Two Time-Scale Cross-Layer Scheduling for Cellular/WLAN Interworking

Amila Tharaperiya Gamage, *Student Member, IEEE*, Hao Liang, *Member, IEEE*, and Xuemin (Sherman) Shen, *Fellow, IEEE*

***Abstract*—In this paper, we investigate uplink resource allocation for wireless local area network and cellular network interworking to provide multi-homing voice and data services. The problem is formulated based on the physical layer and medium access control layer technologies of the two networks to ensure that the resource allocation decisions are feasible and can be executed at the lower layers. Furthermore, to efficiently utilize users' equipment (UEs) battery power, the power distribution among multiple network interfaces of the UEs is included in the problem formulation. The optimal resource allocation problem is a multiple time-scale Markov decision process (MMDP) as the two networks operate at different time-scales and due to voice and data service requirements. We derive decision policies for the upper and the lower levels of the MMDP by decomposing each resource allocation problem over multiple time slots to a set of resource allocation problems for individual time slots and solving the resource allocation problems corresponding to individual time slots using convex optimization techniques. To reduce the time complexity, we further propose a heuristic resource allocation algorithm by deriving the decision policies based on a single system state. The system state consists of average square channel gains for dual variable calculation and instantaneous channel gains for resource allocation based on the calculated dual variables. Simulation results demonstrate the achievable throughput and service quality improvements by employing these two algorithms.**

***Index Terms*—Cellular network, cross-layer scheduling, interworking, multiple time-scale Markov decision process (MMDP), wireless local area network (WLAN).**

I. Introduction

THE monumental growth of smart mobile devices in recent years has exponentially increased the demand for higher data rates with seamless service coverage and support for diverse quality-of-service (QoS) requirements of heterogeneous services [1]. Interworking of wireless networks and multi-homing capability of the users' equipment (UEs) can be utilized to satisfy the capacity, coverage and QoS requirements.

Most of high service demanding areas, such as hotspots, office buildings and airports, are covered by cellular networks and wireless local area networks (WLANs). Also, the recent WLAN technologies have incorporated various mechanisms, such as the hybrid coordination function (HCF) in IEEE 802.11n WLANs, to facilitate QoS support for different services [2]. As interworking mechanisms can be utilized for jointly allocating resources of multiple wireless networks, interworking can be used at such a high service demand environment to improve the network capacity and coverage area. Furthermore, interworking can provide users with enhanced QoS by utilizing resources from multiple networks in an optimal manner based on the user requirements and the attributes of the networks, such as supporting QoS and mobility levels [3], [4]. Multi-homing capability of UEs with multiple radio interfaces allows the UEs to simultaneously communicate over multiple networks, and it can be used for further improving the efficiency of resource utilization in the interworking system [5].

In this work, we investigate uplink resource allocation for cellular/WLAN interworking in the presence of UE multi-homing capability to achieve QoS satisfaction while maximizing the system throughput. The cellular network is based on orthogonal frequency division multiple access (OFDMA) and the WLAN operates on both contention-based and contention-free polling based channel access mechanisms. Resources of the system are subcarriers of the cellular network, transmission opportunities (TXOPs) via the two channel access mechanisms of the WLAN, and UE transmit power. Use of this kind of system model enables the algorithms which are developed in this work to allocate resources for interworking latest/next-generation wireless networks, such as LTE/LTE-A cellular networks and IEEE 802.11n/802.11ac WLANs.

One key challenge for resource allocation for cellular/WLAN interworking is the high complexity due to existence of multiple physical layer (PHY) and medium access control layer (MAC) technologies. The region of feasible transmission rates depends on PHY and MAC technologies of different networks [6]. Therefore, resource allocation should capture diverse PHY and MAC technologies of the networks. Furthermore, resource allocation intervals (i.e., interval between two successive resource allocations) of existing cellular networks are usually shorter than those of existing WLANs [2], [7]. Therefore, the resource allocation algorithms for cellular/WLAN interworking should be designed to periodically allocate cellular network and WLAN resources with shorter and longer periods respectively, where the periods correspond to the resource allocation intervals of the networks. That is, resources of the two networks are allocated at faster and slower time-scales, respectively [8].

Manuscript received January 9, 2014; revised May 18, 2014; accepted July 1, 2014. Date of publication July 10, 2014; date of current version August 20, 2014. This work was presented in part at the IEEE International Conference on Communications, Budapest, Hungary, June 2013. The editor coordinating the review of this paper and approving it for publication was L. Badia.

A. Tharaperiya Gamage and X. Shen are with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: amila.gamage@uwaterloo.ca; sshen@uwaterloo.ca).

H. Liang is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: hao2@ualberta.ca).

Digital Object Identifier 10.1109/TCOMM.2014.2337890

Our contribution in this work is threefold. First, we propose a resource allocation framework for cellular/WLAN interworking operating on two time-scales. Second, we formulate the resource allocation problem in the proposed framework as a multiple time-scale Markov decision process (MMDP) based on the PHY and MAC technologies of the two networks, and derive decision policies for the MMDP. Third, we propose a heuristic resource allocation algorithm for low time complexity.

The remainder of this paper is organized as follows. Section II summarizes related works and Section III describes the interworking system model. Section IV presents the MMDP based resource allocation. Sections V and VI discuss resource allocations at upper and lower levels of the proposed framework, respectively. The heuristic resource allocation algorithm is presented in Section VII, while simulation results and the conclusions are given in Sections VIII and IX, respectively.

II. Related Work

Existing resource allocation schemes can be classified into three categories: schemes using a single network interface of each UE at any given time [9]–[11], schemes utilizing the multi-homing capability of UEs [5], [12]–[14], and schemes that are designed based on different PHY and MAC technologies [4], [15]. A load balancing scheme to improve resource utilization in cellular/WLAN interworking is presented in [9]. New voice and data calls are assigned to a network based on a set of precalculated probabilities. Assigned calls are redistributed whenever necessary to another network by using dynamic vertical handoffs to reduce network congestion and improve QoS satisfaction. To further improve QoS satisfaction, the scheme proposed in [10] allocates voice calls preferably for the cellular network. The resource allocation scheme proposed for WiMAX/WLAN interworking in [11] assigns all streaming calls to the WiMAX network to guarantee QoS satisfaction; data calls that are served by the WiMAX network are preempted to free up bandwidth for the incoming streaming calls when required. The main advantage of these schemes in the first category is that they are easy to deploy as each network can use its own/existing resource allocation scheme to allocate resources. Further, designing an efficient resource allocation scheme is simpler for an individual network than for an interworking system.

When UEs are capable of multi-homing, restricting a UE or a certain traffic type of a user to access only one network limits the flexibility in distributing resources of the interworking system among users. Thus, the resource allocation schemes in the second category take advantage of the multi-homing capability of UEs to efficiently utilize resources of the interworking system. For computational simplicity, it is typically assumed that the WLAN uses a resource reservation protocol to avoid channel contention collisions. Hence, resources of the WLAN are modeled as frequency channels or time slots. Bandwidth allocation algorithms for UEs with different types of traffic requirements are presented in literature. In [5], each network gives more priority to satisfy its own subscribers' QoS requirements, while utility fairness among users in the interworking system is maintained in [12]. A game theoretic approach for bandwidth allocation and admission control is used in [13]. Each network allocates its bandwidth for different service areas on a long-term basis based on the statistics of call arrivals; bandwidths for each service area from different networks are then allocated to users on a short-term basis. To ensure QoS satisfaction, a new call is accepted only if its minimum data rate requirement can be satisfied. Algorithms to allocate time slots in a WLAN and subcarriers in a cellular network subject to a proportional rate constraint are presented in [14].

The third category includes the resource allocation schemes proposed in [4], [15]. These schemes are based on PHY and MAC technologies of the different networks to guarantee the feasibility of resource allocation decisions. Specifically, the effect of transmission collisions caused by the contention-based channel access in the WLAN is considered. In [4], resource allocation and admission control schemes are proposed for an interworking system consisting of a code division multiple access (CDMA) based cellular network and an IEEE 802.11 distributed coordination function (DCF) based WLAN. Maximizing total network welfare ensures QoS satisfaction in the system. In [15], interworking of an OFDMA based femtocell network and an IEEE 802.11 DCF based WLAN is considered. Resources of both femtocell and WLAN are allocated on the same time-scale, and WLAN uses basic access scheme with two-way handshaking.

The existing resource allocation schemes allocate resources of different networks in the interworking system at the same time-scale, and do not fully utilize the QoS support in WLANs. Allocating resources of different networks at the same time-scale is not practical as different networks have different resource allocation intervals. To facilitate QoS in WLANs, recent WLAN standards offer contention-based and contention-free polling based channel access mechanisms. These two channel access mechanisms and their QoS capabilities should be considered to maximize the efficiency of the interworking system. In addition, jointly allocating transmit power levels for different network interfaces at multi-homing capable UEs is essential for an efficient resource utilization. Joint transmit power allocation is studied in [15] without taking the user QoS requirements into account.

In this work, we study the resource allocation for cellular/WLAN interworking to satisfy the QoS requirements of multi-homing UEs. Based on the PHY and MAC technologies of these two networks, the resources are allocated to multi-homing UEs at two time-scales: one time-scale for allocating resources of each network. We consider power allocation for multi-homing UEs, and the two channel access mechanisms of the WLANs.

III. System Model

The interworking system under consideration consists of a single cell of a cellular network and a WLAN within the coverage of the cell, as shown in Fig. 1. We focus on the resource allocation for the uplink. The cellular network is OFDMA based, and the set of subcarriers available at the base station (BS) is denoted by $\mathcal{K}^C$. At any time, each subcarrier is allocated to only one user in order to avoid co-channel interference

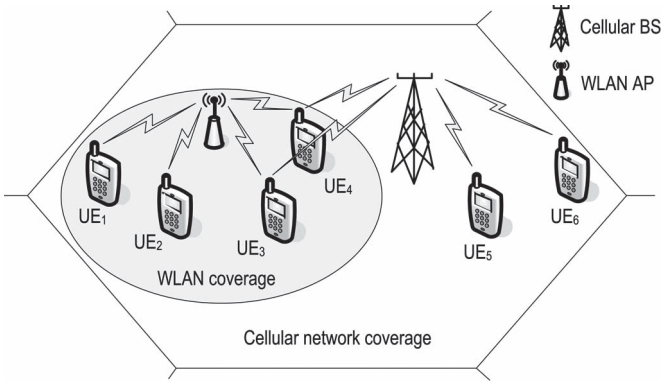


Fig. 1. Cellular/WLAN interworking.

among the users. The WLAN supports both contention-based and contention-free polling based channel access for data and voice services. In the contention-based channel access, four-way handshaking scheme with request-to-send (RTS) and clear-to-send (CTS) messages is used. The time periods during which the two channel access mechanisms operate are referred to as contention period (CP) and contention-free period (CFP), respectively. The CP and CFP alternate over time and they repeat once every T_P. Resources of the cellular network and the WLAN are allocated at two different time-scales.

In the system, there are N users belonging to two groups: high-mobility users and low-mobility users. The set of all the users is denoted by $\mathcal{S}_N$. The set of low-mobility users within the WLAN coverage is denoted by $\mathcal{S}_M$, while the set of remaining users is denoted by $\mathcal{S}_S$. For example, in Fig. 1, UE_1 to UE_4 are in $\mathcal{S}_M$, while UE_5 and UE_6 are in $\mathcal{S}_S$. Each user has voice and data traffic requirements. All the UEs are equipped with WLAN and cellular network interfaces, and have the multi-homing capability. Users in $\mathcal{S}_M$ are allowed to simultaneously communicate over cellular network and WLAN, while users in $\mathcal{S}_S$ are only allowed to communicate over the cellular network.

A. Two Time-Scale Resource Allocation Framework

Resource allocation intervals of existing cellular networks are shorter than those of the existing WLANs, as cellular networks and WLANs are designed to support high mobility and low mobility users, respectively [2], [7]. Therefore, as shown in Fig. 2, resources in the cellular network are allocated at a faster time-scale than in the WLAN. The duration of a time slot in a time-scale is the resource allocation interval of the corresponding network, denoted by T^L and T^U in the fast and slow time-scales ($T^L < T^U$) for the cellular network and the WLAN, respectively. The resource allocation processes at fast and slow time-scales are referred to as lower and upper levels of the resource allocation process, respectively.

As the WLAN resource allocation interval is relatively long, to satisfy the strict delay and jitter requirements of periodically arriving constant bit rate voice traffic, several short CFPs are used within a resource allocation interval of the WLAN instead of using a long CFP [16]. For simplicity, assume $V_L(= T^U/T^L)$ is an integer and the boundaries of the first time slots in the two time-scales are aligned.

B. Symbols and Notations

The lth ($l \in \{0, 1, \ldots, V_L - 1\}$) fast time-scale time slot within the uth slow time-scale time slot is referred to as (u, l)th time slot. Commonly used symbols are written in the form of $X^n_{i,y}$ or $X^n_{i,y}(\cdot)$, where superscript n, $n \in \{C, W, CB, CF, L, U\}$, represents the network or the level of resource allocation process. Superscripts C, W, CB and CF denote the cellular network, WLAN, contention-based channel access and contention-free channel access respectively, while L and U denote lower and upper levels of the resource allocation process respectively. The subscripts denote the user, a particular resource of network n, and a time slot. When $n \in \{W, CB\}$, only one subscript is used representing the user. Boldface letters are used for vectors and matrices, and vector $\mathbf{X}$ is represented as $\mathbf{X} = \{X_1, \ldots, X_{|\mathbf{X}|}\}$ with $|\mathbf{X}|$ being the number of elements in $\mathbf{X}$. The optimum value of variable X is denoted by X^*. The active (or the determined) decision policy $\mathbb{X}$ and the optimal set $\mathbf{X}$ are denoted by $\mathbb{X}^*$ and $\mathbf{X}^*$, respectively. Table I summarizes the important symbols.

C. WLAN

In the WLAN, TXOPs for the UEs are granted using two channel access mechanisms: contention-based channel access during a CP and contention-free polling based channel access during a CFP. In the former, UEs contend for the channel to obtain TXOPs, and each of these TXOPs allows transmitting a data packet of D bits. In the latter, UEs are granted TXOPs using a centralized polling mechanism, and each of these TXOPs is defined as a fixed duration of T_{CF} which allows a UE to transmit. The set of available contention-free TXOPs during a CFP is denoted by $\mathcal{K}^{CF}$. To avoid co-channel interference among the UEs, each contention-free TXOP is allocated for only one UE at any given time. Contention-based channel access is more suitable for variable bit rate data traffic, while contention-free channel access is more suitable for constant bit rate voice traffic [17]. In this work, to optimize resource utilization subject to QoS requirements, voice traffic is served by contention-free channel access, and data traffic is served by both channel access mechanisms. The sets of users communicate using contention-based and contention-free channel access are denoted by $\mathcal{S}^{CB}$ and $\mathcal{S}^{CF}$ respectively, where $\mathcal{S}^{CB}$, $\mathcal{S}^{CF} \subseteq \mathcal{S}_M$ and possibly $\mathcal{S}^{CB} \bigcap \mathcal{S}^{CF} \neq \emptyset$.

