Received December 15, 2017, accepted January 18, 2018, date of publication February 12, 2018, date of current version March 28, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2803258*

# LTE-U and Wi-Fi Coexistence Algorithm Based on Q-Learning in Multi-Channel

**YUHAN SU[1], XIAOJIANG DU[2], (Senior Member, IEEE), LIANFEN HUANG[1], ZHIBIN GAO[1], AND MOHSEN GUIZANI[3], (Fellow, IEEE)**

[1]Department of Communication Engineering, Xiamen University, Xiamen 361005, China
[2]Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122, USA
[3]Department of Electrical and Computer Engineering, University of Idaho, Moscow, ID 83844, USA

Corresponding author: Lianfen Huang (lfhuang@xmu.edu.cn)

This work was supported by the 2015 National Natural Science Foundation of China under Grant 61401381.

**ABSTRACT** Due to the lack of resources in the low spectrum, long term evolution (LTE) in unlicensed spectrum (LTE-U) technology has been proposed to extend LTE to unlicensed spectrum. LTE-U undertakes the task of streaming data traffic for licensed spectrum, which can greatly enhance the capacity of the system. However, the introduction of LTE-U technology also gives rise to the problem of coexistence with Wi-Fi systems. In this paper, an LTE-U and Wi-Fi coexistence algorithm is proposed in multi-channel scenarios based on Q-learning. By taking the idea of alternately transferring data in LTE-U and Wi-Fi, the algorithm takes into account both the fairness and the performance of the system and optimizes the duty cycle. The simulation results show that the proposed algorithm can effectively improve the throughput of the system in the premise of ensuring fairness.

**INDEX TERMS** Long term evolution (LTE)-U, Wi-Fi, Q-Learning, reinforcement learning, coexistence algorithm.

## I. INTRODUCTION

With the development of mobile Internet, intelligent devices and mobile applications grow rapidly, which allow people to carry out communication connections anytime, anywhere, including call, multimedia and cloud services. At the same time, Internet of Things (IoT) equipment also ushers in the era of rapid development for the mobile network and has brought explosive traffic demand. It was reported by Cisco in its published global mobile data traffic forecast white paper [2] that global mobile traffic in 2016 increased by 63% compared to 2015, and had reached up to 18 times growth in the past five years. This figure will grow more than seven times in 2020. In addition, it was predicted that the number of mobile users will reach 5.5 billion and the number of mobile network equipment will be more than 11.6 billion in 2020. Such a huge flow and rate requirements will bring great challenges to the current mobile network capacity and quality of service.

Faced with such huge amounts of equipment access and business growth, people in both industrial and academic communities are trying to find an advanced technology to substantially increase the capacity of network under the premise of ensuring a high level of user experience. To this end, the industrial community has put forward some technologies to improve the efficiency of the existing licensed spectrum through the air interface and physical layer technology innovation, so as to extend the capacity of the system. These measures include aggregating large numbers of small spectrum into a 100 MHz virtual band by CA technology [3], reusing the frequency of macro base stations by deploying Small Cell, and using enhanced cell interference coordination algorithm to eliminate interference and improve spectral efficiency, etc.

### A. PROBLEM DEFINITION AND CHALLENGES

However, in view of the fact that the higher value of the licensed low-frequency resources have been consumed and hence these licensed spectrum are so limited, it is difficult to meet the future mobile network traffic demand by using the underlying technology or reusing the frequency of macro base stations to improve the spectral efficiency and network performance. On the other hand, the bandwidth resource in the unlicensed spectrum is nearly the same as that in the licensed spectrum [4]. In order to alleviate the growing mobile network traffic pressure, the use of unlicensed spectrum for mobile network deployment has become a new research direction. In fact, many operators have chosen to deploy a large number of Wi-Fi in the unlicensed spectrum to relieve the pressure in the licensed frequency. However, Wi-Fi cannot realize the cooperative use of spectrum resources by complicated scheduling process since it is a technology based on competitive access. It is also difficult to provide delay, rate and security like the long term evolution (LTE) system. In addition, the spectrum efficiency is much lower than the LTE system using the same bandwidth.

In order to achieve the full use of unlicensed spectrum, industrial and academic communities proposed the concept of deploying LTE in unlicensed spectrum [6]–[8], known as LTE-U technology. The LTE-U technology extends LTE to the unlicensed spectrum, that is, one can use the LTE standard protocol to communicate on the unlicensed spectrum, and aggregate licensed spectrum and unlicensed spectrum through the CA technology. In this way, part of data transmission in the licensed spectrum is transferred to the unlicensed spectrum [1].

Introducing LTE into a free, common, unlicensed spectrum will inevitably bring the problems of competing and coexisting with other unlicensed communications technologies in the same spectrum. In traditional communication technologies using unlicensed spectrum for data transmission, as represented by Wi-Fi, access channels can only be accessed through a competitive way in order to realize spectrum sharing. The LTE designed in the licensed spectrum has the absolute control over the spectrum. It performs the centralized scheduling of the wireless resource through the base station, so as to obtain higher spectral efficiency. Obviously, if the unlicensed spectrum is used as a new spectrum of LTE, the transmission of LTE-U will cause serious interference to Wi-Fi due to the channel detection and back off mechanism of Wi-Fi. Therefore, for the two systems with completely different time slots and scheduling modes, it is necessary to additionally design reasonable and fair coexistence in order to ensure both in the conditions of good transmission in the unlicensed spectrum.

### *B. CONTRIBUTIONS AND OUTLINE*

In response to the problems and challenges mentioned above, this paper studies the coexistence of LTE-U and Wi-Fi systems in a multi-channel unlicensed spectrum scenario. We design a joint utility function of the system throughput and fairness for the coexistence scenario, and propose a coexistence algorithm for multi-channel LTE-U based on Independent Q-learning [9]–[12] algorithm and Joint Q-learning algorithm. In addition, the coexistence scenarios are modeled by Q-learning and the optimal duty cycle of the system in multiple channels is obtained by learning iterations. The algorithm improves the throughput of system while guaranteeing the fairness [13]–[15] between the two systems.

Meanwhile, a system-level simulation platform for LTE-U and Wi-Fi heterogeneous networks is built and used to simulate the proposed coexistence mechanism algorithm. Through the comparative analysis of the simulation results, it can be seen that the proposed algorithm can improve the system performance while maintaining the fair transmission of LTE-U and Wi-Fi.

