Dynamic Resource Allocation Using Reinforcement Learning for LTE-U and WiFi in the Unlicensed Spectrum

Ying-Ying Liu and Sang-Jo Yoo,
School of Information and Communications
Inha University, Incheon, Korea
yingyingliudora@163.com, sjyoo@inha.ac.kr

***Abstract*— The growing demand for spectrum resources and the limited licensed spectrum have led to widespread concern about the coexistence of LTE and WiFi in the unlicensed spectrum. Since LTE and WiFi are two different systems, many solutions have been proposed to solve their coexistence problems, including the coexistence scheme based on blank subframe allocation. This paper presents a scheme that uses Q-learning algorithm to dynamically allocate blank subframes so that both LTE-U and WiFi system can successfully coexist. To support the proposed scheme, we introduce a new LTE-U frame structure, which can not only allocate blank subframes but also reduce the LTE delay. Simulation results show that the proposed approach can effectively improve the overall system performance in terms of the utilization of spectrum.**

***Keywords*—LTE-U; WiFi; coexistence; Q-Learning; resource allocation.**

I. INTRODUCTION

With the rapid development of the mobile communication devices and smart terminals, there is an exponentially increasing demand for wireless data transmission in cellular networks in recent years [1]. According to a forecast study, by 2020, the wireless data traffic will grow 1000 times compared to that in 2010[2]. Because of the increasing spectrum scarcity in the licensed spectrum bands, a number of companies, different operators, and standardization agencies are interested in the use of LTE in the 5GHz unlicensed bands which are mainly used by WiFi currently[3].

Multiple challenges have to be faced for the coexistence of LTE and WiFi, because they employ different technologies and have evolved independently in the last decade. In order to have a fair coexistence for the two players in the unlicensed bands, many methods have been proposed, for example, LTE unlicensed (LTE-U) and Licensed Assisted Access (LAA). LTE-U uses LTE Release 10-12 carrier aggregation protocols, which does not require the listen-before-talk (LBT) mechanism. Thus, LTE-U is suitable for countries that do not require the LBT mechanism, such as US, South Korea, India and China. LAA is approved by the 3rd Generation Partnership Project (3GPP) as Release-13 and not only meets the LBT regulation, but also other regulations, e.g., channel occupancy bandwidth [4]. The goal is that both LTE-U and LAA do not affect WiFi services, including data, voice, and video services.

In order to solve the problem of spectrum scarcity, dynamic spectrum access (DSA) has also been discussed in the past few years. It is recommended to access the spectrum according to the demand of each system. There are many channel access schemes for LTE-U, including a scheme based on blank frame [5]. The blank frame scheme is inspired by the use of the almost blank subframe (ABS) for enhanced inter-cell interference coordination (eICIC) proposed in LTE Release 10 and is improved on the basis of ABS, then, the coexistence of LTE-U and WiFi is achieved. In [5], blank subframe allocation technique by LAA is described so that LTE and WiFi can apply simultaneously in the unlicensed spectrum. The LAA does not allow transmission during a silent subframe called a blank subframe, so WiFi gets more opportunities to access the channel. A Similar coexistence mechanism is considered in [6], in which LAA allocates silent gaps with a predefined duty cycle to accelerate better coexistence with WiFi. However, this mechanism cannot reach the optimal gain of LTE operation in the unlicensed band because of its discontinuous transmission. In [7], it derives a proportional fair allocation scheme to dynamically adjust the probability of the user access channel and the transmission time according to the throughput of LTE-U and WiFi. However, it is achieved by assigning equal channel time to every competing entity. This method of allocating the same channel time leads to a waste of resources and no improvement in the efficiency of WiFi.

In this paper, we propose a blank subframes dynamic adjustment technique based on reinforcement learning to facilitate the coexistence of LTE-U and WiFi in the unlicensed spectrum. To support this technique, we introduce a new LTE-U frame structure that can not only configure the blank subframes, but also reduce the LTE delay. In particular, we use Q-learning to dynamically configure blank subframes based on its learning from the environment. An advantage of our proposed approach is that resource allocation and delay quality are dynamically controlled under different LTE-U coexistence systems. The simulation results show that based on the Q-learning algorithm, the system can find the optimal allocation of the blank subframes to achieve the maximum capacity satisfying different requirements of the LTE-U and WiFi.

