Doctoral Thesis

Airtime Management for Low-Latency Densely Deployed Wireless Networks

Bo YIN

Graduate School of Informatics, Kyoto University

March 2021

Preface

This thesis proposes airtime management methods for reducing the media access control (MAC) delay in densely deployed wireless networks. Moreover, this thesis also provides performance analysis on these methods focusing on the trade-offs between MAC delay and other performances.

The delay of a wireless communication system refers to the amount of time that information travels from the information source to the destination. Reducing the delay of a wireless communication system is an important topic worth studying. With the fast growth of information technology, there emerge many delay-sensitive applications, e.g., low-latency video streaming. A low-latency wireless communication system is considered necessary to provide high-quality delivery to these delay-sensitive applications.

The studying scope of this thesis is the MAC delay. The MAC delay is a delay that typically occurs in a communication system in which multiple users share the wireless channel. It occurs when the amount of wireless resources is insufficient or when it takes time for wireless resources to be allocated. Note that, the delay caused by other reasons is beyond the scope of this thesis. For example, the delay caused by transmitter signal processing, radio propagation, and receiver signal processing is out of scope. Also, the delay generated by communication overhead is also out of the scope.

In this thesis, the MAC delay in two types of wireless communication systems is studied, contention-based system and allocation-based system.

The first topic is related to wireless local area network (WLAN) system that adopts the contention-based MAC. The latest WLAN standard, i.e., IEEE 802.11ax, introduces a mechanism for concurrent transmission among co-channel neighboring cells with low power. This mechanism increases the airtime of the wireless terminal. This mechanism, however, has the side effect of increasing the packet

collision rate. In this topic, a method is proposed to reduce this side effect by utilizing the reinforcement learning algorithm. Specifically, the identifier in the preamble of the WLAN frame is used to identify the interferer. In doing so, simultaneous transmissions among cells that do not cause frame collision are promoted. On the other hand, simultaneous transmissions among cells that can cause frame collision are suppressed. Furthermore, theoretical performance analysis is performed based on Markov state transition diagrams. Simulation results show that the MAC delay among co-channel neighboring cells is significantly reduced.

The second topic focuses on the flow-in-the-middle (FIM) starvation problem in densely deployed WLANs. It occurs when a transmitter is within the carrier sense range of two exterior transmitters while the two exterior transmitters are not within the carrier sense range of each other. Under heavy traffic conditions, the middle transmitter detects the channel being occupied for a prolonged time and suffers from extremely low transmission opportunities. In this topic, a new distributed frequency channel selection method is proposed that does not cause the starvation problem. Specifically, a contention graph is introduced where the access point (AP) is a node and the availability of carrier detection is an edge. The proposed method reduces the number of linear chain topologies on the contention graph. Furthermore, by proving the existence of the potential function in game theory, the distributed channel selection method is proved to converge. Finally, computer simulation shows that the number of APs for which transmission opportunities are limited can be significantly reduced with limited sacrifice of the efficiency of the entire system.

The final topic is related to the cellular communication system that adopts the allocation-based MAC. In particular, a scenario is considered in which two operators deploy shared base stations (BSs) to reduce capital investment. In this case, it is possible to improve the communication quality by appropriately allocating the user traffic between the existing non-shared BSs and the additional shared BSs. In order to quantitatively evaluate MAC delay, this thesis introduces an evaluation metric called cell load. The physical meaning of the cell load is the normalized airtime of the BS, which is also the expected value of the MAC delay detected by the user equipments (UEs) when the normalized size data is considered. The cell load can be calculated from the expected value of the reciprocal of the data rate of the UEs. In this topic, by assuming that the positional distribution of the BSs

follows the Poisson point process, the cell load is theoretically analyzed. Specifically, the proportional fair load balancing scheme and the round robin type load balancing scheme are compared. The main finding is that the former scheme impairs inter-operator resource isolation but can accommodate more traffic such that the network does not overload. Finally, the correctness of the analysis results was confirmed by numerical evaluation.

The chapters of this thesis are as follows. Chapter 1 explains the background and contributions of this thesis. Chapter 2 introduces related technologies to this work. Chapters 3 to 5 describe the above mentioned topics in detail. Finally, Chapter 6 concludes this thesis.

Acknowledgements

First of all, I would like to express my sincere gratitude to Professor Masahiro Morikura, for his continuous support and guidance in my research. Without his persistent help and constant encouragement, this work would have never been completed.

I would like to offer my special thanks to Associate Professor Koji Yamamoto and Assistant Professor Takayuki Nishio, who are members of Integrated-Media Communications Laboratory of Graduate School of Informatics, Kyoto University. Their constructive feedbacks and discussions have helped my work.

I would like to express my deep appreciation to Professor Hiroshi Harada and Professor Eiji Oki of Graduate School of Informatics, Kyoto University for valuable advice and incisive comments. Their advice and comments have been a great help in improving this thesis.

I wish to express my gratitude to Professor Seong-Lyun Kim of the School of Electrical & Electronic Engineering, Yonsei University, for his constructive comments and suggestions. I am also grateful for the benefits received by the fruitful discussions with members in the NTT Access Service Systems laboratory.

