Effective Radio Resource Allocation for IoT Random Access by Using Reinforcement Learning

Yen-Wen Chen*, Ji-Zheng You

Department of Communication Engineering, National Central University, Taiwan
ywchen@ce.ncu.edu.tw, fers84629@gmail.com

Abstract

Emerging intelligent and highly interactive services result in the mass deployment of internet of things (IoT) devices. They are dominating wireless communication networks compared to human-held devices. Random access performance is one of the most critical issues in providing quick responses to various IoT services. In addition to the anchor carrier, the non-anchor carrier can be flexibly allocated to support the random access procedure in release 14 of the 3rd generation partnership project. However, arranging more non-anchor carriers for the use of random access will squeeze the data transmission bandwidth in a narrowband physical uplink shared channel. In this paper, we propose the prediction-based random access resource allocation (PRARA) scheme to properly allocated the non-anchor carrier by applying reinforcement learning. The simulation results show that the proposed PRARA can improve the random access performance and effectively use the radio resource compared to the rule-based scheme.

Keywords: Internet of Things, Random access, Anchor carrier, LTE, Reinforcement learning

1 Introduction

The internet of things (IoT) concept has been realised in several applications, such as environment monitoring, industrial automation, safety monitoring and control, and disaster prevention. Additionally, ubiquitous sensing enabled by the demand for IoT services (i.e. user equipment (UE) in the 3rd generation partnership project (3GPP)) requires the fast transmission of its sensed data to the server. As IoT devices are deployed over widespread areas, wireless communication technology is more suitable for real applications. The 3GPP proposed NarrowBand IoT (NB-IoT) based on long-term evolution (LTE) for IoT application in release 13. It modifies and simplifies the LTE specification to be compatible with IoT devices and coexist with existing LTE systems [1]. A brief comparision for 4G LTE and 5G communication technologies was presented in [2]. The uplink transmission from many UEs makes random access more critical for IoT service deployment. The random access procedure (RAP) can be supported in non-anchor carriers in release 14 of 3GPP to support the requests from massive UEs. This alleviates the network congestion problem that may occur if UE can only random access via anchor carrier in the narrowband physical random access channel (NPRACH) [3-4]. Allocating more non-anchor carriers for more random access opportunities (RAO) can effectively reduce the collision rate of the random access attempts issued by massive UEs to shorten the uplink transmission delay. However, it also compresses the data transmission resource in the narrowband physical uplink shared channel. Thus, the non-anchor carriers should be carefully allocated for effective resource use.

In addition to the relatively large number of deployed UEs in IoT services, the traffic pattern is quite different from that of the human-held UEs in access behaviour and transmission data size. Generally, the sensing UEs of IoT services always need to transmit data to the server frequently, and the transmitted data size is much smaller than that of traditional human-oriented services. Generally, the number of UEs is much larger than the human-held devices within the coverage of a base station (i.e. eNB). Therefore, the number and frequency of uplink access requests to the base station are much higher than the traditional services. The uplink transmissions of sensing devices can be classified into periodic and aperiodic data. The aperiodic data always arrive unpredictably with emergent information. The IoT UEs must contend with the subcarrier (preamble) for the uplink transmission in the RAP. The performance of random access affects whether timely uplink data is deliverable. If two or more devices use the same subcarrier, then the collision will occur, and eNB will not be able to decode it. The collision probability can decrease with more subcarriers by allocating non-anchor carriers. However, the uplink request is actively issued by IoT UEs, and eNB cannot get the request information in advance. The radio resource will be wasted if the subcarriers are over-provided for contention. In this paper, we predict the uplink request from UEs using reinforcement learning to arrange the non-anchor carrier for effective radio resource use properly.

The remainder of the paper is organised as follows. Section 2 presents the background and related works. Section 3 describes the proposed prediction-based random access resource allocation (PRARA) scheme. In Section 4, we evaluate the performance of the proposed scheme through exhaustive simulations. We also compare the simulation results with the rule-based scheme proposed in [5]. Finally, Section 5 presents the conclusion.

2 Overview of Related Works

*Corresponding Author: Yen-Wen Chen; E-mail: ywchen@ce.ncu.edu.tw
DOI: 10.53106/160792642022092305015

The anchor carrier provides the access medium, i.e. the subcarrier or preamble, for the random access. However, in release 13, only one anchor carrier can be adopted for paging and random access. The total number of subcarriers in the anchor carrier is insufficient to support the high access attempts issued by the IoT UEs. The main access behaviour of UEs in IoT is their high access frequency compared to the traditional hand-held UEs. The limited access medium restricts the success possibility of the random access, resulting in a long uplink delay. The non-anchor carrier can be allocated for random access in release 14 to relax the congestion of random access. Up to 15 non-anchor carriers can be flexibly allocated for random access in addition to the original anchor carrier. Additionally, release 13 UE can only use the anchor carrier for random access, and release 14 UE can use an anchor or non-anchor carrier, as shown in Figure 1.

