

# Wireless/Wired Integrated Transmission for Industrial Cyber-physical Systems: Risk-sensitive Co-design of 5G and TSN Protocols

Yajing ZHANG<sup>1</sup>, Qimin XU<sup>2</sup>, Xinping GUAN<sup>3</sup>, Cailian CHEN<sup>4</sup> & Mingyan LI<sup>5</sup>

*Department of Automation, Shanghai Jiao Tong University,  
Key Laboratory of System Control and Information Processing, Ministry of Education of China,  
Shanghai 200240, P. R. China*

---

**Abstract** In order to realize intelligent manufacturing with a dynamic quality of service (QoS) for communication in industrial cyber-physical systems (ICPS), the deterministic transmission is needed in industrial networks by flexible access and reliable forwarding. For this reason, combining the 5th generation (5G) wireless communication technology with wired time-sensitive networking (TSN) technology is a promising solution. However, wireless communication adopted in the industry has characteristics such as diverse QoS of industrial data and complex signaling process, which bring the severe long tail of delay's distribution. These characteristics lead to the indeterminacy of wireless transmission. To address this issue, we propose a heterogeneous time-sensitive network (HTSN) co-designed by 5G and TSN in this paper. To provide hierarchical deterministic delivery in HTSN, a predictive multi-priority wireless scheduling mechanism is designed based on semi-persistent scheduling (SPS), and an adaptive data injection mechanism for TSN based on per-stream filtering and policing (PSFP). Considering the long tail of delay, we develop a risk-sensitive reinforcement learning strategy for realizing deterministic delivery. Simulations on a hot rolling production scenario demonstrate that HTSN with the proposed mechanisms achieves great performance in terms of the heterogeneous transmission.

**Keywords** industrial cyber-physical systems (ICPS), 5G-TSN codesign, wireless/wired integrated network, risk-sensitive reinforcement learning, deterministic communication

---

**Citation** Zhang Y J, Xu Q M, Guan X P, et al.. Wireless/Wired Integrated Transmission for Industrial Cyber-physical Systems: Risk-sensitive Co-design of 5G and TSN Protocols. *Sci China Inf Sci*, for review

---

## 1 Introduction

With the promotion of the fourth industrial revolution (also known as industry 4.0), the manufacturing industry is converted from traditional industrial automation to intelligent manufacturing, where the latter integrates huge amounts of devices that embedded and networked together in industrial cyber physical system (ICPS) [1]. To improve the performance of intelligent manufacturing, such as real-time monitoring and control, deterministic communication is imperative from the industry field to the data center in order to provide real-time control for the control system. In particular, the automatic control system has a strict time demand, where the timely delivery of data is critical for intelligent manufacturing since violating a deadline may severely damage the control quality. However, the diversity of CPS applications deployed in industry field makes it intractable to meet their QoS demands [2].

As a new generation of wireless communication, the technology of 5G is expected to expand industrial informatics and automation into much broader contexts [3]. Among several typical usage scenarios of 5G, ultra-reliable and low latency communication (URLLC) has the characteristics of high reliability up to 99.9999% and low delay down to millisecond [4], which is appropriate for intelligent manufacturing, smart grid, and other automatic control scenarios. These scenarios require extremely high reliability as well as ultra low latency. Besides, massive machine type communication (mMTC) is another target

---

\* Corresponding author (email: xpguan@sjtu.edu.cn)

† Authors 1 and 2 have the same contribution to this work.

scenario for 5G, which is designed for monitoring and data collection in ICPS. Unfortunately, there are three main problems to apply 5G technology to the ICPS. (1) Due to the complex signaling process of the dynamic access mechanism, the conventional cellular network fails to guarantee the strict timelines of industrial automation. (2) The ICPS has many characteristics such as the complex environment, limited radio resources, and severe interference, which lead to the long tail effect of delay distribution. (3) Data's QoS requirements are diverse and dynamical burst data always carries safety information, which make it intractable to deliver data on demand.

Contrary to the 5G technology, TSN standards ensure the determinacy and reliability of its innovative mechanisms such as gate control list (GCL) and time-aware shaper (TAS), which are suitable for URLLC scenarios. Moreover, TSN gateways can deliver data in a guaranteed time window with bounded latency, small jitter, and extremely low data loss [5]. However, TSN is designed for standard ethernet, which has the inherent shortcomings of wired transmission, such as the limited coverage, the higher cost of maintenance, the complexity of installation, etc.

Therefore, it is inevitable to propose a more flexible and deterministic transmission architecture to meet the requirements of intelligent manufacturing. To address this problem, we propose a heterogeneous architecture integrating 5G and TSN, named HTSN. The contributions of this paper are summarized as follows.

(1) **HTSN: a heterogeneous 5G-TSN integrated architecture:**

To provide flexible and deterministic transmission, we propose a heterogeneous architecture HTSN, in which the data flexibly accesses to the TSN gateway via 5G network and forwarded deterministically within TSN network. Under this architecture, the data is delivered in time after two stages' adjustment in 5G and TSN, respectively.

(2) **Predictive multi-priority wireless scheduling mechanism:**

Considering the limited resources of ICPS, we design a preemption based predictive multi-priority wireless scheduling mechanism on the basis of semi-persistent scheduling (SPS). In particular, limited resources are reserved for sensors according to their trigger correlations. Thus, the handshaking delay is reduced and the resource utilization is improved.

(3) **Adaptive data injection mechanism from 5G to TSN:**

To deal with the interference of industrial wireless environment, we propose an adaptive data injection mechanism based on the per-stream filtering and policing (PSFP) mechanism of TSN to offset the jitter caused by 5G communication. The relationship between the priority of queues and the transmission delay through TSN is given to adjust the priority of data while it is injected into TSN network.

(4) **Risk-sensitive learning strategy for wireless scheduling and data injection:**

In consideration of the tail of the transmission delay, a risk-sensitive utility function is proposed, which consists of the expectation, variance, and third-order central moment of the total delay. Leveraging reinforcement learning to adjust the number of reserved resource blocks (RB) and the TSN injection queue, the delay distribution is more centralized with a shorter tail. Thus, the diverse QoS requirements of data are satisfied and the centralized delay distribution brings a more deterministic transmission.

The remainder of this paper is organized as follows. In Section 2, we introduce the related work of heterogeneous network and real-time scheduling, then we give the overview of our HTSN in Section 3. In Section 4, we introduce the system model and formulate the optimization objective. Then we present the risk-sensitive learning strategy and simulation results in Section 5 and Section 6, respectively. Finally, this paper is concluded in Section 7.

## 2 Related Works

The smart factory has received increasing attention for its strict communication demand of lower latency and higher reliability. The accompanying problem is how to improve the transmission determinacy of end-to-end (E2E) delay, which brings two main issues to focus on: the deterministic transmission via a heterogeneous network, the low access delay of the wireless network. Therefore, the related works are concluded in two aspects as mentioned above.

List of Abbreviations

Full Name	Abbreviation	Full Name	Abbreviation
QoS	Quality of Service	MTC	Machine Type Communication
ICPS	Industrial Cyber Physical System	D2D	Device-to-Device
5G	the 5th Generation wireless communication	IWN	Industrial Wireless Network
TSN	Time-Sensitive Networking	BS	Base Station
HTSN	Heterogeneous Time-Sensitive Networking architecture	NS	Non-Scheduled data
SPS	Semi-Persistent Scheduling	TC	Time-Critical data
PSFP	Per-Stream Filtering and Policing	BE	Best Effort data
URLLC	Ultra-Reliable Low Latency Communication	TTI	Transmission Time Interval
mMTC	massive Machine Type Communication	IPV	Internal Priority Values
GCL	Gate Control List	MAB	Multi-Armed Bandits
TAS	Time-Aware Shaper	RSL	Risk-Sensitive Learning
E2E	End-to-End	GMAB	Gradient Multi-Armed Bandits
SDN	Software-Defined Networking	MI	Mutual Information
SR	Scheduling Request	X2	Chi-Square Test
SG	Scheduling Grant	SSA	Sensor Selection Algorithm

## 2.1 Heterogeneous Network Architecture

### 2.1.1 Wireless/Wired Hybrid Network

In recent years, the combination of wireless and wired networks has attracted significant attentions [6–9]. The authors in [6,7] concentrate on using software-defined networking (SDN) to provide centralized control of the wireless/wired network. [6] proposes a network architecture where wireless and wired transmissions are used in parallel to make up for the shortcomings of the other one. [7] proposes a multi-path transmission mechanism under the architecture from wired to wireless to promote the video's transmission performance. Except for SDN, [8] proposes a protocol to support multimedia data in hybrid wireless/wired networks, which efficiently utilizes the wireless link with coexisting TCP flows and can provide satisfactory QoS for delay-sensitive multimedia applications. As for the industry, [9] points out that since current wireless networks are not suitable to fulfill the applications' requirements, hybrid wired-wireless networks have to be developed in order to support the implementation of industrial CPS. However, these studies only focus on the hybrid transmission based on the traditional TCP/IP network while the coming 5G and TSN are more suitable for industrial communication.

