Study of Uplink Resource Allocation for 5G IoT Services by Using Reinforcement Learning

Yen-Wen Chen, ChengYu Tsai*

Department of Communication Engineering, National Central University, Taiwan
ywchen@ce.ncu.edu.tw, sam850713@gmail.com

Abstract

In order to support real time IoT services, the ultra Reliable and Low Latency Communications (uRLLC) was proposed in 5G wireless communication network. Different from the grant based access in 4G, the grant free technique is proposed in 5G to reduce the random access delay of uRLLC-required applications. This paper proposes the dedicated resource for exclusive access of individual UE and the shared resource pool for the contention of multiple UEs by adopting the reinforcement learning approach. The objective of this paper is to accomplish the uplink successful rate above 99.9% under certain transmission error probability. The proposed Prediction based Hybrid Resource Allocation (PHRA) scheme allocates the access resource in a heuristic manner by referring to the activity of UEs. The dedicated resource is mainly allocated to the high activity UEs and the initial transmission of UEs with medium activity while the shared resource pool is allocated for the re-transmission of medium activity UEs and low activity UEs by using the reinforcement learning model. The burst traffic model was applied during the exhaustive experiments. And the simulation results show that the proposed scheme achieves higher uplink packet delivery ratio and more effective resource utilization than the other schemes.

Keywords: Internet of things, 5G wireless communications, uRLLC, Random access, Resource allocation

1 Introduction

Recently, the international organization 3GPP has continued to customize the standard specifications of the fifth generation mobile communication (5G) especially for the emerging IoT services. The transmission latency and reliable delivery are the most critical requirements for the real time applications such as virtual reality (VR), industry automation, auto-driving, etc.. The enhanced mobile broadband (eMBB), massive machine type communication (mMTC), and ultra-reliable and low latency communications (uRLLC) are the major perspectives of the 5G development when comparing to 4G networks [1]. Among them, uRLLC concerns the delay and reliability issue and is more correlated to the quality of the provided services. For the IoT application scenario, the sensed or collected data by the IoT device is mostly transmitted to the cloud or edge servers for processing. The response time of IoT service includes the transmission delay and processing delay in addition to the sensing time. The transmission delay is more unpredictable than the processing delay due to the limited radio resource. Comparing to the downlink transmission, the uplink has much higher uncertainty because of the contention occurred during the random access procedure. In 4G, the user equipment (UE) shall contend the subcarrier (preamble) of the anchor carrier provided by extend node B (eNB) through handshake and the data can only be uplinked to eNB after the grant from eNB. The handshake procedure needs the message exchanges between UE and eNB and it introduces delay before the transmission of data.

In order to reduce the uplink delay, the 5G system proposes the grant-free random access concept. Thus, UE directly piggybacks the data in its uplink contention message without waiting for the grant from gNB. The gNB can immediately receive the data if the uplink message is not suffered in either collision or transmission error. The uplink request will be repeatedly retransmitted if the collision is occurred for the achievement of reliability. The retransmission procedure will be automatically performed through hybrid automatic repeat transmission (HARQ) mechanism without waiting for the success/fail response from gNB so as to shorten the delivery delay. Thus, the allocation of access resource shall consider not only the number of accessing UEs but also the supported resource of HARQ. Then, based on the above concept, the main issues of the grant-free random access are to effectively allocate the access radio resource and properly arrange the access policy of massive UEs so that the successful rate of random access can be as higher as to approximate 100% and the delivery delay can be shortened to be within several mini-seconds. And it is the main study issue of this paper.

The remainder of the paper is organized as follows. Section 2 presents the background and related works. Section 3 describes the proposed prediction-based hybrid resource allocation (PHRA) scheme. In Section 4, we evaluate the performance of the proposed scheme through exhaustive simulations. We also compare the simulation results with the other approaches. Finally, Section 5 concludes our works.

*Corresponding Author: Yen-Wen Chen; E-mail: ywchen@ce.ncu.edu.tw
DOI: 10.53106/160792642023052403013

2 Overview of Related Works

Comparing to the fixed 15KHz subcarrier spacing in 4G network, 5G adopts the flexible spacing to fit the demand of new radio (NR) [2-3]. Additionally, the mini-slot resource (i.e. the access resource units) allocation provides the density of transmission opportunities to achieve the prompt reaction and low transmission delay. In 5G, the grant-free random access scheme is proposed to speedup the access time. The grant-free procedure simplifies the message exchange between UE and base station when comparing to the grant-based scheme adopted in 4G network as shown in Figure 1.