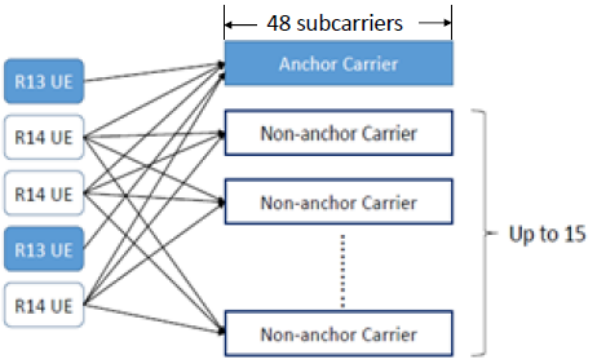


Figure 1. Anchor and non-anchor carriers for random access of release 13 and 14 UEs

The total radio resource is constrained. If more non-anchor carriers are allocated for random access, the radio resource used for data transmission decreases accordingly. The data generated by IoT UE are always much smaller than the traditional human-based services. In [5], the authors designed the collision report flag (CRF) in message 3 to feedback whether the end device is the first access or not and proposed the rule-based scheme to arrange the non-anchor carrier. The proposed small data transmission mechanism allows data to be piggybacked by the random access control message [6]. The parallel and pipelined scheme was proposed to support efficient random access [7]. This scheme divides the subcarriers into subgroups for contention. Additionally, each collided subcarrier is resolved through the first-in, first-out queue. The success ratio of random access can be improved using the sequential resolution concept. In [8], the proposed access class barring (ACB) mechanism would regulate the contention of random access attempts. The original ACB-based scheme can prevent the PRACH overload issue at the cost of a sharp increase in access delay. Several studies have suggested discarding the barring time and dynamically adjusting the barring factor to decrease the delay [9-11]. The author proposed a learning automata-based ACB (LA-ACB) to adjust the ACB value [9]. Generally, the same backoff procedure is applied to all UEs, and it is not easy to support UEs with different quality of services. In addition to using ACB, UEs with different transmission behaviour were differentiated with different backoff times to smooth the access congestion. Most studies analysed the performance of ACB by assuming the barring time is fixed. The recursive access class barring (R-ACB) technique was proposed in [10] to optimally utilize the available resources associated with the random access procedure. It suggested that the ACB factor shall be adjusted through the estimation. In [11], the ACB performance with a randomly selected barring time was analysed concerning several performance indexes. Several traffic models of a massive number of M2M UEs were observed, and the impact of ACB configuration parameters on the network performance was analysed [12]. They proposed the enhanced access algorithm based on the clustering-reuse subcarrier allocation algorithm. The UEs are divided into clusters by referring to the distance between the UE and eNB, and the access intensity.

The random access scheme's design and performance depend highly on the access behaviour. However, there is a lack of information about access statistics and the occurrence of a random collision, and the ideal random access is generally challenging. As the random access behaviour is not easy to predict, the traditional rule-based scheme has its limitations in dealing with this contention issue. In [13], the authors reviewed the existing random access methods based on machine learning (ML) and non-ML techniques. They found that ML-based methods perform better than non-ML-based methods due to their capability to solve high-complexity long-term optimisation problems. The tabular Q-learning (tabular Q), linear approximation-based Q-learning (LA-Q), and deep neural network-based Q-learning (DQN) schemes for random access have been studied and tested. The cooperative multi-agent learning (CMA-DQN) scheme has been proposed in [14]. Their results illustrate that the CMA-DQN scheme achieves better training efficiency and access performance. As mentioned in the previous section, the non-anchor carrier is applicable for random access to relax the access congestion and collision from massive IoT UEs.

The access policy or subcarrier allocation can improve the random access performance. The support of non-anchor carriers in RAP provides a more flexible allocation of sufficient subcarriers for many IoT UEs. However, note that the increase in subcarriers is individual to allocate one more non-anchor carrier results in the provisioning of several subcarriers. Thus, the allocation of non-anchor carriers shall be a tradeoff between the access performance and the effectiveness of resource utilization, which is the main objective of this paper.