I am also very grateful to laboratory secretaries and student members of Integrated-Media Communications Laboratory of Graduate School of Informatics, Kyoto University. Their considerable support and comments were invaluable.

Finally, I would like to thank my parents and friends for their support and encouragement throughout my work.

Contents

List of Figures

List of Tables

Notations

Notation	Description
$\mathbb{R}$	real numbers
$\mathbb{R}_{\geq 0}$	non-negative real numbers
$\emptyset$	empty set
$\setminus$	difference of sets
$\#\{\cdot\}$	number of elements of a set
$\mathcal{G}, \mathcal{G}^{\infty}$	one-stage game, repeated game
$\mathbb{E}[\cdot]$	expectation operator
$\mathbb{E}[\cdot \vert \text{condition}]$	conditional expectation
$\mathbb{P}[A]$	probability of event A
$\mathbb{P}[A \vert B]$	conditional probability of event A given B
${}_2F_1(\cdot,\cdot;\cdot;\cdot)$	hypergeometric function

Abbreviations

Abbreviation	Description
3GPP	third generation partnership project
ACK	acknowledgement
AP	access point
BoE	back-of-the-envelope
BS	base station
CCA	clear channel assessment
CCDF	complementary cumulative distribution function
CapEx	capital expenditure
DC	dual connectivity
DCF	distributed coordination function
DIFS	DCF inter frame space
DSC	dynamic sensitivity control
ECDF	empirical cumulative distribution function
FIM	flow-in-the-middle
MAC	media access control
MCS	modulation and coding scheme
MDP	Markov decision process
MIS	maximal independent state
mmWave	millimeter-wave
MOCN	multi-operator core network
MORAN	multi-operator radio access network
NE	Nash equilibrium
OBSS	overlapping basic service set
OBSS_PD	overlapping basic service set packet detect

Abbreviation	Description
OpEx	operating expenditure
PER	packet error rate
PF	proportionally fair
PG	potential game
PHY	physical
QoS	quality-of-service
RAN	radio access network
RR	round robin
RSSI	received signal strength indicator
RTS/CTS	request to send/clear to send
RUQL	repeated update Q learning
SAP	spatial adaptive play
SIFS	short inter frame space
SINR	signal-to-interference-plus-noise ratio
SIR	signal-to-interference ratio
STA	station
UE	user equipment
WLAN	wireless local area network

Chapter 1

Introduction

1.1 Background

The telecommunication traffic has been constantly growing in recent years. According to the latest research of Japanese ministry of internal affairs and communications, the telecommunication traffic in Japan has reached 4.0 Tbit/s, where the peak traffic is 5.4 Tbit/s and the per-subscriber traffic is 22 kbit/s [1]. Compared to the former year, these values have increased by 1.3, 1.3, and 1.2 times, respectively [1]. Under this background, the technologies of wireless communication and wireless networkings have been constantly evolving. The cellular technologies are switching from 4G to 5G. Meanwhile, Wi-Fi®, defined by a family of IEEE 802.11 standards, is evolving into the IEEE 802.11ax [2] standard.

In the next generation telecommunication technologies, the low latency wireless communications and low latency wireless networkings are important research topics. The delay (or latency) of a wireless communication system is generally defined as the amount of time that the information travels from the information source to the destination [1]. Many emerging applications nowadays require extremely strict delay constraints and high reliability of the wireless communication system, e.g., online video conference. Most of these applications are not only bandwidth-hungry but also delay-sensitive [4]. A robust and reliable low-latency wireless communication system is necessary to support these applications. With the increasing popularity of delay-sensitive applications, reducing the delay of a

[1] More specific definitions are given in the literatures, e.g., [3].

wireless communication system has become an interesting research topic worth studying. At the time of writing this thesis, the low-latency wireless communication system is still an open and hot research topic [4].

The studying scope of this thesis is the media access control (MAC) delay, i.e., the delay due to waiting being granted the access to the physical transmission medium. In a multi-user wireless communication system, the MAC delay occurs when the amount of wireless resources is insufficient and when it takes time for wireless resources to be allocated.

The MAC delay is studied based on the type of multiple access schemes. Specifically, the MAC delay is categorized into the contention delay and the scheduling delay in this thesis. The contention delay is further categorized into the perfect carrier sensing case and the imperfect carrier sensing case. Therefore, the categories of MAC delay studied in this thesis are as follows.

- Contention delay
  - Perfect carrier sensing case
  - Imperfect carrier sensing case
- Scheduling delay

The reason of making this categorization is that similar categorization is commonly provided in literature. For example, multiple access schemes are categorized into fixed assigned, demand-based assigned, and random access schemes in [5]. In [6], they are categorized into conflict-free access protocols, Aloha protocols, and carrier sensing protocols. In [7], they are categorized into distributed, centralized, and hybrid schemes. Note that, there does not exist the only one way of categorizing MAC delay or multiple access schemes. Despite there are some differences in the ways of categorization, the contention-based schemes and the scheduling-based schemes are two basic multiple access schemes. As presented in Section 1.2, the causes of MAC delay in these cases are basically different.