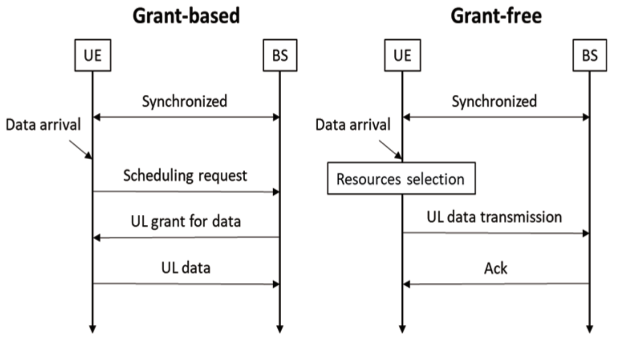


Figure 1. Comparison between grant-based and grant-free random access procedures

5G NR provides flexible downlink, uplink, and mixed downlink and uplink slots configurations. The mixed slot configuration can support the uplink data transmission and HARQ within one time slot. By applying HARQ, UE can either wait for the acknowledgement from base station or automatically and repeatedly retransmitted (also named as the blind HARQ proposed in release 15 of 3gpp) without waiting for the acknowledgement. The blind HARQ can reduce the delivery time by skipping the waiting time of acknowledgement, however, it increases the load of random access requests. For the blind HARQ, UE can terminate the retransmission if it receives the acknowledgement from base station to save the access resource and can provide the access for the next arrival [4-5].

The arrangement of shared allocation of access resource highly depends on the access behavior and frequency of UE. In [6], the performances of reactive non-acknowledgement (NACK) and blind HARQ were compared and showed that although the NACK approach has better spectrum utilization, it cannot satisfy the real time transmission such as URLLC. The hybrid NACK and blind HARQ transmission scheme was proposed to effectively utilized radio resource [7]. In addition to the collision due to accessing the same resource unit, the uplink data may not be successfully received due to transmission error. The concept of providing the dedicated resource and shared resource pool was proposed in [8-9]. In [8], UE is allocated with both dedicated resource unit and shared resource unit. The UE first uses the dedicated resource to uplink its data and will contend the shared resource units if it does not receive the acknowledgement from base station. In [9], the authors classified the behaviors of UEs into periodical and non-periodical uplink transmission. And the proposed scheme reserved the dedicated resource unit for the initial transmission of each UE. The shared resource pool was provided for the retransmission after either collision or transmission error. The authors also investigated the access performance and the resource utilization of UEs with different activities. It is clear from the above literatures that separating dedicated and shared resource units and dividing the UEs by different activities are benefit for the access performance and resource utilization of grant-free uplink transmission in IoT applications. However, the effective allocation of dedicated and shared resource is not well studied. In [10], the authors proposed the prediction-based random access resource allocation (PRARA) scheme to properly allocated the non-anchor carrier by applying reinforcement learning, however, they did not consider the strict performance requirement of URLLC.

3 The Proposed PHRA Scheme

Generally, the information sensed by the IoT device can be either periodically or non-periodical transmitted. The periodical transmission is much easier to allocate resource in advance. The most difficult issue is the uplink of non-periodical data because the uplink request is not predicable and the base station is hard to allocate the access resource prior to the arrival request. Furthermore, the uplink data shall be delivered to the base station within quite limited delay for uRLLC services. As mentioned in [9], the access activity of UEs affect the effectiveness of resource allocation for better access performance. Then the proposed scheme allocates the access resources according to the activities of UEs through the prediction from the reinforcement learning model and design different access policies for UEs with different activities. The activities of UEs are classified by the base station according to their access behaviors in the proposed scheme. In order to examine the UE activities and adaptively arrange the dedicated and shared access resource, the Configured Scheduling-Radio Network Temporary Identifier (CSRNTI) in the Downlink Control Information (DCI) of Transmission Without Grant type 2 (TWG type 2) is applied.

The proposed PHRA scheme can be generally divided into 3 parts. The first part is the activity classification of UEs; the second part is the random access policies of different activity groups; the third part is the adjustment of dedicated and shared resources through the prediction by using the reinforcement learning deep Q-network (DQN) [11-12]. In the first part, the base station re-classify the activities of UEs when the trigger condition is satisfied. The trigger condition is designed as either the success ratio of random accesses is less than the desired objective $target_1$ or the idle ratio of the shared resource is higher than the threshold $target_2$. Once the trigger condition is satisfied, gNB calculate the average interarrival time of requests for each UE. If the average interarrival time of UE is less than the threshold ATS_1 will be classified as the high activity group. And the UE is classified as middle and low activity if the average interarrival is between ATS_1 and ATS_2 and longer than ATS_2, respectively. The basic concept of random access policy is provide the dedicated access resource to the UE with higher activities.

