i on the n-th RB and decoded in the SIC iteration after flow m. The flow number of $\mathcal{N}_{m,n,i}^{\text{sc}}(k)$ is given by $N_{m,n,i}^{\text{sc}}(k)$. Meanwhile, we denote $\mathcal{N}_{n,i}^{\text{cc}}(k)$ 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 $P_{1,n,i}(k) \geq P_{2,n,i}(k) \geq \cdots \geq P_{N_{m,n,i}^{\text{sc}}(k),n,i}(k)$. In each SIC decoding iteration, the singleton CTU having the highest received power is decoded, wherein the received undecoded powers in the same RB are considered as interference noise. Consequently, the signal-to-interference-plus-noise ratio (SINR) for the flow of device m, which is allocated to CG i at RB n is given by:

$$\lambda_{m,n,i}^{\text{SINR}}(k) = \frac{P_{m,n,i}(k)}{B^{\text{RB}} N_0 + \sum_{m' \in \mathcal{N}_{m,n,i}^{\text{sc}}(k)} P_{m'}(k) + \sum_{m'' \in \mathcal{N}_{n,i}^{\text{cc}}(k)} P_{m''}(k)}, \tag{4}$$

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 at the repetition is larger than the SINR threshold λ_{th}; otherwise, SIC decoding fails. During the transmission time, when $\lambda_{m,n,i}^{\text{SINR}}(k) \geq \lambda_{\text{th}}$, SIC decoding is successful.

D. TSN Side Latency Model

In the TSN side, we configure the TSN slot duration to align with the 5G system's slot size. This is because 5G URLLC utilizes the mini-slots to integrate with TSN [11]. The TSN topology is formally modeled as a directed graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ denotes the set of TSN switches and $\mathcal{E}$ represents the set of connections between any pair of connected switches. The transmission path for each flow f_m is described as a set of sequential links as $\mathcal{E}(f_m) = (e_1, \ldots, e_{\text{hop}_m})$, where hop_m is the hop length for flow f_m. Each flow is injected between different switches periodically with the CQF protocol. The TSN side transmission time slot of f_m from the BS to the receiver can be derived from the inferred delaying injection time, which is given by:

$$D_m^{\text{TSN}} = \sum_{e \in \mathcal{E}(f_m)} N_{\text{slot}}, \quad \forall f_m \in \mathcal{F}, e \in \mathcal{E}(f_m), \tag{5}$$

where N_{slot} is the forwarding slot number count for each switch, and $e \in \mathcal{E}(f_m)$ is the set of paths between switches f_m passes through.

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

$$\sum_{m=1}^{M} o_{m,e}(t) \leq Q_{\text{size}}, \quad \forall f_m \in \mathcal{F}, \forall e \in \mathcal{E}(f_m), \forall t, \tag{6}$$

III. Problem Formulation

The target is to increase the maximum successfully transmitted flows, we define δ_m to record whether device m's flow is successfully transmitted within the deadline constraint, where

the flow is generated at device m during subframe τ in the 5G-TSN network given by:

$$\delta_m(\tau) = \begin{cases} 1, & \text{if } D_m t_{\text{slot}} \leq T_m^{\text{D}}, \\ 0, & \text{otherwise}, \end{cases} \tag{7}$$

where t_{slot} denotes the time duration of each time slot.

Then the sum number of successfully served devices during the τ-th subframe is given by $N_\tau^{\text{suc}} = \sum_{m=1}^{M} \delta_m(\tau)$. The optimization problem for joint 5G-TSN design is formulated to maximize the successfully transmitted time-sensitive flows, which is given by

$$\max_{\mathcal{A}_\tau} \quad \sum_{\tau=1}^{\infty} \mathbb{E}(N_\tau^{\text{suc}}) \tag{8}$$

$$\text{s.t.} \quad \sum_{i}^{N_{\text{CG}}} N_i^{\text{CTU}}(\tau) = N_{\text{CTU}}, \forall \tau, \tag{8a}$$

$$\sum_{m=1}^{M} o_{m,e}(t) \leq Q_{\text{size}}, \forall e, \forall t, \tag{8b}$$

where $\mathcal{A}_\tau$ is the sampling indicator for each kind of $\text{CG}_i(\tau) = \{N_i^{\text{CTU}}(\tau), N_i^{\text{s}}(\tau), N_i^{\text{r}}(\tau), N_i^{\text{GB}}(\tau)\}$ at each subframe τ. (8a) indicates that all the CGs are constrained by the spectrum resource, i.e., N_{CTU} is the total number of configured CTUs for the adaptive MCG subframe structure. (8b) is the flow transmission constraint, which denotes the size of the total flows transmitted within each time slot t should be no more than the maximum time slot capacity of each switch.