### 2.1.2 5G-TSN Integrated Network

The TSN industry white paper recently points out that the URLLC scenario of 5G is the key to realizing the industrial internet. How to deeply combine 5G and TSN is the research hotspot. For this issue, Ericsson proposes a 5G-TSN integration conception in [5], where the SDN controller manages the whole network, and TSN protocols are applied in 5G user planes to guarantee strict E2E delay. Furthermore, Ericsson thinks that the integration of URLLC in the manufacturing process has great potential to accelerate the transformation of the manufacturing industry [10]. Nevertheless, these combination of 5G and TSN are only preliminary ideas which still have a long way to be implemented.

## 2.2 Real-time Scheduling of Wireless Communications

### 2.2.1 Risk-sensitive Transmission

As stated in the introduction, the risk is a notion in financial mathematics [11], which we want to resort to measuring the risk of transmission delay. For wireless communication, the risk is equivalent to the loss of valuable information due to the instability and randomness of wireless transmission. For example, the quality of the stochastic channel may cause a variation of latency, which will incur emergency information lost when the variation is higher [12]. What's more, some fine-grained characteristics of delay in queueing networks, such as the delay distribution and probability boundary(the tail of delay), are critical while most works only care about the average delay rather than the worst delay [13]. In existing works, [14,15] study URLLC and low-latency communication to evaluate the performance under the influence of data

dispersion and network density. They all focus on maximizing the average delay of network throughput or minimizing the average latency without providing any guarantee of the higher moment such as variance, skewness, kurtosis, etc. [16] [17] take the mean and variance of delay account to capture the tail of delay and optimize the bandwidth and transmission power without considering frequency diversity. To maximize the throughput of eMBB data, [18] takes burst URLLC data into account and assumes it is a Rayleigh distribution, while the burst data is actually dynamic sporadic.

### 2.2.2 Semi-persistent Scheduling

Due to the complex "scheduling request-scheduling grant(SR-SG)" signaling mechanism, the conventional dynamic access scheme of the cellular network cannot satisfy the strict delay requirements [19]. To this end, a fast uplink access scheme based on SPS is proposed in LTE Release 13 [20]. Thus the uplink resources are assigned in advance to reduce the "SR-SG" overhead, which is suitable for machine-type communications (MTC) in industrial automation [21]. However, SPS was designed for VoIP originally, the transmission of which is fixed and known, while the QoS of MTC devices varies, and the scheduling request is time-varying and even dynamic sporadic [22]. For this, [23] proposes an adaptive SPS scheme to adjust the resources in the next transmission by buffer report. Moreover, to make further use of unused resources, other devices can be assigned partial resources in a semi-persistent way based on device-to-device(D2D), which can lead to extra latency [24].

### 2.2.3 Multi-priority Wireless Transmission

In the industrial wireless networks(IWNs), the priority of data can be established according to the requirements of applications [25]. Thus the essential requirement of IWNs is to support the transmission of mixed-priority data that have different demands of latency and reliability [26, 27]. For example, as mentioned in [28], there are four different types of information in the industry: the safety/emergency information, which requires the highest reliability and the lowest latency, the regulate/monitor control information, which possesses the lenient requirements, the open-loop control information which allows minutes level delay and the monitoring information which has no requirements. To deal with this problem, [29] proposes a TDMA-based multichannel superframe strategy to keep the superiority of each priority with full radio channel reuse while the strategy is implemented for cluster-based IWN without considering burst data. Similarly, a multi-priority scheme named p-persistent is proposed in [30] based on CSMA, which guarantees the transmission of high priority at the expense of longer delay of low priority data and still doesn't take burst data into account.

In conclusion, although optimizing wireless transmission have been well studied, it's still a big challenge to apply it to industrial automation.

## 3 Overview of HTSN

As is illustrated in Figure 1, the architecture of HTSN consists of two coupling stages: the 5G network and the TSN network. The 5G network is composed of a set of  $\mathcal{S} = (s_1, s_2, \dots, s_{|\mathcal{S}|})$  sensors, communication nodes and controllers, etc. Different sensors collect different types of data. The TSN consists of a set of TSN gateways which also work as 5G base stations (BS). The BSs are in charge of the assignment of the RBs  $\mathcal{R} = (r_1, r_2, \dots, r_{|\mathcal{R}|})$ , where one RB denotes a series of time-frequency domains radio resource blocks and is sufficient to transmit most of the data collected by sensors. RB's corresponding time period, denoted by *Subslot*, is the minimum non-divisible time unit in 5G. The frequency band of one RB represents a channel, which is flat fading and homogeneous. In more detail, time-frequency resources of one RB mean that we can use the channel for a period of time (referred to as a *Subslot* mentioned before). RBs are shared indiscriminately by both SPS data and dynamic data. The industrial field data from different sensors (temperature sensors, vibration sensors, etc.) is transmitted through shared RBs to TSN gateways, then is sent to the remote data center via wired TSN.

Due to the diverse QoS requirements of field data, it is divided into three groups: (1) non-scheduled (NS) data, which is sporadically triggered by emergency data, and has the highest priority. (2) time-critical (TC) data, which is the primary type of data, and the amount of it is much larger than NS

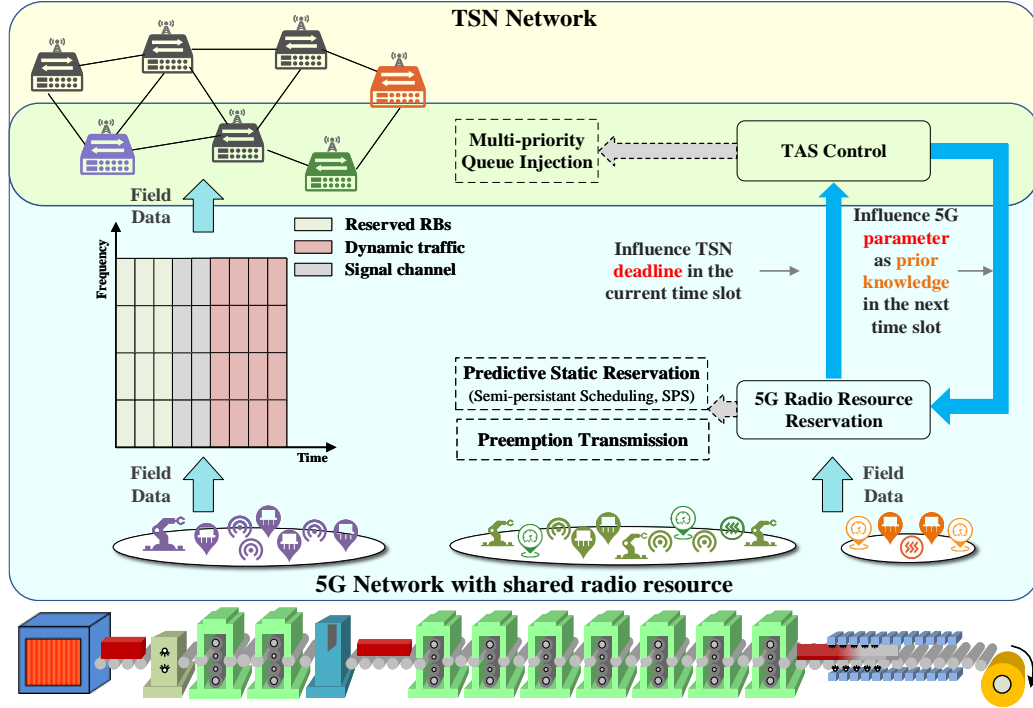


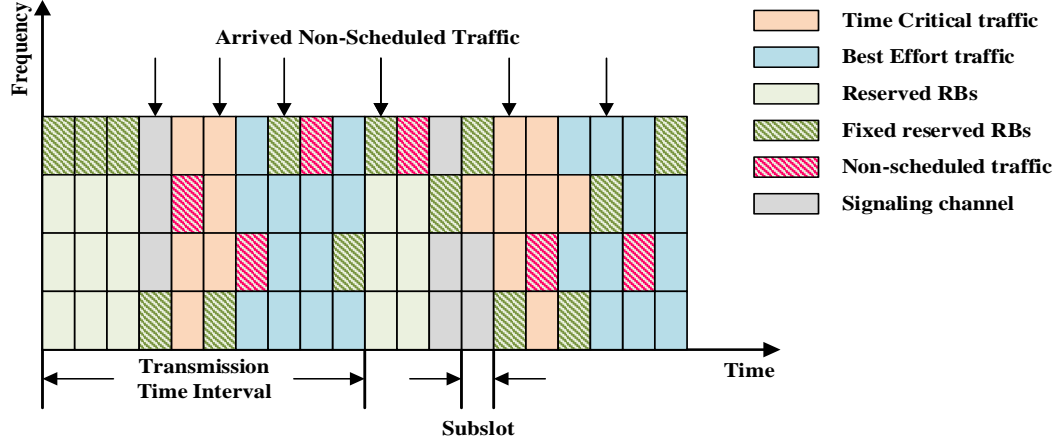
Figure 1 The architecture of HTSN

data. (3) best-effort (BE) data, which has the lowest priority and follows the best-effort forwarding rules. In industry, such as a hot rolling production process, data is transmitted periodically, and so we only consider the transmission in one time period of the 5G network, denoted by transmission time interval (TTI). The whole cycle of industrial automation can be regarded as the accumulation of TTIs. In each TTI, we first select the node that are most likely to trigger, and then we transmit the TC data collected by these nodes on predictive reserved RBs based on SPS technique to omit handshaking delay. The residual TC data and BE data are served on dynamic RBs by the conventional dynamic access procedure orderly, which consists of handshaking in signaling procedure of transmission process. The difference between predictive reserved RBs and dynamic RBs is that the reserved RBs are assigned to sensors in advance without explicit SR-SG signalling procedure. Considering the highest priority and the sporadically triggered feature of NS data, we set a fixed reserved RB at each subslot for possible NS data. To improve the utilization of RBs, the fixed reserved RBs can be used by TC data and BE data with no NS data arriving. And NS data can preempt any data as soon as it arrives. Note that the priority of NS data is the same as TC data after it is embedded in the 5G network.

