# Real-time RL-based 5G Network Slicing Design and Traffic Model Distribution: Implementation for V2X and eMBB Services

**WeiJian Zhou[1], Azharul Islam[1], and KyungHi Chang[1*]**
[1] Department of Electrical and Computer Engineering, Inha University
Incheon 22212, Korea
[e-mail : 2467238208@qq.com, khchang@inha.ac.kr]
[e-mail : azharul.dream007@gmail.com]
[*]Corresponding author: KyungHi Chang

---

## *Abstract*

As 5G mobile systems carry multiple services and applications, numerous user, and application types with varying quality of service requirements inside a single physical network infrastructure are the primary problem in constructing 5G networks. Radio Access Network (RAN) slicing is introduced as a way to solve these challenges. This research focuses on optimizing RAN slices within a singular physical cell for vehicle-to-everything (V2X) and enhanced mobile broadband (eMBB) UEs, highlighting the importance of adept resource management and allocation for the evolving landscape of 5G services. We put forth two unique strategies: one being offline network slicing, also referred to as standard network slicing, and the other being Online reinforcement learning (RL) network slicing. Both strategies aim to maximize network efficiency by gathering network model characteristics and augmenting radio resources for eMBB and V2X UEs. When compared to traditional network slicing, RL network slicing shows greater performance in the allocation and utilization of UE resources. These steps are taken to adapt to fluctuating traffic loads using RL strategies, with the ultimate objective of bolstering the efficiency of generic 5G services.

---

---

## 1. Introduction

$5$G represents not only an advancement in wireless technology, but also a revolution in network service design [1]. According to [2-4], eMBB, massive machine-type communications (mMTC), and ultra reliable low latency communications (uRLLC) are the three key 5G applications. The primary difference between 4G and 5G is that 5G allows huge connections and incredibly low latency. To put it differently, the true strength of 5G is in creating a world where everything is connected. The network speed and latency requirements for mass terminal usage scenarios are changing, and 5G integrates them into a common network architecture. To meet the increasing network demands, 5G network slicing has been introduced as a vital technology [1].

With the use of network slicing, the physical network can be divided up into a number of logical networks to accommodate various applications with different performance and flexibility needs. In this paper, we investigate the RAN slicing issue for 5G universal services, in particular eMBB and V2X. In our suggested strategy, slicing ratios are initially determined using a heuristic-based algorithm with RL. The ratios are then changed to ensure equitable resource distribution and improve network performance. Slices that have more resources than they require to give the surplus to those that are in need, with transfers computed and carried out proportionately based on the resource requirements of each slice. The resources allotted to the slices are dynamically adjusted by RL by utilizing experience gained from network interaction. By doing this, resource usage is improved while maintaining Quality of Service (QoS) [1]. We specifically suggest a 5G network slicing design and traffic model distribution based on real-time RL. Based on the outcomes of our simulations, the suggested design enhances network efficiency in terms of resource use and outage likelihood.

Focus on a Specific Service Type, this work explores RAN slicing for 5G generic services, particularly eMBB and V2X.Here is the unique contribution of our paper as follows:

- We use a heuristic-based approach with RL in our method to allocate resources dynamically and effectively.
- We suggest a special system that will allow slices with surplus resources to proportionately help those in need.
- We present Reinforcement Learning, a novel method in network slicing, to improve resource consumption while maintaining QoS
- We compare a standard network slicing model and an RL-based model in real-time, assessing performance via PRB and outage probability metrics in our simulations.

The rest of the paper is arranged in the same way that Chapter 2 introduces the related work, Chapter 3 explains the system model design for radio resource slicing and the traffic model design for two slicing services, and also analyzes the outage problem of slicing services. Chapter 4 discusses the network slicing strategy based on the RL algorithm, detailing the Q-learning process based on the SoftMax strategy and the optimization solution based on the heuristic algorithm. Chapter 5 sets up the simulation that integrates network deployment and reinforcement learning strategy, and evaluates performance based on whether real-time estimation of online network characteristics is performed.

## 2. Related Work

In the literature, different works proposed solutions to design and control network slices [5-7]. A proper architecture for a 5G system based on network slicing to manage mobility between

different service networks is proposed in [7]. Conversely, functional decomposition and network slicing are proposed in [8] as tools to refine the evolved packet core (EPC). They investigate the feasibility of a mobile core network on the basis of network slicing as well as functional decomposition. Reference [9] provides a customizable, adaptable software-defined (SD)-RAN architecture that focuses on the control plane supporting RAN control applications. Reference [10] focuses on admission control choices for network slice requests as well as network slice traffic analysis and forecast for each network slice. It proposes a measurement-based adaptive adjustment of the estimated load. The combined admission control of virtual wireless networks is also presented in [11] together with a heuristic approach and slicing. A network slicing coordination model based on service demand and resource availability is shown in [12]. They provide a framework for a Markovian decision-making process. [13] proposes a specific slicing strategy based on reinforcement learning and optimizes the results using heuristic algorithms. The choice-making procedure of the agent can be refined using diverse RL methodologies which are mentioned in this research work [14]-[17]. In addition, for the practical network model, this study [18] tackles virtual resource allocation, stochastic distribution, and relative assignment. In this paper, they proposed an offline RL-based 5G network slicing design and traffic model distribution. But in our paper, we propose another real-time RL-based 5G network slicing design and traffic model distribution and evaluate the performance between these two schemes. The simulation results show the improvement in terms of network performance with the proposed scheme.

## 3. System Model Design for DL

### 3.1 System Model and Service Distribution

The scenario considered is that a cellular Next Generation RAN Access Network (NG-RAN) has a gNodeB (gNB) consisting of a single cell. A group of eMBB cellular users are randomly distributed around gNB, numbered as $m = 1, ..., M$. And as shown in **Fig. 1**, several independent vehicle flows travel down a straight route. The highway consists of six lanes, there are three lanes from the left to the right driving direction, the driving direction for the other three lanes is from right to left. According to the length of the highway, it is divided into smaller areas, thus the highway section is separated into clusters. Each vehicle is supposed to have User Equipment (UE) that can communicate with other UEs in the same cluster. The clusters are numbered as $j = 1, ..., C$, the vehicles number in the j-th cluster is $i = 1, ..., V(j)$.

