Computer Communications 150 (2020) 498–505

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

(ReLBT): A Reinforcement learning-enabled listen before talk mechanism for LTE-LAA and Wi-Fi coexistence in IoT☆

R. Ali [a], B. Kim [b], S.W. Kim [c], H.S. Kim [a,*], F. Ishmanov [d,*]

[a] School of Intelligent Mechatronics Engineering, Sejong University, Seoul, South Korea
[b] Department of Computer and Information Communication Engineering, Hongik University, Sejong, 339-701, South Korea
[c] Department of Information and Communication Engineering, Yeungnam University, Gyeongsan-si, 38541, South Korea
[d] Department of Electronics and Communication Engineering, Kwangwoon University, South Korea

ARTICLE INFO

Keywords:
Wi-Fi
Unlicensed band
LTE-LAA
LTE-LAA WiFi coexistence
Listen before talk (LBT)

ABSTRACT

The emergence of Internet of Things (IoT) has increased number of connected devices and consequently transmitted traffic over the Internet. In this regard, Long Term Evolution (LTE) is growing its utilization in unlicensed spectrum as well, and Licensed Assisted Access (LAA) technology is one of the examples. However, unlicensed spectrum is already occupied by other wireless technologies, such as Wi-Fi. The diverse and dissimilar physical layer and medium access control (MAC) layer configurations of LTE-LAA and Wi-Fi lead to coexistence challenges in the network. Currently, LTE-LAA uses a listen-before-talk (LBT) mechanism, and Wi-Fi uses a carrier sense multiple access with collision avoidance (CSMA/CA) as a channel access mechanism. LBT and CSMA/CA are moderately similar channel access mechanisms. However, there is an efficient coexistence issue when these two technologies coexist. Therefore, this paper proposes a Reinforcement Learning-enabled LBT (ReLBT) mechanism for efficient coexistence of LTE-LAA and Wi-Fi scenarios. Specifically, ReLBT utilizes a channel collision probability as a reward function to optimize its channel access parameters. Simulation results show that the proposed ReLBT mechanism efficiently enhances the coexistence of LTE-LAA and Wi-Fi as compared to the LBT, thus improves fairness performance.

1. Introduction

Next-generation wireless networks of 5th Generation (5G) technology are expected to support thousands fold of increased capacity, and at least hundred billion connected devices with tens of Gbps per-user throughput and less than one-millisecond latency [2]. Another upcoming technology, the Internet of Things (IoT) has exponentially increased wireless communication due to massively connected sensors and actuators [3]. Third Generation Partnership Project (3GPP) proposed to extend licensed Long-Term Evolution (LTE) wireless system to unlicensed spectrum [4,5] to effectively support these widely connected devices. For this purpose, LTE-licensed assisted access (LTE-LAA) is introduced in LTE Release 13, which uses 5 GHz unlicensed band to coexist with Wireless-Fidelity (Wi-Fi) wireless local area networks (WLANs) [1,6]. However, other Industrial, Scientific and Medical (ISM) public wireless communication technologies, such as IEEE 802.11 (also known as Wi-Fi or WLAN), ZigBee, and Bluetooth, etc. already occupy the unlicensed spectrum [4]. Thus, LTE-LAA and ISM spectrum will face a considerable interference challenge due to massive channel contention. While, the use of LTE-LAA in an unlicensed spectrum improves the capacity and accomplishes a smooth user quality of experience (QoE), the issues of allowing diverse networks to use a jointly shared spectrum need to be deemed. One important issue is the coordination and management of an interference among the different coexisting technologies [7]. Fig. 1 shows the interference of unlicensed wireless communication devices from LTE-LAA and Wi-Fi.

Currently, medium access control (MAC) layer channel access in Wi-Fi primarily focuses on maximization of the channel utilization using fair MAC layer resource allocation (MAC-RA) schemes [8]. MAC-RA scheme uses a distributed coordination function (DCF)-based Carrier sense multiple access with collision avoidance (CSMA/CA) mechanism to reasonably access the wireless spectrum [8]. On the other hand, standard LTE uses continuous data transmission with a minimum gap. Therefore, to withstand the fair coexistence challenge between LTE-LAA and Wi-Fi, a Listen-Before-Talk (LBT) mechanism is adapted in LTE-LAA. LBT performs CSMA/CA like medium access procedure to access the wireless channel [9,10]. CSMA/CA considers collision if the

☆ The present research has been conducted by the Research Grant of Kwangwoon University, South Korea in 2019. This research was also supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (No.2019R1A4A1023746, No. 2019R1F1A1060799).

* Corresponding authors.
E-mail addresses: rashidali@sejong.ac.kr (R. Ali), jsnbs@hongik.ac.kr (B. Kim), swon@yn.ac.kr (S.W. Kim), hyungkim@sejong.ac.kr (H.S. Kim), farruh.uzb@gmail.com (F. Ishmanov).

https://doi.org/10.1016/j.comcom.2019.11.055
Received 6 October 2019; Received in revised form 20 November 2019; Accepted 30 November 2019
Available online 13 December 2019

Fig. 1. LTE-LAA and Wi-Fi coexistence deployment and interference scenario in the IoT system.

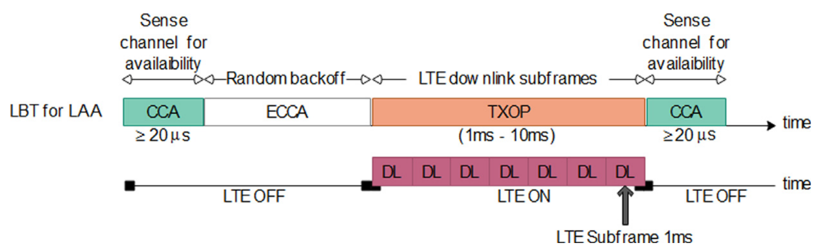


Fig. 2. LBT mechanism of LTE-LAA [1].

sender does not receive an acknowledgment (ACK) control message. However, there is no such frame in LBT. LBT examines collision based on a hybrid automatic repeat request (HARQ) feedback of the current transmission opportunity (TXOP) [11,12]. HARQ shows the number of negative acknowledgments (NACK) from the current TXOP. An 80% of NACK response in HARQ feedback is considered as the collision in the network [13].

LTE-LAA schedules multiple devices in a single transmission frame, where it is usually hard to meet an 80% NACK threshold. According to 3GPP, collision with less than 80% NACK threshold is also neglected [14]. Moreover, due to the integral latencies introduced by the LTE continues transmission protocol stack, the HARQ feedback associated with a specific sub-frame is received at least 4 ms after its transmission time. Therefore, 3GPP proposes only to consider the collisions detected during the first sub-frame of a TXOP [14]. Thus, collisions from the rest of the TXOPs are also ignored.