D. Traffic Model

The traffic generated by each user can be divided into two classes: constant bit rate voice and delay tolerant data. Every user always has at least one packet in the data traffic queue to transmit. The minimum data rates of voice and data traffic classes required by the ith user are denoted by $R_{Vmin,i}$ and $R_{Dmin,i}$, respectively. As voice traffic flows are highly susceptible to delay and jitter, voice traffic requirements are satisfied in average sense over each time slot at the slow time-scale. The data traffic requirements are satisfied in average sense over an infinite time horizon due to their delay tolerance.

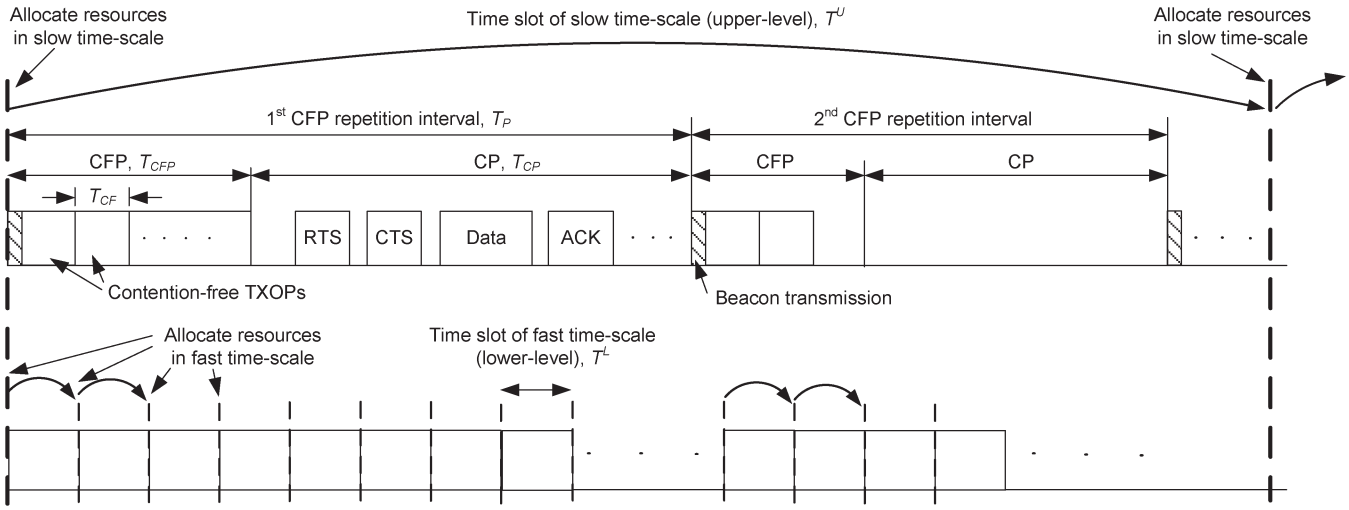


Fig. 2. Resource allocation at slow and fast time-scales.

TABLE I
SUMMARY OF IMPORTANT SYMBOLS

Symbol	Description
$A^L_{u,l}$	Resource allocation decisions made at the lower-level for the (u,l)th time slot
A^U_u	Resource allocation decisions made at the upper-level for the uth time slot
B^X	Bandwidth of network $X,\ X \in \{C, W\}$
$\mathbb{D}$	System policy $\{\mathbb{D}^U, \mathbb{D}^L\}$
$\mathbb{D}^X$	Policy of $X,\ X \in \{U, L\}$
$\mathcal{K}^X$	Set of resources (e.g., OFDM subcarriers or contention-free TXOPs) available at network $X,\ X \in \{C, CF\}$
N	Number of users in the interworking system
N_W	Number of users allocated for the contention-based channel access of WLAN
$P^X_{i,y}$ or P^X_i	Transmit power level of ith user over (yth resource of) $X,\ X \in \{C, CB, CF\}$
$P^X_{avg,i}$ or $P^X_{avg,i}(\cdot)$	Average power usage of ith user over $X,\ X \in \{C, CB\}$ during one time slot in slow time-scale
$r^L_{i,u,l}(\cdot)$	Throughput achieved by the ith user at the lower-level during the (u,l)th time slot
$r^U_{i,u}(\cdot)$	Throughput achieved by the ith user at the upper-level during the uth time slot
$R^X_{i,y}(\cdot)$ or $R^X_i(\cdot)$	Throughput achieved by ith user over (yth resource of) $X,\ X \in \{C, CB, CF\}$
$R^L_{i,u}(\cdot)$	SDT achieved by the ith user at the lower-level over the uth time slot
$R^U_i(\cdot)$	SDT achieved by the ith user at the upper-level over an infinite time horizon
$R_{Dmin,i}$, $R_{Vmin,i}$	Minimum data rate required for data and voice traffic services of ith user
$\mathcal{S}^X$	Set of users communicating through $X,\ X \in \{C, CB, CF\}$
$\mathcal{S}_M$, $\mathcal{S}_S$	Set of low-mobility users within WLAN coverage, and the set of remaining users
$\mathcal{S}_N$	Set of all the users in the system
T^X	Duration of a time slot in $X,\ X \in \{U, L\}$
T_{CP}, T_{CF}, T_{CFP}, T_P	Durations of CP, contention-free TXOP, CFP and CFP repetition period
T_{coh}	Coherence time of a wireless channel
V_L	Number of time slots in fast time-scale within one time slot of the slow time-scale
$\alpha^X_{i,y}$ or α^X_i	SINR of the channel between ith user and BS of $X,\ X \in \{C, W\}$ (over yth resource) with unit transmit power
β, θ	Discount factors of upper and lower levels
Δf	Bandwidth of an OFDM subcarrier
$\rho^X_{i,y}$	$\rho^X_{i,y} = 1$ if yth resource of $X,\ X \in \{C, CF\}$ is allocated for ith user, or else $\rho^X_{i,y} = 0$
$\psi_{u,l}$	System state $\{\psi^U_u, \psi^L_{u,l}\}$ during time slot (u,l)
$\psi^L_{u,l}$	State of the lower-level during the (u,l)th time slot
ψ^U_u	State of the upper-level during the uth time slot
Ω	Average square channel gain

E. Channel Model

We model the wireless channels as a finite-state Markov process to capture the channel time-correlation [18]. The channel gain is time invariant (i.e., quasi-static fading) within each coherence time (T_{coh}) interval. The different wireless channels vary independently from each other. The channel gain domain is partitioned into K_S non-overlapping states. The transition probabilities between different states of a Rayleigh fading channel can be calculated as in [18], assuming that T^U and T^L are not longer than the corresponding channel coherence times to ensure the states do not change within a time slot.

F. Throughputs Per User

1) Via Cellular Network and Contention-Free Channel Access: The maximum achievable error free data rate by the ith user using network $n \in \{C, CF\}$ can be expressed by

$$R^n_{i,y}\left(P^n_{i,y}\right) = \sum_{y \in \mathcal{K}^n} \rho^n_{i,y} B \log_2\left(1 + \alpha^n_{i,y} P^n_{i,y}\right), \qquad (1)$$

where $\alpha^n_{i,y}$ is the signal-to-interference plus noise ratio (SINR) of the channel between cellular BS/WLAN access point (AP) and the ith user over the yth resource of n, i.e., the yth OFDM subcarrier or TXOP, with unit transmit power; $P^n_{i,y}$ is the

transmit power level of the ith user over the yth resource of n; B represents the bandwidth of WLAN (B^W) or bandwidth of an OFDM subcarrier ($\Delta f = B^C/|\mathcal{K}^C|$); B^C is the system bandwidth of the cell; and $\rho^n_{i,y} = 1$ if the ith user is allocated the yth resource of n, and $\rho^n_{i,y} = 0$ otherwise. Furthermore, $\rho^{CF}_{i,y} = 0, \forall y$ if $i \notin \mathcal{S}_M$. As each resource is allocated to only one user to avoid co-channel interference,

$$\sum_{i \in \mathcal{S}_N} \rho^n_{i,y} \leq 1, \quad \forall y \in \mathcal{K}^n. \tag{2}$$

In addition, $\alpha^{CF}_{i,y} = \alpha^{CB}_i = \alpha^W_i, \forall y$ over each channel coherence time interval in the WLAN.

2) Via Contention-Based Channel Access: The average throughput achieved by the ith user during a CP in the WLAN with four-way handshaking scheme is given by [2], [19]

$$R^{CB}_i(\mathbf{L}) = \frac{\tau(1-\tau)^{N_W-1} D}{T_0 + N_W \tau(1-\tau)^{N_W-1} \sum_{j \in \mathcal{S}^{CB}} L_j}, \ \forall i \in \mathcal{S}^{CB} \tag{3}$$

with

$$T_0 = N_W \tau(1-\tau)^{N_W-1}(T_{CTS} + T_{ACK} + 3T_{SIFS}) + (1-(1-\tau)^{N_W})(T_{RTS} + T_{AIFS}) + (1-\tau)^{N_W}\sigma,$$

where L_i is the duration of a packet transmission by the ith user; N_W is the number of users in $\mathcal{S}^{CB}$; and T_{SIFS}, T_{AIFS}, T_{ACK}, T_{RTS}, T_{CTS} and σ are the durations of short interframe space, arbitration interframe space, acknowledgment, RTS message, CTS message and an empty slot, respectively. The probability τ of a user transmitting a packet in a randomly chosen time slot can be calculated as in [19]. The throughput $R^{CB}_i(\mathbf{L})$ given by (3) cannot be directly used in a resource allocation problem as it is not written as a function of the transmit power levels. Therefore, we rewrite L_i in terms of the throughput that is achieved by the ith user during a successful transmission, and substitute it into (3) [15]. Then, $R^{CB}_i(\mathbf{L})$ can be rewritten as

$$R^{CB}_i(\mathbf{P}^{CB}) = \frac{\tau(1-\tau)^{N_W-1} D}{T_0 + \frac{D N_W \tau(1-\tau)^{N_W-1}}{B^W} \sum_{j \in \mathcal{S}^{CB}} \frac{1}{\log_2(1+P^{CB}_j \alpha^W_j)}}, \tag{4}$$

where P^{CB}_i is the transmit power level of the ith UE during a CP over the WLAN interface. Note that $R^{CB}_i(\mathbf{P}^{CB})$ is the same for all the users in $\mathcal{S}^{CB}$ and is a concave function when N_W is fixed (see Appendix A).

G. Power Usage of Multi-Homing Devices

The operating time of a UE is governed by the energy (or average power) consumption of the uplink communications through WLAN and cellular interfaces of the UE [5]. Therefore, we limit the total average power consumption of each UE over each time slot in the slow time-scale to a predefined maximum. We first calculate the average power usage through the WLAN interface for contention-based channel access, and then formulate the constraint on the total average power consumption. With a successful transmission probability of $\tau(1-\tau)^{N_W-1}$ during a CP, the average power consumption of the ith UE through the WLAN interface is $P^{CB}_{avg,i}(\mathbf{P}^{CB}) = \tau(1-\tau)^{N_W-1}(L_i P^{CB}_i / T_{avg})(T_{CP}/T_P)$, where T_{avg} is the average duration of channel occupancy for a successful packet transmission including collision period and empty slot in which no UE transmits [19]. After some simplification, $P^{CB}_{avg,i}(\mathbf{P}^{CB})$ can be expressed as

$$P^{CB}_{avg,i}(\mathbf{P}^{CB}) = \begin{cases} \frac{T_{CP} P^{CB}_i R^{CB}_i(\mathbf{P}^{CB})}{T_P B^W \log_2(1+\alpha^W_i P^{CB}_i)}, & \text{if } P^{CB}_i > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

The constraint on the total average power consumption of each UE over the uth time slot can then be expressed as

$$P^C_{avg,i} + P^{CB}_{avg,i}(\mathbf{P}^{CB}) + \frac{T_{CF}}{T_P} \sum_{j \in \mathcal{K}^{CF}} \rho^{CF}_{i,j} P^{CF}_{i,j} \leq P_{T,i}, \forall i \in \mathcal{S}_N, \tag{6}$$

where $P^C_{avg,i}$ is the average power usage through the cellular interface during the time slot and $P_{T,i}$ is the total average power available for the ith user.

IV. MMDP-Based Optimal Resource Allocation

The objective of resource allocation is to maximize the total throughput of the interworking system subject to the satisfaction of QoS requirements. As discussed in Section III-A, the resource allocation process consists of two (upper and lower) levels operating at slow and fast time-scales respectively, based on the channel state information. Resources of the WLAN and the cellular network are allocated at the beginnings of the uth and the (u,l)th time slots respectively, where $u = \{0, 1, 2, \ldots\}$ and $l = \{0, \ldots, V_L - 1\}$. The set of channel gains of the channels between users in $\mathcal{S}_M$ and the WLAN AP at the beginning of the uth time slot is referred to as the state of the upper-level during the uth time slot (ψ^U_u). The set of channel gains of the channels between all the users and the cellular BS at the beginning of the (u,l)th time slot is referred to as the state of the lower-level during the (u,l)th time slot ($\psi^L_{u,l}$). While the system state $\{\psi^U_u, \psi^L_{u,l}\}$ is denoted by $\psi_{u,l}$, the sets of all the possible states of upper and lower levels are denoted by Ψ^U and Ψ^L, respectively.

An overview of the resource allocation framework is shown in Fig. 3. The optimal resource allocation problem for cellular/WLAN interworking is formulated as an MMDP [8] for three reasons: 1) the resource allocation process operates at two time-scales as explained in Section III-A, 2) state transition of each level is a Markov process due to the Markov channel model, and 3) resource allocations at multiple time slots are jointly optimized to satisfy the user QoS requirements over multiple time slots (see Section III-D).

The MMDP formulation consists of upper and lower level resource allocation policies [8]. As shown in Fig. 3, the decisions of the upper-level are made considering the throughputs achieved through and the power consumed at the lower-level. Therefore, the upper-level policy ($\mathbb{D}^U$) maps system state $\psi_{u,0}$ to a set of resource allocation decisions (A^U_u) at the beginning of uth time slot, $u = \{0, 1, 2, \ldots\}$. The lower-level policy ($\mathbb{D}^L$) maps state $\psi^L_{u,l}$ to a set of resource allocation decisions ($A^L_{u,l}$) at the beginning of (u,l)th time slot, $l = \{0, \ldots, V_L - 1\}$. Decisions in A^U_u and $A^L_{u,l}$ are $\{P^{CB}_i, P^{CF}_{i,j}, \rho^{CF}_{i,j} | \forall i \in \mathcal{S}_M, j \in$

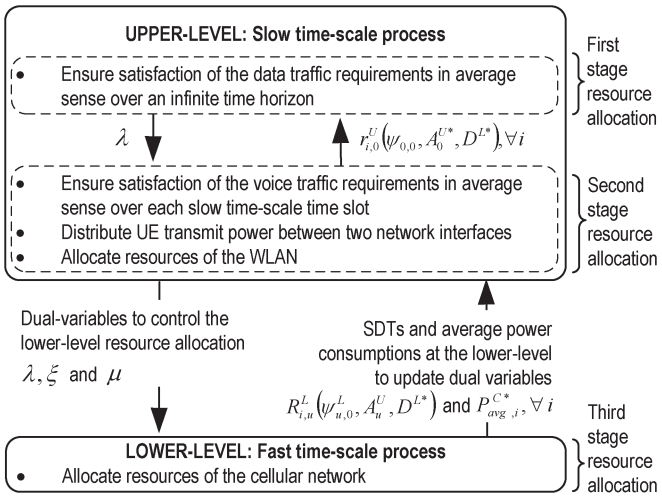


Fig. 3. Overview of the MMDP-based two time-scale resource allocation framework.