The remainder of the paper is organized as follows. Section II mainly explains blank subframe coexistence mechanism and provides specific details about the system model and the proposed LTE-U frame structure. In Section III, we introduce Q-learning algorithm and the proposed blank subframes dynamic adjustment algorithm. Simulation results with various parameter configurations are presented in Section IV. Finally, Section V provides concluding remarks.

II. SYSTEM MODEL

A. Blank Subframe Coexistence Mechanism

In the blank subframe scheme, the uplink or downlink subframe can be silenced and the LTE common reference signals may not be included. In addition, WiFi can reuse the LTE silent frame. If the duration and presence of LTE blank subframes are reported to WiFi during the negotiation phase, the WiFi node may be able to easily limit its transmission within a blank subframe, thereby avoiding interference with LTE. However, if the length of a LTE-U subframe is 1 millisecond, when WiFi want to use this blank subframe, it is necessary to perform carrier monitoring to ensure that other WiFi users do not use the channel. If all the WiFi users are in a retreated state at this time, the blank subframe will undoubtedly be wasted. Even if a certain WiFi user after a period of backoff time access the channel successfully, it may also happen that the reserved blank subframe at that instant is not long enough to provide the WiFi user required transmission time. This may result in loss of data for WiFi users.

In order to avoid the above problems, the blank subframe coexistence scheme have to meet the following two points. First, information sharing, which means LTE-U and WiFi systems must share their own transmission-related information. Through the transmission of information sharing, LTE-U can clear the WiFi transmission status and decide when to reserve blank subframes. Second, dynamic blank subframe adjustment, which means the number of blank subframes must be dynamically adjusted according to the traffic of LTE-U and WiFi. When the WiFi traffic is larger, the blank subframe configuration should be added. Otherwise, the blank subframe configuration should be reduced.

According to the exchange of information between the two coexisting systems, the coexistence mechanism can be divided into two groups: collaborative or non-collaborative (autonomous). Collaboration mechanisms require consistency of parameters used in each network, while non-cooperative mechanisms can be used autonomously to accelerate coexistence with other networks and devices.

In a collaborative coexistence procedure presented in [5], it is supposed that each technique will have two modes of operation: regular mode (RM) and coexistence mode (CM). RM stands for standard operation, which assumes that no other technology uses this spectrum at the same location and time. If a coexistence system is detected, two actions will be done: identification of coexisting systems and synchronization with the identified systems. Then, it enters the negotiation phase. At this stage, the systems reach an agreement on the system parameters for fair coexistence. Each system reconfigures the agreed parameters and then switches to the CM. Once in the CM, the system detects the shared resource in order to check whether the coexistence system is valid. We assume that mobile management entity (MME) can complete these operations here.