In spite of adopting CSMA/CA like protocol in LTE-LAA and Wi-Fi coexistence (that is LBT), the performance of Wi-Fi highly depends on the configurations of LBT channel access parameters [15]. Therefore, in this paper, we propose to intelligently adjust channel access parameters with the help of reinforcement learning (RL). RL is inspired by behaviorist psychology, which allows an agent/device to learn the environment by its interactions and to select an optimum strategy for taking actions [16]. One of the substantial characteristics of RL is that it explicitly imitates the entire problem of learners interacting with an uncertain environment and being directed to its goal [17]. This goal-oriented learner can be the smallest device of a broader behavioral context, such as an LTE-LAA user-equipment (UE), seeking to maximize its performance in terms of fair coexistence with Wi-Fi stations (STAs). Therefore, we propose an RL-enabled LBT (ReLBT) for LTE-LAA and Wi-Fi coexistence in IoT systems. The proposed ReLBT mechanism intelligently optimizes the MAC-RA channel access parameters for LBT. Following are the contributions of our proposed mechanism,

- Our proposed ReLBT mechanism considers channel observation-based collision probability [1] as a reward of its transmission attempts.
- Instead of waiting for 80% HARQ feedback, it exploits the accumulated reward for scaling-up and scaling-down of contention parameters.
- During the exploration phase, ReLBT utilizes a channel observation-based scaled backoff (COSB) [18,19] mechanism, to scale-up and scale-down its contention parameters.
- In particular, the proposed ReLBT mechanism finds optimal actions by interaction and observation from the environment through the RL approach.

The rest of this paper is organized as follows. In the next section, we describe the currently implemented LBT mechanism of LTE-LAA. In Section 3, we discuss the proposed ReLBT mechanism. Section 4 evaluates the performance of the proposed mechanism using an event-driven NS3 simulator with a densely deployed LTE-LAA and Wi-Fi coexistence scenario. Finally, we make conclusions and present our future research in Section 5.

2. Listen Before Talk (LBT)

As mentioned earlier, 3GPP evaluated multiple preferences of LBT before finalizing it to coexist with CSMA/CA. Eventually, the selected LBT flavor is the one that allows channel access most similar to the currently implemented CSMA/CA in Wi-Fi, which uses a binary exponential backoff (BEB) [8] for channel contentions. Specifically, BEB is a DCF-based mechanism and uses a random backoff mechanism to contend among the competing Wi-Fi STAs. In BEB, a random value is selected from a contention window (CW) to observe the channel before transmission [20,21]. An STA exponentially increases the selected CW on collision and resets back to its initial value once transmitted successfully. BEB imposes limits on the TXOP before contention occurs

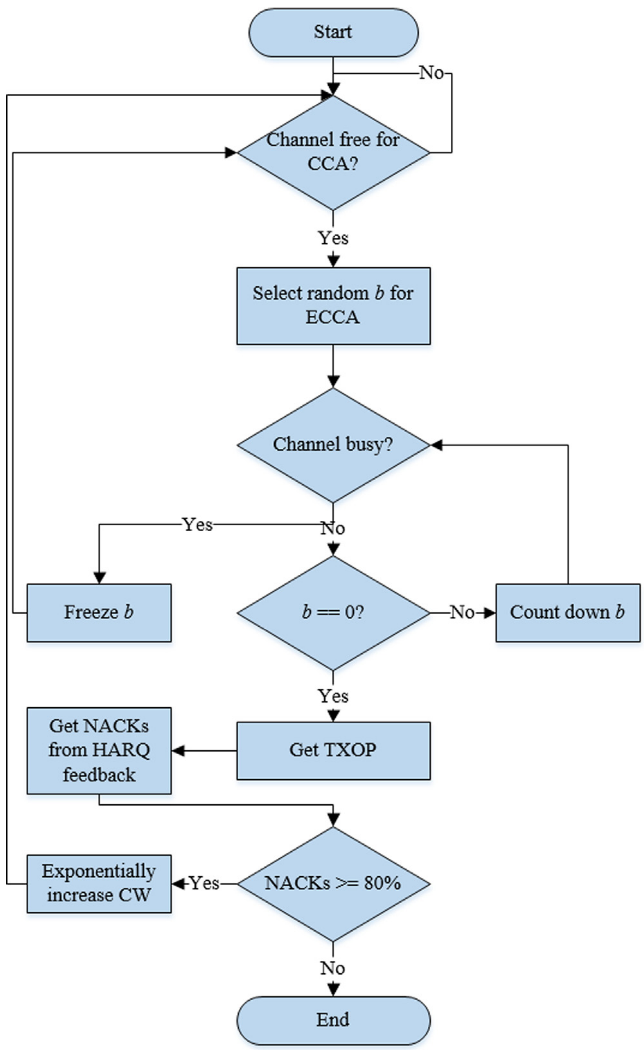


Fig. 3. The flowchart of the LBT mechanism in LTE-LAA.

again, that is, an STA needs to contend for the channel for each data frame.

On the other hand, LTE-LAA UE performs a clear channel assessment (CCA) before LBT implementation. An LTE-LAA UE is aiming to transmit data frames, must observe the channel for a CCA idle period as an initial deferral period. The observing UE performs another deferral-based, known as extended CCA (ECCA) after the successful CCA, as shown in Fig. 2. In ECCA, the UE selects a random backoff value (b), where b describes the number of observed idle slots that need to be sensed before a TXOP, and each of a size of CCA slot-time (σ) [13,22]. Initially, the random backoff value b is selected from a range of $[0, CW_{min}-1]$, where CW_{min} is the minimum backoff contention window. The size of CW is exponentially increased upon 80% NACKs from the current TXOP and is reset to the CW_{min} if the number of NACKs is detected less than 80% [14]. Fig. 3 shows the flowchart of the LTE-LAA LBT mechanism. In the figure, it is shown that a UE increases the size of its CW if it finds NACKs higher than 80%.