IV. Proposed CMA-PPO Algorithm

A. RL Parameters

To solve the optimization problem in equation (8), we propose a CMA-PPO based algorithm where N_{CG} PPO agents are deployed at the BS. These agents select appropriate actions $a \in \mathcal{A}_\tau$ based on the partial observation $s \in \mathcal{S}$. At the start of each subframe τ, each agent in the CGs first records the different CG's state which contains a series of previous state observations $\mathcal{O}_\tau^{\text{old}} = \{s_1, \ldots, s_{\tau-1}\}$. Based on the observation s_τ, each agent chooses $a_\tau \in \mathcal{A}_\tau$, and the environment converts to a new state $s_{\tau+1}$, with a reward $r_{\tau+1}$ as the feedback. Then, the PPO agent updates the action choice function based on the reward to achieve the optimization objective [13]. Then, the proposed CMA-PPO-based algorithm's action, state, and reward design are specified.

1) States: The state of each subframe τ is given by $s_\tau = \{N_\tau^{\text{cc}}, N_\tau^{\text{sc}}, N_\tau^{\text{ic}}, N_\tau^{\text{suc}}, N_\tau^{\text{fail}}\}$. The five parameters are defined as: collision CTU count N_τ^{cc}, singleton CTU count N_τ^{sc}, idle CTU count N_τ^{ic}, the number of flows that have been successfully decoded and transmitted to the receiver under the deadline constraint N_τ^{suc}, and the number of flows fail to be transmitted to the receiver under the deadline constraint N_τ^{fail}.

2) Actions: To increase the long-term number of successfully transmitted time-sensitive flows, the BS collects the CTU observation $\mathcal{O}_\tau^{\text{old}}$ at each subframe τ. At each step τ, the action of each i-th CG is expressed as $a_\tau^i = \{N_i^{\text{CTU}}(\tau), N_i^{\text{s}}(\tau), N_i^{\text{r}}(\tau), N_i^{\text{GB}}(\tau)\}$, where $N_i^{\text{CTU}}(\tau)$ is the number of CTU chosen as i-th CG at subframe τ, $N_i^{\text{s}}(\tau)$ is the start time slot in a subframe of i-th CG at subframe τ, $N_i^{\text{r}}(\tau)$ is the set number of the transmission repetitions of i-th CG at subframe τ, $N_i^{\text{GB}}(\tau)$ is the GB repetitions number of the i-th CG at subframe τ.

3) Rewards: To maximize the number of successfully served flows with the constraints, we denote the reward $r_{\tau+1}$ as:

$$r_{\tau+1} = N_\tau^{\text{suc}}, \tag{9}$$

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

B. CMA-PPO Approach

Numerous actions and states inherently lead to significant computational latency, greatly affecting the performance of RL algorithms. To address this issue, we propose a framework combining a centralized value function with a distributed policy function, enabling distributed execution while maximizing rewards [14].

The Actor selects actions via the policy π, while the Critic θ evaluates the action values. The global rewards are summed as:

$$r_\tau(\theta) = \sum_{i=1}^{N_{\text{CG}}} \sum_{\tau=1}^{\tau_{\text{frame}}} r\left(\omega_i\right) \pi_\theta\left(a_\tau^i \mid s_\tau^i\right), \tag{10}$$

where ω_i denotes the experience sequence of the i-th CG agent's environmental interaction, stored in the experience buffer. The probability of sequence ω_i is calculated as $\sum_{\tau=1}^{\tau_{\text{frame}}} \pi_\theta\left(a_\tau^i \mid s_\tau^i\right)$. To maximize the cumulative reward, the parameter θ of the policy network is updated via the gradient, which is denoted as:

$$\nabla \hat{r}_\theta = \mathbb{E}_{\omega \sim \pi_\theta(\omega)}\left[\nabla \log \pi_\theta\left(a_\tau^i \mid s_\tau^i\right) A_\theta\left(s_\tau, a_\tau\right)\right], \tag{11}$$

where A_θ denotes the advantage function, which is utilized as a substitute for the reward. We consider $V_\phi\left(s_\tau\right)$ as the value function with state s_τ, then, the advantage function is denoted as:

$$A_\theta(s_\tau, a_\tau) = \pi_\theta\left(a_\tau | s_\tau\right) - V_\phi\left(s_\tau\right). \tag{12}$$

PPO's on-policy nature requires the training agent to collect the data samples at the same time. To make adopted data reproducible, importance sampling is introduced, considering the probability ratio between the old policy θ_{old} and the new policy θ.

Meanwhile, to constrain the update of θ within an acceptable range, the PPO objective function is clipped by parameter ϵ, defined as:

$$\begin{aligned} L(\theta) = & \hat{\mathbb{E}}[\min(u_\tau(\theta) A_{\theta_{\text{old}}}(s_\tau, a_\tau), \\ & \text{clip}(u_\tau(\theta), 1-\epsilon, 1+\epsilon) A_{\theta_{\text{old}}}(s_\tau, a_\tau))], \end{aligned} \tag{13}$$

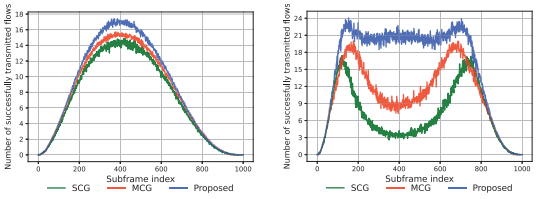
where the clip function restrains extreme policy updates to mitigate catastrophic performance loss.

Algorithm 1 CMA-PPO based adaptive MCG Configuration

Input: Episode time T_{episode}, number of CG N_{CG}, action space $\mathcal{A}$, policy parameters θ, value function parameters ϕ, and the clip parameter ϵ.
Initialize the parameters θ, ϕ, ϵ, and the experience buffer $\mathcal{B}$;
for episode $= 1$ to T_{episode} **do**
  Receive initial state S, and set $r = 0$;
  Initialize the actions of each CG i;
  for $\tau = 1$ to $\tau_{\max}$ **do**
    for $i = 1$ to N_{CG} **do**
      Agent i selects the action a_τ^i based on the state observation o_τ^i;
      Execute action a_τ^i, record $(o_\tau^i, a_\tau^i, o_{\tau+1}^i)$ in sequence of ω_i, and store ω_i in $\mathcal{B}_i$;
      Compute the reward $\hat{r}_\tau(\theta)$;
      Estimate advantage function A_θ^i based on the current value function V_ϕ^i;
      Compute 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 the gradient descent algorithm.
    end for
  end for
end for

TABLE I: Simulation Parameters

Parameters	Values
Devices number M	10000-60000
Bandwidth of subcarrier	60 kHz
Noise power N_0	-132 dBm
Path-loss parameter η	4
Distance from devices to the BS d	0-1000 m
5G NR numerology factor μ	2
Slot number in a subframe	8
Subframe duration	1 ms
SINR decoding threshold λ_{th}	-3 dB
Deadline of flows	1-2 ms



(a) Light-load scenario. (b) Heavy-load scenario.

Fig. 3. Number of successfully transmitted flows with subframe index under light-load and heavy-load scenarios.

The value network's core optimization target for parameter ϕ is defined as:

$$V_\phi^i(\phi) = \min \mathbb{E} \sum_{\mathcal{B}_i} \left(V_\phi^i(s_\tau) - \hat{r}_\tau \right)^2, \tag{14}$$

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

The details of the proposed CMA-PPO algorithm are illustrated in the Algorithm 1.