$\mathcal{K}^{CF}\}$ and $\{P_{i,k}^{C}, \rho_{i,k}^{C} | \forall i \in \mathcal{S}_N, k \in \mathcal{K}^{C}\}$, respectively. For notation simplicity, we use $\mathbb{D}$ to denote the system policy $\{\mathbb{D}^U, \mathbb{D}^L\}$.

We use the summation of discounted throughputs (SDTs) [20], [21] over an infinite time horizon as a reward (objective) function. The SDT based reward function reduces the susceptibility of the determined decision policies to the unpredictable channel changes in the future by giving less importance to those decisions made (and rewards achieved) at far future. The SDTs achieved by the ith user at the upper-level over an infinite time horizon with the initial state of $\psi_{0,0}$ and at the lower-level during the uth time slot with the initial state of $\psi_{u,0}^{L}$ are denoted by $R_i^U(\psi_{0,0}, \mathbb{D})$ and $R_{i,u}^{L}(\psi_{u,0}^{L}, A_u^U, \mathbb{D}^L)$, respectively [8]. As the decision policies are stationary (to be discussed), $R_i^U(\psi_{0,0}, \mathbb{D})$ and $R_{i,u}^{L}(\psi_{u,0}^{L}, A_u^U, \mathbb{D}^L)$ can be interpreted as the average throughputs that are achieved by the ith user over the same periods of time at the upper and lower levels, respectively [21]. They are given by [20], [21]

$$R_i^U(\psi_{0,0}, \mathbb{D}) = \lim_{V_U \to \infty} (1-\theta) \sum_{u=0}^{V_U - 1} \theta^u r_{i,u}^{U}(\psi_{u,0}, A_u^U, \mathbb{D}^L) \quad (7)$$

$$R_{i,u}^{L}\left(\psi_{u,0}^{L}, A_u^U, \mathbb{D}^L\right) = (1-\beta) \sum_{l=0}^{V_L - 1} \beta^l r_{i,u,l}^{L}\left(\psi_{u,l}^{L}, A_u^U, A_{u,l}^{L}\right), \quad (8)$$

where $\theta \in (0,1)$ and $\beta \in (0,1)$ are discount factors; and $r_{i,u}^{U}(\psi_{u,0}, A_u^U, \mathbb{D}^L)$ and $r_{i,u,l}^{L}(\psi_{u,l}^{L}, A_u^U, A_{u,l}^{L})$, which denote the throughputs achieved by the ith user at the upper and lower levels during the uth and (u,l)th time slots respectively, are given by [22]

$$r_{i,u}^{U}\left(\psi_{u,0}, A_u^U, \mathbb{D}^L\right) = \begin{cases} R_{i,u}^{L}\left(\psi_{u,0}^{L}, A_u^U, \mathbb{D}^L\right), & \text{if } i \in \mathcal{S}_S; \\ R_{i,u}^{L}\left(\psi_{u,0}^{L}, A_u^U, \mathbb{D}^L\right) & \\ + \frac{T_{CF}}{T_P} \sum_{j \in \mathcal{K}^{CF}} \rho_{i,j}^{CF} R_{i,j}^{CF}\left(P_{i,j}^{CF}\right), & \text{if } i \in \mathcal{S}_M \setminus \mathcal{S}^{CB}; \\ R_{i,u}^{L}\left(\psi_{u,0}^{L}, A_u^U, \mathbb{D}^L\right) + \frac{T_{CF}}{T_P} \sum_{j \in \mathcal{K}^{CF}} \rho_{i,j}^{CF} R_{i,j}^{CF}\left(P_{i,j}^{CF}\right) & \\ + \frac{T_{CP}}{T_P} R_i^{CB}(\mathbf{P}^{CB}), & \text{if } i \in \mathcal{S}^{CB}; \end{cases} \quad (9)$$

and

$$r_{i,u,l}^{L}\left(\psi_{u,l}^{L}, A_u^U, A_{u,l}^{L}\right) = \sum_{k \in \mathcal{K}^C} \rho_{i,k}^{C} R_{i,k}^{C}\left(P_{i,k}^{C}\right). \quad (10)$$

The data traffic requirements of the users are served through both networks while the voice traffic requirements are served through the contention-free channel access and the cellular network. Therefore, the QoS constraints (see Section III-D), which ensure data and voice traffic requirement satisfaction over an infinite time horizon and over the uth time slot ($u = \{0, 1, 2, \ldots\}$) respectively, can be stated as

$$R_i^U(\psi_{0,0}, \mathbb{D}) \geq R_{Vmin,i} + R_{Dmin,i}, \quad \forall i \in \mathcal{S}_N \quad (11)$$

$$R_{i,u}^{L}\left(\psi_{u,0}^{L}, A_u^U, \mathbb{D}^L\right) + \frac{T_{CF}}{T_P} \sum_{j \in \mathcal{K}^{CF}} \rho_{i,j}^{CF} R_{i,j}^{CF}\left(P_{i,j}^{CF}\right) \geq R_{Vmin,i}, \quad \forall i \in \mathcal{S}_N. \quad (12)$$

As the sum of discounted costs provides a good approximation for the average cost when the policies are stationary [21], $P_{avg,i}^{C}$ over the uth time slot can be calculated by

$$P_{avg,i}^{C} = (1-\beta) \sum_{l=0}^{V_L - 1} \beta^l P_{tot,i,l}^{C}\left(\psi_{u,l}^{L}, A_u^U, A_{u,l}^{L}\right), \quad (13)$$

where $P_{tot,i,l}^{C}(\psi_{u,l}^{L}, A_u^U, A_{u,l}^{L})$ is the total power allocated by the ith user to communicate over the cellular network during the (u,l)th time slot, and is also equivalent to $\sum_{k \in \mathcal{K}^C} \rho_{i,k}^{C} P_{i,k}^{C}$ over the (u,l)th time slot.

The MMDP based optimal resource allocation problem can then be stated as [8], [21]

$$\mathcal{P}1: \max_{\mathbb{D}^U} \max_{\mathbb{D}^L} \sum_{i \in \mathcal{S}_N} R_i^U(\psi_{0,0}, \mathbb{D})$$
$$\text{s.t.} \quad (2) \text{ for } n \in \{C, CF\}, (6), (11) \text{ and } (12).$$

To find the optimal $\mathbb{D}^U$ and $\mathbb{D}^L$ solving problem $\mathcal{P}1$, resource allocation should be optimized over three different time intervals: 1) resource allocation over an infinite time horizon is optimized to satisfy (11), 2) resource allocation over each upper-level time slot is optimized to optimally use upper-level resources while satisfying (6) and (12), and 3) resource allocation over each lower-level time slot is optimized to optimally use lower-level resources. Therefore, problem $\mathcal{P}1$ is solved in three stages, where the first, second and third stages allocate resources over an infinite time horizon, for each upper-level time slot, and for each lower-level time slot, respectively. The resource allocation problem for the mth stage ($m = \{2, 3\}$) is derived by decomposing the $(m-1)$th stage problem into a set of problems, each of which allocates resources over the resource allocation interval of the mth stage, and by imposing constraints that must be satisfied within the resource allocation interval of the mth stage.

The optimality of the solution, which is obtained using the three stage approach, for problem $\mathcal{P}1$ is ensured by iterating the mth stage ($m = \{1, 2\}$) solution until it reaches the optimal while calculating the optimum $(m+1)$th stage solution for each mth stage iteration. During the iteration process, the dual variables of the mth stage are passed to the $(m+1)$th stage while the throughputs/SDTs achieved and power consumed at the $(m+1)$th stage are feedback to the mth stage. At the $(m+1)$th stage, the received dual variables are used for

configuring the objective function of the resource allocation problem such that the $(m+1)$th stage assists maximizing the mth stage objective. At the mth stage, the received information is used for updating the dual variables.

Problem $\mathcal{P}1$ is a non-convex problem. Therefore, we determine the policies by relaxing problem $\mathcal{P}1$ to reduce the computational complexity. Due to the relaxations, the policies determined in this work ($\mathbb{D}^{U*}$ and $\mathbb{D}^{L*}$) are not optimal for problem $\mathcal{P}1$ in certain scenarios. Therefore, we refer to $\mathbb{D}^{U*}$ and $\mathbb{D}^{L*}$ as active (or determined) upper and lower level decision policies, respectively. Derivations of $\mathbb{D}^{U*}$, which is found by solving the first and second stage resource allocation problems, and $\mathbb{D}^{L*}$, which is found by solving the third stage resource allocation problem, are discussed in Sections V and VI, respectively.

Using the state transition probabilities calculated based on the channel statistics, $\mathbb{D}^{U*}$ and $\mathbb{D}^{L*}$ can be determined in advance and applied to the system based on the initial states. The applied policies select A_u^{U*} and $A_{u,l}^{L*}$ for the uth and the (u,l)th time slots respectively, based on the states of the two levels during the time slots. The policies $\mathbb{D}^{U*}$ and $\mathbb{D}^{L*}$ are required to be recalculated when the channel statistics or the number of users in the system or their QoS requirements change.

V. Upper-Level Resource Allocation

To determine $\mathbb{D}^{U*}$, first we maximize the total SDT at the upper-level subject to satisfaction of (11) over an infinite time horizon with the initial system state of $\psi_{0,0}$; this first stage problem is denoted by $\mathcal{P}2$. By further investigating problem $\mathcal{P}2$, we find out that problem $\mathcal{P}2$ is a convex optimization problem and can be solved by solving the dual problem [23]. To find the dual function, the minimum of the Lagrangian is determined by decomposing the Lagrangian into a set of terms, each of which is a negative summation of weighted throughputs of the users corresponding to one time slot. Then, A_u^{U*} for the uth time slot $(u=\{0,1,2,\ldots\})$ is determined such that it maximizes the summation of weighted throughputs corresponding to the uth time slot subject to satisfaction of (2) for $n=CF$, (6) and (12); this second stage problem is denoted by $\mathcal{P}3$. First and second stage problems are solved in Section V-A and B, respectively. In addition, the conditions which the third stage resource allocation at the lower-level should satisfy to ensure the optimality of the three stage solution for problem $\mathcal{P}1$ are derived in Section V-B.

A. First Stage Resource Allocation

First stage resource allocation problem can be stated as

$$\mathcal{P}2: \quad \max_{\mathbb{D}^U} \quad \sum_{i\in\mathcal{S}_N} R_i^U\left(\psi_{0,0},\mathbb{D}^U,\mathbb{D}^{L*}\right) \quad \text{s.t.} \quad \mathrm{C1}:(11).$$

The active policy $\mathbb{D}^{L*}$ is used in problem $\mathcal{P}2$ as for each iteration of the algorithm which solves problem $\mathcal{P}2$, $\mathbb{D}^{L*}$ is calculated by solving the third stage resource allocation problem. From (7)–(10), the objective function of problem $\mathcal{P}2$ is a concave function, and the feasible region is a convex set. Therefore, problem $\mathcal{P}2$ is a convex optimization problem, and is solved by maximizing the dual function which is obtained by minimizing the Lagrangian of problem $\mathcal{P}2$ with respect to $\mathbb{D}^U$ [23]. The Lagrangian of problem $\mathcal{P}2$ is

$$L^U\left(\psi_{0,0},\boldsymbol{\lambda},\mathbb{D}^U,\mathbb{D}^{L*}\right)=\sum_{i\in\mathcal{S}_N}\Big[\lambda_i(R_{Vmin,i}+R_{Dmin,i}) - (1+\lambda_i)R_i^U(\psi_{0,0},\mathbb{D}^U,\mathbb{D}^{L*})\Big], \quad (14)$$

where $\lambda_i,\forall i$ are dual variables.

The iterative algorithm which solves $\mathcal{P}2$ can be summarized as follows. First, $\boldsymbol{\lambda}$ is initialized (e.g., $\boldsymbol{\lambda}\leftarrow\{0,\ldots,0\}$). Second, we find $\mathbb{D}^U$ which minimizes $L^U(\psi_{0,0},\boldsymbol{\lambda},\mathbb{D}^U,\mathbb{D}^{L*})$ for the $\boldsymbol{\lambda}$. To update $\boldsymbol{\lambda}$ for the next iteration, $R_i^U(\psi_{0,0},\mathbb{D}^U,\mathbb{D}^{L*}),\forall i$ are also found in this step. Third, $\boldsymbol{\lambda}$ is adjusted toward $\boldsymbol{\lambda}^*$ using the gradient descent method [5], [24], [25]. The second and the third steps are repeated until $\boldsymbol{\lambda}$ reaches $\boldsymbol{\lambda}^*$. When $\boldsymbol{\lambda}$ reaches $\boldsymbol{\lambda}^*$, each λ_i satisfies the complementary slackness condition [23] and we have found $\mathbb{D}^{U*}$.

To implement the above algorithm, $\mathbb{D}^U$ and $R_i^U(\psi_{0,0},\mathbb{D}^U,\mathbb{D}^{L*}),\forall i$ for any $\boldsymbol{\lambda}$ can be calculated as follows. From (14) and since $\sum_{i\in\mathcal{S}_N}\lambda_i(R_{Vmin,i}+R_{Dmin,i})$ does not depend on $\mathbb{D}^U$, $\mathbb{D}^U$ is determined such that it maximizes $\sum_{i\in\mathcal{S}_N}(1+\lambda_i)R_i^U(\psi_{0,0},\mathbb{D}^U,\mathbb{D}^{L*})$. When $\sum_{i\in\mathcal{S}_N}(1+\lambda_i)R_i^U(\psi_{0,0},\mathbb{D}^U,\mathbb{D}^{L*})$ is maximized, by (7) and using the Bellman optimality equation [21], it is given by the following optimality equation.

$$\begin{aligned}L_{sup}^U(\psi_{0,0},\boldsymbol{\lambda}) &= (1-\theta)\max_{A_0^U}\left[\sum_{i\in\mathcal{S}_N}(1+\lambda_i)r_{i,0}^U\left(\psi_{0,0},A_0^U,\mathbb{D}^{L*}\right)\right] \\ &\quad+\theta\sum_{\psi_1^U\in\Psi^U}\sum_{\psi_{1,0}^L\in\Psi^L}\mathrm{P}_{\psi_0^U\psi_1^U}^U\mathrm{P}_{\psi_{0,0}^L\psi_{1,0}^L}^L L_{sup}^U(\psi_{1,0},\boldsymbol{\lambda}) \quad (15)\end{aligned}$$

with

$$L_{sup}^U(\psi_{u,0},\boldsymbol{\lambda})=\sup_{\mathbb{D}^U}\left[\sum_{i\in\mathcal{S}_N}(1+\lambda_i)R_i^U(\psi_{u,0},\mathbb{D}^U,\mathbb{D}^{L*})\right],$$

where $\mathrm{P}_{\psi_0^U\psi_1^U}^U$ and $\mathrm{P}_{\psi_{0,0}^L\psi_{1,0}^L}^L$ are the probabilities of the upper and lower level states change from ψ_0^U to ψ_1^U and from $\psi_{0,0}^L$ to $\psi_{1,0}^L$ at the end of the 0th time slot, respectively. Equation (15) is a recursive formula, and A_u^{U*} for the uth time slot is determined such that the summation of weighted throughputs given by $\sum_{i\in\mathcal{S}_N}(1+\lambda_i)r_{i,u}^U(\psi_{u,0},A_u^U,\mathbb{D}^{L*})$ is maximized [20]. Furthermore, these resource allocation problems corresponding to different time slots are independent of each other. As $\mathbb{D}^{U*}$ is a stationary policy (to be discussed in Section V-B), finding A_0^{U*} for each $\psi_{0,0}\in\{\Psi^U,\Psi^L\}$ at the 0th time slot is sufficient to find $\mathbb{D}^{U*}$. Next, $R_i^U(\psi_{0,0},\mathbb{D}^U,\mathbb{D}^{L*})$ can be determined by solving the Bellman optimality equation for the ith user written using (7). Bellman optimality equation solving methods, such as Value Iteration algorithm and its variants, are explained in [20].