A Poisson process with an arrival rate of $\lambda_a$ treats vehicles on the roadway as entering the cell. A Poisson process with a generation rate of $\lambda_m$ generates the number of eMBB UEs [19].
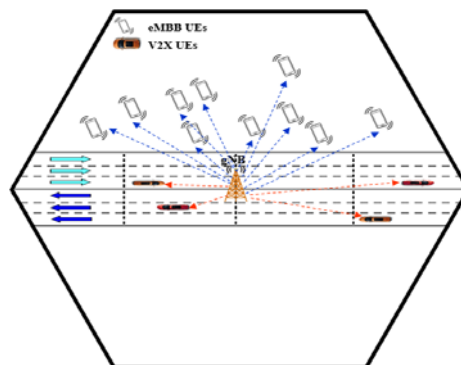


**Fig. 1.** The cellular network system model.

### 3.2 Network Model for Slicing

Due to the fact that eMBB demands a lot of bandwidth for high data rate, V2X services are very latency sensitive, it is necessary to support both V2X and eMBB services simultaneously. The network is conceptually split into two slices, we assume V2X slice_ID = 1 and eMBB slice_ID = 2.

The entire cell bandwidth is arranged into Resource Blocks (RBs) with bandwidth $B$. The number of RBs in downlink is denoted as $N_{DL}$. The slicing process should allocate DL RBs among the two slices. To do this, denote $\alpha_{s,DL}$ as the part of DL resources for slice_ID = $s$, where $s = 1,2$.

The relationship is expressed as follows:

$$\sum_s \alpha_{s,DL} = 1 \tag{1}$$

### 3.3 Traffic Model Design without Slicing Strategy

### 3.3.1 V2X Traffic Model

According to the Poisson arrival model, each vehicle is assumed to generate random packets at a rate of $\lambda_{niu}$ packets/second [20]. The size of the message is $S_m$. The messages are transmitted utilizing DL resources in cellular mode [21]. The average number of PRBs required for each transmission time interval (TTI) for V2X users with slice_ID = 1 in DL, denoted as $\tau_{1,DL}$, can be written as:

$$\tau_{1,x} = \frac{\sum_{t=1}^{T} \sum_{j=1}^{C} \sum_{i=1}^{V(j)} \frac{m(j,i,t) \cdot S_m}{SP_{eff}(j,i,t)}}{T \cdot B \cdot F_d} \tag{2}$$

where x∈{DL}, represents the link type, The number of messages sent by the vehicles is given by $m(j,i,t)$, for the j-th cluster and in the t-th TTI, the spectral efficiency of the downlink is written by $SP_{eff,x}$, TTI duration is denoted as $F_d$, and $T$ is the number of TTIs utilized to determine the time window for calculating the average value.

### 3.3.2 eMBB Traffic Model

Regarding the eMBB service, each eMBB user generates a session that requires a certain guaranteed bit rate. A Poisson process is followed as the session generation model with a rate of $\lambda_e$ (sessions/s) and an exponential distribution of session durations averaged over $T_e$. These users transmit PRBs allocated to the eMBB slice in the downlink. For eMBB users with slice_ID = 2, the average number of PRBs needed in DL to support a specific guaranteed bit rate $R_b$ is denoted as $\tau_{2,DL}$, can be written as:

$$\tau_{2,x} = \frac{\sum_{t=1}^{T} \sum_{m=1}^{M} \rho_x(m,t)}{T} \tag{3}$$

where $x$∈{DL}, $\rho_x(m,t)$ is the required PRBs number by the m-th user in the downlink, in the t-th TTI to get the needed bit rate $R_b$. $\rho_x(m,t) = \frac{R_b}{SP_{eff,x} * B}$. $\tau_{2,DL}$ which is calculated within a time window $T$.

### 3.4 Outage Problem for Slicing

The objective is determining the optimal slice ratios $\alpha_{s,DL}$ to maximize the total resource utilization subject to the constraint of meeting the demands for slice users [22], [23].

There is a constraint for the total DL resource utilization $U_{DL}$, which states that the total resources allocated to this slice by the RAN slicing cannot be exceeded by the sum of the slice, i.e., $\alpha_{s,DL} \cdot N_{DL}$. If not, an outage situation will be caused, and the utilization is limited

to $\alpha_{s,DL} \cdot N_{DL}$. So, it can be defined as:

$$U_{DL} = \sum_s \min(\tau_{s,DL,} \alpha_{s,DL} \cdot N_{DL}) \tag{4}$$

where total downlink resources utilization $U_{DL}$ is the sum across all slices $s$ ,of the minimum between $\tau_{s,DL}$ which seems to be the demand for resources in slice $s$ for DL and the resource allocated to that slice $\alpha_{s,DL}.N_{DL}$.

Accordingly, the downlink optimization problem is defined as maximizing the DL resource utilization while guaranteeing that the outage probability is less than a maximum allowable limit $P_{out}$. Formally, this problem is stated as follows:

$$\max_{\alpha_{s,DL}} U_{DL} \tag{5}$$

$$\text{s.t. } \Pr[\tau_{s,DL} \geq \alpha_{s,DL} \cdot N_{DL}] < p_{out} \ s = 1,2 \tag{5a}$$

$$\sum_s \alpha_{s,DL} = 1 \tag{5b}$$

where $U_{DL}$ is resource utilization, $\alpha_{s,DL}$ represents the ratio of downlink resources allocated to as specific slice (s). The operator indicates that the optimization tries to find the value of $\alpha_{s,DL}$ that will maximize $U_{DL}$. (5a) constraint ensures that the probability (Pr) of $\tau_{s,DL}$ (amount of traffic on slice s in the downlink direction) exceeding the product of the downlink resource allocation ratio $\alpha_{s,DL}$ and the total number of DL resources $N_{DL}$ is less than a given outage probability limit $p_{out}$. This constraint applies to all slices (s=1,2). (5b) constraints ensure that the sum of the Dl resource allocation ratios all slices (s) equals one, implying that all available resources must be allocated.