Note that, the delay generated in other layers is beyond the scope of this thesis. For example, the delay generated in transmitter channel coding, signal processing, transmission, radio propagation, receiver signal processing, and decoding is out of scope. Also, the delay generated by communication overhead, such as request-to-send/clear-to-send (RTS/CTS) is out of scope.

1.2 Research Challenges

In this section, the details of the above-mentioned MAC delay is provided. The research challenges of reducing each type of delay is also discussed.

1.2.1 Contention Delay: Perfect Carrier Sensing Case

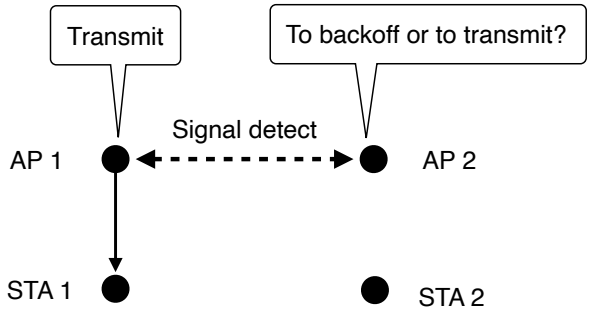


Figure 1.1: An illustration of the contention delay in perfect carrier sensing case.

This type of delay occurs in the wireless communication system where the contention-based MAC protocol is implemented, e.g., the IEEE 802.11 wireless local area network (WLAN). Based on the IEEE 802.11 distributed coordination function (DCF) [8], access points (APs) and stations (STAs) access the channel following a contention-based manner. A random backoff procedure is required before each transmission. Furthermore, during the backoff procedure, the AP or STA will freeze the backoff procedure as long as it detects any on-going transmission in the channel as shown in Fig. 1.1. Finally, the retransmission takes place if packet collisions occur. Overall, the causes of contention delay contains:

- Counting down the backoff counter
- Waiting until the detected transmission finishes
- Retransmissions due to packet collisions

As shown in the simulation in Chapter 3 of this thesis, the MAC delay becomes severer when as the amount of traffic or the number of APs or STAs increases.

Regarding this topic, the research challenge is to reduce the MAC delay in a densely deployed WLANs. This is an important topic because the deployment of WLAN devices has been constantly growing since the 2000s [9]. It is predicted that, globally, total public WLAN hotspots will grow sevenfold from 64.2 million in 2015 to 432.5 million by 2020 [10]. Mobile offload traffic will increase from 3.9 exabytes/month to 38.1 exabytes/month in the meanwhile [10].

This topic is studied in Chapter 3. To reduce MAC delay in the densely deployed WLANs, a channel access scheme is proposed to facilitate concurrent transmissions among co-channel transmitters while reducing retransmissions due to packet collisions. In other words, a channel access scheme is proposed to enhance feasible concurrent transmissions and to suppress the non-feasible ones.

1.2.2 Contention Delay: Imperfect Carrier Sensing Case

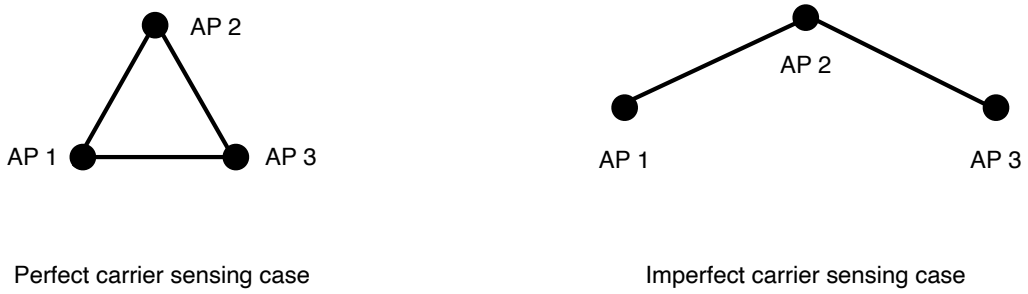


Figure 1.2: An illustration of contention graph.

The previous topic mainly focuses on the small scale wireless networks where perfect carrier sensing operation is assumed. In large scale wireless networks, however, the distance between nodes is too far to perform perfect carrier sensing. In the IEEE 802.11 WLAN, for example, the receiver sensitivity is -82 dBm [8], i.e., the detection of the signal below this level is not guaranteed. Figure 1.2 shows an illustration of contention graph of wireless CSMA networks, where the node represent the AP and the edge represent that two APs detects each other. In this thesis, the imperfect carrier sensing network refers to the network where not all the nodes are within the carrier sensing ranges of each other. Related studies have pointed out that starvation problem [11] may happen in imperfect carrier sensing