B. Second Stage Resource Allocation

We derive A_u^{U*} such that it maximizes $\sum_{i\in\mathcal{S}_N}(1+\lambda_i)r_{i,u}^U(\psi_{u,0}, A_u^U, \mathbb{D}^{L*})$ at system state $\psi_{u,0}$ subject to satisfaction of (2) for $n = CF$, (6) and (12) during the uth time slot. This optimization problem is a non-convex problem due to the integer constraint which is imposed on $\rho_{i,j}^{CF}$ (see Section III-F1). Therefore, to reduce the computational complexity required to solve the problem, we relax the problem to be a convex optimization problem by relaxing the integer constraint such that $\rho_{i,j}^{CF} \in [0,1]$. To calculate power usage and throughputs over partially allocated TXOPs, we define $\bar{P}_{i,j}^{CF} = \rho_{i,j}^{CF} P_{i,j}^{CF}$ and $\bar{R}_{i,j}^{CF}(\bar{P}_{i,j}^{CF}, \rho_{i,j}^{CF}) = (T_{CF}/T_P)\rho_{i,j}^{CF} R_{i,j}^{CF}(\bar{P}_{i,j}^{CF}/\rho_{i,j}^{CF})$, where $\bar{R}_{i,j}^{CF}(\bar{P}_{i,j}^{CF}, \rho_{i,j}^{CF})$ is a concave function [22]. For notation simplicity, we define $\bar{R}_i^{CB}(\mathbf{P}^{CB}) = (T_{CP}/T_P)R_i^{CB}(\mathbf{P}^{CB})$. Then, substituting from (9) to $r_{i,u}^U(\psi_{u,0}, A_u^U, \mathbb{D}^{L*})$, the relaxed problem can be stated as

$$\begin{aligned}
\mathcal{P}3: \max_{A_u^U} & \sum_{i\in\mathcal{S}_N}(1+\lambda_i) \\
& \times \left[R_{i,u}^L\left(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}\right) + \sum_{j\in\mathcal{K}^{CF}} \bar{R}_{i,j}^{CF}(\bar{P}_{i,j}^{CF}, \rho_{i,j}^{CF}) \right] \\
& + \sum_{i\in\mathcal{S}^{CB*}} (1+\lambda_i)\bar{R}_i^{CB}(\mathbf{P}^{CB}) \\
\text{s.t.} \quad & \text{C2}: \sum_{i\in\mathcal{S}_M} \rho_{i,j}^{CF} \le 1, \quad \forall j \in \mathcal{K}^{CF} \\
& \text{C3}: R_{i,u}^L\left(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}\right) + \sum_{j\in\mathcal{K}^{CF}} \bar{R}_{i,j}^{CF}\left(\bar{P}_{i,j}^{CF}, \rho_{i,j}^{CF}\right) \\
& \qquad \ge R_{Vmin,i}, \quad \forall i \in \mathcal{S}_N \\
& \text{C4}: P_{avg,i}^{CB}(\mathbf{P}^{CB}) + P_{avg,i}^{C} \\
& \qquad + \frac{T_{CF}}{T_P} \sum_{j\in\mathcal{K}^{CF}} \bar{P}_{i,j}^{CF} \le P_{T,i}, \quad \forall i \in \mathcal{S}_N \\
& \text{C5}: 0 \le \rho_{i,j}^{CF} \le 1, \quad \forall i \in \mathcal{S}_M, j \in \mathcal{K}^{CF} \\
& \text{C6}: \bar{P}_{i,j}^{CF} \ge 0, P_i^{CB} \ge 0, \quad \forall i \in \mathcal{S}_N, j \in \mathcal{K}^{CF}.
\end{aligned}$$

Problem $\mathcal{P}3$ is a convex optimization problem. Convexity of C4, i.e., $\{P_{avg,i}^C, \bar{P}_{i,j}^{CF}, P_i^{CB} | \text{C4 is satisfied}, i \in \mathcal{S}_N, j \in \mathcal{K}^{CF}\}$, is proved in Appendix B.

Next, we illustrate the relationship between problem $\mathcal{P}3$ and the third stage resource allocation which determines $\mathbb{D}^{L*}$ for the lower-level. Then, we derive A_u^{U*} by solving problem $\mathcal{P}3$ using Karush-Kuhn-Tucker (KKT) conditions [23]. The Lagrangian of problem $\mathcal{P}3$ is shown in (16) (see equation at the bottom of the page) where γ_j, ξ_i and μ_i, $\forall i, j$ are the dual variables associated with C2, C3 and C4, respectively. As the optimal solution for problem $\mathcal{P}3$ minimizes (16) subject to C5 and C6, $\mathbb{D}^{L*}$ is determined such that it maximizes $\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^L)$. Further, $\mathbb{D}^{L*}$ should satisfy the following KKT condition of $\mathcal{P}3$ to ensure the optimality of the solution, which is obtained using the three stages, for problem $\mathcal{P}1$.

$$(1+\lambda_i+\xi_i)\left.\frac{\partial R_{i,u}^L\left(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}\right)}{\partial P_{avg,i}^C}\right|_{P_{avg,i}^C = P_{avg,i}^{C*}} \begin{cases} = \mu_i, & \text{if } P_{avg,i}^{C*} > 0; \\ < \mu_i, & \text{otherwise;} \end{cases} \quad \forall i \in \mathcal{S}_N. \quad (17)$$

Dual variables $\boldsymbol{\mu}$ and $\boldsymbol{\xi}$ couple the upper and the lower level resource allocations to optimally distribute the transmit power available at the UEs among WLAN and cellular network interfaces and to optimally utilize the resources of the two networks to satisfy the users' voice traffic requirements, respectively. Due to this coupling, once resources of the lower-level are allocated, achieved SDTs (i.e., $R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}), \forall i$) and the average power consumptions at the lower-level (i.e., $P_{avg,i}^{C*}, \forall i$) are feedback to the upper-level to solve problem $\mathcal{P}3$, as shown in Fig. 3.

Solution for problem $\mathcal{P}3$ is found as follows. First, $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$ are initialized. Second, the optimal allocations of contention-free TXOPs, UE transmit power levels during contention-free and contention-based TXOPs, $R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}), \forall i$ and $P_{avg,i}^{C*}, \forall i$ are calculated based on $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$. Third, $\boldsymbol{\mu}$ is updated toward $\boldsymbol{\mu}^*$ using the gradient descent method [5], [24], [25]. The second and third steps are repeated until $\boldsymbol{\mu}^*$ is found. Forth, $\boldsymbol{\xi}$ is updated toward $\boldsymbol{\xi}^*$ using the gradient descent method. The last three steps are repeated until $\boldsymbol{\xi}^*$ is found.

In the following, we derive the decision set $A_u^{U*}(= \{P_i^{CB*}, \bar{P}_{i,j}^{CF*}, \rho_{i,j}^{CF*} | \forall i \in \mathcal{S}_M, j \in \mathcal{K}^{CF}\})$ by solving problem $\mathcal{P}3$, and explain how $\mathcal{S}^{CB*}$ is determined. In addition, the optimality of A_u^{U*} for the initial problem (i.e., the problem prior to the relaxation) is also discussed.

$$\begin{aligned}
L^{U(2)}\left(A_u^U, \boldsymbol{\gamma}, \boldsymbol{\xi}, \boldsymbol{\mu}\right) = & -\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)\left[R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}) + \sum_{j\in\mathcal{K}^{CF}} \bar{R}_{i,j}^{CF}(\bar{P}_{i,j}^{CF}, \rho_{i,j}^{CF}) \right] - \sum_{i\in\mathcal{S}^{CB*}} (1+\lambda_i)\bar{R}_i^{CB}(\mathbf{P}^{CB}) \\
& + \sum_{i\in\mathcal{S}_N}\left[\sum_{j\in\mathcal{K}^{CF}} \gamma_j \rho_{i,j}^{CF} + \xi_i R_{Vmin,i} + \mu_i\left(P_{avg,i}^{CB}(\mathbf{P}^{CB}) + P_{avg,i}^C + \frac{T_{CF}}{T_P}\sum_{j\in\mathcal{K}^{CF}} \bar{P}_{i,j}^{CF} - P_{T,i} \right) \right] - \sum_{j\in\mathcal{K}^{CF}} \gamma_j \quad (16)
\end{aligned}$$

1) Allocations of Contention-Free TXOPs and Transmit Power Levels: Based on the KKT conditions for $\mathcal{P}3$, the optimal transmit power levels of the users during contention-free TXOPs are given by

$$\bar{P}_{i,j}^{CF*} = \rho_{i,j}^{CF*}\Theta_i, \quad \forall i \in \mathcal{S}_M, j \in \mathcal{K}^{CF}, \tag{18}$$

where

$$\Theta_i = \left[\frac{B^W}{\ln(2)}\frac{(1+\lambda_i+\xi_i^*)}{\mu_i^*} - \frac{1}{\alpha_i^W}\right]^+$$

and $[x]^+ = \max\{0, x\}$. Next, the optimal contention-free TXOP allocation can be determined as follows. Let

$$\Gamma_{i,j} = (1+\lambda_i+\xi_i^*)\left.\frac{\partial \bar{R}_{i,j}^{CF}\left(\bar{P}_{i,j}^{CF}, \rho_{i,j}^{CF}\right)}{\partial \rho_{i,j}^{CF}}\right|_{\bar{P}_{i,j}^{CF}=\bar{P}_{i,j}^{CF*}}$$

$$= \frac{(1+\lambda_i+\xi_i^*)\,T_{CF}B^W}{T_P} \times \left[\log_2(1+\alpha_i^W\Theta_i) - \frac{1}{\ln(2)}\frac{\alpha_i^W\Theta_i}{1+\alpha_i^W\Theta_i}\right], \quad \forall i \in \mathcal{S}_M, j \in \mathcal{K}^{CF}. \tag{19}$$

Due to the fact that $\Gamma_{i,j}$ is independent of $\rho_{i,j}^{CF}$ and from the KKT conditions, the jth TXOP is allocated to the user with the largest $\Gamma_{i,j}$ [22]. However, when there are multiple users with their $\Gamma_{i,j}$ values equal to the largest $\Gamma_{i,j}$ for the jth TXOP, the optimal solution for the problem $\mathcal{P}3$ allocates fractions of the TXOP among these users allowing them to time-share the TXOP.

Since the channel gain (or α_i^W) and Θ_i remain unchanged over the uth time slot, we can see from (19) that $\Gamma_{i,j}$ of the ith user is the same for all the TXOPs. Consequently, the ith user is allocated the same fraction from each TXOP or is allocated all the TXOPs. When there are N' users $\{i_1, i_2, \ldots, i_{N'}\}$ with their $\Gamma_{i,j}$ values equal to the largest, the optimal fractional values for $\rho_{i,j}^{CF},\ i = \{i_1, i_2, \ldots, i_{N'}\}$ are determined based on the primal feasibility of those $\rho_{i,j}^{CF}$'s with respect to C2, C3 and C4. That is, the optimal set of $\rho_{i,j}^{CF},\ i = \{i_1, i_2, \ldots, i_{N'}\}$ is a solution which satisfies C2 with equality and the following set of linear inequalities

$$\rho_{i,j}^{CF}\left|\mathcal{K}^{CF}\right|\frac{T_{CF}B^W\log_2\left(1+\alpha_i^W\Theta_i\right)}{T_P} \geq R_{Vmin,i} - R_{i,u}^L\left(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}\right), i = \{i_1, i_2, \ldots, i_{N'}\} \tag{20}$$

and

$$\rho_{i,j}^{CF}\left|\mathcal{K}^{CF}\right|\Theta_i\frac{T_{CF}}{T_P} \leq P_{T,i} - P_{avg,i}^{CB}(\mathbf{P}^{CB}) - P_{avg,i}^{C}, \quad i = \{i_1, i_2, \ldots, i_{N'}\}. \tag{21}$$

As the objective of this work is to allocate resources based on the PHY and the MAC technologies of the networks, a near optimal TXOP allocation for the initial problem is found by rounding $\rho_{i,j}^{CF}|\mathcal{K}^{CF}|$ values to the nearest integers. The rounded values indicate the number of TXOPs allocated to each user. Moreover, if $\rho_{i,j}^{CF}|\mathcal{K}^{CF}|, \forall i$ are integers, they are the optimal TXOP allocation for the initial problem.

2) Allocations of Users and Transmit Power Levels for Contention-Based Channel Access: The second stage resource allocation problem should be formulated as a convex optimization problem to reduce the required computational capacity. However, $R_i^{CB}(\mathbf{P}^{CB})$ given by (4) is a non-concave function when N_W varies. Therefore, to formulate the second stage problem as a convex optimization problem, $\mathcal{S}^{CB*}$ should be determined prior to allocating the other upper-level resources. In the MMDP based resource allocation algorithm, $\mathcal{S}^{CB*}$ which achieves the highest total SDT at the upper-level is found via searching over $\mathcal{S}_M$. A low complexity method to find a near optimal $\mathcal{S}^{CB}$ is presented in Section VII.