## 4. Network Slicing Strategy with RL Algorithm

RAN slicing is essentially a problem of optimization of network resource allocation, and we should select the optimal slicing ratio from a series of slicing ratios. That is, the optimal slicing ratio is selected based on the resulting resource utilization performance. For this reason, we decided to use reinforcement learning and low-complexity heuristics to solve this problem. **Fig. 2** shows the overall approach. A slicing controller is considered to decide the slicing ratio, $\alpha_{s,DL}$ for each slice. There are two main parts in the operation of the slicing controller. In the initial segment, an RL algorithm identifies some intermediate slicing ratios, referred to as $\beta_{s,DL}$. The second part, which is a heuristic method, was added afterward. It takes the RL algorithm result as input and alters the slicing ratios to get the final slicing ratios $\alpha_{s,DL}$. In the next sections, we will describe the combination in detail, respectively.
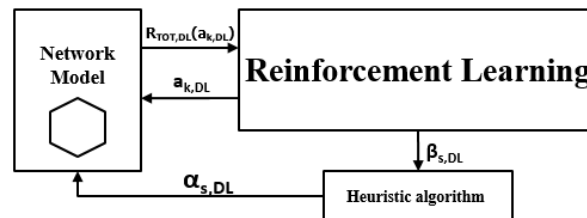


**Fig. 2.** RL-based RAN slicing strategy.

## 4.1 Q-learning Strategy

Assume an RL algorithm is performed for the downlink to determine $\beta_{s,DL}$. The state at the time $'t'$, $s(t)$, for our reinforcement learning approach is made up of the rewards of previous actions ($past_{rewards}$), other significant network performance indicators, and the current and

demanded resource allocations for V2X and eMBB services $current\_(allocation_{V2X})$, $current\_(allocation_{eMBB})$. The state is therefore described as:

$$s(t) = \{current\_(allocation_{V2X}), current\_(allocation_{eMBB}),$$
$$current\_(demand_{V2X}), current\_(demand_{eMBB}),$$
$$Past_{rewards}, other\ network\ performance\ metrics\} \tag{6}$$

Since the paper discusses the problem of slicing based on two service types, we set 20 slicing actions based on V2X and eMBB, the action set is:

$$\begin{cases} \beta_{1,x}(k) = 0.05k \\ \beta_{2,x}(k) = 1 - 0.05k \end{cases} \tag{7}$$

where $k = 1,2,\dots,20, x = \{DL\}$, $\beta_{1,x}(k)$ is the fraction of the PRB reserved for V2X service and $\beta_{2,x}(k)$ is the fraction of the PRB reserved for eMBB service.

In the procedure of RL, trying different action $a_k$ (different slicing ratios) to find an optimal solution. As the result of the action selection, the process of RL can get a reward $R_{TOT,DL}(a_k)$, which evaluates how good the outcome of the action is in regard to the desired optimal goal. On the basis of reward, the decision-making of the RL algorithm can be adjusted, to gradually learn the actions that lead to the highest reward [24]. Selecting an action involves striking a balance between exploitation (doing actions with high rewards) and exploration (trying different actions to gain knowledge from them) [25].

In this paper, the considered RL algorithm is Q-learning, which is based on SoftMax decision and allows for the long-term exploration and exploitation of all feasible actions [26]. Therefore, the reward can be determined [27]. The next section will go through the details of the reward function and how the Q-learning algorithm works.

### 4.1.1 Reward Computation

The appropriate reward mechanism should be in line with the action's potential. On this basis, for a selected action $a_k$ with the corresponding slicing ratio $\beta_{s,DL}(k)$, a normalized resource utilization $\psi_{s,DL}(a_k)$ function can be used to calculate the reward for slice $s$, defined as the proportion of the resources allocated for the related action to the total resources actually consumed.

For the V2X slice (*s*=1):

$$\psi_{1,DL}(a_k) = \frac{\tau_{1,DL}}{\beta_{1,DL}(k) \cdot N_{DL}} \tag{8}$$

For the eMBB slice (*s*=2):

$$\psi_{2,DL}(a_k) = \frac{\tau_{2,DL}}{\beta_{2,DL}(k) \cdot N_{DL}} \tag{9}$$

According to these equations, the reward $R_{s,DL}(a_k)$ as a consequence of the action $a_k$ is described as:

$$R_{s,DL}(a_k) = \begin{cases} e^{\psi_{s,DL}(a_k)} & \psi_{s,DL}(a_k) \leq 1 \\ e^{-\psi_{s,DL}(a_k)} & \psi_{s,DL}(a_k) > 1 \end{cases} \tag{10}$$

In (10), the reward function increases exponentially to a maximum at $\psi_{s,DL}(a_k) = 1$ as long as the value of $\psi_{s,DL}(a_k)$ is between 0 and 1. Thus, the actions that lead to an increase in this value, i.e., lead to higher utilization, also receive greater reward feedback.

On the contrary, if $\psi_{s,x}(a_x) > 1$, this case represents that slice *s* will have an outage. Therefore, the reward will be reduced. Moreover, since the total reward must consider the action effect on all the slices $s = 1, 2$, it is generally described as the geometric mean of the rewards for each slice:

$$R_{TOT,DL}(a_k) = \left(\prod_{s=1}^{2} R_{s,DL}(a_k)\right)^{\frac{1}{s}} \tag{11}$$

where $R_{TOT,DL}(a_k)$ represents the reward associated with each slice 's' for a given action $a_k$. The reward for each slice is calculated based on whether the allocated resources meet the demands of that slice. Total rewards are computed as the geometric mean of the individual slice rewards.

### 4.1.2 Q-values Computation

The ultimate objective of the slice controller's Q-learning method is to identify the course of action that maximizes the anticipated long-term reward for per slice. Q-learning connects with the network model during a discrete time period at a set time and calculates the reward for the selected action to accomplish this aim. The slice controller maintains experience records for actions $a_k$ in accordance with the reward, and the action-value function (also called Q-value) is kept in $Q_{DL}(a_k)$. For each time step, the values of $Q_{DL}(a_k)$ are updated according to the learning method, where it is single-state and has a null discount rate:

$$Q_{DL}(a_k) \leftarrow (1-\alpha)Q_{DL}(a_k) + \alpha \cdot R_{TOT,DL}(a_k) \tag{12}$$

where $R_{TOT,DL}(a_k)$ is the total reward for performing an action $a_k$ for V2X and eMBB slices, and the learning rate is denoted as $\alpha$, where $\alpha \in (0,1)$. $Q_{DL}(a_k)$ is initialized to an arbitrary value at initialization when action $a_k$ has never been used previously.