3 The Proposed PRARA Scheme

The IoT UE contends the subcarrier for random access can be the first access or repeated access after the contention failure due to the collision of its first access. As the collision of the subcarrier means at least two UEs access the same subcarrier simultaneously, the number of UEs that will issue the repeated access can be predicted from the number of collided subcarriers. Then, the basic framework of the proposed PRARA scheme is to allocate the subcarriers into two parts. The subcarriers of part A provide the access for the initial random access, whereas the part B subcarriers support the UEs with repeated access. The base station shall broadcast parts A and B allocated subcarriers to all devices through the system information block message. Meanwhile, estimating the number of devices with initial access is not easy. Therefore, we propose the reinforcement ML approach to allocate the number of subcarriers in part A rewarded by the number of successful subcarriers A_S, the number of collided subcarriers

A_C, and the number of idle subcarriers A_I during the previous predefined RAO interval in part A. The rewarding function will be discussed in the following section. Additionally, the number of subcarriers in part B is determined by referring to the predicted number of collided UEs and the effectiveness of allocating non-anchor carriers. Table 1 presents the notation and description of the proposed scheme.

Table 1. Notations and descriptions

Notations	Descriptions
N_S	Number of subcarriers being detected as success in Part A
N_C	Number of subcarriers being detected as collided in Part A
N_I	Number of subcarriers being detected as idle in Part A
A_S	The average number of subcarriers being detected as a success last 3 RAO.in Part A
A_C	The average number of subcarriers being detected as collided last 3 RAO. in Part A
A_I	The average number of subcarriers being detected as idle last three RAO in Part A
P_A	The prediction of Part A subcarriers
R_A	Number of subcarriers allocated for Part A
R_B	Number of subcarriers allocated for Part B
N_{Sub}	Maximum number of subcarriers in one carrier
M_c	Maximum number of non-anchor carriers can be allocated
M_s	Maximum number of subcarriers can be supported
W_B	The weight to decide Part B resource
W_S	The success weight of the reward function

In the original Q-learning approach, the system calculates the Q values for the states with all possible actions denoted as Q(s, a) in the Q-table and recursively updates the Q-table according to the reward function. The learning rate α and discount γare applied to adjust the weight for the actual value concerning the predicted value when calculating the new Q values. It is worth noting that Q-table may keep expanding in each round, and the calculation of the Q-table needs a quite large computing overhead. Instead, of Q-table, DQN adopts the neural network to determine the Q-value. The use of neural networks can resolve the problem of inflation of the Q-table and speed up the convergence of the model. Figure 2 and Figure 3 show the operational flow and internal architecture with the reward function of DQN, respectively.

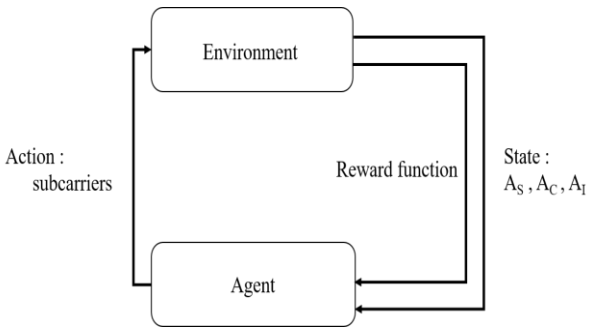


Figure 2. Operational flow of DQN

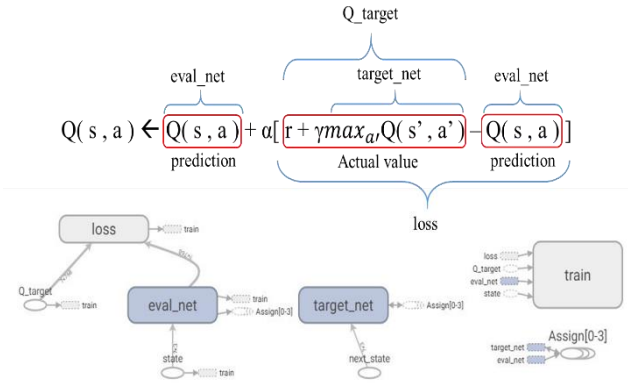


Figure 3. Reward function and internal architecture of DQN

As shown in Figure 2, eNB provides the observed A_S, A_C, and A_I to the DQN to predict required subcarriers in part A. In Figure 3, the *eval_net* and *target_net* of DQN perform the parameter adjustment and the prediction of the learning model, respectively. The eNB keeps observing the attribute [state (*s*), action (*a*), reward value (*r*), and new state (*s'*)] and saves them for reference in the follow-up training procedure. The saved attributes provide the experience replay for learning and are recorded from continuous observations. It is worth noting that the continuous experience may affect the convergence caused by gradient descent towards the same direction. Therefore, the attributes shall be randomly selected for training. During the training, the *Q_target* adopted by *eval_net* is not obtained directly from the reward of the actual state; however, it is obtained from the reward calculation of the *target_net* by applying the actual state. The *eval_net* adjusts the weights and biases of its neurons during the training procedure. The *target_net* updates its weights and biases during the training procedure. It inherits the parameters from the *target_net*, denoted as assign (0–3), after the training. Figure 4 shows the parameters of the assign (0–3) in the neural network. The neural network size applied in the proposed scheme, which is simplified from [14], is one hidden layer with 1536 neurons and one output layer with 768 neurons.