The rest of the paper is organized as follows. In section II, the related works are briefly reviewed. In Section III, the coexistence of LTE-U and Wi-Fi systems in the multi-channel unlicensed spectrum scenario is studied, where we design a joint utility function of system throughput and fairness for coexistence scenarios, and present the LWCA based on Q-learning in multi-channel. In section IV, we design the LTE-U system simulation platform and conduct performance evaluation to the algorithm. The main conclusions are summarized in section V.

## II. RELATED WORK

At present, there are already several mature technologies deployed in unlicensed spectrum, such as Wi-Fi and Bluetooth technologies. The most important problem in the introduction of LTE into the unlicensed spectrum is to consider the coexistence between these technologies. The academia mainly proposes two types of coexistence algorithms: competitive access based algorithm and duty cycle allocation based algorithm.

The coexistence algorithm based on competitive access requires that the LTE-U device access the channel in the same way as the Wi-Fi device in the unlicensed spectrum. This algorithm is based on the Listen-Before-Talk (LBT) scheme mandated by countries such as Europe. In [5], an LTE-U MAC protocol based on LBT was proposed, which requires that the LTE-U device be detected at the end of the Wi-Fi transmission frame. A coexistence algorithm composed of two schemes: periodic detection and fixed perception was proposed in [16]. In [17], an LBT adaptive algorithm was put forward, where it is required that LTE-U be aware of the channel at the edge of the sub-frame and be able to select a new free channel for use. A fair LBT algorithm was developed in [18], which combines the total throughput of the system and the fairness factor between LTE-U and Wi-Fi. The transmission is ensured by allocating the appropriate idle period for Wi-Fi. Several other papers (e.g., [19]–[23]) also studied the related wireless issues.

The coexistence algorithm based on the duty cycle allocation requires that the LTE-U and Wi-Fi systems multiplex and coexist the spectrum resources, that is, the LTE-U and Wi-Fi systems use the channels in ON/OFF mode. Based on this idea, the LTE-U Forum proposed the Carrier-Sensing Adaptive Transmission (CSAT) [24]. In addition, [25] proposed the basic framework of the cooperative coexistence algorithm, which describes the general flow of collaborative coexistence algorithms. In [26], a coexistence algorithm for allocating idle slots by LTE according to a predetermined duty cycle is proposed. In [27], a new algorithm based on duty cycle adaptive algorithm is proposed, which achieves the fast convergence of optimal duty cycle by Q-learning. Reference [28] proposed a fair proportion of resource allocation algorithm, the algorithm requires that the average channel occupancy time of the Wi-Fi device be equal to the average channel occupancy time of the LTE-U device, thus ensuring a fair transmission between LTE-U and Wi-Fi.

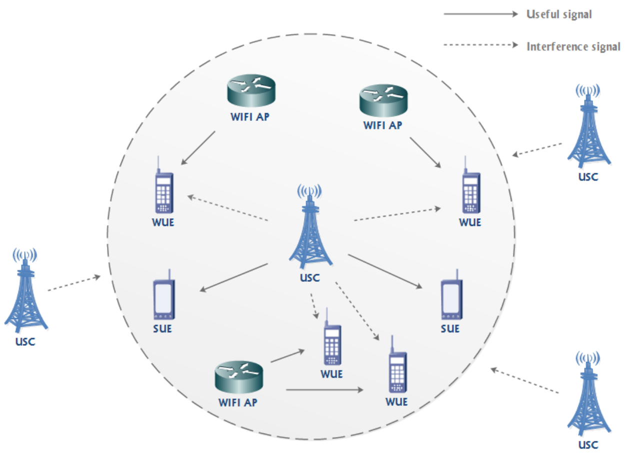



**FIGURE 1.** LTE-U and Wi-Fi coexistence network scenarios.

## III. THE LWCA IN MULTI-CHANNEL

The 5 GHz unlicensed spectrum has a bandwidth of approximately 500 MHz, which is divided into multiple unlicensed channels. In practice, the actual traffic of Wi-Fi for each unlicensed channel is different. Therefore, merely considering the coexistence of LTE-U in a single unlicensed channel is difficult to meet the practical use of the actual scene. In the Traditional algorithm, the idlest channel is usually selected for transmission, and there is rarely research on the coexistence algorithm of multiple unlicensed channels. This chapter analyzes the multi-channel unlicensed spectrum, and designs joint utility function according to system throughput and fairness coefficient. The duty cycle of LTE-U and Wi-Fi is optimized and adaptively adjusted by the Q-learning algorithm, so as to improve the performance of system by allowing LTE-U to occupy the appropriate time slots on multiple unlicensed channels.

### A. SYSTEM MODEL

The scenario of system is shown in Fig. 1, where the LTE-U and the Wi-Fi systems form a heterogeneous network with a common spectrum. The network contains LTE-U nodes and Wi-Fi nodes. The LTE-U nodes include LTE-U Small Cell (USC) and LTE-U Small Cell UE (SUE), and the Wi-Fi node includes an AP and a Wi-Fi device (WUE). Among them, in addition to the mutual interference between the central deployment of LTE-U system and Wi-Fi system, the surrounding deployment of USC will also bring some interference effects.

In the heterogeneous network, $U$ USCs are deployed, which inform the SUEs of the use of one or more unlicensed channels for transmission through the licensed spectrum. In the unlicensed spectrum, a total of $M$ unlicensed channels are available, with one AP and $W_m$ WUEs active in each unlicensed channel. The LTE-U system transmits through the LTE air interface. There are $N$ resource blocks (RBs) available for transmission in each transmission time slot (TTI), where $N$ is related to the current unlicensed channel bandwidth.

If $SINR_{i,n}$ represents the SINR of the USC $i$ in the resource block $n$, then we have

$$SINR_{i,n} = \frac{P_{i,n}G_{i,n}}{\sum_{j=1, j\neq i}^{S} P_{j,n}G_{j,n} + \sigma^2} \tag{1}$$

where $P_{i,n}$ and $P_{j,n}$ respectively represent the transmit power of the USC $i$ and USC $j$ in the RB $n$, $G_{i,n}$ ($G_{j,n}$) is the link gain between the USC $i$ ($j$) and its serving users in the resource block $n$, and $\sigma^2$ is the noise power. LTE-U conforms to the physical layer standard protocol of LTE and implements adaptive coding modulation through CQI of SINR mapping. The higher CQI implies the higher transmission rate [29].