### 4.1.3 Selection Criterion based on SoftMax Policy

The SoftMax strategy is used to select which actions to take depending on $Q_{DL}(a_k)$, where the selection of different actions is probabilistic. In detail, the probability $P(a_k)$ corresponding to the selection of action $a_k$, is described as:

$$P(a_k) = \frac{e^{Q_{DL}(a_k)/\tau}}{\sum_{j=1}^{A_x} e^{Q_{DL}(a_j)/\tau}} \tag{13}$$

where $\tau$ is a positive integer defined as temperature parameter, which affects the selection probability [28]. The higher the value of $\tau$ is, the closer the probability of actions will be to the same. On the contrary, a low value of $\tau$ will result in a large variance in selection probability for actions with various Q values. Exploration and exploitation may be successfully balanced using the SoftMax strategy, that is, selecting actions having a high possibility of yielding a high reward, but also keep exploring new actions with a certain probability, it can lead to better decisions in the future [29].

### 4.1.4. Simulation Result (without heuristic solution)

In the downlink simulation, when running the code, in the MATLAB command window:

episode = 0998: a_k = 007, Gamma1=40.07, Gamma2=91.52, R_DL(V2X) =1.77, R_DL (eMBB)=2.02, R_TOT_DL=1.89, beta_V2X=0.35, beta_eMBB=0.65
episode = 0999: a_k = 003, Gamma1=48.24, Gamma2=92.23, R_DL(V2X) =0.20, R_DL (eMBB)=1.72, R_TOT_DL=0.59, beta_V2X=0.15, beta_eMBB=0.85
episode = 1000: a_k = 011, Gamma1=51.66, Gamma2=81.53, R_DL(V2X) =1.60, R_DL (eMBB)=2.47, R_TOT_DL=1.99, beta_V2X=0.55, beta_eMBB=0.45

where $a_k$ is the action for slicing ratio. "*Gamma*" is the average number of required RBs for each slice. "*R_DL*" is the reward computation for each slice. "*R_TOT_DL*" is the total reward computation for the downlink. "*beta*" is the corresponding slicing ratio for each slice.

Two slicing services are included in the implementation, i.e., the sum of the two slicing ratios is equal to 1. As can be seen from the above, the network resources are successfully allocated to the respective slices.

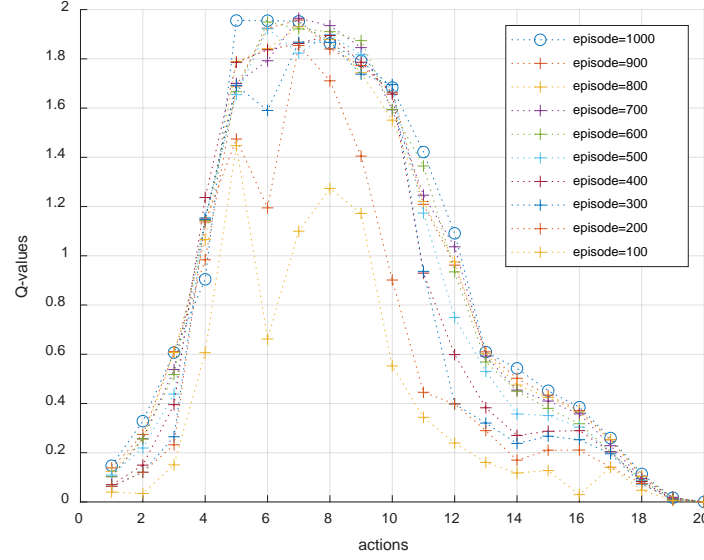The following **Fig. 3**, is the Q-value based on different episodes for V2X and eMBB UEs:



**Fig. 3.** Actions *VS.* Q-values based on different episodes.

In the simulation, the network load is light. In the network, the PRB demand from the vehicular UEs is generally lower than that from the eMBB UEs, as shown in the log from the Matlab console, so a better slicing strategy should allocate more PRBs to the eMBB service.

After 1000 episodes, the final Q-value-table indicates that the action#5/#6/#7 are more preferable than the other actions. Action#5/#6/#7 corresponds to the PRB ratio for V2X service being 0.3, 0.35, 0.4, and the PRB ratio for eMBB being 0.7, 0.65, 0.6, so more PRBs are allocated to the eMBB service.

## 4.2 Optimization Solution

In this section, we utilize a heuristic approach to adjust the intermediary slicing ratio, $\beta_{s,DL}$ selected through RL based on the resource requirements [30].

The idea is based on the actual RB requirements and the slicing ratio $\beta_{s,DL}$. The scheme estimates whether one of the slices has more resources in the downlink than actually needed, i.e., $\psi_{s,DL}(a_{DL\_sel}) < 1$, meanwhile another slice $s'$ has resources that less than required, i.e., $\psi_{s,DL}(a_{DL\_sel}) > 1$. In this case, another slice $s'$ can take use of the extra capacity that slice $s$ leaves behind $\Delta C_{s,DL}$. In detail, the extra capacity is expressed as:

$$\Delta C_{s,DL} = \left(1 - \psi_{s,DL}(a_{DL\_sel})\right) \cdot \omega \tag{14}$$

where $\omega$ is a configuration parameter in the range [0,1], in order to leave some margin for the variation of RBs demand. For slice $s$, the slicing ratio will reduce by $\Delta C_{s,DL}$, i.e., $a_{s,DL} = \beta_{s,DL} - \Delta C_{s,DL}$, Conversely, for another slice $s'$, the slicing ratio will increase by $\Delta C_{s,DL}$, i.e., $a_{s',DL} = \beta_{s',DL} + \Delta C_{s,DL}$.

Following the optimization illustrated in **Fig. 4**, a significant performance improvement can be observed in the values of V2X and eMBB compared to the scenario without the

heuristic method. In both the V2X and eMBB cases, the Q-values now exceed 2 for resource allocation based on demand, representing an increase from the sub-2 levels observed when using the heuristic method.
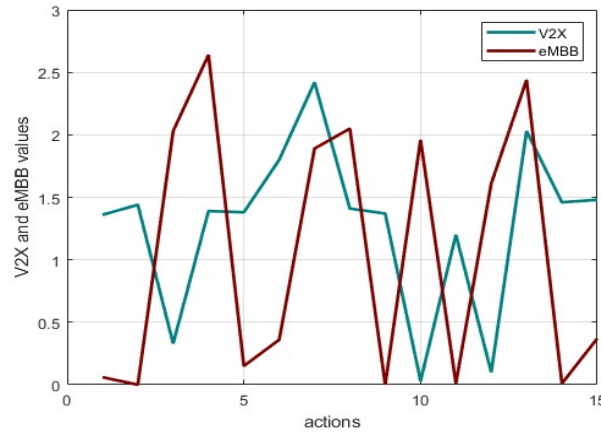


**Fig. 4.** Actions *VS.* optimization V2X and eMBB values.

# 5. Performance Evaluation

In this Section, we employ MATLAB for system-level simulations to examine the efficacy of our proposed RAN slicing strategy.