From (4), it can be seen that $\bar{R}_i^{CB}(\mathbf{P}^{CB})$ depends not only on the ith user's transmit power level, but also on the transmit power levels of the other users in $\mathcal{S}^{CB*}$. Thus, $P_i^{CB*}, \forall i \in \mathcal{S}^{CB*}$ are correlated. Based on the KKT conditions for problem $\mathcal{P}3$, $P_i^{CB*} > 0$ only if

$$\left.\frac{\partial \bar{R}_i^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}}\right|_{\substack{\mathbf{P}_{-i}^{CB}=\mathbf{P}_{-i}^{CB*} \\ P_i^{CB}=0}} > \frac{\mu_i^*}{1+\lambda_i}\left.\frac{\partial P_{avg,i}^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}}\right|_{\substack{\mathbf{P}_{-i}^{CB}=\mathbf{P}_{-i}^{CB*} \\ P_i^{CB}=0}}, \tag{22}$$

where $\mathbf{P}_{-i}^{CB}$ is a vector which consists of the power levels of the users in $\mathcal{S}^{CB*}$ except the ith user. Since these partial derivatives are not defined when $P_i^{CB} = 0$, we rewrite (22) by taking the limits of the partial derivatives as $P_i^{CB} \to 0$. Then, (22) reduces to

$$P_i^{CB*}\begin{cases} > 0, & \text{if } \frac{B^W\alpha_i^W}{\ln(2)} > \frac{\mu_i^*}{1+\lambda_i}; \\ = 0, & \text{otherwise;} \end{cases} \quad \forall i \in \mathcal{S}^{CB*}. \tag{23}$$

Moreover, for $P_i^{CB*} > 0$, the two sides of (22) become equal when the partial derivatives are evaluated at $P_i^{CB} = P_i^{CB*}$. Therefore, value of P_i^{CB*} when $P_i^{CB*} > 0$ can be found by solving (24) using Newton's method if $\mathbf{P}_{-i}^{CB*}$ is known [26].

$$\frac{1+\lambda_i}{\mu_i^*}\cdot\frac{N_W\alpha_i^W\bar{R}_i^{CB}(\mathbf{P}^{CB*})}{1+P_i^{CB*}\alpha_i^W} = \frac{T_{CP}}{T_P}\left[\ln\left(1+P_i^{CB*}\alpha_i^W\right) - \frac{P_i^{CB*}\alpha_i^W}{1+P_i^{CB*}\alpha_i^W}\right] + \frac{\ln(2)N_WP_i^{CB*}\alpha_i^W\bar{R}_i^{CB}(\mathbf{P}^{CB*})}{B^W\cdot\left(1+P_i^{CB*}\alpha_i^W\right)\ln\left(1+P_i^{CB*}\alpha_i^W\right)}. \tag{24}$$

Existence of a solution for (24) is shown in Appendix C.

As P_i^{CB*}, $\forall i \in \mathcal{S}^{CB*}$ are correlated, $\mathbf{P}^{CB*}$ is found using an iterative algorithm. In each iteration, P_i^{CB*}, $\forall i \in \mathcal{S}^{CB*}$ are calculated by (24) using the $\mathbf{P}^{CB*}$ calculated in the previous iteration. The algorithm terminates when the changes to P_i^{CB*}, $\forall i \in \mathcal{S}^{CB*}$ are negligible. Convergence of this iterative algorithm is proved in Appendix D.

From (18)–(24), it can be seen that A_u^{U*} is independent of the time slot (i.e., u). Therefore, A_u^{U*} which is determined for the uth time slot can be used at the u'th time slot ($u' = \{0, 1, 2, \ldots\}$) when states during the uth and the u'th time slots are equivalent (i.e., $\psi_{u,0} = \psi_{u',0}$). Thus, $\mathbb{D}^{U*}$ is a stationary policy [20]. The algorithm which determines $\mathbb{D}^{U*}$ for a given initial state $\psi_{0,0}$ is shown in Algorithm 1.

Algorithm 1 : Upper-Level Policy

input : $\{\psi_0^U, \psi_{0,0}^L\}$, $\mathcal{S}_M$, $\mathcal{S}_S$, and $P_{T,i}, R_{Vmin,i}$ and $R_{Dmin,i}, \forall i$
output : $A_u^{U*} = \{P_i^{CB*}, \rho_{i,j}^{CF*}, \bar{P}_{i,j}^{CF*} | \forall i \in \mathcal{S}_M, j \in \mathcal{K}^{CF}\}$ for every $\{\psi^U, \psi^L\} \in \{\Psi^U, \Psi^L\}$, and $\mathcal{S}^{CB*}$

while $\mathcal{S}^{CB}$ *is not optimal* **do**
  $\boldsymbol{\lambda} \leftarrow \{0, ..., 0\}$.
  while $\boldsymbol{\lambda}$ *is not optimal* **do**
    for *each* $\{\psi^U, \psi^L\} \in \{\Psi^U, \Psi^L\}$ **do**
      $\boldsymbol{\xi} \leftarrow \{0, ..., 0\}$ and $\boldsymbol{\mu} \leftarrow \{\mu_1, ..., \mu_N\}$.
      while $\boldsymbol{\xi}$ *is not optimal* **do**
        while $\boldsymbol{\mu}$ *is not optimal* **do**
          Calculate $\bar{P}_{i,j}^{CF*}, \rho_{i,j}^{CF*}$ and P_i^{CB*} by (18)−(21), (23) and (24) at state ψ^U.
          Allocate resources at the Lower-Level by Algorithm 2 when the initial state is ψ^L.
          Update $\mu_i, \forall i$.
        end
        Update $\xi_i, \forall i$.
      end
    end
    For each user, calculate $R_i^U(\psi_{0,0}, \mathbb{D})$ by solving the Bellman optimality equation written using (7).
    Update $\lambda_i, \forall i$.
  end
  Calculate the total SDT at the upper-level, $\sum_{i=1}^{N} R_i^U(\psi_{0,0}, \mathbb{D})$.
end

VI. Lower-Level Resource Allocation

The objective of the lower-level resource allocation is to maximize $\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^L)$ subject to (2) for $n = C$ and (17) over the uth time slot (see Section V-B). To determine $\mathbb{D}^{L*}$, first we decompose the resource allocation problem over the uth time slot to a set of independent subproblems, each of which allocates resources for one lower-level time slot. Second, $A_{u,l}^{L*}$ for the (u,l)th time slot, $l = \{0, 1, \ldots, V_L - 1\}$, is found by solving the subproblem which corresponds to the same time slot; this third stage resource allocation problem is denoted by $\mathcal{P}4$ (see Section IV). Based on $A_{u,l}^{L*}$ found for the (u,l)th time slot, we show that $\mathbb{D}^{L*}$ is a stationary policy.

To decompose the resource allocation problem which maximizes $\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^L)$ over the uth time slot to a set of independent subproblems, the Bellman optimality equation for the lower-level is written using (8) with the assumption of V_L is very large, which is reasonable since $T^L \ll T^U$, as follows [2], [7].

$$\begin{aligned}
&\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)R_{i,u}^L\left(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*}\right) \\
&= (1-\beta)\max_{A_{u,0}^L}\left[\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)r_{i,u,0}^L\left(\psi_{u,0}^L, A_u^U, A_{u,0}^L\right)\right] \\
&\quad + \beta\sum_{\psi_{u,1}^L\in\Psi^L} \mathrm{P}_{\psi_{u,0}^L\psi_{u,1}^L}^{L(2)} \\
&\quad \times \sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)R_{i,u}^L\left(\psi_{u,1}^L, A_u^U, \mathbb{D}^{L*}\right), \qquad (25)
\end{aligned}$$

where $\mathrm{P}_{\psi_{u,0}^L\psi_{u,1}^L}^{L(2)}$ is the probability of lower-level state changes from $\psi_{u,0}^L$ to $\psi_{u,1}^L$ at the end of the $(u,0)$th time slot. As the left hand side of (25) is maximized when $A_{u,l}^{L*}$ for the (u,l)th time slot maximizes $\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)r_{i,u,l}^L(\psi_{u,l}^L, A_u^U, A_{u,l}^L)$, resources of the lower-level (i.e., subcarriers and transmit power levels) are allocated for the (u,l)th time slot such that $\sum_{i\in\mathcal{S}_N}(1+\lambda_i+\xi_i)r_{i,u,l}^L(\psi_{u,l}^L, A_u^U, A_{u,l}^L)$ is maximized [20]. It should be noted that these resource allocation subproblems corresponding to the lower-level time slots are independent of each other.

Similar to the non-convexity caused by the integer constraint which is imposed on $\rho_{i,j}^{CF}$ (see Section V-B), the integer constraint which is imposed on $\rho_{i,k}^{C}$ makes the subproblem corresponding to the (u,l)th time slot non-convex (see Section III-F1). To reduce the computational capacity required to solve the subproblem, we relax it by following the same relaxation process which is used in Section V-B. That is, we let $\rho_{i,k}^C \in [0,1]$ and define $\bar{P}_{i,k}^C = \rho_{i,k}^C P_{i,k}^C$ and $\bar{R}_{i,k}^C(\bar{P}_{i,k}^C, \rho_{i,k}^C) = \rho_{i,k}^C R_{i,k}^C(\bar{P}_{i,k}^C/\rho_{i,k}^C)$. The relaxed subproblem is considered as the problem $\mathcal{P}4$.

Since problem $\mathcal{P}4$ is solved subject to satisfaction of (17) over the uth time slot, (17) is first translated into a set of constraints, each corresponds to one lower-level time slot, by substituting (8), (10) and (13) into (17). Then, the constraint corresponding to the (u,l)th time slot is given by

$$\begin{aligned}
&(1+\lambda_i+\xi_i)\left.\frac{\partial\bar{R}_{i,k}^C\left(\bar{P}_{i,k}^C, \rho_{i,k}^C\right)}{\partial\bar{P}_{i,k}^C}\right|_{\bar{P}_{i,k}^C=\bar{P}_{i,k}^{C*}} \\
&\quad \begin{cases} = \mu_i, & \text{if } \bar{P}_{i,k}^{C*} > 0; \\ < \mu_i, & \text{otherwise;} \end{cases} \quad \forall i \in \mathcal{S}_N, k \in \mathcal{K}^C. \qquad (26)
\end{aligned}$$

From (26),

$$\bar{P}_{i,k}^{C*} = \rho_{i,k}^{C*}\left[\frac{\Delta f}{\ln(2)}\frac{(1+\lambda_i+\xi_i)}{\mu_i} - \frac{1}{\alpha_{i,k}^C}\right]^+. \qquad (27)$$

Next, the remaining subcarrier allocation problem for the (u,l)th time slot can be stated as

$$\begin{aligned}
\mathcal{P}5: \quad &\max_{\boldsymbol{\rho}^C} \quad \sum_{i\in\mathcal{S}_N}\sum_{k\in\mathcal{K}^C}(1+\lambda_i+\xi_i)\bar{R}_{i,k}^C\left(\bar{P}_{i,k}^C, \rho_{i,k}^C\right) \\
&\text{s.t.} \quad \mathrm{C7}: \sum_{i\in\mathcal{S}_N}\rho_{i,k}^C \le 1, \quad \forall k \in \mathcal{K}^C \\
&\qquad \mathrm{C8}: 0 \le \rho_{i,k}^C \le 1, \quad \forall i \in \mathcal{S}_N, k \in \mathcal{K}^C.
\end{aligned}$$

Problem $\mathcal{P}5$ is a convex optimization problem. Therefore, from (27) and the KKT conditions for problem $\mathcal{P}5$, and using the same approach used for deriving $\rho_{i,j}^{CF*}$ in Section V-B1, the optimal subcarrier allocation can be expressed as [22]

$$\rho_{i',k}^{C*} = \begin{cases} 1, & \text{if } i' = \arg\max_{\forall i}\{\Lambda_{i,k}\}; \\ 0, & \text{otherwise;} \end{cases} \quad \forall i' \in \mathcal{S}_N, k \in \mathcal{K}^C, \tag{28}$$

where

$$\Lambda_{i,k} = (1+\lambda_i+\xi_i)\left[\log_2\left(1+\alpha_{i,k}^C P_{i,k}^{C*}\right) - \frac{1}{\ln(2)}\frac{\alpha_{i,k}^C P_{i,k}^{C*}}{1+\alpha_{i,k}^C P_{i,k}^{C*}}\right].$$

However, when there are multiple users with their $\Lambda_{i,k}$ values equivalent to the largest $\Lambda_{i,k}$ for the kth subcarrier, the optimal solution for the problem $\mathcal{P}5$ requires allocation of fractions, which satisfies C7 with equality, of the kth subcarrier among these users.

When such equality of $\Lambda_{i,k}$ occurs, we randomly allocate the kth subcarrier to one of the users with the largest $\Lambda_{i,k}$ due to the fact that fractional subcarrier allocations are not supported by the PHY. Random subcarrier allocation in this scenario does not significantly deviate the system throughput/QoS performance from the optimum due to two reasons: 1) subcarrier bandwidth (Δf) is small as there is a large number of subcarriers; 2) probability of multiple users having equivalent $\Lambda_{i,k}$ values for more than one subcarrier or for a certain subcarrier over multiple time slots is very small, because the channel gains over different subcarriers are different and varies over time slots.

From (27) and (28), it can be seen that $A_{u,l}^{L*}$ is independent of the time slot. Thus, $A_{u,l}^{L*}$ which is determined for the (u,l)th time slot can be used at the (u,l')th time slot $(l' = \{0,\ldots,V_L-1\})$ when states during these two time slots are equivalent (i.e., $\psi_{u,l}^L = \psi_{u,l'}^L$). Therefore, $\mathbb{D}^{L*}$ is a stationary policy [20].

As $\mathbb{D}^{L*}$ is a stationary policy, calculating $A_{u,0}^{L*}$ for each state $\psi_{u,0}^L \in \Psi^L$ at the $(u,0)$th time slot is sufficient to determine $\mathbb{D}^{L*}$. Next, $R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*})$ and $P_{avg,i}^{C*}, \forall i$ can be found by solving (8) and (13), respectively. Equations (8) and (13) can be solved using the methods explained in [20] by writing them as Bellman optimality equations. Values of $R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^{L*})$ and $P_{avg,i}^{C*}, \forall i$ are then feedback to the upper-level to update $\boldsymbol{\lambda}$, $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$ as shown in Algorithm 2, which determines $\mathbb{D}^{L*}$.

Algorithm 2 : Lower-Level Policy

input : ψ^L, $\mathcal{S}_N$, $\boldsymbol{\lambda}, \boldsymbol{\xi}$ and $\boldsymbol{\mu}$

output : $A_{u,l}^{L*} = \{\rho_{i,k}^{C*}, \bar{P}_{i,k}^{C*} | \forall i \in \mathcal{S}_N, k \in \mathcal{K}^C\}$ for each $\psi^{L(2)} \in \Psi^L$, and $R_{i,u}^L(\psi^L, A_u^U, \mathbb{D}^{L*})$ and $P_{avg,i}^{C*}, \forall i$

for *each* $\psi^{L(2)} \in \Psi^L$ **do**

Calculate $\bar{P}_{i,k}^{C*}$ and $\rho_{i,k}^{C*}$ by (27) and (28) at state $\psi^{L(2)}$.

$r_{i,u,l}^L(\psi^{L(2)}, A_u^U, A_{u,l}^{L*}) \leftarrow \sum_{k\in\mathcal{K}^C} \rho_{i,k}^{C*} R_{i,k}^C(P_{i,k}^{C*})$.