3. Reinforcement learning-enabled LBT (ReLBT)

This section discusses our proposed ReLBT mechanism. Q learning (QL) is one of the RL models, which significantly reflects the whole problem where a learner interacts with an uncertain environment for the purposes of its performance optimization [16]. A specific goal-oriented learner can be a wireless UE in an LTE-LAA environment seeking to maximize its performance in terms of fair coexistence along with Wi-Fi. In ReLBT, the QL-based selection of contention parameters leads to a reduced channel collision and enhances the fairer channel access. One of the major contributions of this paper is the ability to adjust the backoff parameters, such as current CW selection (CW_{cur}) dynamically based on the LTE-LAA and Wi-Fi users' density. Our proposed mechanism requires only a few modifications to the state-of-the-art LBT mechanism in LTE-LAA and maintains full compatibility.

3.1. Replacement of HARQ-based collision detection

Before the implementation of ReLBT, we replace the standard HARQ-based collision detection mechanism of LBT with a more realistic channel observation-based collision probability (p_{obs}) [1] mechanism. In this replaced mechanism, the backoff parameters are scaled based on p_{obs} instead of 80% NACKs in current HARQ. Besides, instead of the exponential increase of CW and reset back to minimum CW_{min} as of LBT, ReLBT scales-up, and scales-down the backoff CW based on p_{obs}.

In the proposed ReLBT mechanism, after the communication medium has been idle for a CCA, an LTE-LAA UE competing for a channel for transmission proceeds to the ECCA procedure by selecting a random backoff value b. ReLBT discretizes the time immediately following an idle CCA into observation time slots (α). The duration of α slot is either a constant slot-time (σ) during an idle period or a variable busy (successful or collided transmission by other devices in the network) period. b decrements by one of the channels is sensed as idle for a period of σ. A TXOP is availed only at the beginning of the slot time when b reaches zero. Also, the UE freezes b and continues sensing the channel for idle CCA if the channel becomes busy. Later, if the channel is again detected to be clear for CCA, b is resumed. Each UE in the network can proficiently measure channel observation-based conditional collision probability p_{obs}. This measured p_{obs} may call a pseudo collision probability as it seems that the estimation of p_{obs} requires each UE to observe and count the number of failed transmissions (NACKs) and divide it by the total number of transmission attempts. However, a more realistic observation for p_{obs} can be achieved if busy and idle time durations of the ECCA procedure are also counted. ReLBT updates p_{obs} at every ECCA backoff contention stage by counting the number of NACKs (S_{nack}) in recent TXOP and the number of busy slots (S_b) during observed time-slot (α).

All observed α time-slots (busy and idle) are represented by B_{obs}, which is $B_{obs} = b + S_b$ between two consecutive ECCA backoff stages. A tagged UE updates p_{obs} from B_{obs} of ECCA backoff stage as follows:

$$p_{obs} = \frac{(S_b + S_{nack})}{(S_{nack} + B_{obs})}, \tag{1}$$

where $S_b = \sum_{k=0}^{B_{obs}-1} S_k$, and for an observation time slot k, $S_k = 0$ if α is empty (idle), while $S_k = 1$ if α is busy due to other device transmissions. The formulation of channel observation-based reasonable collision probability has a prominence to the ReLBT mechanism. ReLBT updates its CW based on the observed practical collision probability, which leads to a more adaptive contention procedure. Thus, brings a fair share between the two coexisting technologies.

As mentioned earlier, in a HARQ feedback-based collision detection mechanism, scaling of contention parameters is based on observed NACKs from the recent TXOP. However, in our proposed replacement, a tagged UE detects a collision and scales current ECCA contention parameters if it finds $p_{obs} > 0$. This process indicates that even if there is no NACK received in the current feedback, still contention parameters can be scaled due to busy slots during observation. Besides, unlike the existing exponential increase for unsuccessful and resetting back to a minimum value of CW in LBT, the ReLBT operates as, scaling-up and scaling-down of the CW. ReLBT scales-up the CW if $p_{obs} > 0$ (that is, there exist busy slots or/and NACKs), and scales-down the CW if

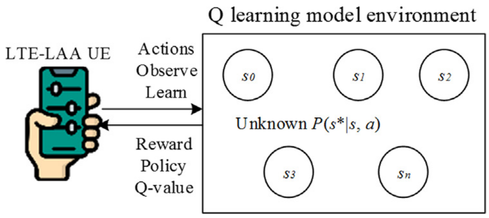


Fig. 4. Q learning-based ReLBT with its elements.

$p_{obs} = 0$ (that is no busy slots and NACKs) [1]. The scaling-up and scaling-down of the ECCA CW operate as follows:

$$CW_{cur} = \begin{cases} min[2 \times CW_{pre} \times \omega^{p_{obs}}, CW_{max}], \forall p_{obs} > 0 \\ max[\frac{CW_{pre} \times \omega^{p_{obs}}}{2}, CW_{min}], \forall p_{obs} = 0, \end{cases} \quad (2)$$

where CW_{cur} is current scaled-up/scaled-down ECCA CW from a previous CW_{pre}. The ω is a constant design parameter to control the adaptive size of the ECCA CW according to the observed p_{obs} and is expressed as $\omega = CW_{min}$.

3.2. Proposed ReLBT mechanism

A QL model consists of a learner (that is an LTE-LAA UE), an environment (that is an LTE-LAA and Wi-Fi coexistence), a policy (that is scale-up and scale-down of CW), a reward (that is p_{obs}), and a Q-value function (an accumulated reward) [16]. An LTE-LAA UE's behavior and learning at a given time depends on the policy it follows. Whereas, a policy is a rule to decide perspective actions to map the perceived states of its environment. A reward is the core objective of a UE, which is a quantitative value determined by the situation at each step. In a QL-based model, a learner's only goal is to maximize the accumulated reward over the long run. While the reward is immediate quantitative value for any single action in a specific state, the Q-value denotes the accumulated reward attained at that state. It is possible that a state always yields a low immediate reward but still has a high Q-value because of continuously followed by other states that produce high rewards.

The proposed ReLBT mechanism contains a set of states (S) (backoff stages in an ECCA), where an intelligent LTE-LAA UE has the ability to act from a set of actions (A) as, $A = \{$increase CW if $p_{obs} > 0$, decrease CW if $p_{obs} = 0\}$. By performing an action a following a policy π in a particular state s, a UE collects a reward r, that is $R(s,a)$ to exploit the collective reward $Q(s,a)$, which is known as a Q-value function. Fig. 4 depicts the model environment with its elements for the proposed ReLBT mechanism.