## 5.1 Simulation Setup

The simulation model we utilized is founded on a single-cell hexagonal structure, equipped with a gNB. This cell configuration accommodates a channel encompassing 200 RBs, 12 subcarriers, and a subcarrier spacing of Δf = 30 kHz, mirroring one of the 5G NR numerologies we reference.

Simulation model considers vehicle UEs that communicate through cellular mode (DL) using slice_ID = 1, and eMBB UEs that communicate in cellular mode (DL) using slice_ID = 2. Vehicular UEs are distributed along two 3-lane highways, eMBB UEs are spread randomly in the cell. In Table 1~4, All relevant parameters are summarized. Since we discuss the problem of slicing based on two service types, the action of RL is defined as: action $a_k$ corresponds to $\beta_{1,DL}(k) = (0.05k)$, $\beta_{2,DL}(k) = (1 - 0.05k)$, where $k = 1,2,...,20, x = \{DL\}$. The evaluation results presented are intended to evaluate and demonstrate the performance for the RAN slicing strategy in the aspects of PRB utilization, network outage probability.

## 5.2 Network Deployment

**Fig. 5**, below shows the network deployment for the downlink, from where we start the simulation, initialize the simulation and the WINNER II channel model, the slicing ratios are initialized as well. In the network model, for every drop, the simulation process includes:
[1] Prepare the simulation for the drop.
[2] Run the WINNER II channel model for the drop.
[3] Collect the network characteristics.
[4] Calculate the PRB utilization.

In this **Fig. 5**, we can set the locations of the gNB, the highways, and every UE. The channel model is generated according to the WINNER II B1 (Urban micro-cell) and the SINR of each link is estimated from the channel gain.
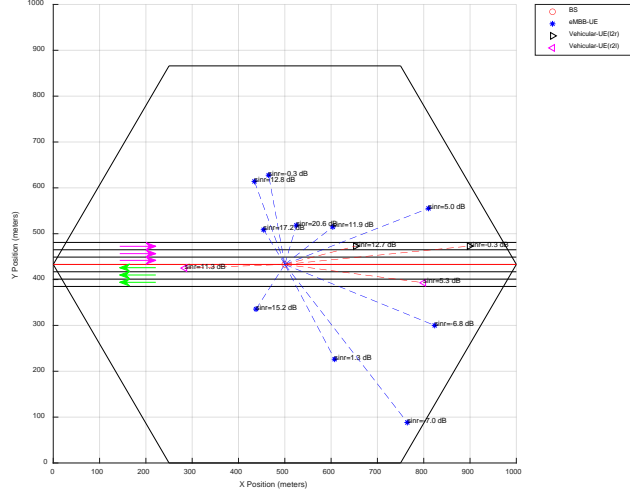


**Fig. 5.** Network Deployment for Downlink.

The specific simulation parameters are as follows:

**Table 1.** General simulation parameters

| Parameters | Values |
|---|---|
| Cell radius | 500m |
| Frequency | 2.6GHz |
| RBs/cell | $N_{DL}$=200RBs |
| Path loss model | WINNER II B1 (Urban micro-cell) |
| Base station antenna gain | 5dB |
| Height of the gNB | 10m |
| Time window T | 10ms |
| TTI duration ($F_d$) | 0.5ms |
| Sim. Num. of drops/round | 20 |
| Sim. Num. of round | 10 |
| Ignore online NW characteristics | False/True |

**Table 2.** Simulation parameters for V2X

| Parameters | Values |
|---|---|
| Highway Length | 1km |
| Num. of lanes | 3 in either direction |
| Lane width | 4m |
| Num. of clusters | 4 |
| Size of cluster | 250m |
| Vehicle speed | 80km/h |
| Vehicular UE height | 1.5m |
| Packet arrival rate $\lambda_{niu}$ | 200packets/s |
| Vehicle arrival rate $\lambda_a$ | 1UE/s |
| Message size ($S_m$) | 300bytes |

**Table 3.** Simulation parameters for eMBB

| Parameters | Values |
|---|---|
| User Equipment arrival rate $\lambda_m$ | 1UE/s |
| UE height | 1.5m |
| Average session generation rate $\lambda_e$ | 300/400/500/600 sessions/s |
| Guaranteed bit rate $R_b$ | 1Mb/s |
| Average session duration | 120s |

**Table 4.** Simulation parameters for RAN slicing algorithm

| Parameters | Values |
|---|---|
| Learning rate α | 0.1 |
| Temperature parameter τ | 0.1 |
| BW of PRB | 30e3*12 |
| Probability of selection (SL vs Cellular) | 0.5 |
| ω | 0.85 |
| Actions of the RL algorithm | 20 actions, $k = 1,2,...,20.$ $\begin{cases} \beta_{1,x}(k) = 0.05k \\ \beta_{2,x}(k) = 1 - 0.05k \end{cases}$ |

## 5.3 Estimation of network characteristics

During each cycle of the simulation, an exchange of updated slicing ratios and network characteristics occurs between the online network model and the slicing controller. The slicing controller consists of the offline network model, the reinforcement learning, and the low-complexity heuristic algorithm. In every round of simulation, the online network run for a number of simulations drops.

During the above simulation, the network characteristics are collected in every simulation drop, these network characteristics consist of:

[1] The number of V2X UEs, eMBB UEs.

[2] The SINR (in dB) values to all the UEs in all of the downlinks.

[3] The frequency at which packets arrive for the V2X service, and the rate at which sessions are generated for the eMBB service.

For the processing of network characteristics data, there are two schemes: whether to estimate the network characteristics.

For the estimation scheme, in the slicing controller, when it gets the network characteristics, it will estimate the parameters for the probability distributions for each category of the characteristics. In the simulation, it will use the "fitdist" function provided by Matlab to fit the samples to some distributions.

[1] For SINR values, it will use the "Normal" distribution for fitting.