Therefore, the high activity UE is designed to blindly transmit its data by using the allocated dedicated resource till the maximum number of the retransmission times reaches. For the medium and low activity groups, the hybrid HARQ scheme is adopted. They use the NACK scheme for the first transmission, and the blind retransmission for the residual retransmissions will be adopted if it is failure in the first transmission. The medium activity UE will be allocated with the dedicated resource for the first transmission and the share resource for the residual retransmission. However, the low activity UE uses the shared resource for all of its transmissions. The proposed UE procedure of PHRA is provided in the following Figure 2.

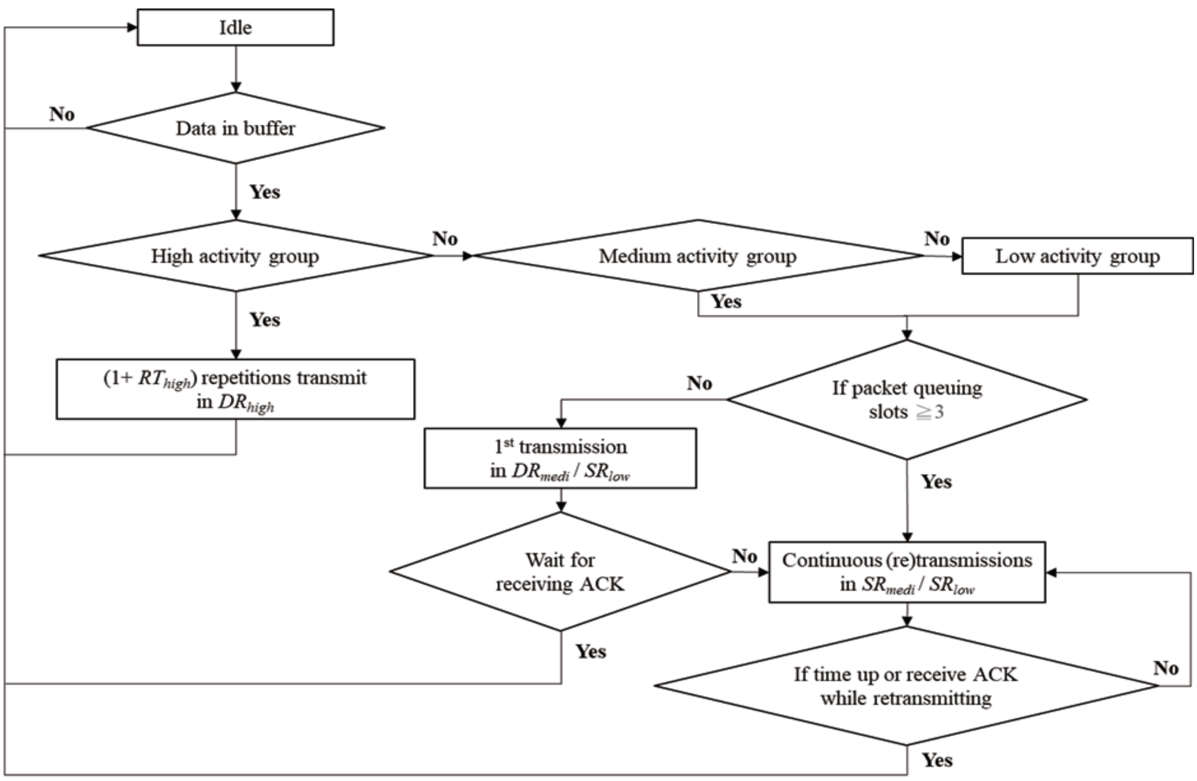


Figure 2. The UE procedure of the proposed PHRA algorithm

The DR_{high} and DR_{medi} represent the dedicated resource allocated for the high and medium activity groups. And the SR_{medi} and SR_{low} are the shared resource allocated for the medium and low activity groups, respectively. It is noted that, the UEs of the medium and low activity groups will transmit their packets and wait for the ACK for the packets when the queuing symbol times (slots) is less than 3. And, for the packet queuing slots being equal to or more than 3 symbols, the UE will blindly retransmit the packet and ignore the waiting for ACK/NACK stage to increase the possibility of successful transmission the packet under the delay constraint.