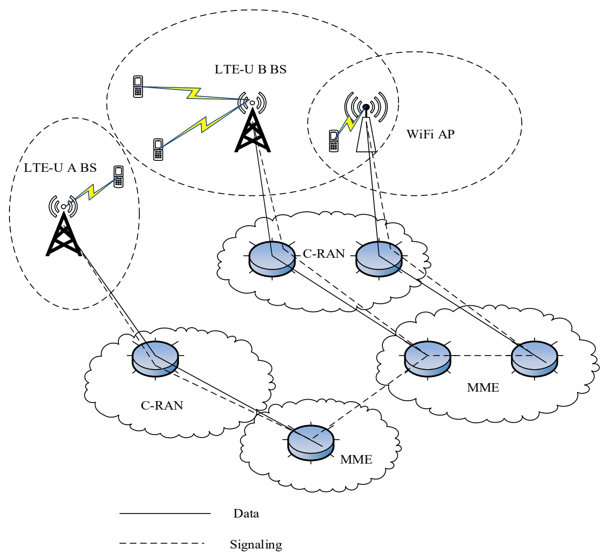

Fig. 1. Network model

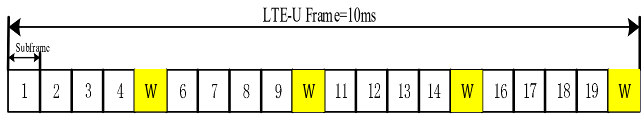

Fig. 2. An example of the LTE-U frame structure

B. Deployment Scenario

For the purpose of sharing information between different LTE-U and WiFi systems, we consider a network model as shown in Fig.1. It includes M LTE-U base stations (BSs) and N WiFi access points (APs). The LTE-U system and the WiFi system share a common unlicensed spectrum band. Both the LTE-U system and the WiFi system are connected to the cloud radio access network (C-RAN). There is a one or more virtual entities in the C-RAN. The mainly functionality of those virtual entities is to separate the data plane and the control plane, and send the control signaling received from the LTE-U and WiFi systems to the MME of the core network. Based on the information contained in the received signal, the MME obtains the traffic load of the LTE-U and WiFi. Then, the coexisting system negotiates how to allocate blank subframes based on the information contained in the received signaling messages.

C. LTE-U Frame Structure

To support the coexistence of the LTE-U system and WiFi system in the network model shown in Fig.1, we introduce a LTE-U frame structure for the channel access of both LTE-U and WiFi users. An example of the LTE-U frame structure proposed in this paper is shown in Fig.2. In this structure, the LTE-U frame still maintains the same frame length as LTE. In Fig. 2, W represents blank subframes which can be used by WiFi AP, and the rest of subframes are used by LTE-U. It is worth mentioning that the frame can be divided into several subframes due to the system requirements. We set the minimum number of subframes in a frame that contains ten subframes, which are the same as LTE. We set the maximum number of subframes in a frame that contains forty subframes, which is consistent with the 3GPP's prediction of the 5G subframe length [8]. In this way, the length of a subframe may be less than 1ms, which signifies less transmission delay. This segmentation instruction can be added to the negotiation phase of LTE-U and WiFi. In case of Fig. 2, a single one frame is divided into 20 subframes and 20% of those subframes are blank subframes.

III. Q-LEARNING DYNAMIC ALLOCATION BLANK SUBFRAME ALGORITHM

A. Q-learning Algorithm

Q learning is one of the reinforcement learning algorithms, and it does not require an environment model. Reinforcement learning is the procedure by which an agent ameliorates its behavior in an environment via experience. It receives some evaluation of its action, but does not know which action is the valid one to achieve its goal. The Q-learning algorithms works by estimating the values of state-action pairs. The value $Q(s,a)$ is defined to be the expected discounted sum of future payoffs obtained by taking action from states and following an optimal policy. After these values have been learned, the optimal action from any state is the action with the highest Q-value. After being initialized, Q-values are estimated on the basis of experience as follows:

1. Initialize the starting state s
2. Select an action a based on the optimal policy from the current state s
3. Apply the action a that will cause a receipt of an immediate payoff R
4. Observe the next state s'
5. Update $Q(s,a)$ based upon this experience as equation (1):
   $$Q(s,a) = R + \gamma Max[Q(s', all\ actions)] \qquad (1)$$
6. Go to 2.

In (1), R represents the future payoffs obtained by taking action from states. The discount factor $\gamma\ (0 \le \gamma \le 1)$ determines the value placed on the future costs. If γ is too small, learning will not rely on future costs much and immediate costs are optimized, on the contrary, learning will be overly dependent on future costs. Then, learning will be achieved in the proposed algorithm.