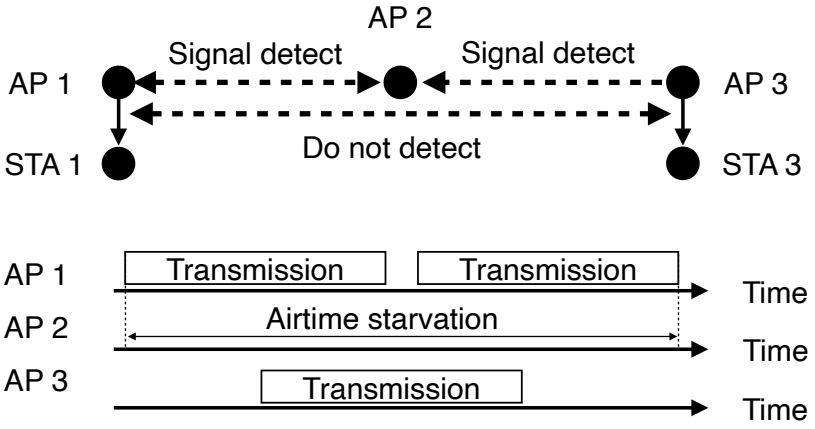


Figure 1.3: Flow-in-the-middle starvation problem.

networks.

Figure 1.3 shows an illustration of the flow-in-the-middle (FIM) starvation problem defined in [11]. Consider three co-channel APs as shown in Fig. 1.3. The middle AP is within the carrier sense range of two exterior APs while the two exterior APs are not within the carrier range of each other. Under heavy traffic conditions, the middle AP detects the channel being occupied for a prolonged time and suffers from extremely low transmission opportunities, because the transmissions of exterior APs might overlap.

Regarding this topic, the research challenge is how to design a distributed algorithm that guarantees convergence to stable solutions. This is because WLAN devices operate in an unlicensed band and some APs are deployed by consumers randomly [9]. A distributed algorithm is necessary since there does not exist a centralized controller to manage all of the APs.

This topic is studied in Chapter 4. To reduce airtime starvation due to imperfect carrier sensing, a channel selection algorithm is proposed to avoid the occurrence of imperfect carrier sensing in the same frequency channel. Game theoretic approach is utilized to study the convergence property of the proposed scheme.

1.2.3 Scheduling Delay

In a scheduling-based wireless communication system, the main cause of MAC delay is due to the intra-cell radio resource scheduling. Consider a cellular com-

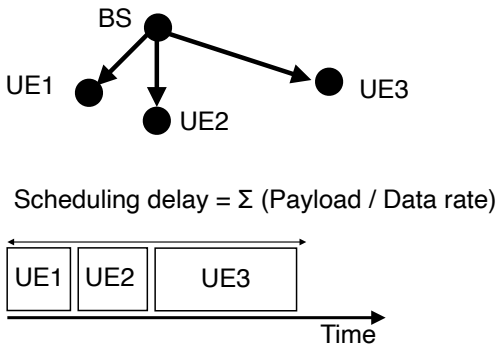


Figure 1.4: An illustration of scheduling delay.

munication model as illustrated in Fig. 1.4, where the base station (BS) schedules the radio resources to each user equipment (UE) in a round-robin manner. Denote the downlink traffic load and data rate at location s as $\omega(s)$ and $R(s)$, respectively. The normalized airtime to transmit the downlink traffic is given as follows.

$$\rho = \int_A \frac{\omega(s)}{R(s)} \, \mathrm{d}s, \tag{1.1}$$

where A is the cell area. This normalized time here can also be interpreted as an expectation of the scheduling delay at the UE [12]. Despite its simplicity, this equation presents intuitions of reducing scheduling delay. From this equation, the approaches of reducing the scheduling delay includes: reducing the traffic $\omega(s)$, increasing the data rate $R(s)$, and reducing the cell area A by shrinking coverage.

The third topic in this thesis mainly studies reducing scheduling delay from a traffic offloading perspective, i.e., reducing $\omega(s)$ in (1.1). This approach is of particular interests under the background of multi-operator infrastructure sharing in the next-generation cellular communication system.

The research challenge at here is that: how to split user traffic over the shared radio access network (RAN) and the non-shared RAN to reduce scheduling delay? This topic is studied in Chapter 5. To provide theoretical guidelines for multi-operator RAN sharing and dynamic offloading, a stochastic geometric analysis is performed that incorporates the cell load equation.