The dedicated resource can be determined by the numbers of UEs in the high and medium activity groups. Thus the allocated dedicated resource, DR, can be obtained as

$$DR = N_{high} * (1 + RT_{high}) + N_{medi}$$

where N_{high} and N_{medi} are the numbers of UEs in the high activity and medium activity groups, respectively, and RT_{high} is the maximum number of retransmissions for high activity group. The allocation of shared resource for low activity group is determined by using DQN. As the UE of medium activity group has been allocated with one dedicated resource to transmission, the shared resource for the medium activity group is allocated with the specific percentage of the number of UE in medium activity group. In the proposed scheme, we adopt it is 20% of N_{medi}. And the total allocated share resource is the summation of the above two resources. The DQN consists of evaluation network for training and target network for target calculation. The system state is defined as the number of UEs in the low activity group N_{low}, its difference to the number after the re-classification of activities, i.e. (N'_{low} - N_{low}), and the resource utilization of the shared resource.

The objective of the proposed DQN model is to provide high reliability and the effectiveness of resource utilization. The proposed reward function shall reflect the achieved reliability for the consideration in the training model. The proposed scheme observed the achieved reliability in consecutive time intervals. The parameters T_{inter} and T_{noerr} represent the number of consecutive time intervals when the reliability is below the desired reliability $target_t$ and the number of the intervals that achieve 100% reliability. The maximum of observed intervals is set to be 5 and we also suggest that we can get the reward if the value of T_{inter} is greater than 3. W_S and W_T represent the weights for the successful rate and the observed T_{inter}, respectively. Figure 3 provides two cases to illustrate the values of T_{inter} and T_{noerr} and the changes of the UE numbers for $target_t = 0.999$.

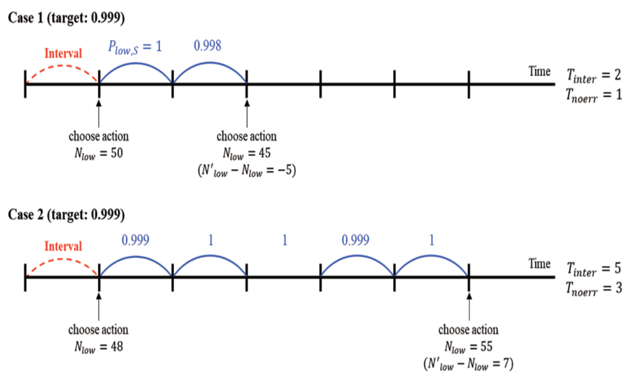


Figure 3. The parameters and the observation intervals

It is noted that the allocated shared resource shall not be larger than the number of UEs. If the allocated resource is greater than the number of UEs, it will get the negative reward value to achieve better resource utilization in the top equation. For example, if the shared resource for the low activity users (SR_{low}) is greater than the number of low activity users (N_{low}), the reward value will be negative. The middle and bottom equations illustrate the cases of $N_{low} > SR_{low}$, however, with the low and high average success rates of the low activity users P_{lows} when comparing to the desired target rate $target_t$. The W_{SR} is applied as the weight for this situation. Then the proposed reward function is provided in the following equation.

$$\text{Reward} = \begin{cases} (-500)\times(SR_{low} - N_{low} + 1), \text{ if } N_{low} \le SR_{low}. \\ \left[\left(\dfrac{P_{low,s}}{\text{target}_t}\right) - 1\right] \times Ws + (T_{inter} - 3)\times W_T + T_{noerr} \times (\dfrac{W_T}{3}) \\ \qquad , \text{if } N_{low} > SR_{low} \text{ and } P_{low,s} < \text{target}_t. \\ \left[\left(\dfrac{P_{low,s}}{\text{target}_t}\right) - 1\right] \times Ws + (T_{inter} - 3)\times W_T + T_{noerr} \times (\dfrac{W_T}{3}) \\ +(N_{low} - SR_{low}) \times W_{SR} \\ \qquad , \text{if } N_{low} > SR_{low} \text{ and } P_{low,s} \ge \text{target}_t. \end{cases}$$

And the gNB operational procedure of the proposed PHRA algorithm is illustrated in Figure 4.