Let $S = \{0, 1, 2, \ldots, m\}$ denotes a finite set of m possible states of the environment and let $A = \{0, 1\}$ represents a finite set of permissible actions allowed to an LTE-LAA UE, where zero indicates decrement, and one indicates increment. At time slot t, a UE observes the current state s_t, that is $s_t = s \in S$ and takes action a_t, that is $a_t = a \in A$ based on policy π_t. A default policy of a UE in ReLBT is to increment its state for collision and decrement for successful transmission. Thus, an action a_t changes the environmental state from s_t to $s_{t+1} = s' \in S$ according to,

$$\pi(a \mid s) = \begin{cases} s' = s + 1, \forall p_{obs} > 0 \\ s' = s - 1, \forall p_{obs} = 0. \end{cases} \quad (3)$$

In a QL model, $Q(s,a)$ estimates the cumulative reward and is updated as follows,

$$Q(s,a) = (1-\gamma) \times Q(s,a) + \gamma \times \Delta Q(s,a), \quad (4)$$

where γ is the learning-rate defined as $0 < \gamma < 1$. A learner learns quickly based on the improved learning estimate $\Delta Q(s,a)$, which is expressed as,

$$\Delta Q(s,a) = \{R(s,a) + \beta \times max'_a Q(s',a')\} - Q(s,a). \quad (5)$$

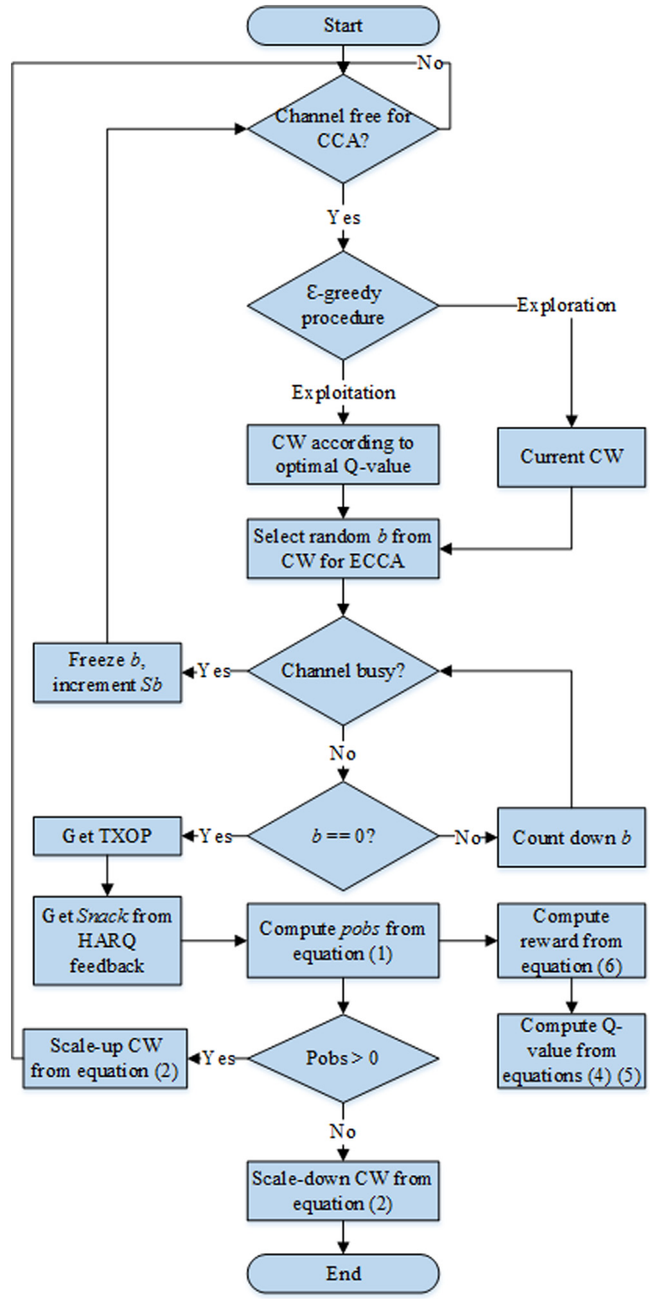


Fig. 5. The flowchart of the proposed ReLBT mechanism.

The reward may easily get unbounded, thus a discounted reward factor β, where $0 < \beta < 1$ is used. In Eq. (5), the $max'_a Q(s',a')$ represents the best estimated Q-value for the prospective state–action pair. In the long run, $Q(s,a)$ converges to the optimal Q-value, that is $\lim_{t\to\infty} Q(s,a) = Q^*(s,a)$. A heuristic policy for action selection can be to exploit the actions with the maximum measured Q-value. However, QL requires a frequent exploration to update learning outcomes dynamically. In QL-based algorithms, one of the methods for exploration and exploitation is known as the ε-greedy method [16]. In this method, a greedy action policy (exploitation) $\pi^*(a^* \mid s) = argmax_a Q(s,a)$ is performed with a probability of ε, where $argmax_a$ represents that $Q(s,a)$ is exploited concerning a. Continuous exploitation leads to the maximization of the instant reward in a greedy manner. A modest substitute is to exploit more often. However, the LTE-LAA UE explores all the legal actions independent of optimal policy (π^*) with a probability of $1-\varepsilon$ (known

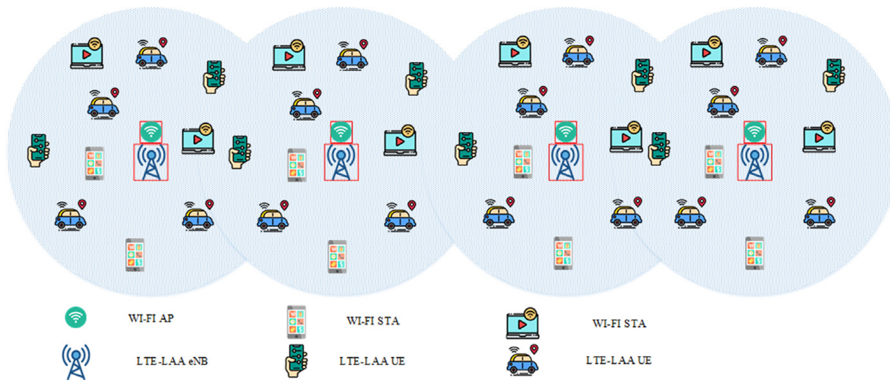


Fig. 6. LTE-LAA and Wi-Fi coexistence scenario for simulation.