Based on the PSFP mechanism in the IEEE 802.1Qci protocol, TSN gateways assign different Gate IDs to data according to their transmission delay in 5G and internal priority values (IPV), which is related to its QoS demands. Data arrived at the gateway are injected into the gateway's sending queue that matches its Gate ID assigned before. Thus, data from the 5G network is scheduled differentially in the TSN to meet their diverse QoS requirements and achieve time-sensitive transmission under HTSN.

Specifically, the delay of the TC data (including arrived NS data) across 5G network and TSN network in TTI  $t$ , denoted by  $T_{5G}(t)$  and  $T_{TSN}(t)$  respectively, are coupled. Note that the E2E delay of TC data from the industrial field to the remote data center can be calculated as  $T_{E2E}(t) = T_{5G}(t) + T_{TSN}(t)$ . In each TTI, the total amount of RBs is fixed, so that the larger amount of predictivel reserved RBs  $|\mathcal{R}_r(t)|$  in TTI  $t$ , the smaller the number of RBs for dynamic access. Note that  $|\mathcal{R}_r(t)|$  is much smaller than the number of sensors in field (i.e.,  $|\mathcal{R}_r(t)| \ll |\mathcal{S}|$ ). If  $T_{5G}(t)$  of TC data is larger than excepted, it should be injected into a high priority queue  $Q(t)$  of TSN gateways to offset the time deviation caused by 5G stage. Otherwise, data that has low 5G delay can be injected into low priority TSN queue to reserve resources for other data. Here, we aim to minimize the E2E delay of TC data, so it is necessary to consider both the  $T_{5G}(t)$  and the  $T_{TSN}(t)$ .

**Example 1.** We take the rolling production process as an illustration. As shown in Figure 1, the 5G BS



**Figure 2** The embedment of multi-priority data in 5G network

is installed in the TSN gateway as a local control center with learning ability. At the beginning of each TTI, the BS decides the number of reserved RBs according to the history data. Then sensors correlated with sensors activated before are assigned to predictive reserved RBs and execute signaling procedure in prior. After that, the residual sensors dynamic access BSs with handshaking delays, thus 5G delay  $T_{5G}(t)$  can be calculated according to the transmission state in 5G. Based on  $T_{5G}(t)$  and the QoS requirement of data, the BS decides the Gate ID of the TSN gateway queue that data will inject into.

## 4 System Model

### 4.1 Multi-priority Transmission Model Based on SPS

As discussed in the previous section, there are three issues of 5G network: (1) predictive assign sensors that are the most correlated, (2) embed sensors obeying their diverse QoS requirements, (3) preemption of NS data. To this end, a preemption based predictive multi-priority scheduling mechanism is proposed in this section. The predictive sensor selection policy will be expounded in Section 5.1 later.

As mentioned before, in intelligent manufacturing, a massive amount of data with diverse QoS requirements is generating all the time, which should be served on their time demand. However, the complex uplink "SR-SG" signaling process brings extra delay for handshaking. As a solution, reserving RBs to correlated sensors in advance without a signaling procedure is a suitable way to reduce transmission time. Even so, the urgent and critical data may still have an additional queueing delay due to the shortage of the time-frequency radio resources in the factory. To this end, a preemption mechanism is proposed in this paper, which permits NS data to preempt TC data and BE data in order to realize no-wait transmission as soon as it arrives.

According to the QoS demands of data, we split radio resources into four parts in time order: predictive reserved RBs  $\mathcal{R}_r$  ( $\mathcal{R}_r \subset \mathcal{R}$ ), referred to as reservation area  $\rightarrow$  RBs for signaling transmission  $\rightarrow$  RBs to transmit TC data  $\rightarrow$  RBs to transmit BE data, as shown in Figure 2. Taking NS data that carries safety information into account, we set one fixed reserved RB dedicated to NS data at each subslot. There comes a problem: what if the NS data arrives while the fixed reserved RB has been allocated to the other data already? That should be discussed according to whether the assigned sensor is triggered in Table 2, where  $r_{brt,i}(t)$  denotes fixed reserved RB at  $i$ -th subslot of TTI  $t$ ,  $s_i(t)$  denotes the sensor assigned to  $r_{brt,i}(t)$ ,  $\varphi(t) = \sum_{i=1}^{|\mathcal{R}_r(t)|/|C|} \varphi_i$  denotes the number of fixed reserved RBs in reservation area, to which the sensors assigned are triggered;  $\zeta(t) = \sum_{i=1}^{|\mathcal{R}_r(t)|/|C|} \zeta_i$  denotes the number of arrived NS data in reservation area and  $\gamma(t) = \sum_{i=1}^{|\mathcal{R}_r(t)|/|C|} \gamma_i$  denotes the number of fixed reserved RBs that is preempted by NS data while sensors assigned to them are triggered as well. Note that  $C = \{c_1, c_2, c_3, \dots\}$  is the set of channels and the number of channels is  $|C|$ .

Then the number of RBs which are neither triggered nor preempted in the reservation area can be:

$$|\mathcal{R}_r(t)| - (|\mathcal{S}_e(t)| + \zeta(t) - \gamma(t)) \quad (1)$$

**Table 2** The Condition of Preemption and Trigger on Fixed Reserved RBs

$\gamma_i = \zeta_i \varphi_i$	1	0
$\zeta_i$	$r_{\text{brt},i}(t)$ is preempted	$r_{\text{brt},i}(t)$ is not preempted
$\varphi_i$	$s_i(t)$ is triggered	$s_i(t)$ is not triggered

where  $|\mathcal{S}_e(t)|$  is the number of the TC data that embedded in the reservation area. Thus the total number of high priority sensors at TTI  $t$ , including TC data and embedded NS data, is given by:

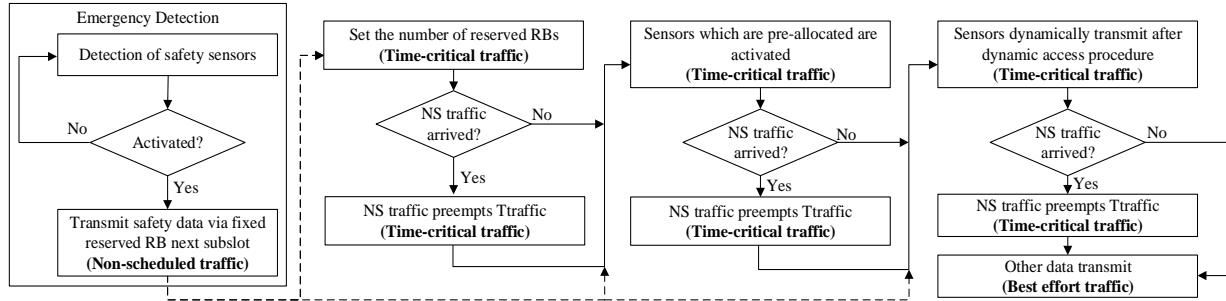
$$\begin{aligned} |\mathcal{S}_h(t)| &= |\mathcal{R}_r(t)| - [|\mathcal{R}_r(t)| - (|\mathcal{S}_e(t)| + \zeta(t) - \gamma(t))] + |\mathcal{S}_c(t)| + |\mathcal{S}_b(t)| - \zeta(t) \\ &= |\mathcal{S}_e(t)| + |\mathcal{S}_c(t)| + |\mathcal{S}_b(t)| + \varphi(t) - \gamma(t) \end{aligned} \quad (2)$$

where  $|\mathcal{S}_h(t)|$  is the total number of high priority sensors (TC data and embedded NS data) scheduled at TTI  $t$ ,  $|\mathcal{S}_c(t)|$  and  $|\mathcal{S}_b(t)|$  is the number of sensors generating TC data for dynamic access and the number of sensors generating NS data respectively.

Ignore the preemption of NS data, sensors pre-allocated to reserved RBs are scheduled firstly, followed by dynamic access sensors with extra signaling delay, the BE data is scheduled in the end if there are any resources left. The whole transmission process is shown in Figure 3. It worth mentioning that we here only focus on the delay of TC data, which is the time difference between the last served TC data and the beginning of the current TTI. The 5G delay can be calculated as follows:

$$T_{5G}(t) = \left\lceil \frac{|\mathcal{R}_r(t)| + (|\mathcal{S}_c(t)| + |\mathcal{S}_b(t)| - (\zeta(t) - \gamma(t)))(1 + \delta)}{|C|} \right\rceil * T_{RB} \quad (3)$$

where  $T_{RB}$  is the time duration corresponding to one RB,  $\delta$  is the proportion of signaling in the overall transmission and  $\lceil * \rceil$  represents the smallest integer which  $*$  is less than or equal to. It can be seen from Figure 2 that if the last TC sensor is left in the last column alone, the 5G delay will be extended for one  $T_{RB}$ .