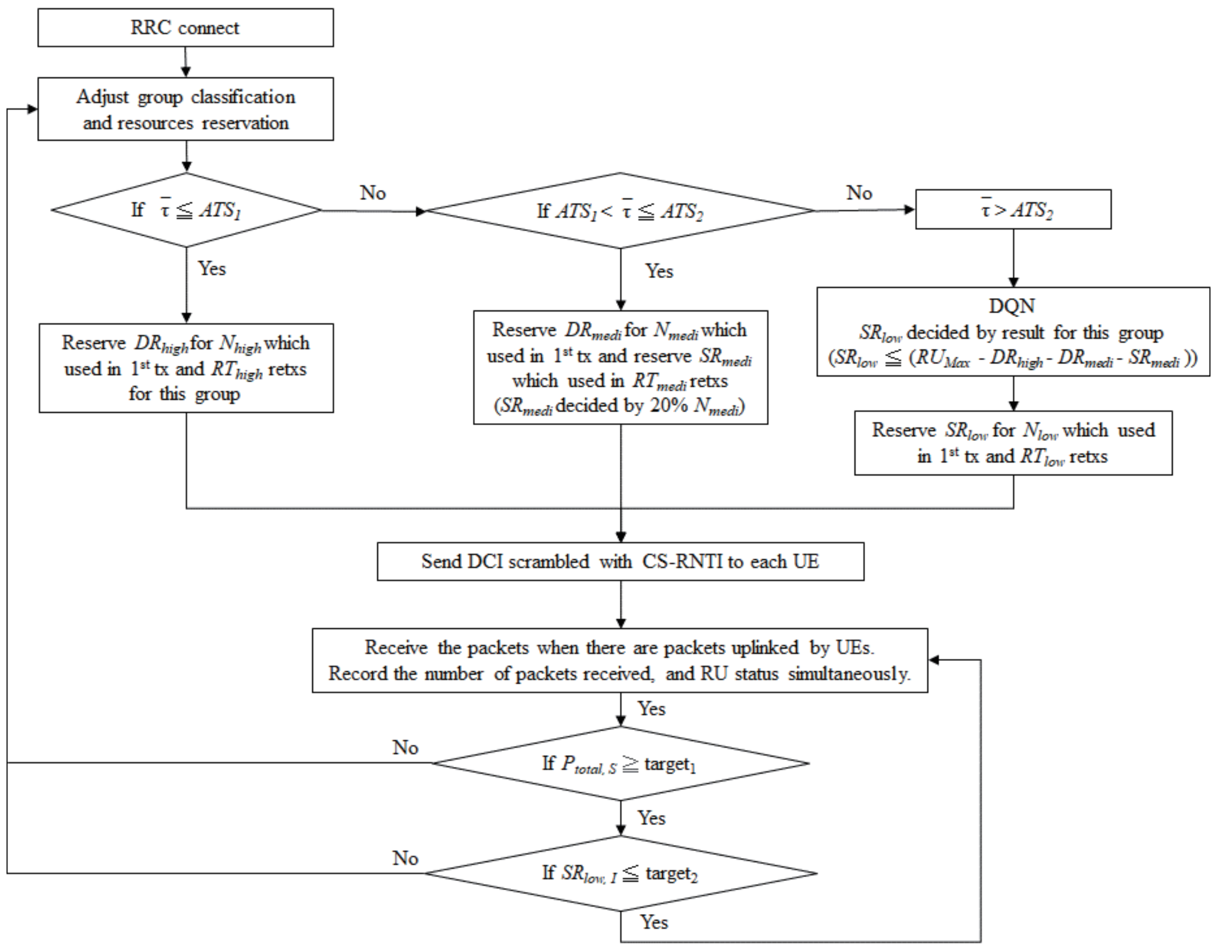


Figure 4. The gNB procedure of the proposed PHRA algorithm

4 Experimental Simulations

In order to examine the performance of the proposed PHRA algorithm, exhaustive simulations were performed. The simulation environment consists of one gNB and several IoT UEs. The gNB may not be able to decide the uplink data due to either the collision or the transmission error in the uplink channel. The proposed scheme is compared to the schemes of fully dedicated resource allocation, fully shared resource allocation, and the fixed but hybrid resource allocation. We carefully compare the performance of the uplink successful rate, the allocated radio resource units for random access, and the utilization for the above schemes. The simulation parameters of the simulations are given in the following Table 1.