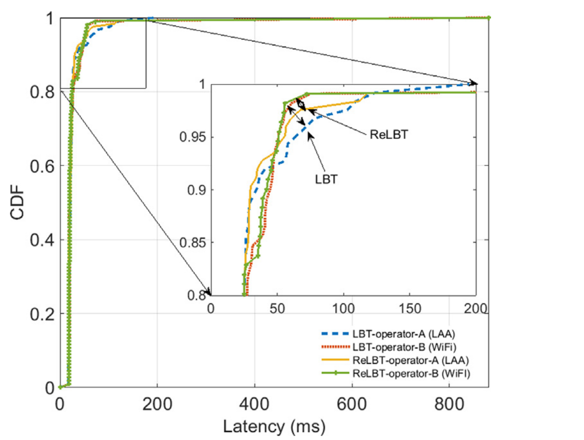


Fig. 7. Latency (ms) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 1.5$ with five UEs/STAs per cell/operator.

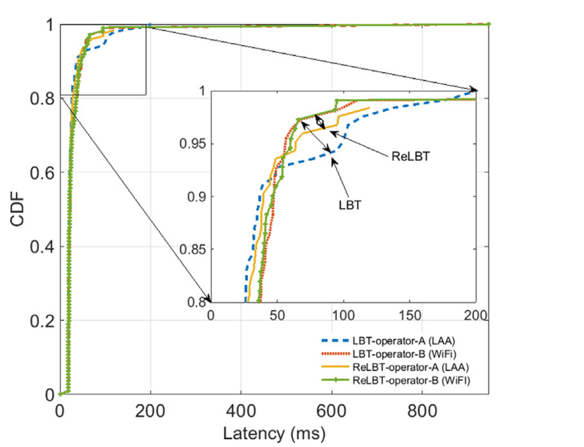


Fig. 8. Latency (ms) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 2.5$ with five UEs/STAs per cell/operator.

as exploration). In the ε-greedy method, over time every action (step) of the learner guarantees the convergence of $Q(s, a)$. A UE exploits to optimize its performance and explores to learn the changes in the LTE-LAA environment.

As the objective of our proposed ReLBT mechanism is to optimize the fair coexistence of LTE-LAA and W-Fi, which is achieved by reducing the unnecessary collisions in the environment. Therefore, we express the reward of the actions performed at any specific state in relation to the channel observation-based collision probability p_{obs}. Therefore, the reward given by an action a_t taken at state s_t in a time

Table 1
List of abbreviations and acronyms used in this paper.

Acronyms	Full description
3GPP	Third Generation Partnership Project
5G	5th Generation
ACK	Acknowledgment
BEB	Binary Exponential Backoff
CCA	Clear Channel Assessment
CDF	Cumulative Distribution Function
COSB	Channel Observation-based Scaled Backoff
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CW	Contention Window
DCF	Distributed Coordination Function
ECCA	Extended CCA
eNB	Evolved Node B
HARQ	Hybrid Automatic Repeat Request
ISM	Industrial, Scientific and Medical
LBT	Listen-Before-Talk
LTE	Long Term Evolution
LTE-LAA	LTE-Licensed Assisted Access
MAC	Medium Access Control
MAC-RA	MAC layer Resource Allocation
NACK	Negative Acknowledgments
QoE	Quality of Experience
QL	Q Learning
ReLBT	RL-enabled LBT
RL	Reinforcement Learning
TXOP	Transmission Opportunity
UE	User Equipment
Wi-Fi	Wireless-Fidelity

slot t is expressed as,

$$R_t(s_t, a_t) = 1 - p_{obs}. \tag{6}$$

The Eq. (6) shows the level of satisfaction (pleasure) of a UE with its action in state s_t. Fig. 5 describes the flowchart of our proposed ReLBT mechanism. As shown in the figure, a UE continuously observes the channel conditions and updates its QL-based parameters. The UE increases or decreases its CW only based on the channel collision probability, that is p_{obs}.

4. Performance evaluation

This section evaluates the performance of our proposed ReLBT mechanism using an event-driven simulator NS3 [23] with an available LTE-LAA and Wi-Fi coexistence scenario [24].

4.1. Simulation deployment scenario

In this scenario, we consider that two operators; operator-A (LTE-LAA), and operator-B (Wi-Fi), and both use the same 20 MHz channel in the 5 GHz frequency spectrum. We evaluate the performance of proposed ReLBT compared to the state-of-the-art LBT mechanism in

Table 2
Parameters used in simulations.

Parameter	Value
Number of cells/operator	4
Number of devices/cell	5, 15
Traffic model	FTP over UDP
Packet arrival rates (λ)	1.5, 2.5
Operating frequency	5 GHz
Channel bandwidth	20 MHz
Physical rate of the channel	MCS 15 (130 Mbps)
Data frame payload	1000 bytes
CW_{min} (LBT/BEB)	15/15
CW_{max} (LBT/BEB)	63/1023
ED threshold (LTE-LAA/Wi-Fi)	−72 dBm
CCA/DIFS (LTE-LAA/Wi-Fi)	60/43 μs
Slot-time σ (LTE-LAA/Wi-Fi)	9 μs
TXOP (LTE-LAA)	8 ms
NACKs feedback (LTE-LAA)	80%
Scaling design factor (ω)	32

terms of the cumulative distribution function (CDF) of both operators with latency (ms) and throughput (Mbps). Fig. 6 shows our deployment scenario, where two operators are deployed in four small cells. The four LTE-LAA eNBs and four Wi-Fi APs are fixed at their locations. Several LTE-LAA UEs and Wi-Fi STAs are randomly distributed around LTE-LAA eNB and Wi-Fi AP, respectively, as shown in Fig. 6. We performed simulations for two set of densities (see Table 2),

1. Five UEs/STAs (N) per cell (that is, $cells/operator = 4$, $N = 5$, $operators = 2$, $total N = 4 \times 2 \times 5 = 40$), and
2. fifteen UEs/STAs (N) per cell (that is, $cell/operator = 4$, $N = 15$, $operators = 2$, $total N = 4 \times 2 \times 15 = 120$).

Table 1 shows detailed simulation parameters.

4.2. Results and discussion

Figs. 7–8 show the latency impact of the LTE-LAA and Wi-Fi coexistence for LBT and ReLBT with five number of devices per cell and two different traffic arrival rates (that is, $\lambda = 1.5$ and $\lambda = 2.5$). The figures show that the proposed ReLBT mechanism performs more fairly between LTE-LAA and Wi-Fi. Besides, an increase in the traffic arrival rate (that is, $\lambda = 2.5$) increases the fairness challenges, as shown in Fig. 8. However, proposed ReLBT enables enhanced fairness between the operators. It becomes more severe for the LTE-LAA and Wi-Fi coexistence when the number of UEs/STAs increases per cell/operator, as shown in Fig. 9 and Fig. 10, where the total number of N is increased to 120 from 40 in the network. These figures show that the proposed ReLBT mechanism enhances the fair coexistence between LTE-LAA and Wi-Fi operators.