B. Q-learning Dynamic Allocation Blank Subframe Alogorithm

When formulating the proposed Q-learning dynamic allocation blank subframe algorithm, we consider a system consisting LTE-U base stations and WiFi APs as the agents of the multi-agent system. Based on the utility of LTE-U and WiFi, MME automatically adjusts frame allocation. We set the action $A = \{a_1, a_2, \dots, a_i, \dots, a_M\}$, where a_i equals $[a_i^1, a_i^2]$, i denotes the serial number of actions, a_i^1 represents the number of total subframes in one frame block and a_i^2 represents the portion of subframes for LTE-U. We set the state$S = \{s_1, s_2, \dots, s_j, \dots s_N\}$, where s_j equals $[s_j^1, s_j^2, s_j^3]$, j denotes the serial number of states, s_j^1 represents the function of LTE-U capacity satisfaction degree (CSD), s_j^2 denotes the function of LTE-U of delay satisfaction degree (CSD) and s_j^3 represents the function of WiFi utilization (U_{WiFi}). In particular, each state has all actions to move to the next states. We consider the action with ε-greedy policy. If the random number is less than ε, the action will be selected randomly, otherwise, the selected action is the action with the highest Q-value.

The CSD can be defined as the degree that the capacity assigned to the LTE-U reaches its expectation. If the LTE-U obtainable capacity meets or exceeds the desired capacity, CSD will be very high, otherwise, it will be low. The CSD can be calculated by equation (2)

$$CSD = \frac{C_o}{C_d}, \qquad (2)$$

where C_d represents the LTE-U desired capacity and can be obtained by MME. C_o represents the LTE-U obtainable capacity and can be computed by equation (3)

$$C_o = C \times a_i^2, \qquad (3)$$

where C denotes the total system capacity, which is calculated using the Shannon equation. a_i^2 is one of the action parameters mentioned above.

The DSD can be described as the ratio of the desired delay and the obtainable delay. That means DSD increases as the obtainable delay is gradually close to the desired delay. The DSD is calculated as shown below (4)

$$DSD = \frac{D_d}{D_o}, \qquad (4)$$

where D_d represents the LTE-U desired average delay and can be set. D_o represents the LTE-U obtainable average delay and can be computed by formula (5)

$$D_o = \frac{5}{a_i^1 \times a_i^2} \text{ millisecond}, \qquad (5)$$

this formula is simplified by the following formula (6)

$$D_o = \frac{1}{2} \times \frac{10}{a_i^1} \times \frac{a_i^1}{a_i^1 \times a_i^2} \text{ millisecond}, \qquad (6)$$

in this formula,$10/a_i^1$ is the duration of a subframe, where the frame time is 10ms. $a_i^1/(a_i^1 \times a_i^2)$ is the number of blank subframes between two LTE-U subframes. Then, their product divided by 2 is the obtainable average delay. We assume that LTE-U subframes are evenly distributed in a frame time.

The U_{WiFi} can be defined as the ratio of the WiFi generated traffic and WiFi obtainable capacity. In order to have a better coexistence with LTE-U, we set the utilization of WiFi is lower than one. The U_{WiFi} is calculated as shown below,

$$U_{WiFi} = min[\frac{T_w}{C_w}, 1], \qquad (7)$$

where T_w represents the WiFi generated traffic and can be obtained by MME. C_w represents the WiFi obtainable capacity is given by,

$$C_w = C - C_o, \qquad (8)$$

it is derived from the total capacity minus the capacity used by LTE-U.

The overall system contains the LTE-U BSs and WiFi APs, so, the utility function for the overall system is computed by

$$R = \alpha U_l + (1-\alpha) U_w, \qquad (9)$$

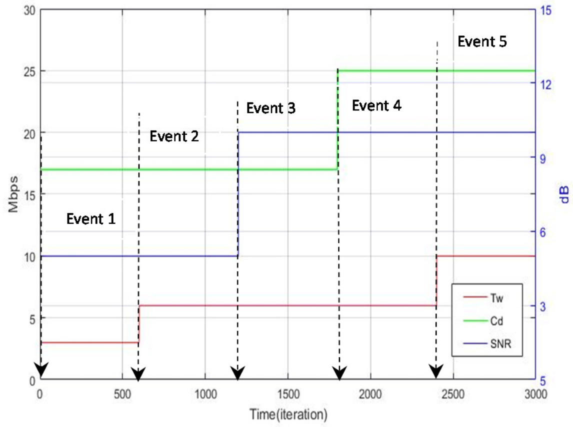

Fig.3. (a) WiFi generated traffic T_w (b) SNR (c) LTE-U desired capacity C_d

where, $0 < \alpha \leq 1$, U_l, U_w are the utilities of LTE-U and WiFi respectively. The utility of LTE-U in this paper is determined by both capacity and delay and it is calculated as shown below