V. Simulation Results

For our simulation, we consider a 5G-TSN network in a circular cell of radius $d_m = 1$ km, with a BS at the center, and M devices randomly distributed within the cell. For 5G wireless transmissions, $N_{\text{sys}} = 7$ OFDM mini-slots are employed with the sub-carrier space (SCS) of 60 kHz (i.e., $\mu = 2$, a SCS of $15 \times 2^\mu$ kHz, and a slot t_{slot} of a subframe can be calculated as $\frac{7}{14 \times 2^\mu}$ ms.), which follows the 3GPP NR guidelines in [15]. According to [16], the deadline constraint of deterministic aperiodic flows is defined as 1-2 ms. A detailed list of additional parameters used can be found in Table I. All performance results are averaged over 1000 subframes. We present our simulation results of multiple CG configurations in the learning-based adaptive MCG configuration in the GF-NOMA system compared with the MCG-GF-NOMA system assisted by the cooperative multi-agent based double deep Q-network (CMA-DDQN) and single configured-grant GF-NOMA (SCG-GF-NOMA) system over a 5G-TSN network. The flows of each device are generated via random inter-arrival processes across subframes, which have Markov properties [17], [18]. We adopt Beta distribution arrival processes to model the arrival intensity of burst flows. We set the Beta distribution parameters as $\alpha = 3$ and $\beta = 4$, then conducted simulations with the number of access devices ranging from 10,000 to 60,000. The light-load and heavy-load scenarios were set with 10,000 and 50,000 devices, respectively, to analyze the number of successfully transmitted flows under different load conditions.

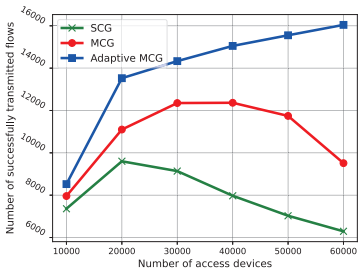


Fig. 4. Successfully transmitted flow numbers vary with different numbers of access devices.

In Fig. 3, we show how the number of successfully transmitted flows in the 5G side network changes as the subframe index varies with Beta distribution in light-load and heavy-load scenarios. From Fig. 3(a), we observe that the proposed learning-based adaptive MCG configuration can improve the number of successfully transmitted flows by 13.71% and 18.29% at the flow peak, compared to the MCG and SCG structures in light-load scenarios, respectively. And from 3(b), the number of successfully transmitted flows increases by more than one and six times significantly in heavy-load scenarios. This is because the flows that failed to be transmitted in the GF stage with the MCG subframe structure and SCG structures can be retransmitted with the proposed learning-based adaptive MCG configuration in the GB stage. The retransmission can improve the number of successfully transmitted flows, especially in the heavy-load scenario.

In Fig. 4, we show how the number of successfully transmitted time-sensitive flows changes as the number of access flows varies. We observe that the proposed learning-

based adaptive MCG configuration can improve the number of successfully transmitted flows compared with the MCG and SCG structures. With the increase in the number of access devices, the benefit can reach 68.66% and 154.97% higher than the MCG and SCG structures, respectively. Both the MCG and SCG structures have the same trend, that the number of successfully transmitted flows first improves and then decreases as the number of access flows increases. This is because as the number of access devices increases, the collision events during random access in the GF stage will surge accordingly. However, the proposed adaptive MCG configuration maintains an increasing trend. This is because the proposed adaptive MCG configuration can retransmit the flows that fail to be transmitted in the GF stage, and the GB manner maintains the flows being successfully transmitted within the deadline.

In Fig. 5, we show how the number of successfully transmitted flows varies over the 5G-TSN network with Beta distribution in light-load and heavy-load scenarios. From Fig. 5, we can see that the proposed learning-based adaptive MCG configuration increases the number of successfully transmitted flows up to 10.71% and 23.81% compared with the MCG structures in light-load and heavy-load scenarios, respectively. Meanwhile, we obtain that the SCG structure can only achieve a bad performance over the 5G-TSN network. However, this performance is lower than Fig. 4, which is only transmitted over the 5G wireless network. This is because a large amount of flows are discarded in the TSN switch due to limitations of maximum forwarding capacity.