Table 1. Simulation parameters

Parameters	Values
Total available bandwidth	37.5 MHz (100 Resource units)
Packet size	100 bits
Subcarrier spacing	15 kHz
Packet mean inter arrival time of each UE	3ms~45ms
Transmission time interval (TTI)	0.144 ms (2 symbols per TTI)
Acknowledgment time	0.288 ms
Latency budget	1 ms
$Target_1$ (Reliability)	0.9995
$Target_2$	0.98
Transmission error probability	10^{-4}
Spectral efficiency	2 bps/Hz (QPSK)
UE max retransmission times	6
(ATS_1, ATS_2)	(8ms, 16ms)
W_S	10000
W_T	30, 60
W_{SR}	1.2~1.5

In practical IoT environment, the occurrences of sensing events are correlated especially for the non-periodical and time critical applications, such as the disaster monitoring. Once a specific event occurs, it may trigger some events in a very short time period. In order to emulate the correlated and burst random uplink packets, we assume that there are two UEs, named as central UE1 and central UE2, that affect the packet generation rate of the other normal UEs. The main objective is to model the correlated phenomenon and the burstiness property. In the designed traffic model, each arrival packet of the central UE1 will increase the arrival rate of the normal UE by 1-3 times; and it will decrease the arrival rate by 1-3 times for each arrival of the central UE2. An example of the histogram of the generated uplink packets of 3 UEs with one normal UE, one central UE1 and one central UE2, is illustrated in Figure 5. The number of generated random access is counted for every 1000ms. It shows the burst arrivals phenomenon of the uplink traffic. For example, at the beginning of the figure, the cumulative arrivals of UE1 (the blue dot) is higher than UE2 (the red dot) and, therefore, the total number of generated arrivals by normal UE (the green dot) increases dramatically.

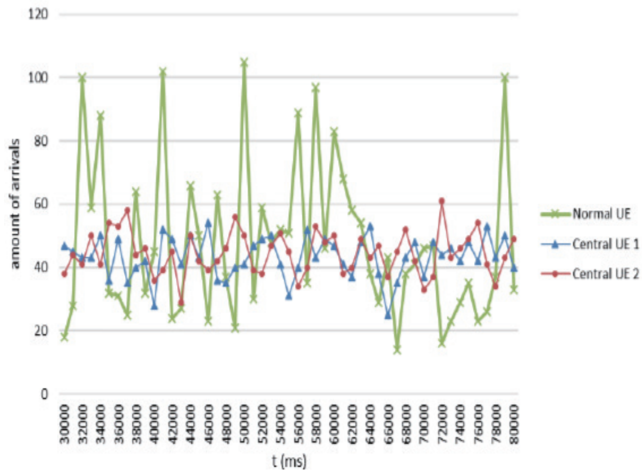


Figure 5. An example of the generated burst uplink packets for simulations

Figure 6 shows the results of the uplink packet rate versus time for 100 UEs. The results compare the proposed PHRA scheme (with different values of W_T) and the shared schemes with 50 and 60 resource units, respectively. The proposed PHRA scheme achieves higher success rate than the shared schemes as illustrated. It is noted that, although the achieved success rate of the all shared resource scheme is as higher as 98.5%, to improve the success rate from 98.5% to near 99.9% is much more difficult when comparing to the situation from low to medium success rates.

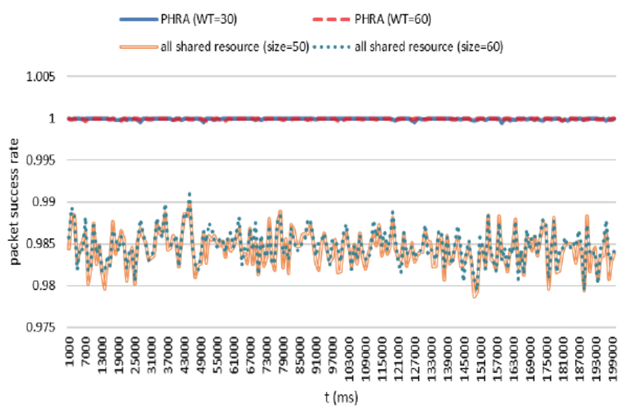


Figure 6. Comparisons of success rate v.s. time

The detail simulation results of the proposed PHRA scheme and the other schemes for the number of UEs being 100 are shown in Table 2. In addition to the all shared resource scheme, the dedicated resource and hybrid schemes are also compared to the proposed scheme. As the number of UEs is 100, the number of allocated resource units of the dedicated scheme is also 100. It is noted that, according to the simulation results, the allocated average resource units of the proposed PHRA scheme for $W_T = 30$ and 60 are 48.32 and 49.12, respectively. The number of access units for the all shared resource scheme is setting to be 50. The resource allocations for the hybrid scheme are assumed to be (10, 40) and (20, 30) for the dedicated resource (dr) and shared resource (sr), respectively. And the dedicated resources are firstly allocated to the high activity UEs. In order to examine the effectiveness of the resource prediction by using DQN, we also propose a simple the rule-based hybrid resource allocation (RHRA) scheme for comparison. The RHRA scheme examines the status of the shared resource, it decreases the resource units by one if the decreased number of low activity UEs is greater than 4 or the idle rate is higher than $target_2$. And it will increase the resource units by one if the increased number of UEs in low activity group is greater than 4 or the increased number of the failure packets is equal to or more than 3 during the observation interval. The parameters used in RHRA are properly determined after several experiments.