$$U_l = \beta min[CSD, T_l^C] + (1-\beta)min[DSD,\ T_l^D]\ , \quad (10)$$

where, $0 < \beta \leq 1$, T_l^C, T_l^D are the target of LTE-U CSD and DSD respectively, which operator can set. The utility of WiFi can be derived from WiFi utilization and is given by

$$U_w = \begin{cases} 1 - 2 \times (T_W^{down} - U_{WiFi}) & U_{WiFi} < T_W^{down} \\ 1 & T_W^{down} < U_{WiFi} < T_w^{up} \\ \max[0, 1 - 2 \times (U_{WiFi} - T_w^{up})] & U_{WiFi} > T_w^{up} \end{cases} \quad (11)$$

where, T_W^{down}, T_w^{up} are the target of WiFi channel utilization lower limit value and upper limit value respectively.

After the modification of the basic formula of the Q-learning, the Q-learning update equation in our algorithm is given in equation (11).

$$Q(s,a) = (1-\alpha)Q(s,a) + \alpha\{R + \gamma max_{a'} Q(s',a')\}, \quad (12)$$

where, s' denotes the next state, a' represents the optimal action from the next state and it is the one with the highest Q-value. α $(0 \leq \alpha \leq 1)$ is the learning rate, which commands how quickly the learning can occur. If α is too small, it will take long time to accomplish the learning process, on the other hand, if it is too high, algorithm may not converge. γ is still the discount factor.

IV. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed algorithm. In simulations, we consider one LTE-U BS and one WiFi AP and they share a same unlicensed spectrum. We set 12 actions and 27 states that are shown in the Table I and Table II, respectively. The requirements of the LTE-U and WiFi data changes randomly. In order to investigate the effects of the changes in LTE-U and WiFi traffic demands, we made five adjustments to the three parameters that are SNR, LTE-U desired capacity C_d and WiFi generated traffic T_w as shown in Fig.3. For simulations, we set $\alpha = 0.5, \gamma = 0.9, \varepsilon = 0.2$. All the configuration parameters used in simulations are given in Table III.

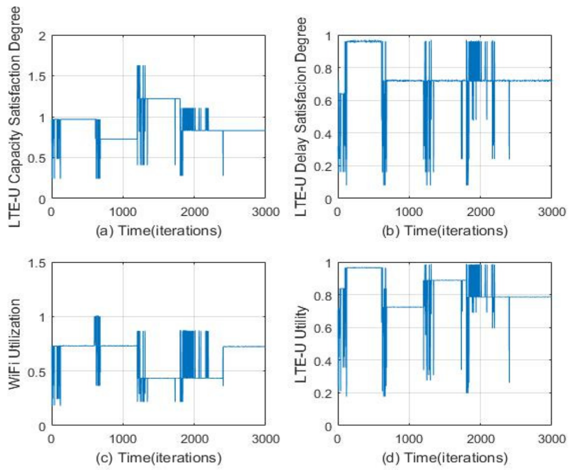

Fig.4. (a) LTE-U Capacity Satisfaction Degree. (b) LTE-U Delay Satisfaction Degree. (c) WiFi utilization. (d) LTE-U utility

TABLE I: ACTIONS

A	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}	a_{12}
a_i^1	10	10	10	10	20	20	20	20	30	30	30	30
a_i^2	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8