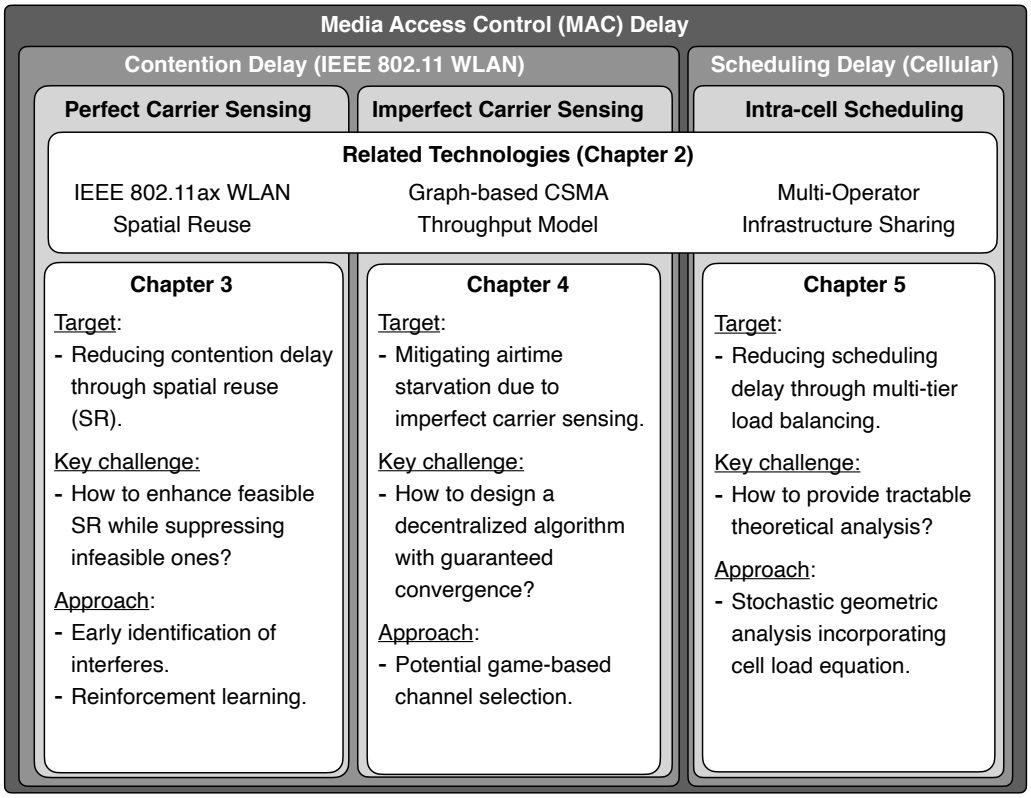


Figure 1.5: Chapter overview.

1.3 Outline and Contributions of This Thesis

As categorized above, this thesis focuses on reducing the MAC delay from three different subtopics separately, i.e., the delay due to resource contention under perfect carrier sensing, resource contention under imperfect carrier sensing, and resource scheduling. The chapter overview is summarized in Fig. 1.5.

The related technologies are presented in Chapter 2. Preliminary knowledges and some related works on the similar topics are reviewed in this chapter. Specifically, the IEEE 802.11ax spatial reuse, the graph-based throughput model for CSMA networks, and the multi-operator infrastructure sharing are presented.

The first subtopic is presented in Chapter 3. It is related to IEEE 802.11ax [2] WLAN system that adopts the contention-based MAC. In this chapter, a reinforcement learning-based spatial reuse scheme for WLANs is proposed and analyzed. In this scheme, when an AP (or an STA) overhears an on-going transmission, it

decodes the information in the frame header to identify the transmitter and decides whether or not to exploit spatial reuse accordingly. Specifically, it decides whether to stop receiving the remaining part of the frame and start its own transmission or to refrain from channel access until the detected transmission finishes. Through reinforcement learning algorithms, e.g., the repeated update Q-learning (RUQL) algorithm, the agent learns the optimal decision in the sense of reducing the media access control layer delay. Moreover, the proposed scheme with the spatial reuse operation in IEEE 802.11ax is compared, which makes the spatial reuse decision only based on a binary identification of the detected interferer, i.e., whether it is in my cell or neighboring cells. The proposed scheme, however, treats different interferers differently for exploiting spatial reuse. From a theoretical perspective, a theoretical bound on the gains in the value function is derived, i.e., the discounted sum of delay, due to making non-binary identifications. Simulation evaluations confirm that the proposed scheme achieves high throughput by reducing the time of freezing backoff counter while not increasing the time of failed transmissions.

The second subtopic is presented in Chapter 4. Distributed channel selection schemes are proposed to mitigate the FIM starvation in dense WLANs. The FIM starvation occurs when the middle transmitter is within the carrier sense range of two exterior transmitters, while the two exterior transmitters are not within the carrier sense range of each other. Since an exterior transmitter sends a frame regardless of the other, the middle transmitter has a high probability of detecting the channel being occupied. Under heavy traffic conditions, the middle transmitter suffers from extremely low transmission opportunities, i.e., airtime starvation. The basic idea of the proposed schemes is to let each AP select the channel which has less three-node-chain topologies within its two-hop neighborhood. The proposed schemes are formulated in strategic form games. Payoff functions are designed so that they are proved to be potential games. Therefore, the convergence is guaranteed when the proposed schemes are conducted in a distributed manner by using unilateral improvement dynamics. Moreover, evaluations are conducted through graph-based simulations and the ns-3 simulator. Simulations confirm that the FIM starvation has been mitigated since the number of three-node-chain topologies has been significantly reduced. The 5th percentile throughput has been improved.

The third subtopic is presented in Chapter 5. The infrastructure sharing offers

great potential for the cost-efficient deployment of the next-generation RANs. By sharing the usage of RANs, mobile network operators can rapidly develop new services with substantially reduced expenditures. In such a background, the following problem has become apparent: From an operator's perspective, how to split user traffic over the shared RAN and the non-shared RAN to accomplish load balancing? In order to provide guidelines for multi-operator load balancing, in this chapter, stochastic geometric analysis is performed that incorporates the cell load equation. The theoretically conditions on multi-operator traffic density is derived such that the networks do not overload. It is theoretically proved that the proportionally fair (PF) load balancing can accommodate more traffics than the round robin (RR) load balancing. Moreover, an incomplete information load balancing game is formulated to study the scenario in which multiple operators do not hope to exchange any private information with other operators, such as the amount of incoming traffic or base station densities. Simulation results show that, despite the information is incomplete, the reinforcement learning-based solutions achieve as good performance as the complete information scenario under certain traffic conditions.