$P_{tot,i,l}^C(\psi^{L(2)}, A_u^U, A_{u,l}^{L*}) \leftarrow \sum_{k\in\mathcal{K}^C} \bar{P}_{i,k}^{C*}$.

end

For each user, calculate $R_{i,u}^L(\psi^L, A_u^U, \mathbb{D}^{L*})$ by solving the Bellman optimality equation written using (8).

Calculate $P_{avg,i}^{C*}, \forall i$ by (13).

The MMDP based resource allocation algorithm (i.e., $\mathbb{D}^{U*}$ and $\mathbb{D}^{L*}$) efficiently allocates resources of the interworking system. However, it has a high time complexity as it requires to find A_0^{U*} and $A_{u,0}^{L*}$ for each system state, where the total number of system states in a system model is given by $(K_S)^{N|\mathcal{K}^C|+|\mathcal{S}_M|}$. Therefore, we propose a heuristic resource allocation algorithm with low time complexity when the number of system states is significantly higher due to the large number of users and/or OFDM subcarriers.

VII. Heuristic Resource Allocation

The heuristic algorithm consists of two steps. The first step is executed only once at the beginning, and it calculates the dual variables which correspond to data and voice traffic constraints (i.e., $\boldsymbol{\lambda}$ and $\boldsymbol{\xi}$) based on the average square channel gains (Ω's), where $\Omega = \mathrm{E}\{h^2\}$, $\mathrm{E}\{\cdot\}$ is the ensemble average operator and h is the channel gain. The second step uses the dual variables calculated in the first step, and allocates upper and lower level resources based on the instantaneous channel gains subject to total power constraints of the users (i.e., C4). Since these two steps are executed based on a single system state which consists of either Ω's or instantaneous channel gains, solving the Bellman optimality equations is not required in the heuristic algorithm for calculating $R_i^U(\psi_{0,0}, \mathbb{D})$, $R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^L)$ and $P_{avg,i}^C$. In the MMDP based algorithm, the Bellman optimality equations are solved by determining A_0^U and $A_{u,0}^L$ for each possible system state. In addition, $\mathcal{S}^{CB}$ is determined using a simple method in the heuristic algorithm (to be discussed). Due to these two reasons, the heuristic algorithm has low time complexity.

In the first step, $\boldsymbol{\lambda}$ and $\boldsymbol{\xi}$ are found by solving problems $\mathcal{P}2$, $\mathcal{P}3$ and $\mathcal{P}4$. Therefore, the solutions which we obtained for problems $\mathcal{P}3$ and $\mathcal{P}4$ in Sections V-B and VI are used in this step with modifications to utilize Ω's as follows. Average throughput over a Rayleigh fading channel is given by [18], [28]

$$\mathrm{E}\{R\} = \int_0^\infty \frac{2Bh}{\Omega}\log_2\left(1+\frac{h^2 p}{n}\right)\mathrm{e}^{-\frac{h^2}{\Omega}}\,\mathrm{d}h$$

$$= \frac{B}{\ln(2)}\mathrm{e}^{\frac{n}{\Omega p}}\mathrm{E}_1\left(\frac{n}{\Omega p}\right), \tag{29}$$

where p is the transmit power level, B is the bandwidth, n is the total noise plus interference power, and $\mathrm{E}_1(x) = \int_x^\infty \mathrm{e}^{-x}x^{-1}\,\mathrm{d}x$. Since $0.5\mathrm{e}^{-x}\ln(1+2x)$ provides a tight lower bound for $\mathrm{E}_1(x)$ [28], by (29)

$$\mathrm{E}\{R\} > \frac{B}{2}\log_2\left(1+\frac{2\Omega p}{n}\right). \tag{30}$$

Thus, the solutions for the problems $\mathcal{P}3$ and $\mathcal{P}4$ are modified to calculate the throughput over each wireless channel by $(B/2)\log_2(1+(2\Omega p/n))$. That is, the equations in Sections V-B and VI are modified with the substitutions of $B/2$ to B, and

2Ω to h^2. The latter is also equivalent to the substitution of $2\mathrm{E}\{\alpha_{i,y}^n\}$ to $\alpha_{i,y}^n$. Moreover, as Ω's are used in this step, the number of possible system states in the MMDP reduces to one. Consequently, $R_i^U(\psi_{0,0}, \mathbb{D}) = r_{i,0}^U(\psi_{0,0}, A_0^U, \mathbb{D}^L)$, $R_{i,0}^L(\psi_{0,0}^L, A_0^U, \mathbb{D}^L) = r_{i,0,0}^L(\psi_{0,0}^L, A_0^U, A_{0,0}^L)$ and $P_{avg,i}^C = \sum_{k\in\mathcal{K}^C} \rho_{i,k}^C P_{i,k}^C$.

Furthermore, $\log_2(1+x) > (1/2)\log_2(1+2x),\ \forall x \in \mathbb{R}^+$. Therefore, $\boldsymbol{\lambda}$ and $\boldsymbol{\xi}$ calculated in the first step will satisfy the QoS requirements with an additional margin if the throughputs are given by $B\log_2(1+(\Omega p/n))$. Therefore, when the resources are allocated in the second step, QoS requirements of the users are satisfied with a higher satisfaction though the instantaneous channel gains are used in this step. First step of the heuristic algorithm is shown in Algorithm 3.

Algorithm 3 : First step of the heuristic algorithm

input : Average square channel gains (Ω's), $\mathcal{S}_M$, $\mathcal{S}_S$, and $P_{T,i}, R_{Vmin,i}$ and $R_{Dmin,i}, \forall i$

output : $\boldsymbol{\lambda}^*, \boldsymbol{\xi}^*$ and $\mathcal{S}^{CB*}$

Form ψ_0^U and $\psi_{0,0}^L$ using $\sqrt{2\Omega}$ values of the channels.

$\alpha_{i,k}^C \leftarrow 2\mathrm{E}\{\alpha_{i,k}^C\}$ and $\alpha_i^W \leftarrow 2\mathrm{E}\{\alpha_i^W\}$.

Sort users in $\mathcal{S}_M$ in the descending order of their $\mathrm{E}\{\alpha_i^W\}$.

while $\mathcal{S}^{CB}$ *corresponding to maximum* $\sum_{i\in\mathcal{S}_N} R_i^U(\psi_{0,0})$ *is not found* **do**

  $\mathcal{S}^{CB} \leftarrow \mathcal{S}_M(|\mathcal{S}^{CB}|+1)$.

  $\boldsymbol{\lambda} \leftarrow \{0,...,0\}$, $\boldsymbol{\xi} \leftarrow \{0,...,0\}$ and $\boldsymbol{\mu} \leftarrow \{\mu_1,...,\mu_N\}$.

  while $\boldsymbol{\lambda}$ *and* $\boldsymbol{\xi}$ *are not optimal* **do**

    while $\boldsymbol{\mu}$ *is not optimal* **do**

      Calculate $\bar{P}_{i,j}^{CF*}, \rho_{i,j}^{CF*}, P_i^{CB*}, \rho_{i,k}^{C*}$ and $\bar{P}_{i,k}^{C*}$ by (18)−(21), (23), (24), (27) and (28) substituting $\frac{B^W}{2}$ to B^W and $\frac{\Delta f}{2}$ to Δf.

      Update $\mu_i, \forall i$.

    end

    $R_{i,0}^L(\psi_{0,0}^L) \leftarrow \sum_{k\in\mathcal{K}^C} \bar{R}_{i,k}^C(\bar{P}_{i,k}^C, \rho_{i,k}^C)$.

    $P_{avg,i}^C \leftarrow \sum_{k\in\mathcal{K}^C} \bar{P}_{i,k}^C$.

    Calculate $r_{i,0}^U(\psi_{0,0})$ by (9), and $R_i^U(\psi_{0,0}) \leftarrow r_{i,0}^U(\psi_{0,0})$.

    Update λ_i and $\xi_i, \forall i$.

  end

end

In the second step, resources of both upper and lower levels are jointly allocated at the beginning of uth time slot ($u = \{0,1,2,\ldots\}$) subject to C4 and assuming that the current lower-level state remains unchanged during the uth time slot (i.e., $\psi_{u,0}^L = \psi_{u,l}^L,\ \forall l \in \{1,\ldots,V_L-1\}$). With this assumption, $R_{i,u}^L(\psi_{u,0}^L, A_u^U, \mathbb{D}^L) = r_{i,u,0}^L(\psi_{u,0}^L, A_u^U, A_{u,0}^L)$ and $P_{avg,i}^C = \sum_{k\in\mathcal{K}^C} \rho_{i,k}^C P_{i,k}^C$. The algorithms which solve problems $\mathcal{P}3$ and $\mathcal{P}4$ are used for allocating resources while using $\boldsymbol{\lambda}$ and $\boldsymbol{\xi}$ from the first step. Note that these algorithms need to calculate $\boldsymbol{\mu}$ only and that UE power is distributed between WLAN and cellular network interfaces at this point. At the beginnings of the remaining (u,l)th time slots (i.e., $l = \{1,\ldots,V_L-1\}$), lower-level resources of subcarriers and the amount of power dedicated for the cellular network interfaces are reallocated based on the current state $\psi_{u,l}^L$ to fully exploit the multi-user diversity in the cellular network. The second step of the heuristic algorithm is shown in Algorithm 4.

Algorithm 4 : Second step of the heuristic algorithm

input : Instantaneous channel gains during (u,l)th time slot, $\boldsymbol{\lambda}^*, \boldsymbol{\xi}^*$ and $\mathcal{S}^{CB*}$

output : $A^{U*} = \{P_i^{CB*}, \rho_{i,j}^{CF*}, \bar{P}_{i,j}^{CF*} | \forall i \in \mathcal{S}_M, j \in \mathcal{K}^{CF}\}$ and $A^{L*} = \{\rho_{i,k}^{C*}$ and $\bar{P}_{i,k}^{C*} | \forall i \in \mathcal{S}_N, k \in \mathcal{K}^C\}$

Form ψ_u^U and $\psi_{u,l}^L$ using instantaneous channel gains during (u,l)th time slot.

while $\boldsymbol{\mu}$ *is not optimal* **do**

  if $l = 0$ **then**

    Calculate $\bar{P}_{i,j}^{CF*}, \rho_{i,j}^{CF*}, P_i^{CB*}, \rho_{i,k}^{C*}$ and $\bar{P}_{i,k}^{C*}$ by (18)−(21), (23) and (24), (27) and (28) based on ψ_u^U and $\psi_{u,l}^L$.

  else

    Recalculate $\rho_{i,k}^{C*}$ and $\bar{P}_{i,k}^{C*}$ by (27) and (28) based on $\psi_{u,l}^L$.

  end

  Update $\mu_i, \forall i$

end

The optimal $\mathcal{S}^{CB}$ consists of only a few users with strong channel conditions due to two reasons: 1) allocation of too many users for contention-based channel access degrades the aggregated throughput of the users due to increased number of collisions [2], and 2) allocation of a user with a weaker channel degrades the throughputs of all the users as the weak user takes a longer time to transmit a packet [27]. Based on these characteristics, users for the contention-based channel access are allocated in the first step of the heuristic algorithm as follows. First, the users in $\mathcal{S}_M$ are sorted in the descending order of their $\mathrm{E}\{\alpha_i^W\}$. Next, the first step of the heuristic algorithm is repeated, each time adding the next user in $\mathcal{S}_M$ to $\mathcal{S}^{CB}$, until the total throughput achieved at the upper-level reaches the maximum.

VIII. Simulation Results

Wireless channels are modeled as Rayleigh fading channels, and their path loss is proportional to d^{-4}, where d denotes the distance between users and the WLAN AP or the cellular BS. Further, the channels over the cellular network are generated at the carrier frequency of 2.1 GHz and mobile speed of 50 kmh^{-1}, while those over the WLAN are generated at the carrier frequency of 2.4 GHz and mobile speed of 3 kmh^{-1}. Based on the coherence times of the channels, T^L and T^U are selected to be 4.23 ms and 63.45 ms, respectively [29], [30]. The radiuses of the WLAN and the cellular coverage areas are 50 m and 1000 m respectively, and users are uniformly distributed over the coverage areas. The total power available at each user is uniformly distributed between 0 and 1 watt. Table II shows the remaining parameters.

First we evaluate the performance of the MMDP based resource allocation algorithm (MM) and the heuristic algorithm (HM) in a small-scale system, denoted by system-1, and compare the performance with that of a benchmark algorithm (BM1) which resembles the first category resource allocation algorithms (see Section II). Algorithm BM1 allocates the resources as follows. First, it assigns users for the two networks via exhaustive search such that the total average system throughput is maximized. In this step, average users'

TABLE II
SIMULATION PARAMETERS

Parameter	Value (unit)
B^W	20MHz
D	4095 octets
$\|\mathcal{K}^C\|$	4 or 128
$\|\mathcal{K}^{CF}\|$	2 or 10
T^L	4.23ms
T^U, T_P	63.45ms
T_{ACK}	24.5μs
T_{AIFS}	34μs
T_{CF}	$31.72/\|\mathcal{K}^{CF}\|$ ms
T_{CP}	31.72ms
T_{CTS}	24.5μs
T_{RTS}	24.7μs
T_{SIFS}	16μs
Δf	$5/\|\mathcal{K}^C\|$ MHz
σ	9μs
Additive white Gaussian noise density	-174 dBm/Hz
Initial window size of WLAN	16
Maximum number of backoff stages in WLAN	6

throughputs are calculated using $(B/2)\log_2(1+(2\Omega p/n))$, as in the first step of HM. Next, each network individually allocates its resources among the assigned users to maximize the network throughput. This step utilizes instantaneous channel gains and is repeated at the every time slot. It should be noted that BM1 does not allow UE multi-homing and it allocates resources at two time-scales based on the PHY and MAC technologies of the networks.

In the system-1, there are four users ($|\mathcal{S}_S| = |\mathcal{S}_M| = 2$), four subcarriers and two contention-free TXOPs. Two-state Markov channels are used [31], and the boundary between the two states of each channel is determined such that the steady state probability of each state is 0.5. For each channel, the channel is at the first state if $h < \sqrt{\Omega \ln(2)}$; otherwise, it is at the second state. When a channel is at first and second states, the channel gains are considered to be $\sqrt{\Omega(1-\ln(2))}$ and $\sqrt{\Omega(1+\ln(2))}$, respectively. These channel gains are determined by averaging the square of the channel gain of a continuous-envelope Rayleigh fading channel within the boundaries of the respective state. Transition probabilities of the states are calculated as in [18]. Discount factors θ and β are set to be 0.9 in MM.

Fig. 4 compares the throughputs achieved by MM, HM and BM1 in system-1 for different QoS requirements. Algorithm MM provides throughput improvement of at least 10.7% compared to HM, and both MM and HM provide higher throughputs than BM1 as they enable multi-homing. In BM1, each user is allowed to access one network only. When multi-homing is enabled, users achieve higher throughputs due to efficient resource utilization, which is a result of catering user QoS requirements utilizing multiple network resources and of dynamically adjusting resource allocation including UE power distribution for the two network interfaces based on the instantaneous channel gains. In addition, MM outperforms HM due to the fact that MM allocates the resources statistically considering the future state changes using an MMDP, whereas HM allocates the resources based on the current states only.