The Wi-Fi system accesses channels in the way of competing. In DCF protocol mode [31], WUE throughput can be expressed as

$$T_{wi-fi} = R \times t_r \tag{2}$$

where $t_r$ is the statistical value obtained by simulating the flow of WUE access channel. The physical layer rate of the WUE is mapped by (2) according to the SINR of the WUE.

### B. LWCA BASED ON Q-LEARNING

The current coexistence algorithm of LTE-U commonly accesses the system by simply choosing the most free access channel. It lacks the support for multi-channel and is likely to cause a waste of spectrum resources. Moreover, the Traditional algorithm lacks a dynamic learning process and cannot adjust the system parameters according to the actual situation.

In view of the fact that it has model-independent features, Q-learning is used in this paper to solve the problem of the coexistence of LTE-U and Wi-Fi. Q-learning does not have to get the environment transfer function and reward the expected conditions to obtain the system optimal policy, which is in line with the actual situation of the LTE-U network. In addition, by independent learning or joint learning of multi-channel, Q-learning can adjust the duty cycle of USC in different channels, and approach the optimal solution through a certain number of iterations.

In the process of learning iteration, the agent perceives the state $s \in S$ of the current environment, and selects the corresponding action $a \in A$ according to the policy $\pi$. The agent receives the reward value $r$ from the environment after the action, and changes its state from $s \in S$ to $s' \in S$. Then it updates the policy based on the current state and the reward value, and proceeds to the next round of learning iterations.

The purpose of Q-learning is to find the optimal policy $\pi^*$ that maximises some notion of cumulative reward of the agent. Every agent observes his current state, discovers his environment, then takes the best decision concerning his next action/state. It has been demonstrated that for any given Markov decision process, Q-learning can be used to obtain the optimal policy $\pi^*$ for each state $\tilde{s}$ which maximises the observed rewards over time [9]–[11].

The discounted cumulative reward $V(\tilde{s})$ of a state $\tilde{s}$ can be expressed as follows [11]:

$$V(\tilde{s}) = R(\tilde{s}, \tilde{a}) + \gamma \sum\nolimits_{s' \in S} P\left(s'|\tilde{s}, \tilde{a}\right) V\left(s'\right) \tag{3}$$

where $R(\tilde{s}, \tilde{a})$ is the agent's immediate reward when he selects action $\tilde{a}$ at a state $\tilde{s}$, $0 < \gamma < 1$ is a discount factor, and $P\left(s'|\tilde{s}, \tilde{a}\right)$ is the transition probability from state $\tilde{s}$ to $s'$ when the agent chooses action $\tilde{a}$. The optimal policy $\pi^*$ is obtained when the total discounted expected reward is maximal according to the Bellman's theory [11]. That is when

$$V^*(\tilde{s}) = \max_{\tilde{a} \in A} [R(\tilde{s}, \tilde{a}) + \gamma \sum_{s' \in S} P\left(s'|\tilde{s}, \tilde{a}\right) V(s')] \tag{4}$$

where $V^*(\tilde{s})$ is the maximum discounted cumulative reward that an agent can obtain if he starts from a state $\tilde{s}$ and follows the optimum policy $\pi^*$.

Since the decision policy is $\pi: s \to a$, its optimum value can be obtained as follows [11]:

$$\pi^*(\tilde{s}) = \arg\max_{\tilde{a} \in A} [R(\tilde{s}, \tilde{a}) + \gamma \sum_{s' \in S} P\left(s'|\tilde{s}, \tilde{a}\right) V(s'|)] \tag{5}$$

Since the reward $R(\tilde{s}, \tilde{a})$ and the transition probability $P(s'||\tilde{s}, \tilde{a})$ are unknown, the use of Q-learning is important to learn those values over time. For a given policy $\pi$, the Q-value is expressed as follows [9]:

$$Q^\pi(\tilde{s}, \tilde{a}) = R(\tilde{s}, \tilde{a}) + \gamma \sum\nolimits_{s' \in S} P(s'||\tilde{s}, \tilde{a}) V^\pi(s'|) \tag{6}$$

It is clear that the Q-value represents the expected discounted reward when the agent follows the policy $\pi$ starting from a state $\tilde{s}$ and executing an action $\tilde{a}$. Therefore, our objective is to estimate the Q-value for an optimal policy $\pi^*$.

In Q-learning, the agent calculates $Q_t(\tilde{s}, \tilde{a})$ value at each instant $t$ in recursive manner in order to learn which action is optimal for each state. The Q-value update depends on $\tilde{s}$, $\tilde{a}$, $s'$, and $R_t(\tilde{s}, \tilde{a})$ information and is estimated based on the following formula:

$$Q_t(\tilde{s}, \tilde{a}) = Q_{t-1}(\tilde{s}, \tilde{a}) + \alpha[R_t(\tilde{s}, \tilde{a}) + \gamma \max_{\hat{a}} Q_{t-1}(s', a') - Q_{t-1}(\tilde{s}, \tilde{a})] \tag{7}$$

where $0 < \alpha < 1$ is the learning rate and $\max_{\hat{a}} Q_{t-1}\left(s', a'\right)$ is the estimation of the optimal future value.

According to the theory of Q-learning and the network model under the coexistence scenario, two schemes are considered here: Independent Q-learning based algorithm and Joint Q-learning based algorithm.

### C. ALGORITHM 1: INDEPENDENT Q-LEARNING BASED ALGORITHM

In Independent Q-learning based scenario, the USC proceeds learning for each unlicensed channel independently. The agents, states, actions, and reward function for the Independent Q-learning based algorithm are defined as follows:

***Agents:*** USC $u \in \{1, 2, \ldots, U\}$ serves as an agent, in which an unlicensed channel is used as the allocation unit for the duty cycle, and the set of unlicensed channels in each USC is $m \in \{1, 2, \ldots, M\}$.