Figs. 11–14 evaluate the proposed mechanism for LTE-LAA and Wi-Fi coexistence (Figs. 11–12 for five number of devices per cell per operator, and Figs. 13–14 for fifteen number of devices per cell per operator) for LBT and ReLBT schemes. Fig. 11 shows that for the LBT mechanism, there exists a prominent amount of throughput degradation for the Wi-Fi STAs, and it is more miserable for the higher data arrival rate, as shown in Fig. 12. However, the proposed ReLBT mechanism enables LTE-LAA UEs to perform their transmissions intelligently, which increases the channel access chances for Wi-Fi STAs. Hence improves the overall systems performance. The fairness between LTE-LAA and Wi-Fi becomes more noticeable for the dense network environment, where the number of Nodes/STAs (N) increases per cell, as shown in Figs. 13–14. This is because of the increase in channel occupancy probability and time by LTE-LAA UEs. Currently implemented LTE-LAA LBT mechanism only considers the HARQ feedback for contention parameters update, which is received with much more delay that is after 7 ms [14]. As shown in Figs. 13–14, the Wi-Fi operator faces noticeable performance degradation due to the LTE-LAA

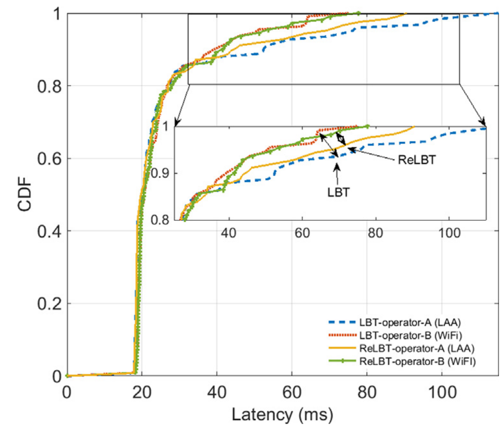


Fig. 9. Latency (ms) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 1.5$ with fifteen UEs/STAs per cell/operator.

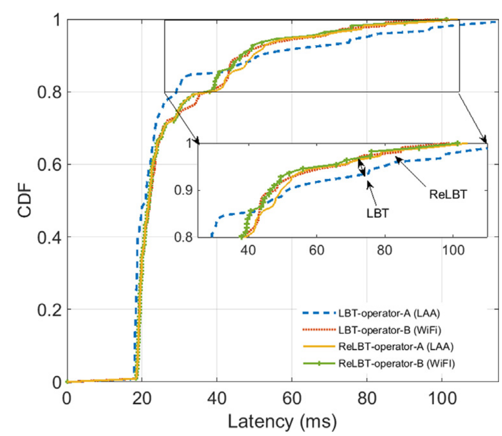


Fig. 10. Latency (ms) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 2.5$ with fifteen UEs/STAs per cell/operator.

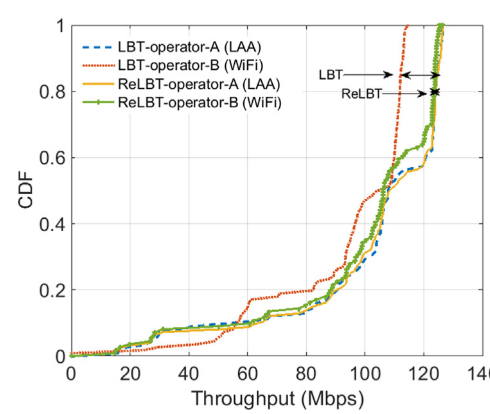


Fig. 11. Throughput (Mbps) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 1.5$ with five UEs/STAs per cell/operator.

LBT mechanism. Our proposed ReLBT allows LTE-LAA UEs to access channel resources more efficiently and fairly. Therefore, the performance of Wi-Fi STAs is improved. Since the LTE-LAA ReLBT adjusts the CW based on the channel inference by using channel collision probability, thus the throughput degradation due to an increase of the number of contenders has a small effect on ReLBT as compared to LBT as shown in Figs. 13 and 14. The reinforcement learning-based channel access of ReLBT enhances the fair channel occupancy for both LTE-LAA and Wi-Fi devices in the network.

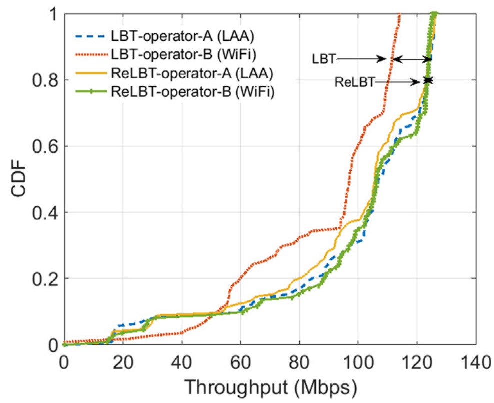


Fig. 12. Throughput (Mbps) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 2.5$ with five UEs/STAs per cell/operator.

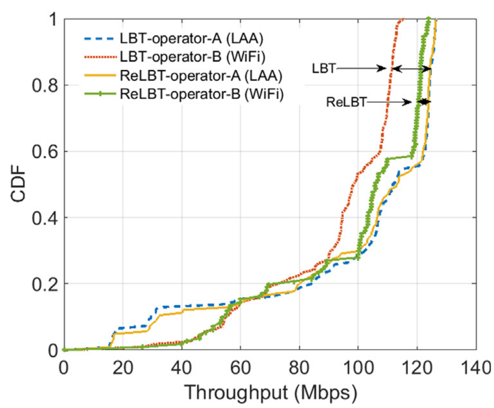


Fig. 13. Throughput (Mbps) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 1.5$ with fifteen UEs/STAs per cell/operator.

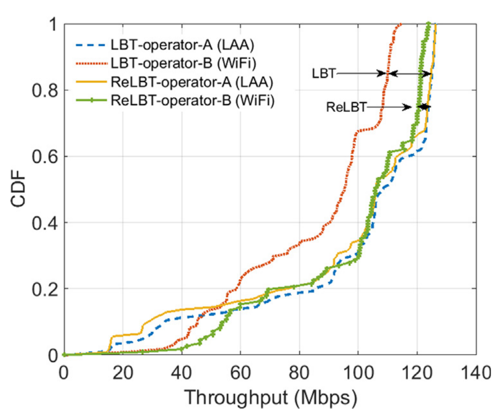


Fig. 14. Throughput (Mbps) performance of LTE-LAA and Wi-Fi coexistence with proposed ReLBT for traffic arrival rate $\lambda = 2.5$ with fifteen UEs/STAs per cell/operator.