**Figure 3** The process of Multi-priority data Transmission with Preemption Considered

## 4.2 Data Injection Model of TSN Gateways

As mentioned in Section 3, there are several queues of each port of TSN gateways, which have different priorities while transmitting via TSN. The coordination of TAS and GCL can make sure the data forwarded in TSN can be deterministically delivered within its time demand while keeping its priority.

PSFP mechanism proposed by IEEE 802.1Qci points out that each data arrived owns a priority number, named IPV, to be a reference to transmit data heterogeneously in TSN. In more detail, every data will be assigned a Gate ID to match its IPV number, where each Gate ID represents a TSN sending queue, so the data with different IPV will be injected into different TSN queues. It can be seen that the key to deciding which queue to inject is the value of IPV, so we propose a data injection mechanism of TSN gateways based on IPV to offset the time deviation such as jitters caused by 5G network. Thus, the TSN delay is reduced according to the 5G network delay. Note that the data with lower IPV number has higher priority.

Outlined in Figure 4, the data from the 5G network will be firstly gathered into a frame pool, where data generated within a TTI is classified into different priorities (referred to as IPV numbers) according to its 5G delay, and then it is injected into different TSN queue dynamically.

First, we divide the forwarding delay of TC data via TSN into  $x$  parts:



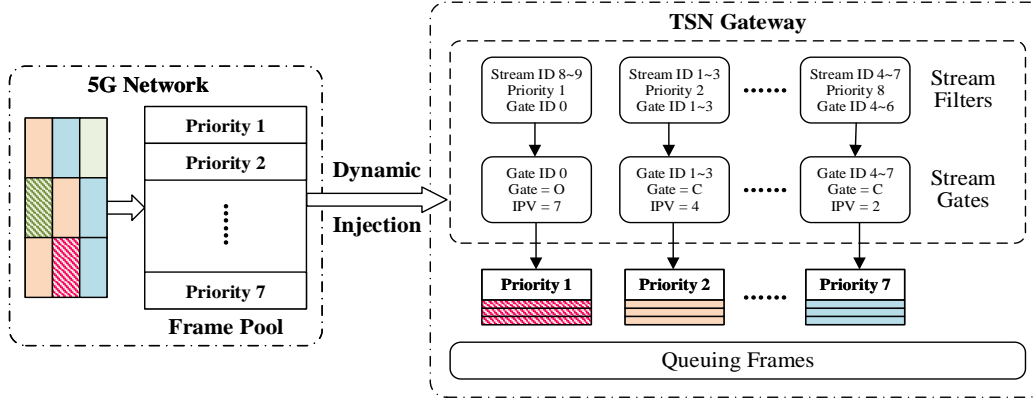


Figure 4 Injection Policy of TSN Queue Based on PSFP

$$\Delta = \frac{T_{\text{TSN}}^{\max} - T_{\text{TSN}}^{\min}}{x}, x = 1, 2, \dots, \dots \quad (4)$$

where  $T_{\text{TSN}}^{\max}$  and  $T_{\text{TSN}}^{\min}$  denotes the maximum and the minimum forwarding delay through TSN respectively,  $x$  is the number of queues of a TSN port, which is usually eight.

Thus we can get the forwarding deadline of each queue as follows:

$$\Lambda(Q(t)) = T_{\text{TSN}}^{\min} + Q(t) * \Delta, Q(t) = 1, 2, \dots, x \quad (5)$$

where  $Q(t)$  is the IPV number of arriving data from 5G at TTI  $t$ . Therefore we get the forwarding delay function of  $Q(t)$ , which can be used to calculate the delay of TSN and get the number of  $Q(t)$  in reverse.

The TSN delay with different queue is calculated by:

$$T_{\text{TSN}}(t) = H \left\lceil \frac{|\mathcal{S}_h(t)| * D * \Lambda(Q(t))}{\theta_{\text{low}} T_{\text{cyc}}} \right\rceil T_{\text{cyc}} \quad (6)$$

where  $|\mathcal{S}_h(t)|$  has been given in Section 4.1,  $H$  is the number of hops within TSN with fixed beginning and terminal,  $D$  is the fixed amount of data that one RB transmit,  $T_{\text{cyc}}$  is the forwarding cycle of TSN gateways and  $\theta_{\text{low}}$  is the lowest data rate of TSN.

Based on the TSN delay we get above, the value of  $Q(t)$  is obtained by  $\Lambda(Q(t))$  as follows:

$$Q(t) = \arg \min_{1 \leq Q(t) \leq x} (T_{\text{TSN}}(t) - (T_{\text{ddl}}(t) - T_{5G}(t))) \quad (7)$$

where  $T_{\text{ddl}}(t)$  is the delay demand of TC data arrived at TTI  $t$ , and  $T_{5G}(t)$  is the transmission delay through previous 5G network.

Thus the whole transmission delay under HTSN at TTI  $t$  is given by:

$$T_{\text{E2E}}(T_{5G}(t), T_{\text{TSN}}(t)) = T_{5G}(t) + T_{\text{TSN}}(Q(t), \mathcal{F}(T_{5G}(t))) \quad (8)$$

Obviously, the delay of 5G network  $T_{5G}(t-1)$  and the delay of TSN  $T_{\text{TSN}}(t-1)$  is coupling and interact with each other as shown in Figure 5. In particular, we first take the cumulative transmission delay  $\sum_{m=1}^{t-2} T(m)$  as the prior information to obtain the number of reserved RBs  $\mathcal{R}_r(t-1)$  at TTI  $t-1$ . Then according to the 5G delay calculated by  $\mathcal{R}_r(t-1)$  and the coupling relationship between 5G delay and TSN delay, we can get the IPV number of TC data based on the data injection model above. Finally, on the basis of the newly get  $T_{5G}(t-1)$  and  $T_{\text{TSN}}(t-1)$ , we can get the whole transmission delay  $T_{\text{E2E}}(t-1)$  under HTSN at TTI  $t-1$ , whereby it can be added into the cumulative transmission delay  $\sum_{m=1}^{t-1} T(m)$  and become the prior information of TC data in TTI  $t$ .

### 4.3 Risk-sensitive Utility Formulation

Considering the unreliability of 5G network, the higher-order quantity of wireless delay should be involved in the optimal problem as mentioned in Section 1. In this regard, we apply entropic risk measure  $\frac{1}{\rho} \ln(\mathbb{E}[\exp(\rho T)])$  and formulate a risk-sensitive minimization utility function of the whole transmission delay under HTSN with as follows [13]:



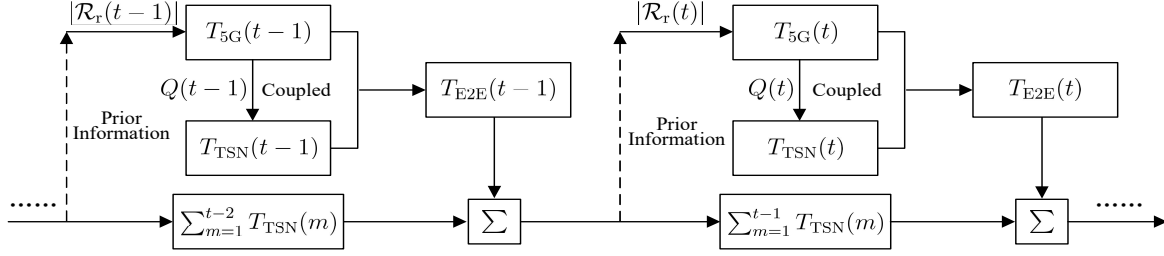


Figure 5 The flow of TC data under HTSN

$$\begin{aligned}
 \mathcal{P1} : \quad & \min_{\{|\mathcal{R}_r(t)|, Q(t)\}} \frac{1}{\rho} \ln(\mathbb{E}[\exp(\rho \sum_{m=1}^{t-1} T(m))]) \\
 \text{s.t.} \quad & |\mathcal{R}_r(t)| \leq |\Pi^{\text{sel}}| \\
 & 1 \leq Q(t) \leq 8
 \end{aligned} \tag{9}$$

where  $\mathbb{E}[*]$  is the expectation operator,  $|\Pi^{\text{sel}}|$  represents the number of predictive selected candidate sensors which is described in detail in Section 5.1.

The entropic risk measure  $\frac{1}{\rho} \ln(\mathbb{E}[\exp(\rho \mathcal{T})])$  can be expanded out as  $\frac{1}{\rho} \ln(\mathbb{E}[\exp(\rho \mathcal{T})]) = \mathbb{E}[\mathcal{T}] + \frac{\rho}{2!} \text{Var}(\mathcal{T}) + \frac{\rho^2}{3!} \mathbb{E}[(\mathcal{T} - \mathbb{E}[\mathcal{T}])^3] + \dots$ , where  $\mathcal{T}$  denotes the cumulative time  $\sum_{m=1}^{t-1} T(m)$ . It is obvious that the optimal object takes into account the variance  $\text{Var}(\mathcal{T})$  and the third central moment  $\mathbb{E}[(\mathcal{T} - \mathbb{E}[\mathcal{T}])^3]$  of  $\mathcal{T}$ . Note that the skewness of  $\mathcal{T}$  equals to  $\mathbb{E}[(\mathcal{T} - \mathbb{E}[\mathcal{T}])^3] / \text{Var}(\mathcal{T})^{\frac{3}{2}}$ . In other words, we formulate the optimization problem in the view of the mean, variance, skewness and other high-order quantity of the cumulative time  $\sum_{m=1}^{t-1} T(m)$ . Additionally, the parameter  $\rho > 0$  reflects the weight of high-order statistics.