***States:*** The set of states is expressed as

$$\vec{s}_u = \left\{T_u^m, \overline{F_u}\right\} \tag{8}$$

where $T_u^m$ is the total throughput obtained in the unlicensed channel $m$ and $\overline{F_u}$ is the average of the fairness coefficients calculated in each of the unlicensed channels $m$. The fairness factor $F$, which is used to calculate the fairness between the parameters, is given by

$$F = \frac{(\sum x_i)^2}{N(\sum x_i^2)} \tag{9}$$

where $x_i$ is a calculation factor denoting the normalized LTE-U system and Wi-Fi system throughput value, and $N$ is the number of factors. The factor $F$ lies in the range of [0,1] and its value closer to 1 represents the better the fairness. According to the pre-set throughput thresholds $T_{th}^m$ and $\overline{F_{th}}$, the agent $u$ is divided into four states: low profitability low fairness, high throughput low fairness, low throughput high fairness and high throughput, and high fairness. The state set is $I$ and the elements are represented by

$$s_i^m \begin{cases} 1, & T_u^m \le T_{th}^m \text{ and } \overline{F_u} \le \overline{F_{th}} \\ 2, & T_u^m > T_{th}^m \text{ and } \overline{F_u} \le \overline{F_{th}} \\ 3, & T_u^m \le T_{th}^m \text{ and } \overline{F_u} > \overline{F_{th}} \\ 4, & T_u^m > T_{th}^m \text{ and } \overline{F_u} > \overline{F}_{th} \end{cases} \tag{10}$$

***Actions:*** The duty cycle set $a_j^m \in \{a_1^m, a_2^m, \ldots, a_k^m\}$, which represents the set of optional actions for the agent $u$ on the unlicensed channel $m$.

***Reward:*** The reward function is defined as

$$r_u = \begin{cases} 0, & if\ \bar{F}_u < F_{min}\ or\ T_u^m \& < T_{min} \\ \dfrac{T_u^m}{T_{min}} \times e^{-|1-\bar{F}_u|}, & \text{others} \end{cases} \tag{11}$$

In the four quantities defined above, $F_{\min}$ is the minimum requirement of system fairness coefficient and $T_{\min}$ is the minimum requirement of system throughput. By taking into account the overall network throughput performance and network fairness factors, the reward function makes the fairness factor close to 1 as far as possible on the condition that the system throughput is larger than the minimum threshold of throughput. Under this reward function, USC chooses the policy to iterate in the direction of high-throughput and high fairness.

For both Independent and Joint Q-learning based coexistence algorithm, the Boltzmann algorithm [30] is used as the action policy of the agent selection. The Boltzmann algorithm is a common algorithm for balancing the accumulation of experience and exploring. The algorithm calculates the probability of different actions according to (12), and then select the action according to the probability,

$$\Pr(\mathrm{a}|\mathrm{s}) = \frac{e^{\frac{Q(s,a)}{T}}}{\sum_{a' \in A} e^{\frac{Q(s,a')}{T}}} \tag{12}$$

where Pr (a|s) is the probability that the agent will select the action $a$ in the state $s$, and $T$ is the temperature value. We reduce the $T$ value by (13), to reduce the number of explorations of the optimal policy of the agent,

$$T = \frac{T_0}{log_2(1+N)} \tag{13}$$

where $T_0$ is the initial temperature value, and $N$ the times that the action is selected for the agent.

The Independent Q-learning based algorithm is summarized as follows:

1) *At t = 0, initialize the Q-value of states and actions to be 0 for the agent u in the unlicensed channel m.*
2) *Compute the initial state $s_t$ of the agent u in unlicensed channel m.*
3) *Calculate the probabilities for the agent u in the unlicensed channel m with different actions $a_j^m$ in the state $s_t$ according to Equations (12) and (13),and perform the action with the maximum probability in current state, randomly selecting one of them if there are multiple identical probabilities.*
4) *Perform the action $a_j$, to get the corresponding environmental reward value $\gamma_t$, and then enter the next state $s_{t+1}$;*
4) *Update the corresponding action Q-value of the agent u;*
5) *$t \leftarrow t+1$, and jump to step 1.*
6) *After completing learning of unlicensed channel m, $m \leftarrow m+1$, and jump to the beginning of the cycle.*

### D. ALGORITHM 2: JOINT Q-LEARNING BASED ALGORITHM

In the Joint Q-learning algorithm, the USC needs to select the duty cycle for all the unlicensed channels in one action selection, taking the total throughput of the LTE-U and Wi-Fi systems as the environmental feedback value. The corresponding agents, states, actions, and reward function for the Joint Q-learning based algorithm are defined as follows:

***Agents:*** Set USC $u \in \{1, 2, \ldots, U\}$ as an agent.

***States:*** The set of states is expressed as

$$s_u = \{\textstyle\sum_M T_u^m, F_u\} \tag{14}$$

where $F_u$ is the system fairness factor.

Similar to the Independent Q-learning based algorithm, the agent of Joint Q-learning based algorithm is divided into four states according to the default system throughput threshold $T_{th}$ and fairness threshold $F_{th}$, and the elements state set $I$ are given by

$$s_i \begin{cases} 1, & \sum_M T_u^m \leq T_{th} \text{ and } F_u \leq F_{th} \\ 2, & \sum_M T_u^m > T_{th} \text{ and } F_u \leq F_{th} \\ 3, & \sum_M T_u^m \leq T_{th} \text{ and } F_u > F_{th} \\ 4, & \sum_M T_u^m > T_{th} \text{ and } F_u > F_{th} \end{cases} \tag{15}$$

***Actions:*** The set of actions is $a_j \in \{a_1, a_2, \ldots, a_l\}$. In the Joint Q-learning based algorithm, the agent $u$ needs to select a duty cycle for each of the unlicensed channels in an action selection, i.e., each action is a combination of $m$-dimensional duty cycle. In the case of three unlicensed channels, if the set of duty cycles for a single unlicensed channel is {20%, 40%, 60%, 80%}, the total number of actions of the agent $u$ is $3^4 = 64$.