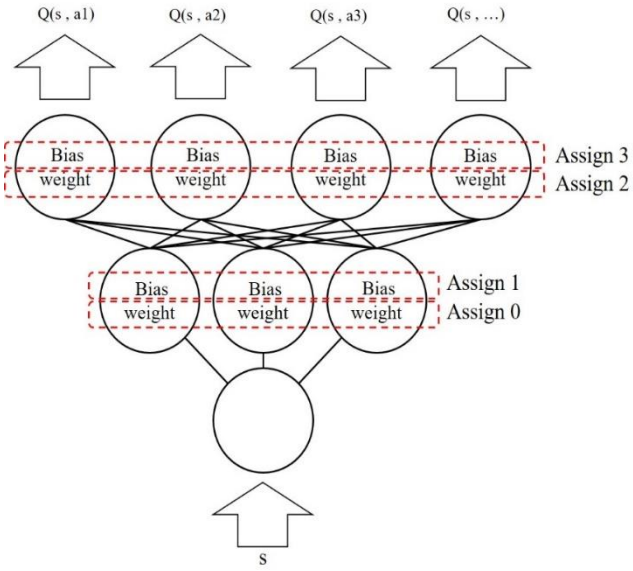


Figure 4. "Assign" parameter in the neural network

As mentioned above, the allocation of subcarriers is per anchor carrier basis. Thus, more N_{Sub} subcarriers will be provided for allocating one non-anchor carrier. After predicting the subcarriers of part A (P_A), the proposed algorithm considers whether it is worth allocating more non-anchor carriers. The decision shall consider the predicted P_A and the collision status to determine the actual subcarriers to be allocated in part A (R_A) and part B (R_B). Figure 5 presents the decision procedure of the proposed algorithm. In addition to P_A, the number of collided subcarriers N_C reflects the congestion status in RAP. The collision of each subcarrier means at least two UEs access the same subcarrier. Therefore, we can assume that there are N_C*W_B ($W_B \geq 2$) collided UEs in the previous RAO. As shown in part inside the dashed rectangle in Figure 5, the value of ($P_A+N_C*W_B$) represents the number of subcarriers to be allocated in part A if it is less than the maximum number of subcarriers that can be offered (i.e. M_s). For each newly allocated non-anchor carrier, its N_{Sub} subcarriers will be totally allocated for part B or partially shared with part A according to whether the above ($P_A+N_C*W_B$) value is divisible by N_{Sub} or not.