Due to the function  $\frac{1}{\rho} \ln(*)$  is monotonically increasing, we remove it and focus on an equivalent utility problem as follows:

$$\begin{aligned}
 \mathcal{P2} : \quad & \min_{\{|\mathcal{R}_r(t)|, Q(t)\}} \mathbb{E}[\exp(\rho \sum_{m=1}^{t-1} T(m))] \\
 \text{s.t.} \quad & |\mathcal{R}_r(t)| \leq |\Pi^{\text{sel}}| \\
 & 1 \leq Q(t) \leq 8
 \end{aligned} \tag{10}$$

which can be expanded by Maclaurin series expansion analogously, i.e.,  $\mathbb{E}[\exp(\rho \mathcal{T})] = 1 + \rho \mathbb{E}[\mathcal{T}] + \frac{\rho^2}{2!} \mathbb{E}[\mathcal{T}^2] + \frac{\rho^3}{3!} \mathbb{E}[\mathcal{T}^3]$  for  $\mathcal{T} = \sum_{m=1}^{t-1} T(m)$ .

It is challenging to solve the minimization problem because of TSN's dynamic network state and the unknown environmental factors of each gateway. Thus, we leverage the principles of multi-armed bandits(MAB) in reinforcement learning to optimize the long-term transmission delay in a decentralized manner.

## 5 Risk-sensitive Learning Strategy for Predictive Scheduling

For the minimization problem we formulated above, a risk-sensitive learning (RSL) strategy is proposed to solve the problem in a static stage and a dynamic stage. The static stage attempts to improve the prediction precision and constrain the number of reserved RBs by selecting sensors through the access history. Then the total delay and the long delay tail of  $T_{\text{E2E}}$  are considered in the dynamic stage based on the gradient multi-armed bandits (GMAB) scheme.

### 5.1 Static Stage: Sensor Selection Algorithm Based on Naive Bayesian Model

As discussed above, the main task of the static stage is to select the most correlated sensors to compose the reservation candidates set. Obviously, the utilization of reserved RBs is determined by whether assigned sensors trigger. So a high prediction precision is necessary to reduce 5G transmission delay. Besides, a salient feature of industrial automation is that the trigger of one event-triggered MTC device

may increase the probability that other devices in the vicinity also generate data in quick succession [3,31]. Based on this, calculate the correlation between sensors at two adjacent TTIs is the key to predictive select sensors to be reserved. Here we don't care about the relationship between sensors but only concern the correlation between their trigger. In the literature [3], we apply the Naive Bayesian model to learn the correlation of sensors since it simplifies the assumption of conditional independence.

Let  $\mathbf{x}(t) = \{x_1(t), x_2(t), \dots\}$  denotes the set of sensors triggered at TTI  $t-1$ , we select sensors to make up the reservation candidates set  $\mathbf{y}(t) = \{y_1(t), y_2(t), \dots\}(\mathbf{x}(t) \cap \mathbf{y}(t) = \emptyset)$ , which may be triggered at the TTI  $t$  with a high probability. Note that the scale of  $\mathbf{y}(t)$  lies on the threshold  $\alpha$  we set and is not related to the scale of  $\mathbf{x}(t)$ .

In order to explore the correlation between sensors, the access history is used to get the trigger probability of sensor  $y_i(t)$  as  $\mathbb{P}(y_i(t))(y_i(t) \in \mathcal{S})$ , then three metrics are used to measure the correlation between the set  $\mathbf{x}(t)$  and the sensor  $y_i(t)$  as follows [32]:

1) **Conditional probability**: Conditional probability at TTI  $t$  is the probability that sensors out of  $\mathbf{x}(t)$  will trigger after any sensors in  $\mathbf{x}(t)$  that has been triggered, which can be obtained by:

$$\mathbb{P}_C(y_i(t), \mathbf{x}(t)) = \mathbb{P}(y_i(t) | x_j(t)) = \sum_{x_j(t) \in \mathbf{x}(t)} \frac{\mathbb{P}(y_i(t), x_j(t))}{\mathbb{P}(x_j(t))} \quad (11)$$

where  $\mathbb{P}(y_i(t), x_j(t))$  is the joint probability of  $y_i(t)$  and  $x_j(t)$ , and it is calculated as  $\mathbb{P}(y_i(t), x_j(t)) = \mathbb{P}(y_i(t) | x_j(t)) * \mathbb{P}(x_j(t))$ .

2) **Mutual information (MI)**: As described in information theory, MI can measure the development of the trigger probability of  $y_i(t)$  after  $\mathbf{x}(t)$  is triggered, which is given by:

$$\mathbb{P}_{MI}(y_i(t), \mathbf{x}(t)) = \sum_{x_j(t) \in \mathbf{x}(t)} \mathbb{P}(y_i(t), x_j(t)) \log_2 \frac{\mathbb{P}(y_i(t), x_j(t))}{\mathbb{P}(y_i(t))\mathbb{P}(x_j(t))} \quad (12)$$

where  $\mathbb{P}(y_i(t), x_j(t)) = \mathbb{P}_C(y_i(t), x_j(t)) * \mathbb{P}(x_j(t))$ .

3) **Chi-square  $\chi^2$** : Chi-square ( $\chi^2$ ) test is a suitable way to estimate the relevance of two sensors by comparing  $\mathbb{P}(y_i(t), x_j(t))$  and  $\mathbb{P}(y_i(t)) * \mathbb{P}(x_j(t))$ , so to judge whether these two sensors are dependent, we use  $\chi^2$  test metric as below:

$$\mathbb{P}_{\chi^2}(y_i(t), \mathbf{x}(t)) = \sum_{x_j(t) \in \mathbf{x}(t)} \frac{(\mathbb{P}(y_i(t), x_j(t)) - \mathbb{P}(y_i(t)) * \mathbb{P}(x_j(t)))^2}{\mathbb{P}(y_i(t)) * \mathbb{P}(x_j(t))} \quad (13)$$

The aim of these three metrics is to evaluate the correlation between  $y_i(t)$  and  $\mathbf{x}(t)$ . Thus we can select the most relevant sensors to  $\mathbf{x}(t)$  to be reserved predictively at TTI  $t$  for higher resource utilization and lower transmission delay. Thus we use a policy  $\Pi^{\text{sel}} = \{\pi_1, \pi_2, \dots\}$  to choose which sensor to reserve as:

$$\pi_i = \begin{cases} 1, & \text{if } \mathbb{P}(y_i(t), \mathbf{x}(t)) \geq \alpha \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

As we can see from the equation (14), the number of the reserved RBs depends on the threshold  $\alpha$  we set, since  $\pi_i = 1$  when  $\mathbb{P}(y_i(t), \mathbf{x}(t)) \geq \alpha$ . In fact, there exists a trade-off of the number of reserved RBs  $|\mathcal{R}_r(t)|$  for the following reasons. If  $|\mathcal{R}_r(t)|$  is too small, most of the sensors still need to access the 5G network dynamically with a high handshaking cost, which may violate the time demand of NS data and TC data. On the other hand, if  $|\mathcal{R}_r(t)|$  is too big but the prediction precision cannot be guaranteed, the time-frequency resources left will be insufficient for for dynamic access sensors. Besides, the reserved resources will also be wasted a lot. Therefore,  $\alpha$  is a critical factor to balance this trade-off.

Form [3] we know that sensors with lower trigger frequency will obtain a higher ranking in the  $\chi^2$  test than in the MI. Based on this, we propose a sensor selection algorithm that is outlined in Algorithm 2 for good practicability. Note that the  $i$ -th elements of  $n$  history access samples set  $\mathbf{H}_n$  is boolean set  $\mathbf{h}_i$ . And  $\beta$  is the threshold of trigger time used to select a metric.

In static stage, we use three different metrics to measure the correlation of sensors according to different scenarios, which can improve the precision of sensor selection, and then reduce the delay of 5G network by multi-priority scheduling based on SPS while considering preemption.