TABLE II: STATES

	s_j		
	s_j^1	s_j^2	s_j^3
$0 \leq CSD(DSD)(U_{WiFi}) \leq 0.5$	1	1	1
$0.5 < CSD(DSD)(U_{WiFi}) \leq 1$	2	2	2
$CSD(DSD)(U_{WiFi}) > 1$	3	3	3

TABLE III: SIMULATION PARAMETERS

Parameter	Value
Bandwith	10MHz
LTE-U desired delay	0.2ms
Target of LTE-U capacity utilization	1
Target of LTE-U delay utilization	1
Target of WiFi upper utilization	0.7
Target of WiFi lower utilization	0.5

The overall system performance varies with the three changed parameters. In each configuration of the parameters, through the Q-learning algorithm, it can automatically find the best distribution about the blank subframes for the maximum total capacity. The simulation results about the LTE-U CSD, LTE-U DSD, U_{WiFi} and LTE-U utility are shown in Fig.4. Fig.5 shows the simulation results about the WiFi utility. Fig.6 captures the final utility about the overall utility.

According to the initial values of the three parameters: $SNR = 5dB$, $C_d = 17Mbps$ and $T_w = 3Mbps$, it can be seen from these graphs that the system finds the optimal allocation and segmentation of LTE-U frame based on the Q-learning algorithm. Then, the final values of LTE-U CSD and DSD are close to their target values and the final result of U_{WiFi} is also close to the expected range. So the utility of LTE-U and WiFi get a relative good condition respectively, at the same time, the max total utility of the overall system are also reached.

The second event causes the total utility decrease, because larger WiFi generated traffic leads the condition that the portion

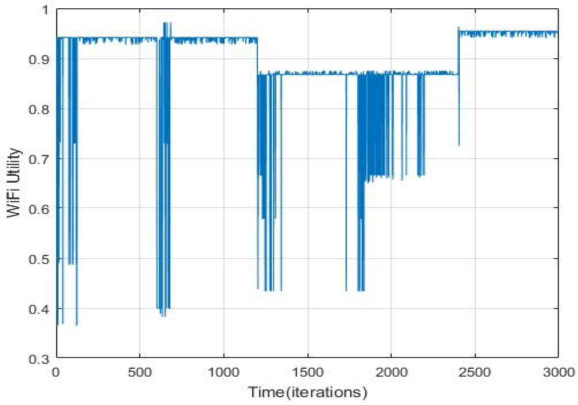

Fig.5. WiFi utility

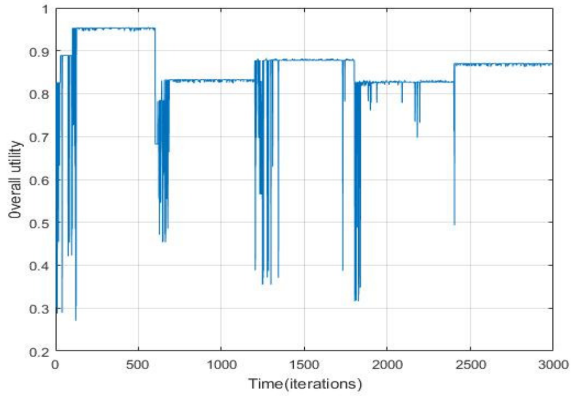

Fig. 6. Overall utility

of the blank subframes allocated to WiFi have increased, which means WiFi have more obtainable capacity. So the utility of WiFi does not change. SNR keeps the same value with the first event, which means that the total capacity has not changed. These result in the reduction of the capacity for LTE-U, however, the LTE-U desired capacity have not changed, so the LTE-U CSD decrease. In addition, the subframe ratio allocated to LTE-U is reduced, resulting in an increase of the LTE-U obtainable delay, so the LTE-U DSD decrease. Both LTE-U CSD and LTE-U DSD are reduced, giving rise to the decline of the LTE-U utility. Based on the utility of the LTE-U and WiFi, the total utility presents the final decrease.