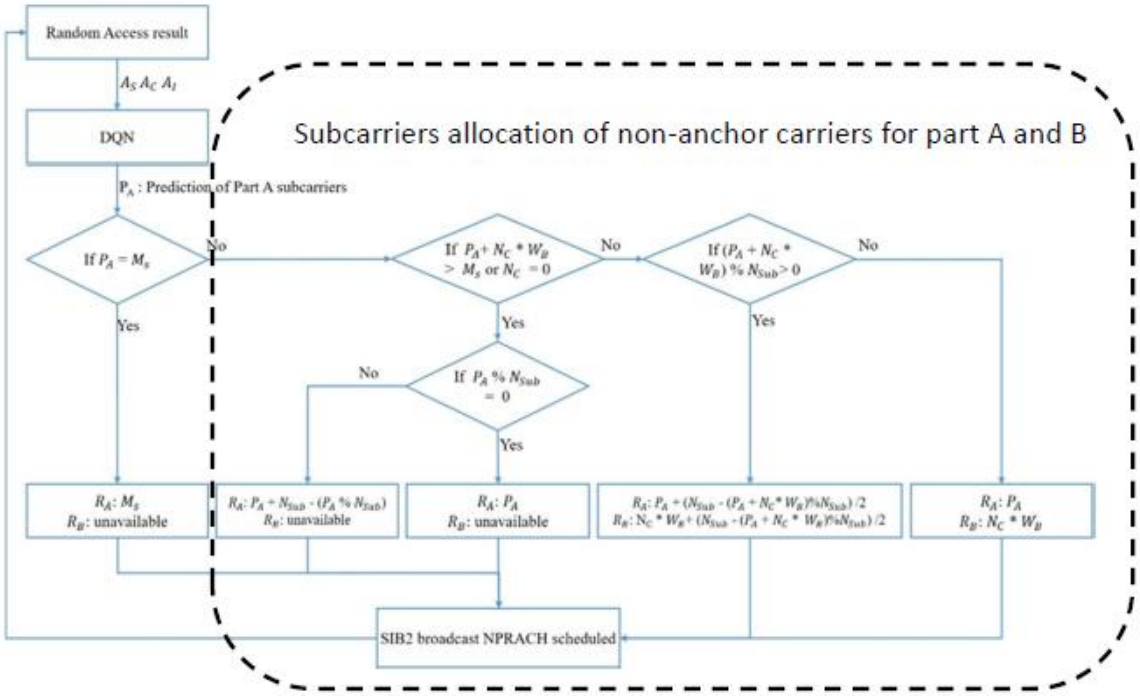


Figure 5. Procedure of subcarriers allocation for parts A and B in PRARA

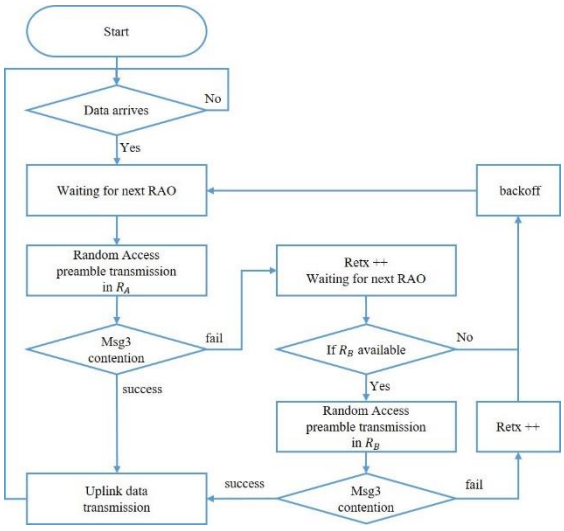


Figure 6. Random access procedure of devices in PRARA

The UEs issue the initial random access using the subcarriers provided in part A and examine whether there is any subcarrier provided in part B in the next RAO or not. It will contend with the subcarriers of part B if R_B is available if it collides with part A; otherwise, it will enter the backoff stage before the arrival of its next RAO. Figure 6 shows the operational procedure of the UE. Note that UE can still randomly access one of the subcarriers in part A after backoff. This is the main reason why the proposed algorithm first allocates the subcarriers for part A.

4 Experimental Results

To evaluate the performance of the proposed scheme, we performed exhaustive simulations and compared it with the rule-based scheme in [5]. Table 2 presents the parameters adopted during the simulations.

Table 2. Simulation parameters

Parameters	Values
Traffic model	Periodical Beta (3,4) distribution
Duration of traffic period (T)	600 s
RA period	1280 ms
W_B	2
Number of anchor carrier	1
Max. non-anchor carrier	15
N_{Sub}	48
Max. retransmission times (rtx)	10
Backoff time	U (0,256*2^{retx})
α	0.0001
γ	0.3
ε	0

The traffic model is assumed to be periodically beta distributed with parameters 3 and 4. Figure 7 shows the histogram of the traffic model.

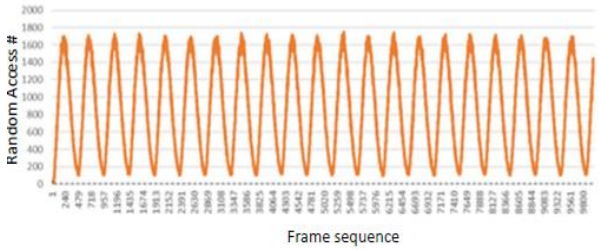


Figure 7. Traffic model of simulations

(1) Rewarding Function Design

The proposed scheme seeks the tradeoff between the low collision ratio and high resource utilization. The reward function is designed to benefit the number of successful subcarriers A_S, and decrease the reward for the number of collided subcarriers A_C, and the number of idle subcarriers A_I as shown in equation (1).

$$W_s = N_s * S - N_C - N_I \quad \textbf{(1)}$$

The parameter S is adopted to control the elastic of the subcarrier allocation. The higher value of S increases the preference for success subcarrier and results in allocating more radio resources for random access. Note that the above reward function tends to provide one subcarrier to achieve a higher reward value when S is 2, especially in heavy loads. Most subcarriers collide in heavy loads, and the reward from the successful subcarrier is only two times that of the collided subcarrier. The better way to obtain a higher reward value is to provide fewer subcarriers to reduce the number of collided subcarriers in equation (1). The beta distribution is adopted as the traffic model. Additionally, the low load (30K UEs/period), medium load (50K UEs/period), and high load (70K UEs/period) are provided during the simulations. The experiment results compare the proposed schemes to the results of the rule-based scheme [5] by adjusting the parameter S of the reward function. However, the case of S = 2 (i.e. $W_S = 2$) is not applied to the heavy load for the above reason. Figure 8 shows a preliminary comparison of the reinforcement learning (with $W_S = 3$) and rule-based approaches. To observe the allocation in detail, we examine the numbers of the random access arrivals and the allocated subcarriers in one cycle of the traffic model. The reinforcement learning scheme was adopted to allocate the subcarriers of part A in the proposed PRARA scheme. Therefore, only the RA arrivals and associated allocations in part A are observed. It illustrates that the number of subcarriers allocated by the proposed scheme is very close to the actual arrivals. However, the rule-based scheme allocates much more subcarriers than the actual arrivals. Additionally, the proposed scheme allocates fewer subcarriers than the actual arrivals in some frames because of the conservative allocation by applying S = 3.