Chapter 2

Related Technologies

2.1 Spatial Reuse in IEEE 802.11ax WLAN

The significant growth in the number of WLAN devices in recent years [13–15] has resulted in the common occurrence of overlapping basic service sets (OBSSs), i.e., co-located WLAN cells operating in the same frequency channel. In dense OBSSs scenario, the throughput degradation becomes a severe problem because concurrent transmissions among OBSSs were not allowed in previous IEEE 802.11 standards, e.g., IEEE 802.11n, 11ac. In particular, an AP or an STA has to defer its channel access when it detects the transmission of any other APs or STAs.

The IEEE 802.11ax standard has approved a new operation called the *OBSS packet detect (OBSS_PD)-based spatial reuse operation* to improve spatial frequency reuse in high-density scenario [2, 13–15]. This operation improves spatial reuse by allowing concurrent transmissions among OBSSs. As shown in Fig. 2.1, once an AP or an STA has detected an on-going transmission, it immediately identifies whether this transmission is in OBSS or not. This identification is done by checking the basic service set (BSS) color field in the frame header. If the detected transmission is in OBSS and its interference power is lower than a predefined threshold, i.e., OBSS_PD, the AP or STA stops receiving the remaining part of the frame and regards the wireless medium as idle, i.e., it is feasible to start transmission.

This OBSS_PD-based spatial reuse operation, however, has a major challenge.

The challenge is that comparing the interference power with the predefined threshold OBSS_PD has limited predictive value in determining the success or failure of concurrent transmissions. In other words, a transmitter cannot guarantee that its receiver receives its transmission successfully under the interference from OBSS, even if the interference is less than OBSS_PD. This results in performance degradation since if packet loss occurs, the concurrent transmission would be nonsense and should not be performed.

There are at least two reasons causing this unreliability of the OBSS_PD-based spatial reuse operation. First of all, this operation only identifies whether a detected transmission is in my BSS or in OBSS. It, however, treats interferers in different OBSSs indifferently. Consider the example in Fig. 2.2, where AP 0 tries to send a packet to STA 0. OBSS AP 1 is close to STA 0 and OBSS AP 2 is far from STA 0. In case (a), AP 0 detects the transmission of OBSS AP 1 whose interference power is I. In case (b), AP 0 detects the transmission of OBSS AP 2 whose interference power is also I. Although the measured interference at AP 0 is identical in cases (a) and (b), the receiver experiences different amounts of interference power. Concurrent transmission with OBSS AP 1 is more likely to fail. Hence, it is desirable if AP 0 can identify which interferer is transmitting before deciding whether or not to transmit concurrently with that interferer. Besides the first reason, another essential reason is that it is hard to establish an accurate and universal model of the one-to-one correspondence between interference power and packet loss probability in the actual wireless channel [16, 17].

2.2 Graph-based CSMA Throughput Model

This section introduces a graph-based throughput model for wireless carrier sense multiple access (CSMA) network. It is the back-of-the-envelop (BoE) throughput model proposed in [18].

The throughput model proposed in [19] assumes that all the wireless stations are within the carrier sensing ranges of each other. This assumption does not always hold in dense WLANs, especially when parameters such as transmission power or carrier sensing threshold are adjusted [20]. The BoE model provides model of throughput when not all the wireless stations sense each other. This model has been used to calculate the throughput of WLAN in previous works

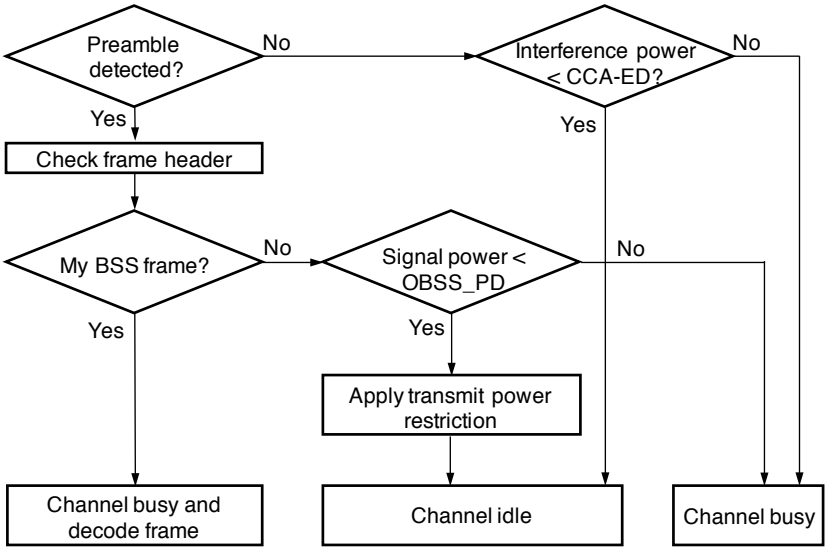


Figure 2.1: The OBSS_PD-based spatial reuse operation in the IEEE 802.11ax. ©2020 IEEE.