**Algorithm 1** Sensor Selection Algorithm(SSA)

---

**Input:** Access samples at last TTI  $\mathbf{x}(t) = \{x_1(t), x_2(t), \dots\}$ ;  
History access samples  $\mathbf{H}_n = \{\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n\}$ ;  
Trigger time set  $T = \{0\}$  with length  $n$ ,  $\mathbf{y}(t) = \emptyset$ ;  
**Output:** Predictive select sensors sets  $\mathbf{y}(t) = \{y_1(t), y_2(t), \dots\}$ ;

```

1: while  $j \leq n$  do
2:    $T = T + h_j$ 
3: end while
4: while  $i \leq |\mathbf{x}(t)|$  do
5:   if  $T(i) \leq \beta$  and  $y_i(t) \notin \mathbf{x}(t)$  then
6:      $\mathbb{P}(y_i(t), \mathbf{x}(t)) = \mathbb{P}_{\chi^2}(y_i(t), \mathbf{x}(t))$ ;
7:   else  $\{T(i) > \beta$  and  $y_i(t) \notin \mathbf{x}(t)\}$ 
8:      $\mathbb{P}(y_i(t), \mathbf{x}(t)) = \mathbb{P}_{\text{MI}}(y_i(t), \mathbf{x}(t))$ ;
9:   end if
10:  if  $\Pi_{\text{sel}}$  then
11:     $\mathbf{y}(t) = \mathbf{y}(t) \cup \{y_i(t)\}$ ;
12:  end if
13: end while

```

---

**5.2 Dynamic Stage: RSL Algorithm Based on GMAB**

As mentioned in Section 5.1, the number of reserved RBs  $|\mathcal{R}_r(t)|$  brings a trade-off. There are three parameters that are influenced by this trade-off such as the 5G delay, the IPV number of data when it is injected into TSN gateway, and the whole transmission delay under HTSN. Each TSN gateway needs to make decisions based on a limited state information every TTI to find out the optimal policy to dynamic reserve RBs and inject data into its sending queue while guaranteeing the time demand as well. This predictive pre-allocation problem with no prior knowledge is skin to the famous multi-armed bandits' problem [3, 33, 34], which also concerns the balance of multiple arms' exploration and exploitation to gain the long-term rewards via trying to choose different arms.

We here leverage the GMAB tool to solve the predefined problem since it only focuses on the relative preferences between actions rather than the value of the action itself. In particular, each TSN gateway acts as an agent that selects an action to maximize the long-term rewards. The action set is defined as  $k = (k_1, k_2, \dots, k_{|k|})$ , where  $k_i = (|\mathcal{R}_r(i)|, Q_i)$  denotes the corresponding action of the  $i$ -th arm. The policy gateway made at TTI  $t$  is given by  $\varpi(t) = \{\varpi_1(t), \varpi_2(t), \dots, \varpi_{|k|}(t)\}$ , which means the agent chooses the  $i$ -th arm with probability  $\varpi_i(t)$  at TTI  $t$ . Here we denote the long-term delay in (10) as a utility function  $U(t) = -\exp(\rho \sum_{m=1}^{t-1} T(m))$ . The main step of GMAB algorithm are outlined as follows:

- (1) Each TSN gateway is given an initial policy:  $\varpi(0) = \{\varpi(0, 1), \varpi(0, 2), \dots, \varpi(0, |k|)\}$ , and the preference function  $H(t, k_i)$  of each action is the same (i.e.  $H(0, k_i) = 0$  for all  $i$  in  $|k|$ ).
- (2) At every TTI  $t$ , each TSN gateway selects actions according to the policy updated at last TTI  $t-1$ , then obtain an utility function  $U(t) = -\exp(\rho \sum_{m=1}^{t-1} T(m))$  as the reward of the actions selected.
- (3) Then, TSN gateways calculate each action's probability by a soft-max distribution based on the preference function got before:

$$\Pr\{k_i = (|\mathcal{R}_r(i)|, Q_i)\} \doteq \frac{e^{H(t, k_i)}}{\sum_{j=1}^{|k|} e^{H(t, k_j)}} \doteq \varpi(t, k_i) \quad (15)$$

and get the new policy  $\varpi(t)$ .

- (4) Subsequently, each TSN gateway updates the preference of actions by:

$$\begin{aligned} H(t+1, k_i) &\doteq H(t, k_i) + \eta (U(t) - \bar{U}(t)) (1 - \varpi(t, k_i)), \quad \text{and} \\ H(t+1, k_j) &\doteq H(t, k_j) - \eta (U(t) - \bar{U}(t)) \varpi(t, k_j), \quad \text{for all } k_j \neq k_i \end{aligned} \quad (16)$$

where  $\eta > 0$  is a step-size parameter, and  $\bar{R}(t)$  is the average rewards up to TTI  $t-1$ , which can be computed incrementally [35].

- (5) TSN gateways iteratively update its policy  $\varpi(t)$  severally and make new decisions abide by the latest policy, so that the likelihood of choosing the optimal action is proportional to its rewards.

Since we take high order quantity into account when we formulate the minimization problem, the tail of transmission delay is optimized as well as the optimal policy is learned via reinforcement learning.

## 6 Numerical Results

### 6.1 Sensor Prediction Simulation

In this section, we evaluate the proposed HTSN architecture's performance based on data with a deterministic deadline, which generated by monitoring sensors deployed along the hot rolling line. Sensors for monitoring temperature, humidity, pressure are randomly distributed to sense the manufacturing process to provide robust control via closed-loop feedback. Sensors active periodically and generate TC data, which is the main type of data we care about. Event-triggered sensors such as camera and vibration sensors only active as long as there occur safety emergencies suddenly. They generate burst NS data, which has the highest priority and can preempt TC data.

Here, data from different processes with its particular transmission deadline is transmitted in turns. What we need to do is to learn the correlation between sensors and despite the interference brought by sensors that have no relationship with the arrivals of steels. We consider the probability and proportion of field sensors as Table 3.

The performance of the predictive sensor selection algorithm (SSA) is evaluated in terms of prediction accuracy(successful prediction ratio). We compare the performance of three correlation metrics(X2: Chi-square test, MI: Mutual Information, Cond: Conditional Probability) and the sensor selection algorithm we proposed on three sizes(3000, 4500, 6000) of data. Figure 6 shows how different correlation metrics diverse in the performance of prediction accuracy. It is obvious that the  $\chi^2$  metric, as well as the conditional metric has the best performance while the iteration times are few. The prediction accuracy of Mutual Information is much lower and flat than the other two. But after several times of iteration, the difference between these three metrics' performance is inconspicuous, and the prediction accuracy of them all converges to 0.8. Since the SSA algorithm we proposed chooses the metric by threshold  $\alpha$  and  $\beta$ , the performance of it is as good as the best metric. The high prediction accuracy guarantees the resource utilization of 5G and lays the foundation for the HTSN architecture.

### 6.2 HTSN Simulation

Since the 5G network and TSN are coupling and interactive, the main influence factor of HTSN is the number of reserved RBs  $|\mathcal{R}_r(t)|$ . Thus, we simulate the relationship between  $|\mathcal{R}_r(t)|$  and delay of 5G/TSN. Figure 7 shows that as  $|\mathcal{R}_r(t)|$  growing, the delay of the 5G network decrease firstly and then increases. The lowest point of the delay curve is the trade-off point we mentioned before, which is because the prediction accuracy cannot reach 100%, thus assigning sensors in prior would cause the waste of radio resources. Here we set the total number of *Subslot* to 15, and when  $|\mathcal{R}_r(t)| = 7$ , the performance of the pre-allocation mechanism is optimal. What's more, the reason why the delay of 5G network increases rapidly after the trade-off point is the limitation of radio resource, since the more  $|\mathcal{R}_r(t)|$  is reserved, the

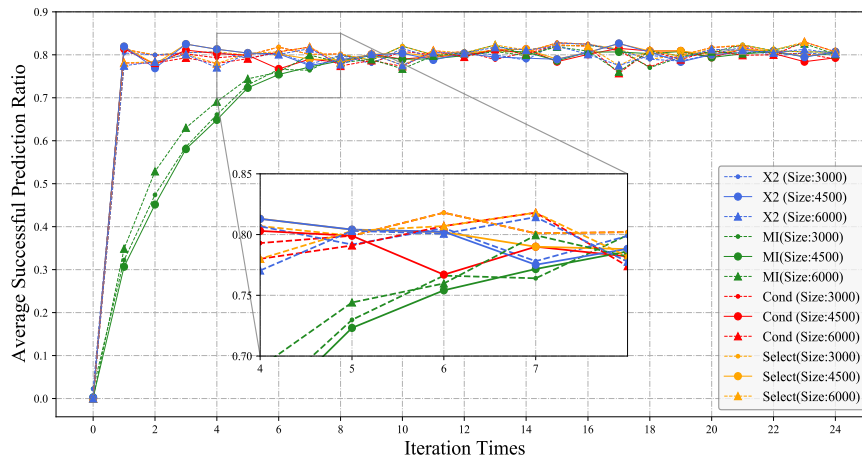
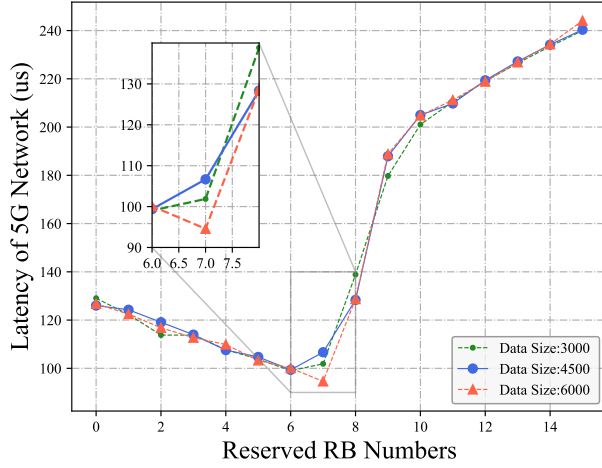
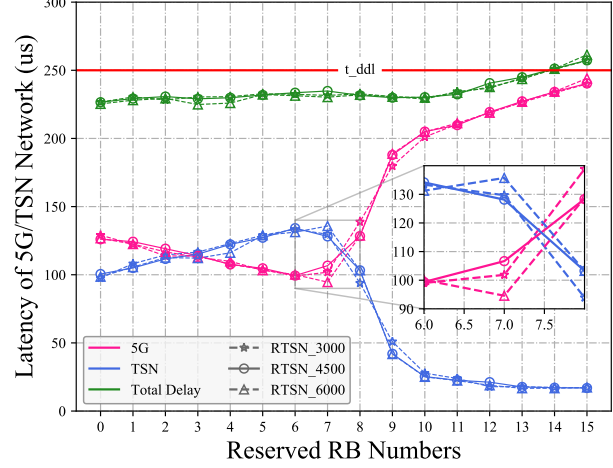


Figure 6 The prediction performance of different correlation metrics

Figure 7 The varies of 5G delay with  $|\mathcal{R}_r(t)|$ Figure 8 The varies of 5G and TSN delay with  $|\mathcal{R}_r(t)|$ 

more resource is wasted accordingly, the fewer resources used for dynamic access and the lower reliability of transmission.