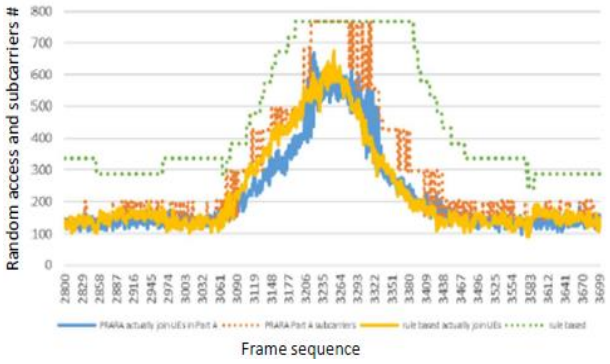
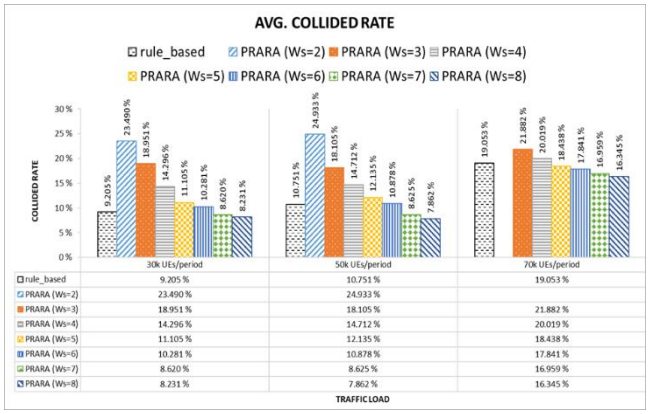


Figure 8. Comparison of subcarrier allocation of the proposed scheme ($W_S = 3$) and rule-based scheme

(2) Simulation Results

Figure 9 shows the performance comparison of subcarrier collided rates for different traffic loads. In the low traffic load, the proposed PRARA scheme performs similarly to the rule-based scheme when S is greater than 6 (i.e. $W_S = 7$ and $W_S = 8$). Additionally, the proposed scheme can achieve lower collided rates than the rule-based scheme for medium and high traffic loads. It is worth noting that the proposed scheme achieves a lower collided rate when S is greater than or equal to 5 in high traffic load.



	30k UEs/period	50k UEs/period	70k UEs/period
rule_based	9.205 %	10.751 %	19.053 %
PRARA (Ws=2)	23.490 %	24.933 %	
PRARA (Ws=3)	18.951 %	18.105 %	21.882 %
PRARA (Ws=4)	14.296 %	14.712 %	20.019 %
PRARA (Ws=5)	11.105 %	12.135 %	18.438 %
PRARA (Ws=6)	10.281 %	10.878 %	17.841 %
PRARA (Ws=7)	8.620 %	8.625 %	16.959 %
PRARA (Ws=8)	8.231 %	7.862 %	16.345 %

Figure 9. Comparisons of the subcarrier collided rates

Figure 10 compares the performance of the average random access delay. The results show that the proposed scheme achieves better delay performance than the rule-based scheme when S is greater than or equal to 4 in the three different traffic loads. This is because the design of the subcarriers in part B provides the UE to reissue RA without

entering the backoff procedure, as shown in Figure 6. Therefore, it can reduce the delay time if UE successfully accesses the subcarrier in part B.

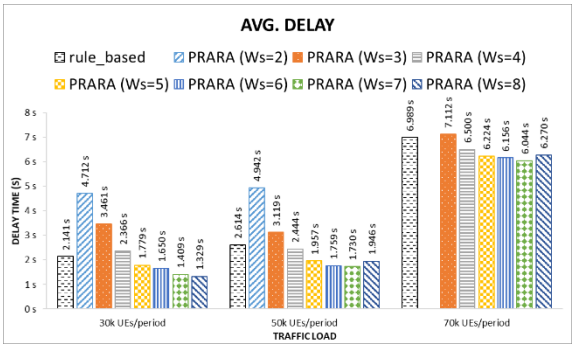
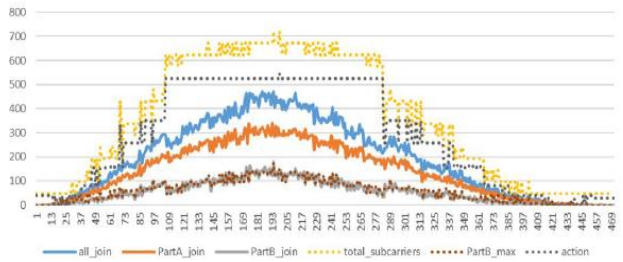
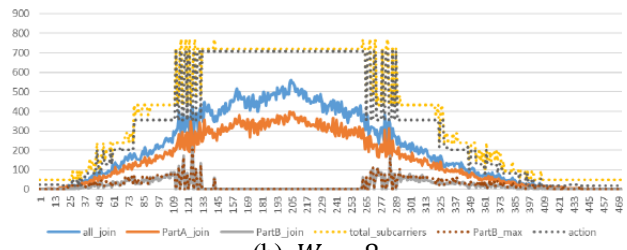