The third event results in the total utility with a slight raise due to the increase of SNR. The growth SNR means the total capacity increase. Since the previous LTE-U utility was low, there are more capacity can be used by LTE-U at this time. So the LTE-U CSD get satisfied and have increased. Owing to the LTE-U CSD is so high at this time, for a fair coexistence with WiFi, the splits of LTE-U frame have not changed. So the LTE-U DSD value is maintained and the LTE-U utility shows a larger increase. These lead to the WiFi obtainable capacity shows a little less, so the utility of WiFi was slightly reduced. A slight reduction of WiFi utility and a significant increase of LTE-U utility lead to the overall utility behaves a slight increase and reaches to the maximum at this condition.

The fourth event is that the expected capacity of LTE-U increase. As the previous LTE-U gained too much capacity, the LTE-U CSD value exceeds one. Currently, the desired capacity increase and the obtainable capacity is still so much, as a result, the ratio of them is close to the target value of the LTE-U CSD. Therefore, the portion of the subframes and the splits of the frame do not need to alter, then, the LTE-U DSD remain unchanged again. Eventually, the reduction of the LTE-U utility and the invariable WiFi utility cause the total utility cut back.

The fifth event makes WiFi generated traffic increase again. As can be seen from the figure, the WiFi utilization is a little lower in the last configuration. With the increase of the WiFi generated traffic, the obtainable capacity do not need to change, which can also lead to the WiFi utilization shows a raise. So the subframe allocation condition remain constant. The increasing WiFi utility and the changeless LTE-U utility together result in the increase of the overall utility.

V. CONCLUSIONS

In this paper, we have proposed a Q-learning dynamic allocation blank subframes, so as to effectively coexist with WiFi systems in the unlicensed spectrum. First, we proposed a new LTE-U frame structure. Second, we evaluated WiFi and LTE-U performance with a series of changes. Then, the overall system performance with the proposed Q-learning approach is evaluated. Simulation results show that the proposed Q-learning dynamic allocation blank subframes approach can effectively find the best allocation way for the coexisting system and enhance the overall utility performance.

ACKNOWLEDGMENT

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2016-8501-16-1019) supervised by the IITP (Institute for Information & communications Technology Promotion). This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and future Planning (2017R1A2B4003512).

REFERENCES

[1] R. Yin, G. Yu, A. Maaref, and G. Ye Li, “LBT-Based Adaptive Channel Access for LTE-U Systems” , *IEEE Transactions on Wireless Communications*. vol. 15, no. 10, pp.6585-6897, Oct. 2016.

[2] NSN Whitepaper,“Enhance mobile networks to deliver 1000 times more capacity by 2020,”J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73,2013.

[3] T. Nihtilä *et al.*, “System performance of LTE and IEEE 802.11 coexisting on a shared frequency band,” in *Proc. IEEE WCNC*, pp. 1038–1043, Apr. 2013.

[4] Xuyu Wang, Shiwen Mao，Michelle X.Gong “A Survey of LTE WIFI coexisrence in unlicensed bands”, in GetMoblie: Mobile Computing and Communications, vol.20, no.3, pp.17-23, July 2016.

[5] E. P. L. Almeida, A. M. Cavalcante, R. C. D. Paiva, F. S. Chaves, F. M.Abinader Jr., R. D. Vieira, S. Choudhury, E. Tuomaala, and K. Doppler,“Enabling LTE/WiFi coexistence by LTE blank subframe allocation,”in *Proc. IEEE ICC*, pp.5083-5088, Jun. *2013.*

[6] CableLabs, “Cable Labs perspective on LTE-U Coexistence with Wi-Fi and Operational Modes for LTE-U,” 3GPP RAN1 standard contribution - RWS-140004, Sophia Antipolis, France, Jun. 2014.

[7] Cano C, Leith D J. “Coexistence of WiFi and LTE in unlicensed bands: A proportional fair allocation scheme” in *Proc. IEEE ICC*, pp.2288-2293, Jun. 2015.

[8] 3GPP R1-162167. 3GPP TSG RAN WG1 Meeting #84bis – Overview of 5G Framestructure.