Based on the dynamic injection of the TSN queue, the indeterminacy of transmission caused by the 5G network can be made up by the TSN to meet the strict transmission requirements of intelligent manufacturing. Figure 8 shows the interaction of 5G and TSN as  $|\mathcal{R}_r(t)|$  growing. It can be seen that the delay of 5G and TSN has a near-ideal negative correlation via adjusting the TSN queue priority self-adaptively. Given the deadline =  $250\mu s$ , the TSN tries to save as many resources as it can under the premise of not violating the deadline of the total delay under HTSN. Obviously, the data injection mechanism we proposed makes up for the shortcoming of the 5G network perfectly and improves the determinacy since it nearly not violates the deadline.

To further analyze the relationship between TSN queue and 5G delay, we give Figure ??, from which we can see that the IPV number of data and the 5G network delay also have a similar negative correlation. Since the IPV number of data is directly proportional to its TSN delay, the general trend of the queue priority curve and TSN delay curve is analogous. However, there are some step-changes of the IPV number as it is an integer. It can also be concluded that although we arrange a IPV number between 0 and 8, we can't assign 5G data to the queue corresponding to IPV 8 since the scheduling capability of the 5G network is not enough to realize low latency high-reliability transmission by itself.

Except for  $|\mathcal{R}_r(t)|$ , the ratio of signaling transmission is also a critical factor that influences the SPS within the 5G network and the total delay of HTSN. Here, we formulate the extra delay caused by signaling as a signal ratio and plot the relationship between it and 5G delay. As Figure ?? shows, we select  $|\mathcal{R}_r(t)|$  from 6 to 9 according to Figure 7 to find the influence among signal ratio,  $|\mathcal{R}_r(t)|$  and delay of 5G. It can be seen that no matter  $|\mathcal{R}_r(t)|$  is larger, equal, or less than the trade-off point, the delay of 5G network increases with the increase of signal ratio. This is because  $|\mathcal{R}_r(t)|$  has nothing to do with the number of TC data; there always exists TC data to be dynamically scheduled due to the inaccuracy of prediction, so the increase of signal ratio will lead to the increase of the 5G network's delay.

### 6.3 Risk-sensitive Learning Simulation

To deal with the tail of 5G delay, we use a risk-sensitive utility function to optimize expectations and high order quantities of accumulative integral delay of HTSN. The proposed HTSN architecture is

Table 3 Trigger of Sensors Setting

Proportion	NS data	TC data	BE data
Proportion 1	6%	40%	54%
Proportion 2	4%	40%	56%
Proportion 3	2%	40%	58%
Proportion 4	2%	50%	48%
Proportion 5	2%	30%	68%

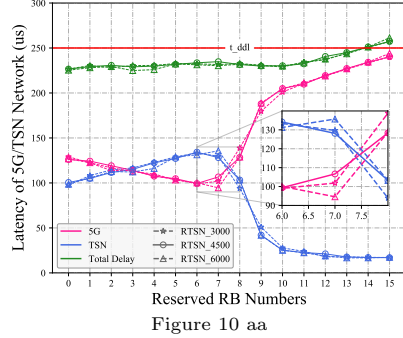


Figure 10 aa

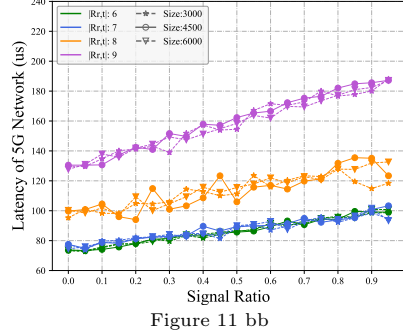


Figure 11 bb

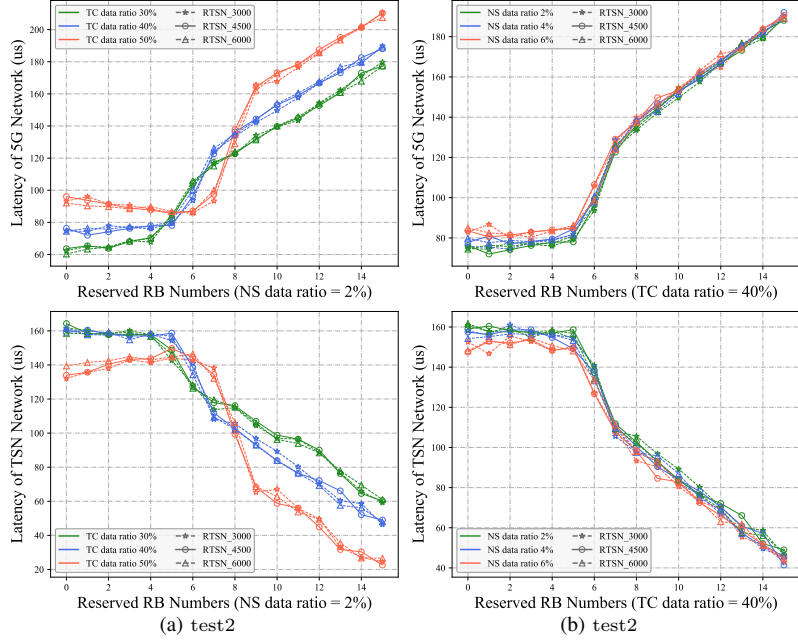


Figure 11 subfigure test

evaluated as an integral in Figure 12, which delineates the improvement brought by the risk-sensitive utility. We repeat GMAB learning for three sizes (20,100,500) of independent runs, and each run we measure its performance with experience over 1000 time steps. With the increase of repetitions, the risk-sensitive learning curve is more stable, and the learning reward is getting closer to the normalized true reward. It means that the learning rate of risk-sensitive learning is faster than classical learning. Besides, with the increase of iteration times, the convergence rate of risk-sensitive learning is faster.

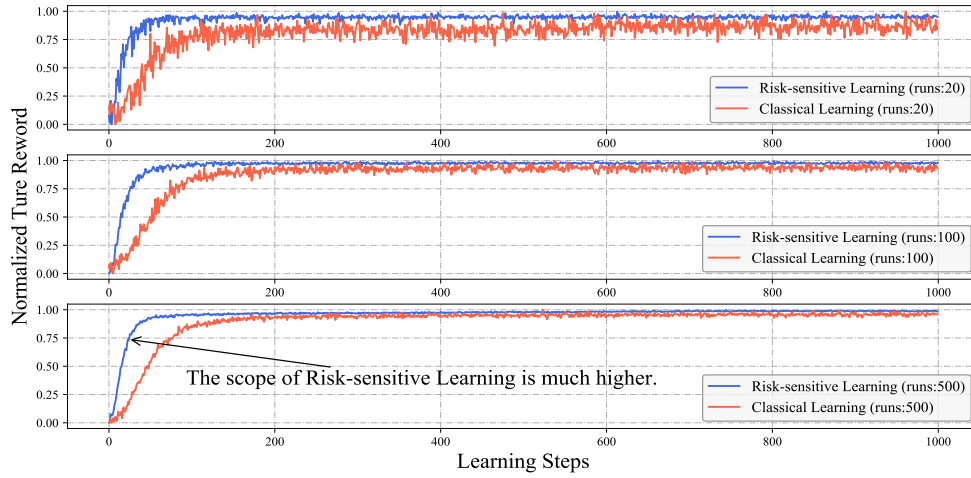
As for the performance of the total delay shown in Figure 13, the curve of Figure 12 and Figure 13 are analogous with the same parameter settings. It can be seen that the curve of risk-sensitive learning is steeper than the classical one, which means that the high order quantities are optimized while learning, and the delay distribution of HTSN is more centralized with a shorter tail. It proves the risk-sensitive reinforcement learning strategy we proposed works and develops the transmission reliability.