Table 2. Comparisons of the proposed PHRA and the other schemes

Schemes	Resource allocated	Ave. resource utilizations	Ave. Success rate
All dedicated resource	100	0.008322	0.999998
All shared resource	50	0.086059	0.984308
Hybrid	50 (dr=10, sr=40)	0.083010	0.989494
	50 (dr=20, sr=30)	0.062296	0.994423
PHRA (W_T = 30)	48.32	0.053256	0.999905
PHRA (W_T = 60)	49.12	0.055373	0.999926
PHRA	53.73	0.040929	0.999885

As each UE of the all dedicated resource scheme can be allocated with the resource unit for its exclusive use, there is no collision and the failure rate are mainly due to the transmission error and it achieves the highest success rate, however, it needs much more resource with relatively very low utilization. The proposed PHRA scheme achieves the near success rate when comparing to that of the all dedicated resource scheme, however, only half of the radio resource is required. The success rate of the RHRA scheme is very close to the proposed scheme, however, its average allocated resource is higher and with lower utilization. We also checked the uplink success rates of the PHRA scheme for the high, medium, and low activity groups. The results are shown in Figure 7. It clearly indicates that the high activity group has the highest success rate because UEs in this group are allocated with dedicated resources for the initial uplink transmission and retransmissions. However, all of the three groups can achieve the successful rate of 99.9%. The results also indicate that the difference between different W_T values is not very significant.

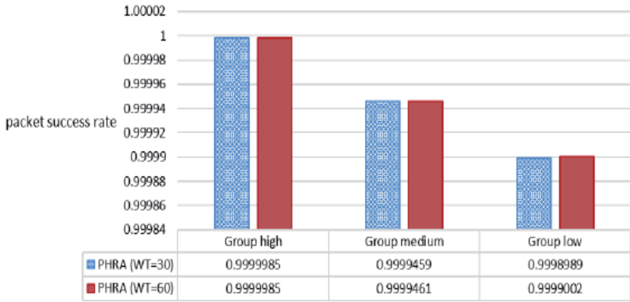


Figure 7. Uplink success rates of the three groups

The comparison of the average retransmission times is provided in Figure 8. It is noted that, for the all dedicated resource scheme, each UE is designed to repeatedly transmitted two times (i.e. one retransmission) and, therefore, it has one initial transmission and one retransmission. The other schemes retransmit the packet depending on the transmission status, and the maximum number of retransmission times is 6. According to the results, the proposed PHRA scheme and the RHRA scheme has much less retransmission times than the other schemes. The main reason is mainly due to the allocation of dedicated resource. Although the hybrid scheme arranges the dedicated resource for exclusive use, the number of dedicated resource is fixed and cannot be flexibly arranged especially for the burst arrivals.