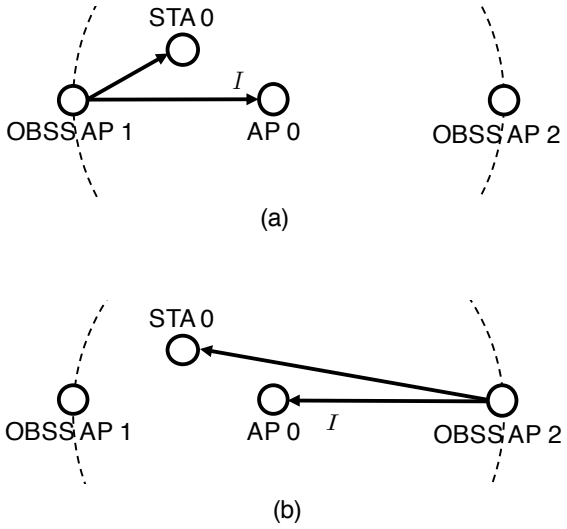


Figure 2.2: An example showing the flaw of the OBSS_PD-based spatial reuse operation. ©2020 IEEE

such as [9].

The calculation is conducted on the contention graph $C := (\mathcal{N}, \mathcal{E})$, where the node set $\mathcal{N}$ denotes an index set of APs. The set of undirected edges $\mathcal{E}$ on the contention graph denotes the carrier sensing availability of two APs. Specifically,

an undirected edge ij denotes that AP i and AP j are within the carrier sense range of each other.

Before introducing the procedure of throughput calculation using the BoE model, some definitions of terms in the graph theory are given as follows.

Definition 1 (Independent Set)**.** *Independent set is defined as a set of nodes,* $\mathrm{IS} \subseteq \mathcal{N}$*, on the contention graph* $\mathcal{C}$*, that for any two nodes* $\{i, j\} \subseteq \mathrm{IS}$*,* $ij \notin \mathcal{E}$*.*

Definition 2 (Maximum Independent Set)**.** *An maximum independent set,* $\mathrm{MIS} \subseteq \mathcal{N}$*, is an independent set of largest possible number of nodes on a given graph* $\mathcal{C}$*.*

This model ignores packet collisions and assumes saturated traffic conditions, that is, each transmitter is back-logged with packets. Given the contention graph $\mathcal{C}$ of a certain network, the normalized throughput achieved by node i can be calculated as follows.

$$\mathrm{TR}_{\mathrm{norm}\,i} = \frac{\sum_{\mathcal{S}\in\mathrm{MIS}} \mathbb{1}_{\{i\in\mathcal{S}\}}}{\sum_{\mathcal{S}\in\mathrm{MIS}} 1}. \tag{2.1}$$

The numerator represents the number of MISs in which AP i is included and the denominator represents the number of the MISs on the whole contention graph.

A normalized throughput of 1 corresponds to the throughput an AP achieves in isolation of other APs, as if it were the only AP in the whole network. The actual throughput of AP i can be computed as,

$$\mathrm{TR_i} = \mathrm{TR}_{\mathrm{norm}\,i} \cdot \mathrm{TR}_{\mathrm{single\,link}\,i}. \tag{2.2}$$

A typical value of $\mathrm{TR}_{\mathrm{single\,link}}$, for example in the case of IEEE 802.11a, is about 30.91 Mbps in the UDP session [18].

Consider the following example whose contention graph is shown in Fig. 2.3. The independent set on the graph are $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$, $\{1,4\}$ and $\{2,4\}$. Among them, the maximum independent set are $\{1,4\}$ and $\{2,4\}$. According to (2.1), the normalized throughput of the four APs are 1/2, 1/2, 0/2 and 2/2. This indicates that AP 3 will suffer from airtime starvation problem whereas AP 4 achieves near highest throughput.

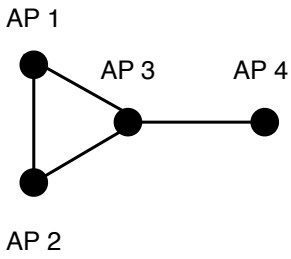


Figure 2.3: An example of contention graph with four APs.

2.3 Multi-Operator Infrastructure Sharing

The infrastructure sharing among mobile network operators offers great potential for cost-efficient deployment and operation of the next-generation RAN [21, 22]. Compared to independent deployments of wireless access infrastructure, the inter-operator infrastructure sharing allows operators to share a part of site components among each other. Hence, it can rapidly expand the coverage area while substantially reducing the capital expenditures (CapEx) and operational expenditures (OpEx) for operators [23,24]. In the standardization society, two types of architectures are defined in the 3GPP TS 23.251 [25]. Their sharing components are compared in Table 2.1. The multi-operator RAN (MORAN) architecture exploits only the infrastructure sharing, whereas the multiple operator core network (MOCN) architecture further exploits the spectrum sharing to avoid under-utilization of the spectrum. As the frequency spectrum is an extremely limited resource, it is crucial to make the best utilization of its resources to obtain maximum values and benefits [24, 26]. Moreover, intelligent spectrum sharing is also considered to play important roles in future 6G [27]. To this end, this chapter focuses on studying the joint infrastructure and spectrum sharing architecture, i.e., MOCN-like architecture.