Figure 10. Comparison of the average access delay

Note that the average delay of the proposed scheme decreases when *S* becomes larger, except for the cases of $W_S = 8$ in the medium and high traffic loads. The average delay of $W_S = 8$ is slightly higher than that of $W_S = 7$. This is because the number of subcarriers allocated in part B is the residual subcarriers after the allocation for part A under the constraint of the maximum number of subcarriers that can be provided for RA, as illustrated in Figure 5. Additionally, the number of subcarriers allocated in part A increases as the value of S becomes large. We examine the subcarrier allocation in parts A and B in one traffic period. Figure 11(a) and Figure 11(b) provide the allocation results of the cases $W_S = 7$ and $W_S = 8$, respectively. It shows that part B can still be allocated with some subcarriers when W_S equals 7. However, there is no subcarrier allocated in part B for the case of $W_S = 8$ because all subcarriers are allocated in part A. If there is no subcarrier allocated in part B from frame 145 to 265, UE can not reissue the RA immediately after the collision. Without the subcarrier in part B, the collided UE is required to perform backoff procedure, and, therefore, results in a longer delay.



(a) $W_S = 7$



(b) $W_S = 8$

Figure 11. Allocations of subcarriers in parts A and B for different W_S

Figure 10 shows that the proposed PRARA achieves a lower average delay than the rule-based scheme when *S* is greater than or equal to 4, especially for the medium and high traffic loads. The larger the value of *S*, the more easy the prediction for the required subcarriers. Although allocating more subcarriers can reduce the collision rate and delay time, whether the subcarriers are effectively used is a critical issue from the resource allocation perspective. Table 3 presents the results of the average allocated subcarriers, non-anchor carriers, and the use of the rule-based and the proposed PRARA scheme (S = 4 to 8).

Table 3. Simulation results of subcarrier allocation and utilization

Schemes	50K UEs/period			70K UEs/period		
	Ave. allocated subcarriers	Ave. allocated non-anchor carriers	Ave. utilization of subcarrier	Ave. allocated subcarriers	Ave. allocated non-anchor carriers	Ave. utilization of subcarrier
Rule based	359.858	6.497	29.624%	465.578	8.700	32.020%
PRARA (*Ws* =4)	325.636	5.784	32.724%	433.788	8.037	34.341%
PRARA (*Ws* =5)	342.787	6.141	31.097%	445.188	8.275	33.477%
PRARA (*Ws* =6)	354.252	6.380	30.095%	449.632	8.367	33.154%
PRARA (*Ws* =7)	375.648	6.826	28.385%	457.508	8.531	32.593%
PRARA (*Ws* =8)	383.524	6.990	27.803%	467.487	8.739	31.899%

The results show that the proposed scheme allocates fewer subcarriers and has higher utilization than the rule-based scheme in medium traffic load when *S* equals 4, 5, or 6. For the high traffic load, the proposed scheme also performs better than the rule-based scheme when *S* equals 4, 5, 6, or 7. According to the simulation results, the selection for the value of S shall consider the traffic load and the allowable radio resource for random access. This parameter can be applied as the tradeoff role in radio resource management.

5 Conclusions

In this paper, we propose the reinforcement learning-based subcarrier allocation scheme to deal with random access from mass IoT devices in the 3GPP NB-IoT network. The proposed PRARA scheme allocates the subcarriers in two parts. The number of subcarriers provided in part A is determined from the prediction of the reinforcement learning model and the number of collided subcarriers in the previous RAO. Additionally, subcarriers allocated in part B allow the collided device to reissue the access without a backoff procedure to reduce the random access delay. The proposed learning model is rewarded by the access conditions of subcarriers, i.e. success, idle, or collision. The parameter *S* is applied to adjust the elasticity of the prediction. We conducted exhaustive simulations to evaluate the effectiveness of the proposed scheme. The results show that the proposed PRARA scheme achieves a lower collision rate, delay time, and allocated subcarriers than the rule-based scheme [5]. It is noted that the subcarrier allocation is in per carrier basis, not in per subcarrier, therefore, this issue is very critical from radio resource management point of view. The simulation results also show that the larger value of *S* leads the learning model to allocate more subcarriers in part A to reduce the collision rate. However, it results in a shortage of the number of subcarriers allocated in part B, and the access delay time may increase. The design of an adjustable reward function, such as

a deep deterministic policy gradient [15], could be a possible approach to solving this issue, which is one of our ongoing research directions.