***Reward:*** The reward function is defined as

$$r_u = \begin{cases} 0, & \text{if } F_u < F'_{min} \text{ or } \sum_M T_u^m < T'_{min} \\ \dfrac{\sum_M T_u^m}{T'_{min}} \times e^{-|1-F_u|} & \text{others} \end{cases} \tag{16}$$

The Joint Q-learning based algorithm is summarized as follows:

1) *At t = 0, initialize Q-value of states and actions to be 0.*
2) *Compute the starting state $s_t$ of the agent u.*
3) *Calculate the probabilities of the current agent u in state $s_t$ under different actions $a_j$ according to Equations (12) and (13), and perform the action with the maximum probability in current state, randomly selecting one of them if there are multiple identical probabilities.*
4) *Perform the action $a_j$, to get the corresponding environmental reward value $\gamma_t$, and then enter the next state $s_{t+1}$.*
5) *Update the corresponding action Q-value of the agent u.*
6) *$t \leftarrow t+1$, and jump to the step 2.*

## IV. PERFORMANCE EVALUATION

### A. SIMULATION RESULT PARAMETER SETTINGS

The LTE-U system-level simulation platform built in this paper was developed using MATLAB. The system simulation uses the snapshot mode, that is, the statistics is made by taking the average over random user locations. Without loss of generality, we set a total of unlicensed channels in the simulation, an AP is deployed in each unlicensed channel and different WUE are used to respectively simulate idle, general, and busy Wi-Fi transmission strength. USC is allowed to use multiple unlicensed channels at the same time, and inform the UE to complete the communication process in the corresponding channel through the licensed spectrum.

The simulation scene parameters are set as shown in TABLE 1. The simulation parameters for the LTE-U and Wi-Fi systems are shown in TABLE 2.

### B. SIMULATION RESULTS AND ANALYSIS

In the simulation process, the Q-learning algorithm is compared with the Traditional algorithm, Average algorithm, LBT algorithm, and CSAT algorithm. The USC of the Traditional algorithm chooses the channel with the least interference to transmit, without taking any coexistence measures. The Averaging algorithm requires the USC to take up available channels in a polling manner during the transmission time, so that the SUEs are balanced in different channels.

**TABLE 1.** Simulation parameter.

| | | | |
|---|---|---|---|
| Simulation Scene | SDL | Number of Channels | 3 |
| USC Transmit Power | 23dBm | AP Transmit Power | 7dBm |
| SUE Number | 6 | WUE Number | 2/6/10 |
| SUE Packet Size | 50 | WUE Packet Size | 20 |
| Spectrum | 5.8GHz | Noise Power Density | 174dBm/Hz |
| Packet Arrives Model | Poisson | Simulation Time | 10s |
| PL Model | 36.7log10(d[m])+22.7+26log10(frq[GHz]) | | |

**TABLE 2.** Wi-Fi and LTE-U system parameter.

| WI-FI | | LTE-U | |
|---|---|---|---|
| MAC Protocol | DCF | Bandwidth | 20MHz |
| Time Slot | 20us | Scheduling Method | Polling |
| CW | 32-256 | Discount Factor $\beta$ | 0.04 |
| SIFS | 10us | Learning Factor $\alpha$ | 0.5 |
| DIFS | 50us | $T_0$ | 0.15 |

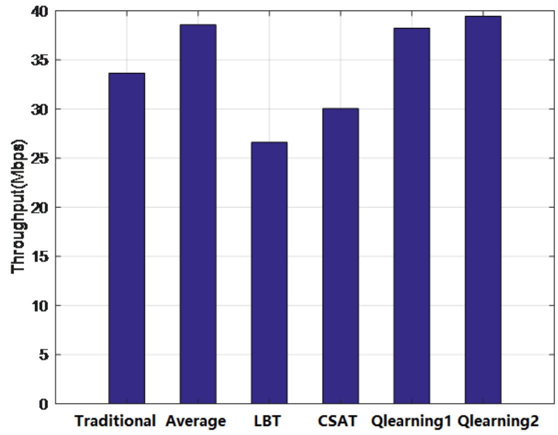



**FIGURE 2.** Comparison of average throughput of different algorithms.

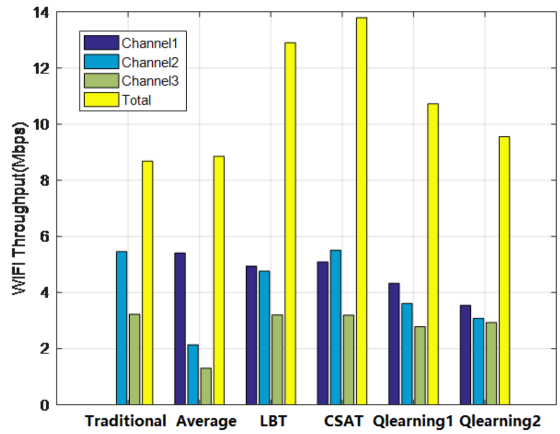



**FIGURE 3.** Comparison of Wi-Fi throughput of different algorithms.

The LBT algorithm achieves coexistence through idle channel selection and carrier sense avoidance. The CSAT algorithm first selects the idlest channels in the currently available channel, and, by listening to the Wi-Fi data transmission, dynamically adjusts the percentage of the LTE-U in the slot of the unlicensed channel to achieve coexistence.

The abscissa of the simulation below, Traditional represents the Traditional algorithm, Average represents the Average algorithm, Qlearning1 is the Independent Q-learning based algorithm, the single channel action step is 10%, Qlearning2 is the Joint Q-learning based algorithm, the single channel action step is 20%.

Fig. 2 presents the simulation results for the system user's average throughput, where Qlearning1 represents the Independent Q-learning based algorithm with the single channel action step 10% and Qlearning2 represents the Joint Q-learning based algorithm with the single channel action step 20%. It can be seen from the figure that the Q-learning based algorithm can achieve better throughput performance. In contrast, although the Traditional algorithm and Average algorithm can achieve higher throughput, they do not consider the coexistence between devices, and consequently will squeeze more channel resources. The LBT algorithm demonstrates poor performances due to the "degeneration" of LTE-U to the way of competing access, which may result in more channel time wasted. The throughput performance of CSAT algorithm is slightly better than the LBT algorithm. However, due to its more parameters, it is difficult to achieve the best duty cycle settings in the current environment, since the different parameter settings can easily affect the performance of the algorithm. The reason why the Q-learning based algorithm can get the increase of throughput is that, through the system learning, the currently available channel is occupied by a reasonable duty cycle, which reduces the backlash caused by the channel collision and improves the channel utilization.