[2] For the packet arrival rates in V2X service, it will use "Poisson" distribution for fitting.

[3] For the session arrival rates in eMBB service, it will use "Poisson" distribution for fitting.

The above process of estimating network characteristics denotes as "*ignore_online_nw_characteristics=False*", it will use the fitted network characteristics to the following reinforcement learning.

Such a fitting behavior can be disabled by a flag. If disabled, i.e., no estimation of network characteristics, some predefined probability distributions with fixed parameters will be used instead. This scheme denotes as "*ignore_online_nw_characteristics=True*", it will use the default network characteristics to the following reinforcement learning.

It should be noted that all the above simulations are based on using the default network characteristics scheme, i.e., "*ignore_online_nw_characteristics=True*". In the following simulation, we will compare the performance based on the two schemes.

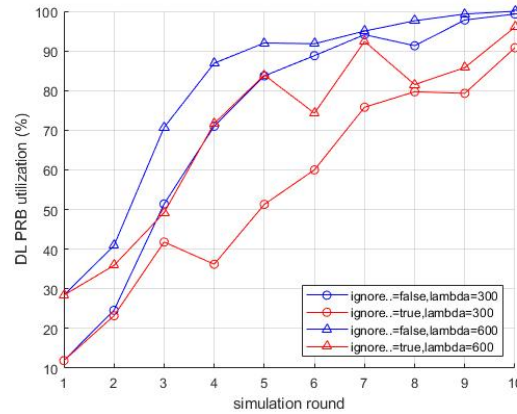## 5.4 Performance in terms of PRB Utilization



**Fig. 6.** Simulation round *VS*. PRB utilization.

In **Fig. 6**, blue lines are the ones when ignore_online_nw_characteristics=False (i.e., using the network characteristics from the online network model for parameters estimation), while the red lines are the ones when ignore_online_nw_characteristics=True (i.e., no estimation of network characteristics).
It shows that:
(1) When using the network characteristics from the online network model for parameters estimations before applying the RL algorithms, the PRB utilization is better than when ignoring the network characteristics. It demonstrates the benefits in using the fitted network characteristics for reinforcement learning.
(2) The PRB utilization will increase when more rounds of simulation are performed, because more network characteristics are available for better parameter estimation, and more samples for better RL learning algorithm.
(3) There are some fluctuations in the curves, which indicates that the simulation time is not enough. In this simulation, *sim_nr_drops_per_round=20* is too short. In order to make the following result smoother, *sim_nr_drops_per_round* should be larger than, e.g., 100, however, it will take longer simulation time.
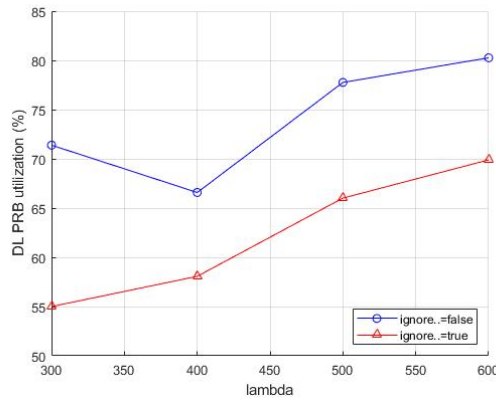
**Fig. 7.** eMBB session generation rate *VS*. PRB utilization.

**Fig. 7** shows the evolution of the DL PRB utilization when the session generation rate of the eMBB service is increasing from 300 sessions/second to 600 sessions/second.
It shows that:
(1) Using the online network characteristics for parameters estimation (i.e., *ignore_online_nw_characteristics=False*) will noticeably improve the PRB utilizations, by almost 5% - 10%.
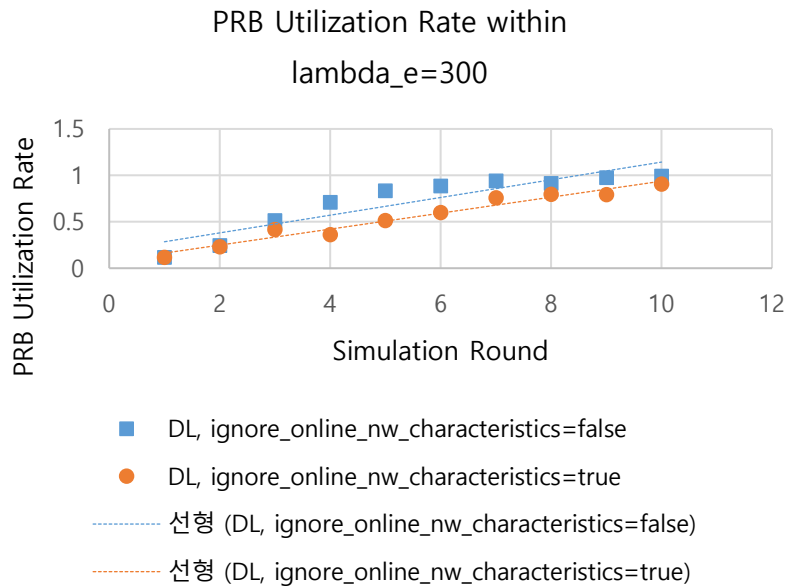(2) There are some fluctuations in the curves, which indicate that the simulation time is not enough.



**Fig. 8.** Simulation round *VS*. PRB utilization rate.

In the above **Fig. 8**, it shows the evolution of the PRB utilization rate (in a range between 0 and 1) along with the simulation round. The increase of the PRB utilization rate is due to the increase of the eMBB session number generated in an average session generation rate $\lambda_e = 300$ sessions/second according to a Poisson distribution. The blue and the orange dotted lines are the linear fitting results, which clearly demonstrate the gain from utilizing the online

network characteristics for traffic pattern and SINR measurement estimations.