## 7 Conclusion

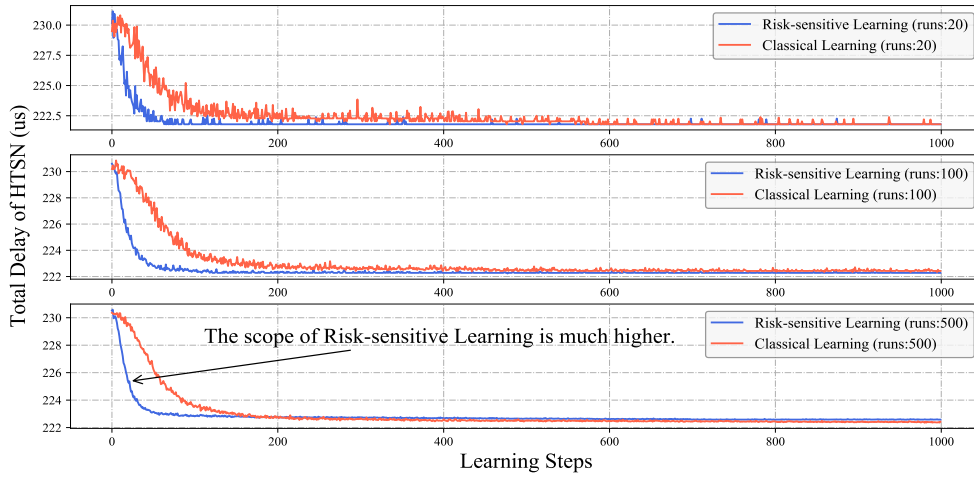
In this paper, we proposed HTSN, a heterogeneous architecture that co-designs 5G and TSN to provide deterministic transmission from the industrial field to the remote data center. Within the 5G network, a preemption considered multi-priority wireless scheduling mechanism was proposed to satisfy the diverse QoS requirements of field data by exploring the trigger correlation between uplink sensors in industrial process automation. To offset the indeterminacy caused by the 5G network, an adaptive data injection mechanism of the TSN gateway queue was developed here to reduce the total delay under HTSN. Then we formulated a risk-sensitive utility function, which considers the high order quantity of the total delay into account. Finally, a risk-sensitive reinforcement learning based on GMAB is executed and thus the determinacy of HTSN is improved.

## References

- 1 J. Zhou, L. Li, A. Vajdi, X. Zhou, and Z. Wu, "Temperature-constrained reliability optimization of industrial cyber-physical systems using machine learning and feedback control," *IEEE Transactions on Automation Science and Engineering*, 2021.
- 2 X. Zhou, W. Liang, S. Shimizu, J. Ma, and Q. Jin, "Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5790–5798, 2020.
- 3 M. Li, X. Guan, C. Hua, C. Chen, and L. Lyu, "Predictive pre-allocation for low-latency uplink access in industrial wireless networks," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pp. 306–314, IEEE, 2018.
- 4 M. Li, C. Chen, C. Hua, and X. Guan, "A learning-based pre-allocation scheme for low-latency access in industrial wireless networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 650–664, 2019.



**Figure 12** The normalized reward of risk-sensitive reinforcement learning with different run times



**Figure 13** The total delay of HTSN with different learning run times

- 5 J. F. B. V., G. M., and J. Sachs, "5g-tsn integration meets networking requirements," *Ericsson Technology Review*, 2019.
- 6 S. Fu, J. Wu, H. Wen, Y. Cai, and B. Wu, "Software defined wireline-wireless cross-networks: Framework, challenges, and prospects," *IEEE Communications Magazine*, vol. 56, no. 8, pp. 145–151, 2018.
- 7 C.-H. Ke, Y.-S. Chen, and Y.-S. Yu, "Improving video transmission in software defined wired and wireless networks using multi-path transmission," *Journal of Communications and Networks*, vol. 19, no. 6, pp. 587–595, 2017.
- 8 L. Cai, X. S. Shen, J. W. Mark, and J. Pan, "Qos support in wireless/wired networks using the tcp-friendly aimd protocol," *IEEE transactions on wireless communications*, vol. 5, no. 2, pp. 469–480, 2006.
- 9 L. Underberg, R. Kays, S. Dietrich, and G. Fohler, "Towards hybrid wired-wireless networks in industrial applications," in *2018 IEEE Industrial Cyber-Physical Systems (ICPS)*, pp. 768–773, IEEE, 2018.
- 10 J. Sachs, K. Wallstedt, F. Alriksson, and G. Eneroth, "Boosting smart manufacturing with 5g wireless connectivity," *Ericsson Tech. Rev.*, 2019.
- 11 H. Föllmer, A. Schied, and T. J. Lyons, "Stochastic finance. an introduction in discrete time," *The Mathematical Intelligencer*, vol. 26, no. 4, pp. 67–68, 2004.
- 12 S. Batewela, C.-F. Liu, M. Bennis, H. A. Suraweera, and C. S. Hong, "Risk-sensitive task fetching and offloading for vehicular edge computing," *IEEE Communications Letters*, vol. 24, no. 3, pp. 617–621, 2019.
- 13 M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834–1853, 2018.
- 14 G. Yang, M. Xiao, and H. V. Poor, "Low-latency millimeter-wave communications: Traffic dispersion or network densification?," *IEEE Transactions on Communications*, vol. 66, no. 8, pp. 3526–3539, 2018.
- 15 T. K. Vu, C.-F. Liu, M. Bennis, M. Debbah, and M. Latva-Aho, "Path selection and rate allocation in self-backhauled mmwave networks," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2018.
- 16 T. K. Vu, C.-F. Liu, M. Bennis, M. Debbah, M. Latva-Aho, and C. S. Hong, "Ultra-reliable and low latency communication in mmwave-enabled massive mimo networks," *IEEE Communications Letters*, vol. 21, no. 9, pp. 2041–2044, 2017.
- 17 M. Assaad, A. Ahmad, and H. Tembine, "Risk sensitive resource control approach for delay limited traffic in wireless networks," in *2011 IEEE Global Telecommunications Conference-GLOBECOM 2011*, pp. 1–5, IEEE, 2011.
- 18 M. Alsenwi, N. H. Tran, M. Bennis, A. K. Bairagi, and C. S. Hong, "embb-urllc resource slicing: A risk-sensitive approach," *IEEE Communications Letters*, vol. 23, no. 4, pp. 740–743, 2019.
- 19 B. Hofeld, D. Wieruch, T. Wirth, L. Thiele, S. A. Ashraf, J. Huschke, I. Aktas, and J. Ansari, "Wireless communication for factory automation: An opportunity for lte and 5g systems," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 36–43,



- 2016.
- 20 3GPP, "Evolved universal terrestrial radio access (e-utra) and evolved universal terrestrial radio access network (e-utran)," *3gpp Ts*, 2010.
- 21 P. Schulz, M. Matthe, H. Klessig, M. Simsek, G. Fettweis, J. Ansari, S. A. Ashraf, B. Almeroth, J. Voigt, I. Riedel, et al., "Latency critical iot applications in 5g: Perspective on the design of radio interface and network architecture," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 70–78, 2017.
- 22 J.-B. Seo and V. C. Leung, "Performance modeling and stability of semi-persistent scheduling with initial random access in lte," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4446–4456, 2012.
- 23 N. Afrin, J. Brown, and J. Y. Khan, "Design of a buffer and channel adaptive lte semi-persistent scheduler for m2m communications," in *2015 IEEE International Conference on Communications (ICC)*, pp. 5821–5826, IEEE, 2015.
- 24 D. M. Soleymani, A. Puschmann, E. Roth-Mandutz, J. Mueckenheim, and A. Mitschele-Thiel, "A hierarchical radio resource management scheme for next generation cellular networks," in *2016 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 416–420, IEEE, 2016.
- 25 M. Raza, H. Le-minh, N. Aslam, and S. Hussain, "A novel mac proposal for critical and emergency communications in industrial wireless sensor networks," *Computers & Electrical Engineering*, vol. 72, pp. 976–989, 2018.
- 26 H. Farag, E. Sisinni, M. Gidlund, and P. Österberg, "Priority-aware wireless fieldbus protocol for mixed-criticality industrial wireless sensor networks," *IEEE Sensors Journal*, vol. 19, no. 7, pp. 2767–2780, 2018.
- 27 P. Gaj, J. Jasperneite, and M. Felser, "Computer communication within industrial distributed environment—a survey," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 182–189, 2012.
- 28 P. Zand, S. Chatterjea, K. Das, and P. Havinga, "Wireless industrial monitoring and control networks: The journey so far and the road ahead," *Journal of sensor and actuator networks*, vol. 1, no. 2, pp. 123–152, 2012.
- 29 F. Lin, W. Dai, W. Li, Z. Xu, and L. Yuan, "A framework of priority-aware packet transmission scheduling in cluster-based industrial wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5596–5606, 2019.
- 30 N. T. T. Hang, N. C. Trinh, N. T. Ban, M. Raza, and H. X. Nguyen, "Delay and reliability analysis of p-persistent carrier sense multiple access for multi-event industrial wireless sensor networks," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 12402–12414, 2020.
- 31 M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, and J. Wang, "Large-scale measurement and characterization of cellular machine-to-machine traffic," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1960–1973, 2013.
- 32 R. Sanasam, H. Murthy, and T. Gonsalves, "Feature selection for text classification based on gini coefficient of inequality," in *Feature Selection in Data Mining*, pp. 76–85, PMLR, 2010.
- 33 P. Arora, C. Szepesvári, and R. Zheng, *Sequential learning for optimal monitoring of multi-channel wireless networks*. IEEE, 2011.
- 34 Q. Xu and R. Zheng, "When data acquisition meets data analytics: A distributed active learning framework for optimal budgeted mobile crowdsensing," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pp. 1–9, IEEE, 2017.
- 35 R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.