Fig. 3 and Fig. 4 present the simulation results for the Wi-Fi and LTE-U throughput, respectively. For the Traditional and Average algorithms, the performances of both Wi-Fi and LTE-U are extremely unbalanced, and therefore they are suitable for the actual scene.

In view of the fact that the LBT and CSAT algorithms can dynamically adjust the channel resources according to the actual transmission situation of Wi-Fi, the two algorithms can get a high Wi-Fi throughput, as shown in Fig. 3. This indicates that both two algorithms have realized the protection of

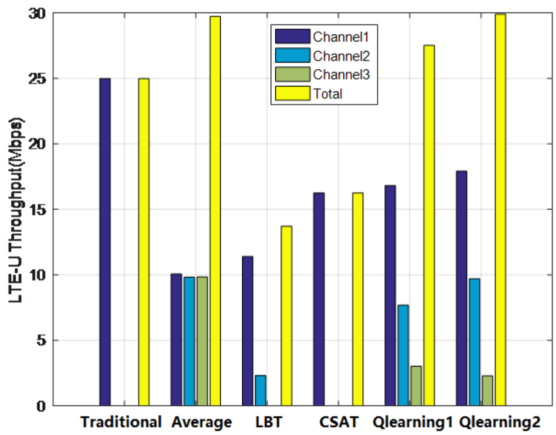



**FIGURE 4.** Comparison of LTE-U throughput of different algorithms.

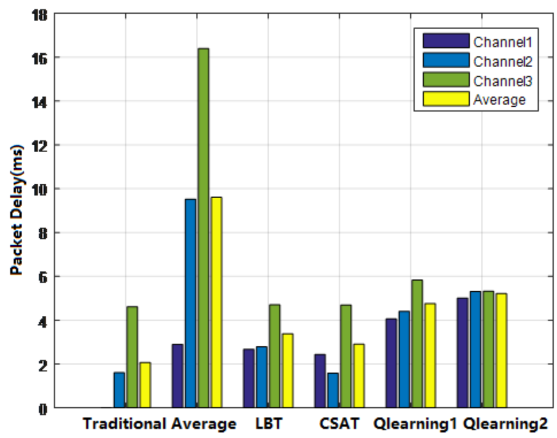



**FIGURE 6.** Comparison of Wi-Fi packet delays of different algorithms.

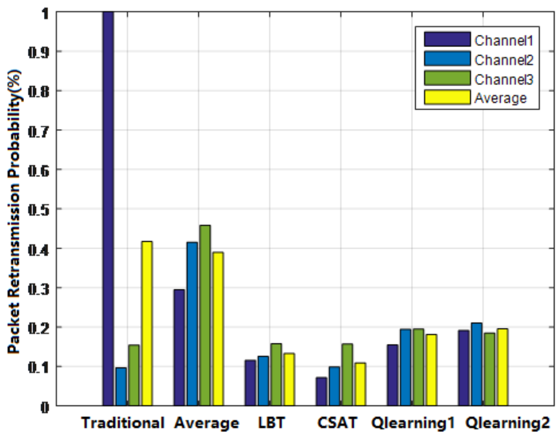



**FIGURE 5.** Comparison of Wi-Fi retransmission probability of different algorithms.

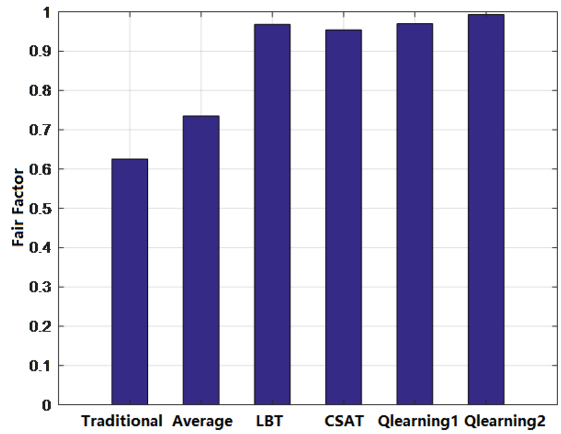



**FIGURE 7.** Comparison of fairness factor of different algorithms.

Wi-Fi system performance, which may facilitate the deployment of LTE-U. However, one can find from Fig. 4 that this protection greatly sacrifices the performance of LTE-U. In addition, according to Fig.2, both two coexistence algorithms possess low utilization of the spectrum, which may slightly degenerate the comprehensive performance. For the Q-learning based algorithm, however, the distribution of Wi-Fi and LTE-U throughput on different channels is highly correlated with the busy degree of the channel, that is, the throughput of LTE-U is large in the idle channels and relatively small in the busy channels. It can also be seen from Fig. 4 that the average throughput of Wi-Fi users in different channels are more balanced and thus the fairness between users is higher.

Fig. 5 and Fig. 6 respectively show the Wi-Fi user average retransmission probability and packet delay. It can be seen that the performances based on two sets of simulation data are similar. The retransmission probability and delay of the Traditional algorithm approach infinite in Channel 1 due to being completely occupied by LTE-U channel. Two kinds of Q-learning based algorithms can achieve better performance both in average retransmission probability and in packet delay. In addition, the data values in the each channels are more balanced, reflecting the presentation of the fairness of the previous algorithm.

Fig. 7 presents the simulation results of the system's fairness factor. Because the coexistence technology is not considered in the Traditional and Average algorithms, the fairness factors obtained by both algorithms are low, which does not conform to the original intention of LTE-U to introduce the unlicensed spectrum. The LBT algorithm aims at access fairness, but exhibits a poor performance in the system fairness coefficient. Limited by too many parameter settings, the CSAT algorithm cannot stably converge at the system optimal time slot ratio. In contrast, both two Q-learning based algorithms can get better performances in the system fairness.

From the above simulation results, we can see that the Q-learning based coexistence algorithm proposed in this paper can greatly enhance the channel utilization by improving the reasonable occupancy ratio in different channels and reducing the retreat caused by collisions. Meanwhile, by taking the fairness factor into the reward function of Q-learning, the proposed algorithm can converge on a high fairness index, which ensures the performance of the Wi-Fi system.