Acknowledgment

This work was supported in part by the Ministry of Science and Technology (MOST) (grant number: 109-2221-E-008 -052 -MY2), Taiwan.

References

[1] Y.-P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, H. S. Razaghi, A Primer on 3GPP Narrowband Internet of Things, *IEEE Communications Magazine*, Vol. 55, No. 3, pp. 117-123, March, 2017.

[2] M. Jamalzadeh, L.-D. Ong, M. N. B. M. Nor, 5G Technologies: A New Network Architectures and Design, *Journal of Internet Technology*, Vol. 19, No. 7 , pp. 1983-1991, December, 2018.

[3] 3GPP, *Release 14 Description; Summary of Rel-14 Work Items, V14.0.0*, TR 21.914, May, 2018.

[4] Rohde & Schwarz, *Narrowband Internet of Things: White Paper*, 1MA266_0e, August, 2016. https://cdn.rohde-schwarz.com/pws/dl_downloads/dl_application/application_notes/1 ma266/1MA266_0e_NB_IoT.pdf.

[5] Y.-W. Chen, G.-Y. Xue, The Adaptive Random Access Carrier Allocation Scheme in NB-IoT Networks, *Communications and Network*, Vo. 14, No. 1, pp. 1-11, February, 2022.

[6] S.-M. Oh, J. S. Shin, An Efficient Small Data Transmission Scheme in the 3GPP NB-IoT System, *IEEE Communications Letters*, Vol. 21, No. 3, pp. 660-663, March, 2017.

[7] K. Lee, J. W. Jang, An Efficient Contention Resolution Scheme for Massive IoT Devices in Random Access to LTE-A Networks, *IEEE Access*, Vol. 6, pp. 67118-67130, 2018.

[8] Y. Zhao, K. Liu, H. Yan, L. Huang, A Classification Back-off Method for Capacity Optimization in NB-IOT Random Access, *2017 11th IEEE International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, Xiamen, China, 2017, pp. 104-108.

[9] C. Di, B. Zhang, Q. Liang, S. Li, Y. Guo, Learning Automata-Based Access Class Barring Scheme for Massive Random Access in Machine-to-Machine Communications, *IEEE Internet of Things Journal*, Vol. 6, No. 4, pp. 6007-6017, August, 2019.

[10] H. S. Jang, H. Jin, B. C. Jung, T. Q. S. Quek, Resource-Optimized Recursive Access Class Barring for Bursty Traffic in Cellular IoT Networks, *IEEE Internet of Things Journal*, Vol. 8, No. 14, pp. 11640-11654, July 2021.

[11] L. Tello-Oquendo, I. Leyva-Mayorga, V. Pla, J. Martinez-Bauset, J.-R. Vidal, V. Casares-Giner, L. Guijarro, Performance Analysis and Optimal Access Class Barring Parameter Configuration in LTE-A Networks With Massive M2M Traffic, *IEEE Transactions on Vehicular Technology*, Vol. 67, No. 4, pp. 3505-3520, April 2018.

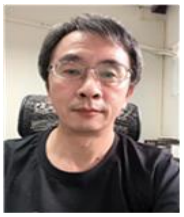
[12] F. Wu, B. Zhang, W. Fan, X. Tian, S. Huang, C. Yu, Y. Liu, An Enhanced Random Access Algorithm Based on the Clustering-Reuse Preamble Allocation in NB-IoT System, *IEEE Access*, Vol. 7, pp. 183847-183859, 2019.

[13] N. Jiang, Y. Deng, A. Nallanathan, Traffic Prediction and Random Access Control Optimization: Learning and Non-learning-Based Approaches, *IEEE Communications Magazine*, Vol. 59, No. 3, pp. 16-22, March, 2021.

[14] N. Jiang, Y. Deng, A. Nallanathan, J. A. Chambers, Reinforcement Learning for Real-Time Optimization in NB-IoT Networks, *IEEE Journal on Selected Areas in Communications*, Vol. 37, No. 6, pp. 1424-1440, June, 2019.

[15] D. Wu, X. Dong, J. Shen, S. C. H. Hoi, Reducing Estimation Bias via Triplet-Average Deep Deterministic Policy Gradient, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 31, No. 11, pp. 4933-4945, November, 2020.

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).

Ji-Zheng You 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).