VI. Conclusion

In this paper, we propose a novel adaptive MCG subframe structure enhanced 5G-TSN uplink framework aimed at addressing the challenges inherent in integrating 5G and TSN for massive access IIoT applications. The proposed 5G-TSN uplink framework aims to maximize the number of flows successfully transmitted to the receiver. We present a detailed description of the adaptive MCG subframe structure. Subsequently, we propose a CMA-PPO algorithm to enhance the successful transmission count of time-sensitive flows. Finally, simulations are conducted for light and heavy access flow scenarios under Beta distribution models within a GF-NOMA-assisted 5G-TSN network. The results demonstrate that the CMA-PPO-based adaptive MCG configuration in GF-NOMA enabled 5G-TSN significantly outperforms the MCG assisted by CMA-DDQN and SCG in terms of system performance. Future work may focus on integrating this framework with emerging 6G technologies and exploring its applicability in other time-sensitive networking scenarios.

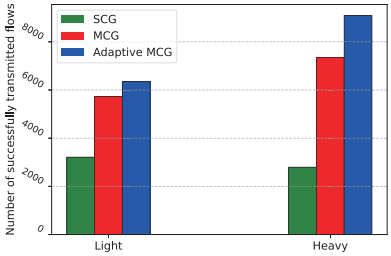


Fig. 5. Successfully transmitted flow number over 5G-TSN network.

References

[1] 3rd Generation Partnership Project (3GPP), "5G NR physical layer procedures for data," 3GPP, Tech. Rep. TS 38.214, Mar. 2020.

[2] Y. Liu, Y. Deng, M. Elkashlan, A. Nallanathan, and G. K. Karagiannidis, "Analyzing grant-free access for URLLC service," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 741–755, Aug. 2020.

[3] J. Zhao, T. Wang, H. Tong, X. He, and C. Yin, "Deadline-aware hybrid HARQ scheme for multi-devices deterministic transmission over 5G-TSN systems," in *Proc. 2023 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2023, pp. 511–516.

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

[5] X. Chen, T. Wang, H. Tong, and C. Yin, "Reconfigurable intelligent surface-assisted grant-free NOMA transmission," in *Proc. 2023 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2023, pp. 460–465.

[6] 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, Dec. 2020.

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

[8] Y. Chen, L. Xu, L. Li, J. Lan, Y. Tang, R. Lu, and C. Chen, "CFS: a CQF-based fast scheduling scheme for time-sensitive networking," in *Proc. 42nd Chinese Control Conference*, Jul. 2023, pp. 6221–6226.

[9] A. Nasrallah, A. S. Thyagaturu, Z. Alharbi, C. Wang, X. Shao, M. Reisslein, and H. ElBakoury, "Ultra-low latency (ULL) networks: The IEEE TSN and IETF DetNet standards and related 5G ULL research," *IEEE Communications Surveys Tutorials*, vol. 21, no. 1, pp. 88–145, Sept. 2018.

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

[11] 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 Transactions on Network Science and Engineering*, vol. 11, no. 1, pp. 543–555, Aug. 2024.

[12] S. Sun, T. S. Rappaport, S. Rangan, T. A. Thomas, A. Ghosh, I. Z. Kovacs, I. Rodriguez, O. Koymen, A. Partyka, and J. Jarvelainen, "Propagation path loss models for 5G urban micro- and macro-cellular scenarios," in *Proc. IEEE Vehicular Technology Conference*, May 2016, pp. 1–6.

[13] M. Xu, D. Niyato, J. Kang, Z. Xiong, and M. Chen, "Learning-based sustainable multi-user computation offloading for mobile edge-quantum computing," in *Proc. ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 4045–4050.

[14] Y. Guan, S. Zou, H. Peng, W. Ni, Y. Sun, and H. Gao, "Cooperative UAV trajectory design for disaster area emergency communications: A multiagent PPO method," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 8848–8859, Mar. 2024.

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

[16] J. Garc ́ı a-Morales, M. C. Lucas-Esta, and J. Gozalvez, "Latency-sensitive 5G RAN slicing for industry 4.0," *IEEE Access*, vol. 7, pp. 143 139–143 159, Sept. 2019.

[17] 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 Communications Surveys Tutorials*, vol. 22, no. 2, pp. 905–929, Feb. 2020.

[18] 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 Communications Surveys Tutorials*, vol. 22, no. 1, pp. 426–471, May 2020.