5. Conclusion and future work

LTE-LAA and Wi-Fi coexistence performance are very delicate to the channel access mechanisms of these two diver technologies that are LBT of LTE-LAA and BEB of Wi-Fi. Essentially, the reason is the contention parameter choices in LTE-LAA LBT, such as HARQ feedback. Thus, the different LTE-LAA procedures for parameter selection affect coexistence performance. With the realization of this problem statement, this paper proposes an intelligent reinforcement learning-enabled LBT (ReLBT) to enhance the fairness of LTE-LAA and Wi-Fi coexistence. This paper evaluates the influence of the parameters associated with the LBT access protocol to improve the fairness between LTE-LAA and Wi-Fi networks. In particular, we assessed the sensitivity to HARQ feedback-based collision detection and the exponential backoff mechanism used by LTE-LAA LBT for scaling of the CW. Our proposed ReLBT mechanism utilizes Q learning, one of the reinforcement-learning (RL) paradigm for performance optimization of LTE-LAA and Wi-Fi coexistence. The potential applications of RL to the channel access to unlicensed wireless networks have already been increasingly recognized. RL is a behaviorist learning technique, which uses experience from the environment to optimize its performance. ReLBT uses channel observation-based collision probability to learn the environment and optimize its performance. Simulation results show that the proposed ReLBT mechanism scales the contention parameters more intelligently and effectively to enhances the LTE-LAA and Wi-Fi coexistence as compared to the state-of-the-art LBT mechanism. The future work expects to extend the applications of ReLBT for QoS-based traffic in the network, such as voice and video applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

R. Ali: Conceptualization, Formal analysis, Invetsigation, Writing - origional draft. **B. Kim:** Writing - review & editing. **S.W. Kim:** Project administration, Validation. **H.S. Kim:** Supervision, resources, funding acquisition. **F. Ishmanov:** Supervision, resources, funding acquisition.

References

[1] R. Ali, N. Shahin, A. Musaddiq, B. Kim, S.W. Kim, Fair and efficient channel observation-based listen-before talk (CoLBT) for LAA-WiFi coexistence in unlicensed LTE, in: International Conference on Ubiquitous and Future Networks (ICUFN), Prague, Czech Republic, 2018, pp. 162–167, http://dx.doi.org/10.1109/ICUFN.2018.8436776.

[2] A. Al-Dulaimi, S. Al-Rubaye, Q. Ni, E. Sousa, 5G communications race: Pursuit of more Capacity triggers LTE in unlicensed band, IEEE Veh. Technol. Mag. 10 (1) (2015) 43–51, http://dx.doi.org/10.1109/MVT.2014.2380631.

[3] N. Shahin, R. Ali, Y. Kim, Hybrid slotted-CSMA/CA-TDMA for efficient massive registration of IoT devices, IEEE Access 6 (2018) 18366–18382, http://dx.doi.org/10.1109/ACCESS.2018.2815990.

[4] R. Bajracharya, R. Shrestha, Y.B. Zikria, S.W. Kim, LTE in the unlicensed spectrum: A survey, IETE Tech. Rev. 35 (1) (2016) 78–90, http://dx.doi.org/10.1080/02564602.2016.1251344.

[5] R. Bajracharya, R. Shrestha, R. Ali, A. Musaddiq, S.W. Kim, LWA In 5G: State-of-the-art architecture, opportunities, and research challenges, IEEE Commun. Mag. 56 (10) (2018) 134–141, http://dx.doi.org/10.1109/MCOM.2018.1701177.

[6] R. Bajracharya, R. Shrestha, Y.B. Zikria, S.W. Kim, LTE or LAA: Choosing network mode for my mobile phone in 5G network, in: 2017 IEEE 85th Vehicular Technology Conference (VTC) Sydney, Australia, 2017, pp. 1-4. http://dx.doi.org/10.1109/VTCSpring.2017.8108640.

[7] R. Bajracharya, R. Shrestha, S.W. Kim, Q-learning based fair and efficient coexistence of LTE in unlicensed band, Sensors 19 (13) (2019) 2875, http://dx.doi.org/10.3390/s19132875.

[8] R. Ali, S.W. Kim, B.-S. Kim, Y. Park, Design of MAC layer resource allocation schemes for IEEE 802.11ax: Future directions, IETE Tech. Rev. 35 (1) (2016) 28–52, http://dx.doi.org/10.1080/02564602.2016.1242387.

[9] R. Bajracharya, R. Shrestha, S.W. Kim, Impact of contention based LAA on Wi-Fi network, Information 20 (2) (2017) 827–836.

[10] J. Jeon, H. Niu, Q. Li, A. Papathanassiou, G. Wu, LTE with listen-before-talk in unlicensed spectrum, in: 2015 IEEE International Conference on Communication Workshop (ICCW) London, 2015, pp. 2320–2324, http://dx.doi.org/10.1109/ICCW.2015.7247527.

[11] Y. Kim, Y. Song, Y. Choi, Y. Han, Nonsaturated throughput analysis of coexistence of Wi-Fi and cellular with listen-before-talk in unlicensed spectrum, IEEE Trans. Veh. Technol. 66 (12) (2017) 11425–11429, http://dx.doi.org/10.1109/TVT.2017.2734109.

[12] S. Saadat, D. Chen, K. Luo, M. Feng, T. Jiang, License assisted access-WiFi coexistence with TXOP backoff for LTE in unlicensed band, China Commun. 14 (3) (2017) 1–14, http://dx.doi.org/10.1109/CC.2017.7897317.

[13] in: 3GPP RP-151977 Status Report of WI Licensed-Assisted Access using LTE; rapporteur Ericsson, Huawei, 3GPP TSG RAN meeting, Spain, Dec. (2015) 7-10, Online At: https://www.3gpp.org/DynaReport/TDocExMtg--RP-70--31199.htm.

[14] 3GPP TR 36.889 V13.0.0 (2015-06). 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Study on Licensed- Assisted Access to Unlicensed Spectrum; (Release 13) Online At: https://www.tech-invite.com/3m36/tinv-3gpp-36-889.html.

[15] M. Faheem, V.C. Gungor, Energy efficient and qos-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. (68) (2018) 910–922, http://dx.doi.org/10.1016/j.asoc.2017.07.045.