## 5.5 Performance in terms of Outage Probability

Outage events consider when the demanded PRBs exceed the reserved PRBs for each of the services (V2X and eMBB). We can define outage events $O\_V2X$ and $O\_eMBB$ for each service as follows:

$$O_{V2X} = 1 \, if \, slice \, for \, V2X > slice \, for \, V2X \, is \max else \, 0$$
$$O_{eMBB} = 1 \, if \, slice \, for \, eMBB > slice \, for \, eMBB \, is \max else \, 0 \qquad (15)$$

Then overall network outage $O_{NW}$ is then given by:

$$O_{NW} = 1 \, if \, (O_{V2X} + O_{eMBB}) \geq 1 else \, 0 \qquad (16)$$

Where $O_{NW}$ calculates the overall network outage. If there is an outage event in either the V2X or the eMBB service (or both) is greater than or equal to 1, then $O_{NW}$ is set to 1, indicating a network outage. If there is no outage event in both services, then $O_{NW}$ is set to 0, indicating no network outage.
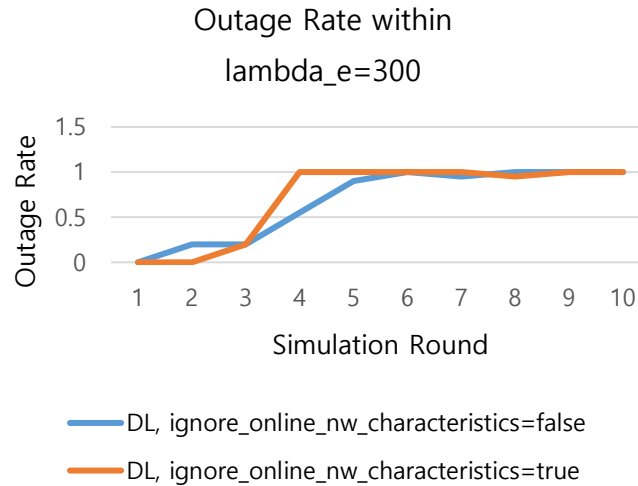


**Fig. 9.** Simulation round *VS*. Outage rate.

The above **Fig. 9** compares the outage rates whether using the online network characteristics for parameters estimations. It also demonstrates the advantage of using the online network characteristics (the blue line), which has less probability of outage on average.

## 6. Conclusions

In this paper, we have studied the problem of allocating radio resources among multiple RAN slices in the case of V2X and eMBB services that involve downlink communications. A RAN slicing strategy has been applied that is based on reinforcement learning and a low-complexity heuristic algorithm to determine the resource allocation for the eMBB and V2X slices. In particular, we evaluated the performance of real-time RL-based 5G network slicing design and traffic model distribution with and without the scheme. Based on eMBB 300 session generation rate without the scheme, the cell performance traffic model shows PRB utilization rate of 55%, but 72% with the scheme. When we increase the session generation rate from 300 to 600 for PRB utilization, it increases from 71% to 80% with the slicing strategy model. The simulation results also show that it has a less average outage probability with the scheme. The

above comparisons in terms of the network performance demonstrate the advantage of using real-time RL-based 5G network slicing design and traffic model distribution.

## Acknowledgements

## References

[1]     W. Zhou, C. J. Pawase, and K. Chang, "A Deep Reinforcement Learning based 5G-RAN Slicing Strategy for V2X Services," *Korean Institute of Communications and Information Sciences (KICS)*, pp. 944-945, Jun. 2021. Article (CrossRef Link)

[2]     ITU-R, "Minimum Technical Performance Requirements for IMT-2020 radio interface(s)," document ITU-R M.2410-0, Nov. 2017. [Online]. Available: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2410-2017-PDF-E.pdf

[3]     3GPP, "Study on New Radio (NR) Access Technology," Release 15, Jun. 2018. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3059

[4]      NGMN Alliance, "Description of Network Slicing Concept," Accessed: Apr. 5, 2019. [Online]. Available: https://www.ngmn.org/publications/description-of-network-slicing-concept.html

[5]     5G PPP Architecture Working Group, "View on 5G Architecture," Dec. 2017. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2020/02/5G-PPP-5G-Architecture-White-Paper_final.pdf

[6]     Emblasoft, "5G Network Slicing," Hammarbybacken, Stockholm, Sweden. Accessed: Apr. 5, 2019. [Online]. Available: https://emblasoft.com/use-cases/5g-slicing?campaignid=18204368630&adgroupid=147536485611&keyword=network%20slicing%20in%205g&creative=624885213779&gclid=CjwKCAjw8symBhAqEiwAaTA__MysurS2Kw2pWBOI8t5jcaFKzFA0xsY4-H9AXgG3VTvzP2KZgJ4ChhoCso0QAvD_BwE

[7]     3GPP, "Service Requirements for the 5G System; Stage 1," Release 15, Jun. 2018. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3107

[8]     3GPP, "Management of Network Slicing in Mobile Networks; Concepts, Use Cases and Requirements," document 3GPP TS 28.530 v0.3.0, Aug. 2018. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3273

[9]     3GPP, "System Architecture for the 5G System; Stage 2," Release 15, Sep. 2018. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144

[10]    Ericsson, "White Paper: 5G Radio Access Technology and Capabilities," White Paper Uen 284 23-3204 Rev C, 2016. Accessed: Apr. 5, 2019. [Online]. Available: https://gsacom.com/paper/5g-radio-access-technology-and-capabilities/

[11]    3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN)," Release 15, Dec. 2017. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2430

[12]  3GPP, "NR; Study on Vehicle-to-Everything," Release 16, Nov. 2018. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3497

[13]  H. D. R. Albonda and J. Pérez-Romero, "An Efficient RAN Slicing Strategy for a Heterogeneous Network With eMBB and V2X Services," *IEEE Access*, vol. 7, pp. 44771-44782, 2019. Article (CrossRef Link)

[14]  R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction," *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 1054-1054, Sept. 1998. Article (CrossRef Link)

[15]  M. Raza, C. Natalino Da Silva, P. Öhlen, et al., "Reinforcement Learning for Slicing in a 5G Flexible RAN," *Journal of Lightwave Technology*, vol. 37, no. 20, pp. 5161-5169, 2019. Article (CrossRef Link)

[16]   F. Mason, G. Nencioni, and A. Zanella, "Using Distributed Reinforcement Learning for Resource Orchestration in a Network Slicing Scenario," *IEEE/ACM Transactions on Networking*, vol. 31, no. 1, pp. 88-102, Feb. 2023. Article (CrossRef Link)

[17]  F. Rezazadeh, H. Chergui, L. Christofi, and C. Verikoukis, "Actor-Critic-Based Learning for Zero-touch Joint Resource and Energy Control in Network Slicing," in *Proc. of ICC 2021 - IEEE International Conference on Communications*, Montreal, QC, Canada, pp. 1-6, 2021. Article (CrossRef Link)