The satisfaction index (SI), which can be used for quantifying the ability of a resource allocation algorithm to satisfy

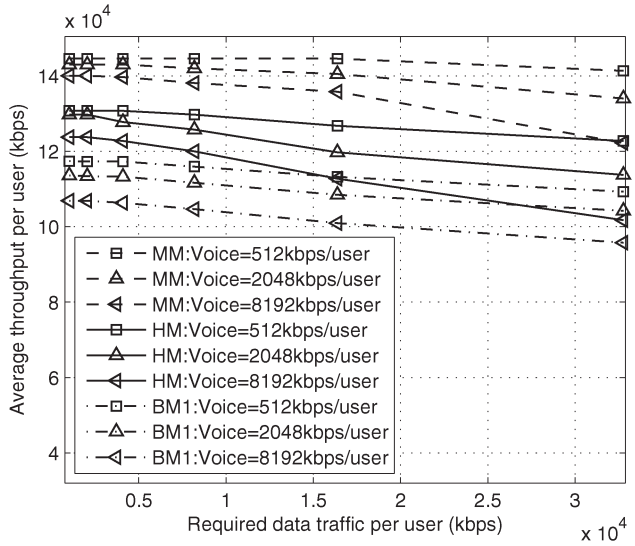


Fig. 4. Throughputs achieved by different algorithms in system-1.

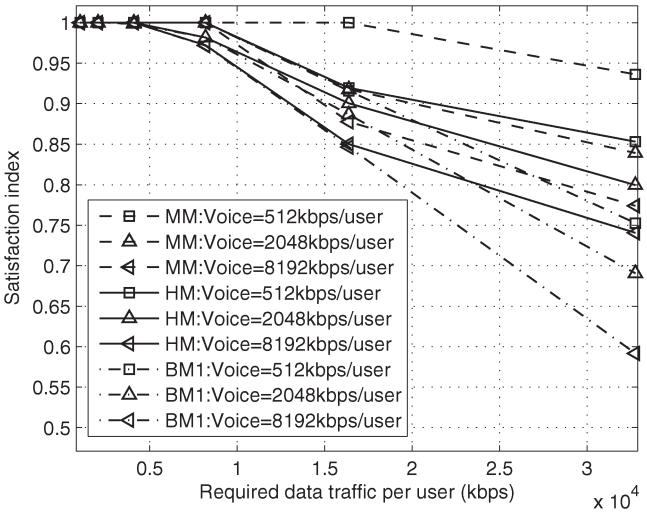


Fig. 5. Satisfaction index achieved by different algorithms for data traffic in system-1.

the QoS requirements [25], is defined for a particular traffic class as

$$\mathrm{SI} = \mathrm{E}\left\{1_{R \geq R_{\min}} + 1_{R_{\min} > R} \cdot \frac{R}{R_{\min}}\right\}, \qquad (31)$$

where R and $R_{\min}$ are the achieved and the required throughputs, and $1_{x \geq y} = 1$ if $x \geq y$ but it is zero otherwise. All three algorithms have achieved SI for voice traffic (SI_V) of one in system-1, and the achieved SI's for data traffic (SI_D) by them in system-1 are shown in Fig. 5. Similar to the throughput performance, MM and HM achieve higher SI_D's compared to BM1, providing users with better QoS. Difference between these SI_D's is significant at the higher data traffic requirements as multi-homing is particularly useful for catering higher user requirements via multiple networks.

Complexities of the algorithms are measured in terms of the required number of iterations in the inner most loop per user per time slot in fast time-scale, and Fig. 6 compares them in

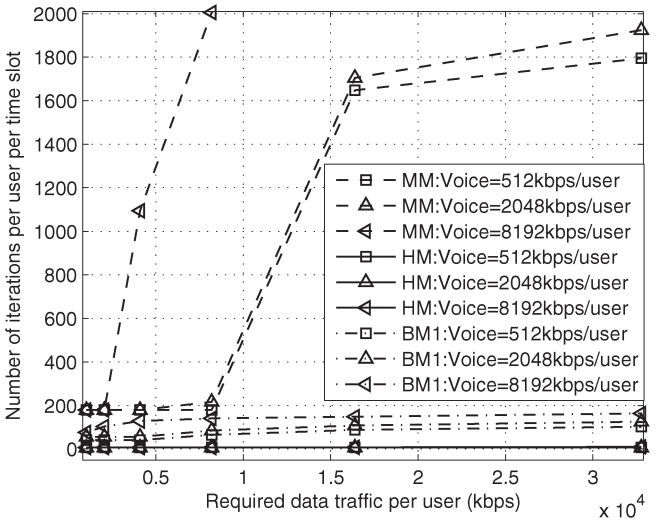


Fig. 6. Complexities of the different algorithms in system-1.

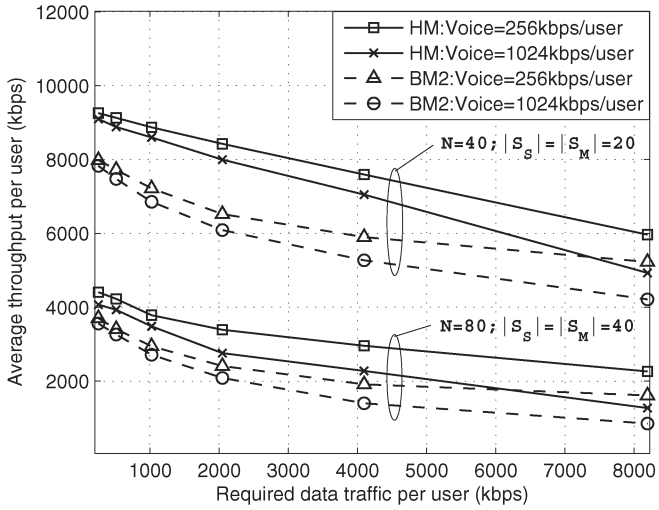


Fig. 7. Throughputs achieved by different algorithms in system-2.

system-1. As MM solves an MMDP based resource allocation problem with 2^{18} system states, it requires a large number of iterations. Algorithm BM1 requires a higher number of iterations than HM as BM1 recalculates $\boldsymbol{\lambda}$ and $\boldsymbol{\xi}$ at each time slot, whereas HM calculates $\boldsymbol{\lambda}$ and $\boldsymbol{\xi}$ only once in the first step.

Next, the performance of HM is evaluated in a large-scale system, denoted by system-2, and is compared with the performance of a benchmark algorithm (BM2). Since the highest number of resource blocks per cell in a LTE system is 110, system-2 consists of 128 subcarriers, 10 contention-free TXOPs and 40 or 80 users. This system uses continuous envelope Rayleigh fading channels generated at the same carrier frequencies and mobile speeds as in system-1. The performance of MM is not evaluated in this system due its high complexity. Algorithm BM2 uses a simpler user allocation mechanism than exhaustive search which is used in BM1, because exhaustive search is not feasible when there is a large number of users. It allocates users for the networks as follows. First, users in $\mathcal{S}_M$ are sorted in the descending order of their $\mathrm{E}\{\alpha_i^W\}$. Second, resources of the two networks are individually allocated $|\mathcal{S}_M|$ times while calculating users' average throughputs using $(B/2)\log_2(1+(2\Omega p/n))$; at the jth resource allocation, first j users in $\mathcal{S}_M$ are allocated to the WLAN while the remaining users in $\mathcal{S}_M$ and all the users in $\mathcal{S}_S$ are allocated to the cellular network. Finally, the user allocation which resulted in the highest total average throughput is selected. Once the user allocation is completed, BM2 allocates resources of the two networks at each time slot similar to BM1.

Throughput, SI_V and SI_D performance of HM and BM2 in system-2 are shown in Figs. 7–9, respectively. Due to the advantages of user multi-homing, HM provides better throughput and SI performance than BM2. The performance of the algorithms decreases with the number of users, because the resources are distributed among more users as each user has a certain QoS requirement. When there are 80 users in the system, HM provides at least 14.5%, 8.4% and 8.1% of improvements compared to BM2 in average throughput per user, SI_V and SI_D, respectively.

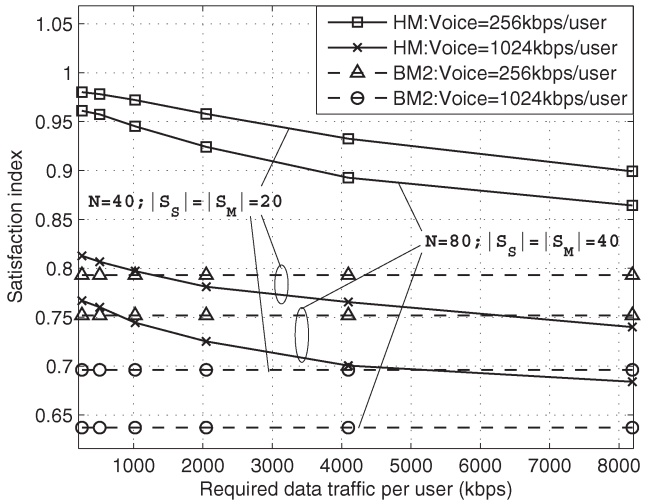


Fig. 8. Satisfaction index achieved by different algorithms for voice traffic in system-2.

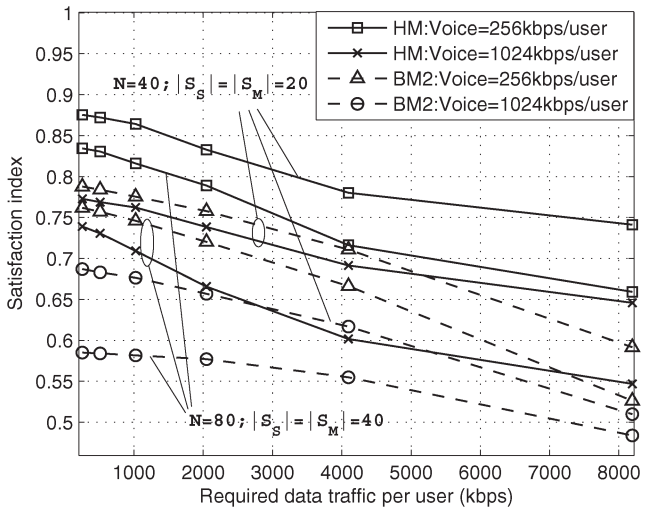


Fig. 9. Satisfaction index achieved by different algorithms for data traffic in system-2.

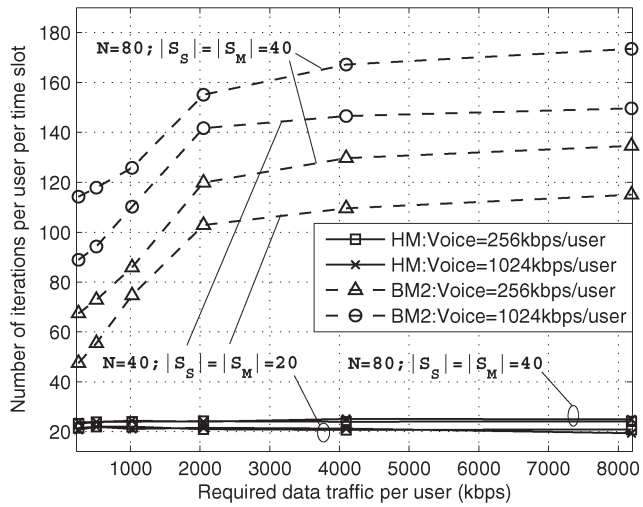


Fig. 10. Complexities of different algorithms in System-2.

According to the complexities of the algorithms shown in Fig. 10, HM converges within 25 iterations per user per time slot while BM2 requires more than 47 iterations per user per time slot. The complexity of HM does not considerably vary with the QoS requirements and the number of users as HM recalculates only $\boldsymbol{\mu}$ at each time slot in the second step of the algorithm, whereas the complexity of BM2 increases with the QoS requirements and the number of users as BM2 recalculates $\boldsymbol{\lambda}$, $\boldsymbol{\xi}$ and $\boldsymbol{\mu}$ at each time slot.

IX. Conclusion

This paper has presented an MMDP based resource (subcarrier, TXOPs of WLAN, and power) allocation for cellular/WLAN interworking considering two time-scales; underlying PHY and MAC layer technologies of an OFDMA based cellular network and a WLAN which operates on contention-based and contention-free channel access mechanisms; and multi-homing capable users with voice and data traffic requirements. Further, by eliminating the requirement to solve Bellman optimality equations, a low time complexity heuristic resource allocation algorithm has been proposed. Simulation results have shown that the MMDP based algorithm provides 10.7% of throughput improvements than the heuristic algorithm, and that the MMDP based and the heuristic algorithms provide higher throughputs and satisfaction indexes (i.e., QoS) than the benchmark algorithms (BM1 and BM2) which do not consider user multi-homing. The MMDP based algorithm has a high time complexity due to large number of states in the system model. The low time complexity heuristic algorithm converges within 25 iterations per user per time slot in practical systems, which enables it to allocate resources online based on the instantaneous channel gains.

Appendix A
Proof of Convexity of $R_i^{CB}(\mathbf{P}^{CB})$

Let

$$f(\mathbf{x}) = \frac{-1}{k_0 + \sum_{i=1}^{N} \frac{k_i}{x_i}}, \tag{32}$$

where $k_i \in \mathbb{R}^+, \forall i \in \{0, \ldots, N\}$; $\mathbf{x} = [x_1, \ldots, x_N]$; and $x_i \in \mathbb{R}^+,\ \forall i \in \{1, \ldots, N\}$. Now consider

$$f(\mathbf{z}) - f(\mathbf{x}) - \nabla f(\mathbf{x})[\mathbf{z} - \mathbf{x}]^T = \left(\frac{1}{k_0 + \sum_{i=1}^{N} \frac{k_i}{z_i}}\right)\left(\frac{1}{k_0 + \sum_{i=1}^{N} \frac{k_i}{x_i}}\right)^2 g(\mathbf{x}), \tag{33}$$

where

$$g(\mathbf{x}) = \left(k_0 + \sum_{i=1}^{N} \frac{k_i}{z_i}\right)\left(k_0 + \sum_{i=1}^{N} \frac{k_i z_i}{x_i^2}\right) - \left(k_0 + \sum_{i=1}^{N} \frac{k_i}{x_i}\right)^2 \tag{34}$$

and $\mathbf{z} = [z_1, \ldots, z_N]$ with $z_i \in \mathbb{R}^+,\ \forall i \in \{1, \ldots, N\}$. Furthermore, $g(\mathbf{x})$ can be rewritten as

$$\begin{aligned} g(\mathbf{x}) &= k_0 \sum_{i=1}^{N} k_i \left(\frac{1}{\sqrt{z_i}} - \frac{\sqrt{z_i}}{x_i}\right)^2 \\ &\quad + \sum_{i=2}^{N} \sum_{j=1}^{\lfloor \frac{i}{2} \rfloor} k_{i-j+1} k_j \left(\sqrt{\frac{z_j}{z_{i-j+1} x_j^2}} - \sqrt{\frac{z_{i-j+1}}{z_j x_{i-j+1}^2}}\right)^2 \\ &\quad + \sum_{j=2}^{N-1} \sum_{i=1}^{\lfloor \frac{j}{2} \rfloor} \left[k_{N-i+1} k_{N-j+i} \left(\sqrt{\frac{z_{N-j+i}}{z_{N-i+1} x_{N-j+i}^2}} - \sqrt{\frac{z_{N-i+1}}{z_{N-j+i} x_{N-i+1}^2}}\right)^2\right]. \end{aligned} \tag{35}$$

Since $k_i \in \mathbb{R}^+,\ \forall i \in \{0, \ldots, N\},\ g(\mathbf{x}) \geq 0$. Thus, by (33), $f(\mathbf{x})$ is a convex function as it satisfies the first order condition [23]. Moreover, $f(\mathbf{x})$ is a non-increasing function as

$$\frac{\partial f(\mathbf{x})}{\partial x_i} = \frac{-k_i}{x_i^2}\left(\frac{1}{k_0 + \sum_{i=1}^{N} \frac{k_i}{x_i}}\right)^2 \leq 0, \quad \forall i \in \{1, \ldots, N\}. \tag{36}$$

Let $x_i = \log_2(1 + y_i),\ \forall i \in \{1, \ldots, N\}$. Since $x_i = \log_2(1 + y_i)$ is a concave function with respect to $y_i \in \mathbb{R}^+$, and $f(\mathbf{x})$ is non-increasing in each of its argument and it is a convex function, by vector composition theory [23],

$$f(\mathbf{y}) = \frac{-1}{k_0 + \sum_{i=1}^{N} \frac{k_i}{\log_2(1+y_i)}} \tag{37}$$

is a convex function. Thus, $R_i^{CB}(\mathbf{P}^{CB})$ is a concave function.