Compared with the Traditional and Average algorithms, or with the current mainstream algorithms, such as LBT and CSAT algorithms, Q-learning based algorithm has better and more balanced performances in the system throughput, fairness coefficient, transmission retransmission probability, transmission delay and other indicators.

The main distinction of the two Q-learning based algorithms is that they have different actions granularity and convergence time. The Joint Q-learning based algorithm can learn according to the performance of the whole network and the fairness coefficient, and consequently the overall performance of the network is better. However, because of the need to simultaneously set a number of unlicensed channel duty cycle, the action set is too large and it will take a long time to search and get convergence. The Independent Q-learning based algorithm performs reinforcement learning on each of the unauthorized channels independently, so that the action set is small and it is easy to achieve convergence. However, the overall performance of the system is slightly worse. Therefore, it is necessary to select the appropriate Q-learning based algorithm according to the actual simulation scene. In the case that the numbers of the unlicensed channels and the network variation are small, the Joint Q-learning based algorithm can be selected. On contrary, when the unlicensed number is large and the network changes quickly, or when rapid convergence is needed, the use of Independent Q-learning based algorithm is more effective.

## V. CONCLUSION

In this paper, the LWCA based on Q-learning is proposed based on the scene of LTE-U. The algorithms calculate and optimize the duty cycle of multiple unlicensed channels by Q- learning. We set up the system-level simulation platform of LTE-U and Wi-Fi heterogeneous network, and the co-existence mechanism algorithms is simulated accordingly. Through the comparison and analysis of the simulation results, we can see that the proposed algorithms can improve the performance of the system while maintaining the fair transmission of LTE-U and Wi-Fi. In our future work, we will consider improving the accuracy of adjustment of actions, such as considering Fuzzy Q-learning. On the other hand, we will further study on convergence speed of learning algorithm.

## REFERENCES

[1] R. Zhang, M. Wang, L. X. Cai, Z. Zheng, X. Shen, and L.-L. Xie, "LTE-unlicensed: The future of spectrum aggregation for cellular networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 150–159, Jan. 2015.

[2] Cisco. (Feb. 7, 2017). *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast, 2016–2021*. Accessed: Apr. 4, 2017. [Online]. Available: http://www.cisco.com/c/en/us/solutions/colla-teral/service-provider/visual-networking-index-vni/vni-forecast-qa.pdf

[3] R. Zhang, Z. Zheng, M. Wang, X. Shen, and L.-L. Xie, "Equivalent capacity in carrier aggregation-based LTE-A systems: A probabilistic analysis," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6444–6460, Nov. 2014.

[4] Qualcomm. (Feb. 2015). *The Prospect of LTE and Wi-Fi Sharing Unlicensed Spectrum*. Accessed: Apr. 4, 2017. [Online]. Available: https://www.qualcomm.com/media/documents/files/the-prospect-of-lte-and-WiFi-sharing-unlicensed-spectrum-srg-whitepaper.pdf

[5] R. Zhang, M. Wang, L. X. Cai, X. Shen, L.-L. Xie, and Y. Cheng, "Modeling and analysis of MAC protocol for LTE-U co-existing with Wi-Fi," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.

[6] LTE-U Forum. (Feb. 2015). *LTE-U Technical Report Coexistence Study for LTE-USDL V1.0*. Accessed: Apr. 4, 2017. [Online]. Available: http://www.lteuforum.org/uploads/3/5/6/8/356-8127/lte-u_forum_lte-u_technical_report_v1.0.pdf

[7] LTE-U Forum. (Feb. 2015). *UE Minimum Requirements for LTE-U SDL V1.0*. Accessed: Apr. 4, 2017. [Online]. Available: http://www.lteuforum.org/uploads/3/5/6/8/3568127/lte-u_forum_ue_minimum_requirements_for_lte-u_sdl_v1.0.pdf

[8] LTE-U Forum. (Oct. 2015). *LTE-U SDL Coexistence Specifications V1.3*. Accessed: Apr. 4, 2017. [Online]. Available: http://www.lteuforum.org/uploads/3/5/6/8/3568127/lte-u_forum_lte-u_sdl_coexistence_specifications_v1.3.pdf

[9] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, 1992.

[10] M. Bennis and D. Niyato, "A Q-learning based approach to interference avoidance in self-organized femtocell networks," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2010, pp. 706–710.

[11] A. El Shafie, T. Khattab, H. Saad, and A. Mohamed, "Optimal cooperative cognitive relaying and spectrum access for an energy harvesting cognitive radio: Reinforcement learning approach," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2015, pp. 221–226.

[12] A. Schwartz, "A reinforcement learning method for maximizing undiscounted rewards," in *Proc. 10th Int. Conf. Mach. Learn.*, vol. 298. 1993, pp. 298–305.

[13] Qualcomm. (Jun. 22, 2016). *Best Use of Unlicensed Spectrum*. Accessed: Apr. 4, 2017. [Online]. Available: https://www.qualcomm.com/media/documents/files/making-the-best-use-of-unlicensed-spectrum-presentation.pdf

[14] *Study on Licensed-Assisted Access to Unlicensed Spectrum (Release 13)*, document 3GPP TR 36.889 V13.0.0, Jun. 2015.

[15] *Study on Licensed-Assisted Access using LTE*, document 3GPP Study Item-RP-141397, Ericsson, Qualcomm, Huawei, Alcatel-Lucent, Edinburgh, Scotland, Sep. 2014.

[16] B. Jia and M. Tao, "A channel sensing based design for LTE in unlicensed bands," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 2332–2337.

[17] A. Al-Dulaimi, S. Al-Rubaye, Q. Ni, and E. Sousa, "5G communications race: Pursuit of more capacity triggers LTE in unlicensed band," *IEEE Veh. Technol. Mag.*, vol. 10, no. 1, pp. 43–51, Mar. 2015.

[18] H. Ko, J. Lee, and S. Pack, "A fair listen-before-talk algorithm for coexistence of LTE-U and WLAN," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10116–10120, Dec. 2016.

[19] X. Du and F. Lin, "Maintaining differentiated coverage in heterogeneous sensor networks," *EURASIP J. Wireless Commun. Netw.*, vol. 5, no. 4, pp. 565–572, Sep. 2005.