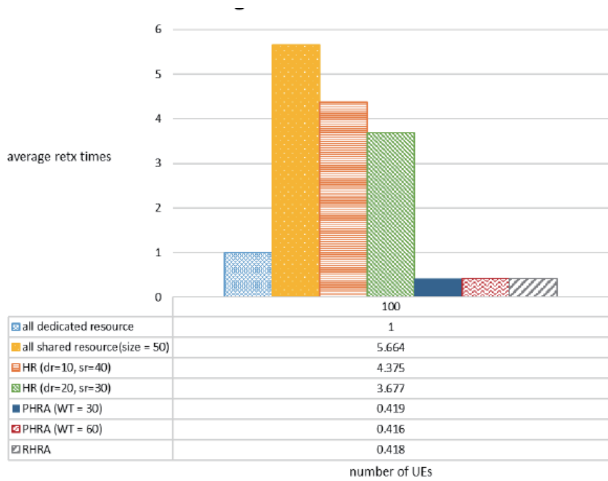


Figure 8. Comparisons of retransmission times

5 Conclusions

In this paper, the scheme of uplink resource allocation for random access is proposed to achieve the rigorous demand of uRLLC IoT applications. The resource units are divided into dedicated part and shared part. UEs are classified into three groups according to their activities measured by gNB. The dedicated resource can guarantee the exclusive use to prevent the access collision, however, it needs much more resource when there are massive number of access devices. In the proposed scheme, the number of dedicated resource units is determined according to the UEs in the high and medium activity groups and the number of shared resource units are allocated through the reinforcement learning model. The simulation results demonstrate that the proposed scheme accomplishes higher successful rate than the other schemes can achieve almost the similar success rate of the all dedicated resource scheme. Thus the proposed scheme provides the adjustment of resource according to the traffic behaviours in heuristic manner and achieves the average successful access rate to be higher than 99.99%. By applying the learning model, the shared resources can be properly arranged under the reliability demand. In this paper, we only consider the situation of the access from uRLLC devices without other service requirements. The resource allocation and access policy for the coexistence of uRLLC and eMBB services shall be carefully studied in real deployment and this is one of our ongoing works.

Acknowledgment

This work was supported in part by the National Science and Technology Council, Taiwan (grant number: 109-2221-E-008-052-MY2, and 111-2221-E-008-044-).

References

[1] ITU, *IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond*, Rec. ITU-R M.2083-0, September, 2015.

[2] 3GPP, *Study on scenarios and requirements for next generation access technologies*, 3GPP Tech. Rep. TR 38.913 V15.0.0, July, 2018.

[3] 3GPP, *Study on New Radio (NR) access technology*, 3GPP Tech. Rep. TR 38.912, V15.0.0, July, 2018.

[4] 3GPP, *UL grant-free transmission for URLLC*, 3GPP TSG RAN WG1 #88, R1-1705654, April, 2017.

[5] S. Xing, X. Xu, Y. Chen, Y. Wang, L. Zhang, Advanced Grant-free Transmission for Small Packets URLLC Services, *International Conference on Communications Workshops (ICC Workshops)*, Shanghai, China, 2019, pp. 1-5.

[6] N. H. Mahmood, R. Abreu, R. Bohnke, M. Schubert, G. Berardinelli, T. H. Jacobsen, Uplink Grant-Free Access Solutions for URLLC services in 5G New Radio, *International Symposium on Wireless Communication Systems*, Oulu, Finland, 2019, pp. 1-6.

[7] L. Buccheri, S. Mandelli, S. Saur, L. Reggiani, M. Magarini, Hybrid Retransmission Scheme for QoS-defined 5G Ultra-Reliable Low-Latency Communications, *IEEE Wireless Communications and Networking Conference*, Barcelona, Spain, 2018, pp. 1-6.

[8] Z. Zhou, R. Ratasuk, N. Mangalvedhe, A. Ghosh, Resource Allocation for Uplink Grant-Free Ultra-Reliable and Low Latency Communications, *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Porto, Portugal, 2018, pp. 1-5.

[9] S. E. Elayoubi, P. Brown, M. Deghel, A. Galindo-Serrano, Radio Resource Allocation and Retransmission Schemes for URLLC Over 5G Networks, *IEEE Journal on Selected Areas in Communications*, Vol. 37, No. 4, pp. 896-904, April, 2019.

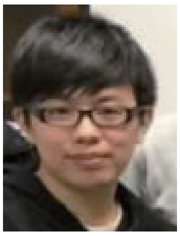
[10] Y.-W. Chen, J.-Z. You, Effective Radio Resource Allocation for IoT Random Access by Using Reinforcement Learning, *Journal of Internet Technology*, Vol. 23, No. 5, pp. 1069-1075, September, 2022.

[11] K. Arulkumaran, M. P. Deisenroth, M. Brundage, A. A. Bharath, Deep Reinforcement Learning: A Brief Survey, *IEEE Signal Processing Magazine*, Vol. 34, No. 6, pp. 26-38, November, 2017.

[12] B. Gu, X. Zhang, Z. Lin, M. Alazab, Deep Multiagent Reinforcement-Learning-Based Resource Allocation for Internet of Controllable Things, *IEEE Internet of Things Journal*, Vol. 8, No. 5, pp. 3066-3074, March, 2021.

Biographies

Yen-Wen Chen received the Ph.D. degree in Electronic Engineering from NTUST in 1997. During 1983 to 1998, he worked at Chunghua Telecommunication Laboratories. Currently, Dr. Chen is a professor of the National Central University. His research interests include 4G/5G networks, radio random access, and internet of things (IoT).



ChengYu Tsai received the MS degree in Communication Engineering from National Central University in 2019. His current research interests include the 4G/5G wireless communication networks, machine learning, and internet of things (IoT).