APPENDIX B
PROOF OF CONVEXITY OF C4

From (5), the derivative of $P_{avg,i}^{CB}(\mathbf{P}^{CB})$ when $P_i^{CB} > 0$ is

$$\frac{\partial P_{avg,i}^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}} = \frac{T_{CP}}{T_P B^W \cdot \left(\log_2\left(1+P_i^{CB}\alpha_i^W\right)\right)^2} \times \left[R_i^{CB}(\mathbf{P}^{CB}) \left(\log_2\left(1+P_i^{CB}\alpha_i^W\right) - \frac{P_i^{CB}\alpha_i^W}{\ln(2)\left(1+P_i^{CB}\alpha_i^W\right)} \right) + \frac{\partial R_i^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}} P_i^{CB} \log_2\left(1+P_i^{CB}\alpha_i^W\right) \right]. \quad (38)$$

Next, we define

$$g(x) = \log_2(1+x) - \frac{x}{(1+x)\ln(2)}, \qquad x \geq 0. \quad (39)$$

We can conclude that $g(x) \geq 0$ since $g(0) = 0$ and $dg(x)/dx \geq 0,\ \forall x \in \mathbb{R}^+$. Moreover, $R_i^{CB}(\mathbf{P}^{CB})$ is a positive non-decreasing concave function. Therefore, by (38), $\partial P_{avg,i}^{CB}(\mathbf{P}^{CB})/\partial P_i^{CB} \geq 0$. Thus, $P_{avg,i}^{CB}(\mathbf{P}^{CB})$ is a non-decreasing function of P_i^{CB}. Therefore, we can show that $P_{avg,i}^{CB}(\mathbf{P}^{CB}) \leq P_{T,i}$ is a convex set [23], and hence C4, i.e., $\{P_{avg,i}^{C}, \bar{P}_{i,j}^{CF}, P_i^{CB} | \text{C4 is satisfied},\ i \in \mathcal{S}_N, j \in \mathcal{K}^{CF}\}$, is a convex set. Since $P_{avg,i}^{C}$ is a linear combination of $\bar{P}_{i,k}^{C},\ \forall k$ (see (13)), $\{\bar{P}_{i,k}^{C}, \bar{P}_{i,j}^{CF}, P_i^{CB} | \text{C4 is satisfied}, i \in \mathcal{S}_N, j \in \mathcal{K}^{CF}, k \in \mathcal{K}^{C}\}$ is also a convex set.

APPENDIX C
PROOF OF EXISTENCE OF A SOLUTION FOR (24)

By differentiating (5) and then removing the non-negative term, we have

$$\frac{\partial \bar{R}_i^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}} \begin{cases} = \frac{B^W \alpha_i^W}{\ln(2)} \cdot \frac{\partial P_{avg,i}^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}}, & \text{at } P_i^{CB} = 0; \\ < \frac{B^W \log_2\left(1+\alpha_i^W P_i^{CB}\right)}{P_i^{CB}} \cdot \frac{\partial P_{avg,i}^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}}, & \text{for } P_i^{CB} > 0. \end{cases} \quad (40)$$

Partial derivatives of $\bar{R}_i^{CB}(\mathbf{P}^{CB})$ and $P_{avg,i}^{CB}(\mathbf{P}^{CB})$ with respect to P_i^{CB} are positive and monotonically decreasing functions of P_i^{CB}, because $\bar{R}_i^{CB}(\mathbf{P}^{CB})$ and $P_{avg,i}^{CB}(\mathbf{P}^{CB})$ are concave increasing functions with respect to P_i^{CB}. Furthermore, the value of $(B^W \log_2(1+\alpha_i^W P_i^{CB}))/P_i^{CB}$ decreases from $B^W \alpha_i^W / \ln(2)$ to 0 as P_i^{CB} goes from 0 to ∞. Therefore, there exist a p_i, $(p_i > 0)$ such that at each $P_i^{CB} > p_i$,

$$\frac{\partial \bar{R}_i^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}} < \frac{\mu_i^*}{1+\lambda_i} \cdot \frac{\partial P_{avg,i}^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}}. \quad (41)$$

Thus, from (22), (23) and (41), there exist a P_i^{CB*} which is the solution for (24), and $P_i^{CB*} \in (0, p_i)$.

APPENDIX D
PROOF OF CONVERGENCE OF THE ITERATIVE ALGORITHM WHICH CALCULATES $\mathbf{P}^{CB*}$

At the nth iteration, P_i^{CB*} is calculated using (24) such that when $P_i^{CB} = P_i^{CB*}$, the left hand and the right hand side terms of (41) are equal. Since

$$\frac{\left(\frac{\partial^2 \bar{R}_i^{CB}(\mathbf{P}^{CB})}{\partial P_j^{CB} \partial P_i^{CB}}\right)}{\left(\frac{\partial \bar{R}_i^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}}\right)} > \frac{\left(\frac{\partial^2 P_{avg,i}^{CB}(\mathbf{P}^{CB})}{\partial P_j^{CB} \partial P_i^{CB}}\right)}{\left(\frac{\partial P_{avg,i}^{CB}(\mathbf{P}^{CB})}{\partial P_i^{CB}}\right)}, \quad \forall P_j^{CB} > 0, i \neq j, \quad (42)$$

if P_j^{CB*} is increased at the nth iteration, the left hand side of (41) becomes larger than the right hand side of (41). Thus, P_i^{CB*} increases at the $(n+1)$th iteration. Therefore, by induction, $P_i^{CB*},\ \forall i$ increase in each iteration. However, each P_i^{CB*} is upper bounded by p_i, where $(B^W \log_2(1+\alpha_i^W p_i))/p_i = \mu_i^*/(1+\lambda_i)$. Therefore, the algorithm converges. Furthermore, it should be noted that each P_i^{CB*} converges to a value which is smaller than p_i, because if $P_i^{CB*} = p_i$, it violates (40).

REFERENCES

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2012–2017," Cisco Syst., San Jose, CA, USA, Cisco White Paper, Feb. 2013.

[2] *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE P802.11-REVmb/D12, Mar. 2012.

[3] R. Ferrus, O. Sallent, and R. Agusti, "Interworking in heterogeneous wireless networks: Comprehensive framework and future trends," *IEEE Wireless Commun.*, vol. 17, no. 2, pp. 22–31, Apr. 2010.

[4] X. Pei, T. Jiang, D. Qu, G. Zhu, and J. Liu, "Radio-resource management and access-control mechanism based on a novel economic model in heterogeneous wireless networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 3047–3056, Jul. 2010.

[5] M. Ismail and W. Zhuang, "A distributed multi-service resource allocation algorithm in heterogeneous wireless access medium," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 425–432, Feb. 2012.

[6] M. J. Neely, E. Modiano, and C.-P. Li, "Fairness and optimal stochastic control for heterogeneous networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 2, pp. 396–409, Apr. 2008.

[7] "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Stage 2," Sophia-Antipolis, France, 3GPP TS 36.300 Rep. V11.6.0, 2013.

[8] H. S. Chang, P. J. Fard, S. I. Marcus, and M. Shayman, "Multitime scale Markov decision processes," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 976–987, Jun. 2003.

[9] W. Song, W. Zhuang, and Y. Cheng, "Load balancing for cellular/WLAN integrated networks," *IEEE Netw.*, vol. 21, no. 1, pp. 27–33, Jan./Feb. 2007.

[10] W. Song and W. Zhuang, *Interworking of Wireless LANs and Cellular Networks*. New York, NY, USA: Springer-Verlag, 2012.

[11] J. Xu, Y. Jiang, and A. Perkis, "Multi-service load balancing in a heterogeneous network," in *Proc. WTS*, 2011, pp. 1–6.

[12] C. Luo, H. Ji, and Y. Li, "Utility-based multi-service bandwidth allocation in the 4G heterogeneous wireless access networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2009, pp. 1–5.

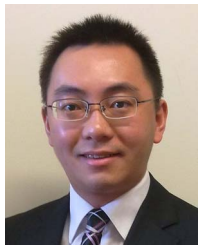
[13] D. Niyato and E. Hossain, "A noncooperative game-theoretic framework for radio resource management in 4G heterogeneous wireless access networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 3, pp. 332–345, Mar. 2008.

[14] P. Xue, P. Gong, J. H. Park, D. Park, and D. K. Kim, "Radio resource management with proportional rate constraint in the heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1066–1075, Mar. 2012.

[15] A. T. Gamage and X. Shen, "Uplink resource allocation for interworking of WLAN and OFDMA-based femtocell systems," in *Proc. IEEE ICC*, Budapest, Hungary, Jun. 2013, pp. 6071–6075.

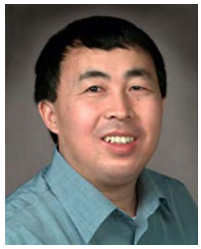
[16] P. Wang, H. Jiang, and W. Zhuang, "Capacity improvement and analysis for voice/data traffic over WLAN," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1530–1541, Apr. 2007.
[17] J. Zhu and A. O. Fapojuwo, "A new call admission control method for providing desired throughput and delay performance in IEEE 802.11e wireless LANs," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 701–709, Feb. 2007.
[18] H. S. Wang and N. Moayeri, "Finite-state Markov channel—A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.
[19] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535–547, Mar. 2000.
[20] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley, 1994.
[21] E. Altman, *Constrained Markov Decision Processes*. Boca Raton, FL, USA: CRC, 1999.
[22] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
[23] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[24] D. P. Bertsekas, *Non-Linear Programming*. Belmont, MA, USA: Athena Scientific, 2003.
[25] M. S. Alam, J. W. Mark, and X. Shen, "Relay selection and resource allocation for multi-user cooperative OFDMA networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2193–2205, May 2013.
[26] C. Kelley, *Solving Nonlinear Equations With Newton's Method*. Philadelphia, PA, USA: SIAM, 2003.
[27] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. S. Panwar, "CoopMAC: A cooperative MAC for wireless LANs," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 340–354, Feb. 2007.
[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Function with Formulas, Graphs, and Mathematical Tables*. Washington, DC, USA: NBS, 1964, ser. Applied Mathematics Series 55.
[29] "WINNER II Interim Channel Models," WINNER II, Tech. Rep. D1.1.1 v1.0, Dec. 2006. [Online]. Available: https://www.ist-winner.org/
[30] T. S. Rappaport, *Wireless Communications*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.
[31] H. Shen, L. Cai, and X. Shen, "Performance analysis of TFRC over wireless link with truncated link level ARQ," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1479–1487, Jun. 2006.

Amila Tharaperiya Gamage (S'07) received the B.E. degree in electronics and telecommunications engineering from Multimedia University, Cyberjaya, Malaysia, in 2008 and the M.E. degree in telecommunications engineering from the Asian Institute of Technology, Pathumthani, Thailand, in 2011. He is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. From 2008 to 2009, he was a Solutions Architect with Dialog Telekom PLC, Sri Lanka. His current research interests include resource management for interworking heterogeneous wireless networks, cooperative communication, and cloud computing. He is a corecipient of the Best Paper Awards from the IEEE International Conference on Communications 2014.



Hao Liang (S'09–M'14) received the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2013. From 2013 to 2014, he was a Postdoctoral Research Fellow with the Broadband Communications Research Laboratory and Electricity Market Simulation and Optimization Laboratory, University of Waterloo. Since July 2014, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Alberta, Canada. His research interests are in the areas of smart grid, wireless communications, and wireless networking. He is the recipient of the Best Student Paper Award from IEEE 72nd Vehicular Technology Conference (VTC Fall-2010), Ottawa, ON.

Dr. Liang was the System Administrator of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY (2009–2013). He served as a Technical Program Committee Member for major international conferences in both information/communication system discipline and power/energy system discipline, including IEEE International Conference on Communications (ICC), IEEE Global Communications Conference (Globecom), IEEE VTC, IEEE Innovative Smart Grid Technologies Conference (ISGT), and IEEE International Conference on Smart Grid Communications (SmartGridComm).



Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the B.Sc. degree from Dalian Maritime University, Dalian, China, in 1982, and the M.Sc. and Ph.D. degrees from Rutgers University, Camden, NJ, USA, in 1987 and 1990, all in electrical engineering. He is a Professor and University Research Chair with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. From 2004 to 2008, he was the Associate Chair for Graduate Studies.

His research focuses on resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular *ad hoc* and sensor networks. He is a coauthor/editor of six books and has published more than 600 papers and book chapters in wireless communications and networks, control, and filtering. Dr. Shen served as the Technical Program Committee Chair/Cochair for IEEE Infocom'14, IEEE VTC'10 Fall, the Symposia Chair for IEEE ICC'10, the Tutorial Chair for IEEE VTC'11 Spring and IEEE ICC'08, the Technical Program Committee Chair for IEEE Globecom'07, the General Cochair for Chinacom'07 and QShine'06, the Chair for IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He also serves/served as the Editor-in-Chief for IEEE Network, Peer-to-Peer Networking and Application, and IET Communications; a Founding Area Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS; an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, Computer Networks, and ACM/Wireless Networks, etc.; and the Guest Editor for IEEE JSAC, IEEE WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS MAGAZINE, and ACM Mobile Networks and Applications, etc. He received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004, 2007, and 2010 from the University of Waterloo, the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. He is a Registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, and a Distinguished Lecturer of IEEE Vehicular Technology Society and Communications Society.