[16] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, second ed., MIT Press, Cambridge, MA, 1998, ISBN: 0262193981.

[17] R. Ali, N. Shahin, Y.B. Zikria, B. Kim, S.W. Kim, Deep reinforcement learning paradigm for performance optimization of channel observation-based MAC protocols in dense WLANs, IEEE Access 7 (1) (2019) 3500–3511, http://dx.doi.org/10.1109/ACCESS.2018.2886216.

[18] R. Ali, N. Shahin, Y. Kim, B. Kim, S.W. Kim, Channel observation-based scaled backoff mechanism for high-efficiency WLANs, Electron. Lett. 54 (10) (2018) 663–665, http://dx.doi.org/10.1049/el.2018.0617.

[19] R. Ali, N. Shahin, R. Bajracharya, B. Kim, S.W. Kim, A self-scrutinized backoff mechanism for IEEE 802.11ax in 5G unlicensed networks, Sustainability 10 (4) (2018) 1–15, http://dx.doi.org/10.3390/su10041201.

[20] M. Faheem, S.B.H. Shah, R.A. Butt, B. Raza, M. Anwar, M.W. Ashraf, Md.A. Ngadi, V.C. Gungor, Smart grid communication and information technologies in the perspective of industry 4.0: Opportunities and challenges, Comp. Sci. Rev. (30) (2018) 1–30, http://dx.doi.org/10.1016/j.cosrev.2018.08.001.

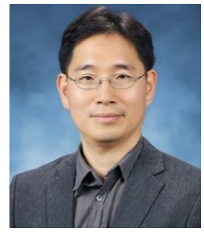
[21] M. Faheem, R.A. Butt, B. Raza, M.W. Ashraf, Md.A. Ngadi, V.C. Gungor, A multi-channel distributed routing scheme for smart grid real-time critical event monitoring applications in the perspective of Industry 4.0, Int. J. Ad Hoc Ubiquitous Comput. 32 (4) (2019) 236–256, http://dx.doi.org/10.1504/IJAHUC.2019.103264.

[22] M. Faheem, R.A. Butt, B. Raza, M.W. Ashraf, S. Begum, Md.A. Ngadi, V.C. Gungor, Bioinspired routing protocol for WSNbased smart grid applications in the context of industry 4.0, Trans. Emerg. Telecommun. Technol. 30 (8) (2019) e3503, http://dx.doi.org/10.1002/ett.3503.

[23] The Network Simulator — ns-3. Online AT: https://www.nsnam.org/.

[24] L. Giupponi, T. Henderson, B. Bojovic, M. Miozzo, Simulating LTE and wi-fi coexistence in unlicensed spectrum with NS-3, 2016, arXiv preprint arXiv:160406826, Available At: https://arxiv.org/abs/160406826.

Rashid Ali (Ph.D.) received his B.S. degree in Information Technology from Gomal University, Pakistan in 2007 and his M.S. degree under supervision of Dr. Stanislav Belenki in Computer Science (Advanced Network Design) in 2010 and M.S. degree in Informatics under the supervision of Dr. Maria Spante in 2012-13 from University West, Sweden. He received his Ph.D. diploma in Information and Communication Engineering from the Department of Information and Communication Engineering, Yeungnam University, Korea in February 2019. Between 2007 and 2009, he worked for Wateen Telecom Pvt. Ltd. Pakistan as WiMAX Executive Engineer in Operations & Research Department. From July 2013 to June 2014, he worked for COMSATS University (Vehari), Pakistan, as a Lecturer. Currently, he is working as a PostDoc researcher at the School of Intelligent Mechatronics Engineering, Sejong University, Seoul, Korea. His research interests include next-generation wireless local area networks (IEEE 802.11ax/ah), unlicensed wireless networks in 5G, Internet of Things (IoT), performance evaluation of wireless networks, Reinforcement Learning techniques for wireless networks, and Blockchain technology for IoT trustworthiness.



Byung-Seo Kim received his B.S. degree in Electrical Engineering from In-Ha University, Korea in 1998 and his M.S. and Ph.D. degrees under supervision of Dr. Yuguang Fang in Electrical and Computer Engineering from the University of Florida in 2001 and 2004, respectively. Between 1997 and 1999, he worked for Motorola Korea Ltd., Korea as CIM Engineer in ATR & D. From January 2005 to August 2007, he worked for Motorola Inc., Schaumburg Illinois, as a Senior Software Engineer in Networks and Enterprises. He is currently an Associate Professor at the Department of Computer and Information Communication Engineering in Hongik University, Korea.

Sung Won Kim received his B.S. and M.S. degrees from the Department of Control and Instrumentation Engineering, Seoul National University, Korea, in 1990 and 1992, respectively, and his Ph.D. degree from the School of Electrical Engineering and Computer Sciences, Seoul National University, Korea, in August 2002. From January 1992 to August 2001, he was a Researcher at the Research and Development Center of LG Electronics, Korea. From August 2001 to August 2003, he was a Researcher at the Research and Development Center of AL Tech, Korea. From August 2003 to February 2005, he was a Postdoctoral Researcher in the Department of Electrical and Computer Engineering, University of Florida, Gainesville, USA. In March 2005, he joined the Department of Information and Communication Engineering, Yeungnam University, Gyeongsangbuk-do, Korea, where he is currently a Professor. His research interests include resource management, wireless networks, mobile networks, performance evaluation, and embedded systems.

Hyung Seok Kim received the B.S. degree from the School of Electrical Engineering, Seoul National University, South Korea, in 1996, and the M.S. and Ph.D. degrees from the School of Electrical Engineering and Computer Engineering, Seoul National University, in 1998 and 2004, respectively. In 2003 and 2004, he held visiting researcher positions with the University of Virginia. From 2004 to 2006, he was with the Telecommunication R & D Center, Samsung Electronics. He is currently a Faculty Member with the Department of Information and Communication Engineering, Sejong University. His research interests include networking with special emphasis on wireless communication and visual light communication.

Farruh Ishmanov received his B.S in Information Systems in 2007 from Tashkent State University of Economics, Uzbekistan. At this University, he studied and worked in the Multimedia Lab during the undergraduate years. He received his M.S. and Ph.D. degrees from the Department of Information and Communication Engineering, Yeungnam University, Korea, in 2009 and 2014, respectively. He was awarded Korean Government IITA Scholarship for pursing M.S degree. In March 2015, he joined the Department of Electronics and Communication Engineering, Kwangwoon University, Korea, where he is currently an Assistant Professor. His research interests include resource management and security in wireless sensor networks and IoT.