[20] X. Du, Y. Xiao, M. Guizani, and H.-H. Chen, "An effective key management scheme for heterogeneous sensor networks," *Ad Hoc Netw.*, vol. 5, no. 1, pp. 24–34, Jan. 2007.

[21] L. Wu, X. Du, J. Wu, and B. Song, "Achieving fair spectrum allocation for co-existing heterogeneous secondary user networks," presented at the Int. Conf. Comput. Commun. Netw. (ICCCN), Vancouver, BC, Canada, Jul. 2017.

[22] L. Huang, G. Zhu, X. Du, and K. Bian, "Stable Multiuser Channel Allocations in Opportunistic Spectrum Access," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013, pp. 1715–1720.

[23] Y. Li, Z. Gao, L. Huang, Z. Tang, and X. Du, "A wideband spectrum data segment compression algorithm in cognitive radio networks," in *Proc. IEEE WCNC*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.

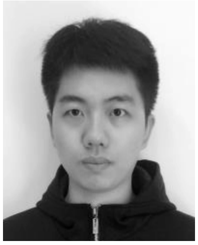
[24] LTE-U Forum. (May 2015). *LTE-U Technology and Coexistence*. Accessed: Apr. 4, 2017. [Online]. Available: http://www.lteuforum.org/uploads/3/5/6/8/3568127/lte-u_coexistence_mechansim_qua-lcomm_may_28_2015.pdf

[25] F. M. Abinader *et al.*, "Enabling the coexistence of LTE and Wi-Fi in unlicensed bands," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 54–61, Nov. 2014.

[26] *Cable Labs Perspective on LTE-U Coexistence With Wi-Fi and Operational Modes for LTE-U*, document 3GPP RAN1 Standard Contribution-RWS-140004, CableLabs, Sophia Antipolis, France, Jun. 2014.

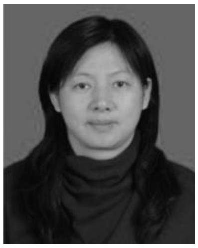
[27] N. Rupasinghe and I. Güvenç, "Reinforcement learning for licensed-assisted access of LTE in the unlicensed spectrum," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 1279–1284.

[28] C. Cano and D. J. Leith, "Coexistence of WiFi and LTE in unlicensed bands: A proportional fair allocation scheme," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 2288–2293.
[29] *Physical layer Procedures (Release 9), Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA)*, doucment 3GPP TS 36.213, V9.3.0, Sep. 2010.
[30] A. Barto and R. S. Sutton, *Reinforcement Learning: An Introduction*, 1st ed. Cambridge, Ma, USA: MIT Press, 1998.
[31] Y.-B. Chen, G.-Y. Lin, and H.-Y. Wei, "A dynamic estimation of the unsaturated buffer in the IEEE 802.11 DCF network: A particle filter framework approach," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5397–5409, Jul. 2016.



**YUHAN SU** received the B.S. degree from Huaqiao University, Xiamen, China. He is currently pursuing the Ph.D. degree in communication and information system with Xiamen University, China. His research interests include wireless communication, congestion control, and network coding.

**XIAOJIANG DU** (SM'09) received the B.S. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1996 and 1998, respectively, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland at College Park in 2002 and 2003, respectively. He is currently a Professor with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA. He received over U.S. $5 million research grants from the U.S. National Science Foundation, the Army Research Office, Air Force, NASA, the State of Pennsylvania, and Amazon. His research interests include security, wireless networks, and systems. He has authored over 250 journals and conference papers in these areas, as well as a book published by Springer. He is a Life Member of the ACM. He received the Best Paper Award at the IEEE GLOBECOM 2014 and the Best Poster Runner-Up Award at the ACM MobiHoc 2014. He served as the Lead Chair of the Communication and Information Security Symposium of the IEEE International Communication Conference 2015 and a Co-Chair of the Mobile and Wireless Networks Track of the IEEE Wireless Communications and Networking Conference 2015.



**LIANFEN HUANG** received the B.S. degree in radio physics and the Ph.D. degree in communication engineering from Xiamen University, Xiamen, China, in 1984 and 2008, respectively.

She was a Visiting Scholar with Tsinghua University, Beijing, China, in 1997, and with The Chinese University of Hong Kong, Hong Kong, in 2012. She is currently a Professor of communication engineering with Xiamen University. Her research interests include wireless communication, wireless network, and signal process.

**ZHIBIN GAO** received the B.S. degree in communication engineering, the M.S. degree in radio physics, and the Ph.D. degree in communication engineering from Xiamen University, Xiamen, China, in 2003, 2006, and 2011, respectively.

He is currently a Senior Engineer of communication engineering with Xiamen University. His research interests include wireless communication, wireless network resource management, and signal processing.

**MOHSEN GUIZANI** (S'85–M'89–SM'99–F'09) received the B.S. (Hons.) and M.S. degrees in electrical engineering and the M.S. and Ph.D. degrees in computer engineering from Syracuse University, Syracuse, NY, USA, in 1984, 1986, 1987, and 1990, respectively. He served as the Associate Vice President of graduate studies with Qatar University, the Chair of the Computer Science Department, Western Michigan University, and the Chair of the Computer Science Department, University of West Florida. He also served in academic positions at the University of Missouri–Kansas City, University of Colorado-Boulder, Syracuse University, and Kuwait University. He is currently a Professor and the ECE Department Chair with the University of Idaho, USA. He is the author of nine books and over 500 publications in refereed journals and conferences. His research interests include wireless communications and mobile computing, computer networks, mobile cloud computing, security, and smart grid. He is a Senior Member of the ACM. He received the Wireless Technical Committee's Recognition Award in 2017. He received the Teaching Award multiple times from different institutions, as well as the best research award from three institutions. He was the Chair of the IEEE Communications Society Wireless Technical Committee and the TAOS Technical Committee. He also served as a member, the chair, and the general chair of a number of international conferences. He served as the IEEE Computer Society Distinguished Speaker from 2003 to 2005. He currently serves on the editorial boards of several international technical journals and the Founder and the Editor-in-Chief of the *Journal of Wireless Communications and Mobile Computing* (Wiley). He guest edited a number of special issues in IEEE journals and magazines.