[18]  L. Tang, Q. Tan, Y. Shi, C. Wang, and Q. Chen, "Adaptive Virtual Resource Allocation in 5G Network Slicing Using Constrained Markov Decision Process," *IEEE Access*, vol. 6, pp. 61184-61195, 2018. Article (CrossRef Link)

[19]  H. Zhang, N. Liu, X. Chu, K. Long, A.-H. Aghvami, and V. C. M. Leung, "Network slicing based 5G and future mobile networks: Mobility, resource management, and challenges," *IEEE Commun. Mag*., vol. 55, no. 8, pp. 138-145, Aug. 2017. Article (CrossRef Link)

[20]  J. Mei, X. Wang, and K. Zheng, "Intelligent Network Slicing for V2X Services Toward 5G," *IEEE Network*, vol. 33, no. 6, pp. 196-204, Nov.-Dec. 2019. Article (CrossRef Link)

[21]  M. R. Sama, X. An, Q. Wei, and S. Beker, "Reshaping the mobile core network via function decomposition and network slicing for the 5G Era," in *Proc. of IEEE Wireless Communications and Networking Conference*, Doha, Qatar, pp. 1-7, 2016. Article (CrossRef Link)

[22]  X. Foukas, N. Nikaein, M. M. Kassem, M. K. Marina, and K. Kontovasilis, "FlexRAN: A flexible and programmable platform for software-defined radio access networks," in *Proc. of Int. Conf. Emerg. Netw. Exp. Technol.*, Irvine, CA, USA, pp. 427-441, Dec. 2016. Article (CrossRef Link)

[23]  V. Sciancalepore, K. Samdanis, X. Costa-Perez, D. Bega, M. Gramaglia, and A. Banchs, "Mobile traffic forecasting for maximizing 5G network slicing resource utilization," in *Proc. of IEEE INFOCOM*, Atlanta, GA, USA, pp. 1-9, 2017. Article (CrossRef Link)

[24]  J. Wang and J. Liu, "Secure and Reliable Slicing in 5G and Beyond Vehicular Networks," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 126-133, Feb. 2022. Article (CrossRef Link)

[25]  H. M. Soliman and A. Leon-Garcia, "QoS-Aware Frequency-Space Network Slicing and Admission Control for Virtual Wireless Networks," in *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, USA, pp. 1-6, 2016. Article (CrossRef Link)

[26]  D. T. Hoang, D. Niyato, P. Wang, A. de Domenico, and E. C. Strinati, "Optimal cross slice orchestration for 5G mobile services," *arXiv*, Dec. 2017. Article (CrossRef Link)

[27]  "Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced," ITU-R M.2135, 2009. [Online]. Available: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2135-1-2009-PDF-E.pdf

[28]  "WINNER II Channel Models," D1.1.2 V1.2, Apr. 5, 2019. [Online]. Available: https://www.cept.org/files/8339/winner2%20-%20final%20report.pdf

[29]  P. Caballero, A. Banchs, G. de Veciana, and X. Costa-Pérez, "Network slicing games: Enabling customization in multi-tenant networks," in *Proc. of IEEE INFOCOM*, Atlanta, GA, USA, pp. 1-9, 2017. Article (CrossRef Link)

[30]  M. I. Kamel, L. B. Le, and A. Girard, "LTE Wireless Network Virtualization: Dynamic Slicing via Flexible Scheduling," in *Proc. of IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, Vancouver, BC, Canada, pp. 1-5, 2014. Article (CrossRef Link)

**WiJian Zhou** received his B.S. degree in Automation from Heilongjiang University of Science and Technology in Harbin, China in 2016, and a M.S. degree in Department of Electrical and Computer Science Engineering (ECE) from Inha University, Incheon, Korea in 2023. His current research interests include control systems, robotics, artificial intelligence, machine learning, radio resource management, and vehicular communication networks.
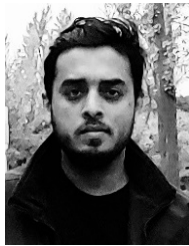
**Azharul Islam** received his B.Sc. degree in Computer Science and Software Engineering (CSSE) from the American International University of Bangladesh (AIUB) in 2015 and later earned his M.Sc. degree in Electronics Engineering (EE) at Inha University, South Korea, 2021. He is currently pursuing his Ph.D. degree at the Department of Electrical and Computer Science Engineering (ECE) at Inha University, South Korea. Since 2019, he has been serving as a research assistant at the Mobile Telecommunication Research Lab (MTRL). His research interests include 5G technologies, heterogeneous wireless networks, artificial intelligence, resource allocation management, and vehicular communication networks.

**KyungHi Chang** (Senior Member, IEEE) received the B.S. and M.S. degrees in electronics engineering from Yonsei University, Seoul, South Korea, in 1985 and 1987, respectively, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 1992. From 1989 to 1990, he was a Member of the Research Staff with the Samsung Advanced Institute of Technology (SAIT), where he was involved in digital signal processing system design. From 1992 to 2003, he was with the Electronics and Telecommunications Research Institute (ETRI) as a Principal Member of the Technical Staff, where he led the design teams involved in the WCDMA UE modem and 4G radio transmission technology (RTT). He is currently with the Electrical and Computer Engineering Department, Inha University. His research interests include radio transmission technology in 3GPP LTE and 5G NR systems, public safety, and mobile ad hoc networks (especially for UAV), cellularV2X technology, non-terrestrial networks (NTNs), network intelligence for 6G, and applications of AI technologies. He was a recipient of the LG Academic Award, in 2006, the Haedong Best Paper Award, in 2007, the IEEE ComSoc Best Paper Award, in 2008, the Haedong Academic Award, in 2010, the SKT SafeNet Best Idea Award, in 2015, the KICS Outstanding Academic Society Award, in 2018 and 2019, the MSIT Minister's Commendation and a KICS Fellow, in 2020, and the Presidential Commendation, in 2021. He is currently the Chairman of the 6G Forum, the Executive Committee, the Technology Committee for National Integrated Public Network, and the Expert Committee in SafeNet Forum. He served as the Editor-in-Chief and the Executive Director for the Journal of Korean Institute of Communications and Information Sciences (KICS), from 2010 to 2012 and in 2013, respectively, and the Vice President for KICS, from 2017 to 2018, and since 2021. He also served as the Editor for ITU-R TG8/1 IMT.MOD.