Depending on the circumstances, shared RANs are less likely to be used in a stand-alone manner. They are rather deployed as a complement of the non-shared RANs for mobile data traffic offloading. For instance, Rebato *et al.* [28] studied a hybrid sharing architecture, where data packets are scheduled through two millimeter-wave (mmWave) carriers with different characteristics: a mmWave band with exclusive access at 28 GHz and a mmWave band at 73 GHz, where spectrum is pooled between multiple operators. In such case, dual connectivity

Table 2.1: Sharing components comparison.

Component	MORAN	MOCN
Core network	Independent	Independent
Base station	Shared	Shared
Carriers	Independent	Shared

Table 2.2: Previous works on WLANs spatial reuse.

Reference	Adaptive CCA?	Identifying interferers?	Learn-based?
[29–35]	No	No	No
[36–40]	Yes	No	No
[41]	Yes	Yes	No
This work	Yes	Yes	Yes

(DC) plays an important role in aggregating the radio resources at both shared RANs and non-shared RANs. DC enables UEs to simultaneously communicate with a shared BS and an exclusive BS via different radio interfaces. With DC, UEs can flexibly split the traffic between two BSs for traffic offloading. In such a case, the DC-based load balancing acts as a key enabler for achieving robust and low-latency networking [4].

2.4 Literature Review

Several previous works have been performed to improve spatial reuse in dense WLANs. A brief comparison between related works and this work is presented in Table 2.2.

A category of previous works [29–32] discusses the optimal setting of the clear channel assessment (CCA) threshold, i.e., OBSS_PD. Recent works [33–35] that use stochastic geometry approach to study the optimal CCA threshold also belong to this category. This category of works mainly aims at providing theoretical insights. They develop theoretical models of achievable throughput and study the

optimal trade-off between spatial reuse and interference mitigation. The optimal CCA threshold is usually derived by considering homogeneous network density or regular topologies.

Another category of works proposes to dynamically adapt CCA threshold to network conditions. Two survey papers of adaptive CCA methods can be found in [15, 40]. An adaptive CCA method in which an AP adjusts CCA threshold according to the packet error rate (PER) is proposed in [42]. Dynamic sensitivity control (DSC) that adjusts CCA threshold based on the communication distance also belongs to this category. Evaluations of DSC-like methods are presented in [36–39]. All these schemes, however, treat interferers in different OBSS interferers indifferently. As shown in Fig. 2.2, more desirable operations can be achieved by identifying which interferer is transmitting.

Reference [41] is similar to this work in that it proposes to adapt the CCA threshold to individual OBSS interferer. The scheme in [41], however, is model-based and requires the prior knowledge of the correspondence between interference power and packet loss probability. For example, it assumes that STAs are aware of the required signal-to-interference ratio (SIR) for successful receptions. In practice; however, it is often hard to establish an accurate and universal model of such correspondence [16, 17]. In contrast, the proposed scheme is learning-based and does not require such prior knowledge. Moreover, the scheme in [41] requires periodically information reports between the transmitter and the receiver, which is not supported in the IEEE 802.11ax standard. In contrast, the proposed scheme only needs modifications on the transmitter side whereas the receiver does not require any modifications beyond the IEEE 802.11ax standard.

The impact of multi-operator RAN sharing was investigated using the stochastic geometry approach in [21, 28, 43]. The stochastic geometry is a popular theoretical approach. It uses spatial point process models to account for the spatial randomness of BSs so that tractable theoretical results can be obtained. Kibilda *et al.* [43] analyzed the performance of spectrum sharing and RAN sharing in microwave bands. They conducted theoretical formulations of the coverage probability and the maximum achievable downlink data rate. They also conducted a tractable theoretical analysis of cross-operator interference in spectrum sharing networks, and showed that an uncoordinated use of the shared spectrum leads to severe performance degradation. In [21, 44, 45], the authors presented stochas-

tic geometric models of a mmWave channel and analyzed the resource sharing in mmWave bands. The abovementioned studies laid the foundation for stochastic geometric analysis of multi-operator resource sharing. Most of these works, however, ignored the effect of the cell load factor, which is an important measure that related to the QoS satisfaction of UEs. Although [21] considered the partial load case, the load of a cell was modeled as the number of active UEs per BS.

The problem of splitting traffic over the DC was investigated in [46–48]. A data rate optimization problem for DC-enabled downlink traffic splitting was studied in [46]. Singh *et al.* [47] proposed a proportional fair-based traffic splitting scheme for DC networks with non-ideal backhaul links. Wu et al. [48] studied the optimal uplink traffic offloading scheme based on DC. They optimized the overall network radio resource usage by optimizing the BS bandwidth allocation as well as the traffic scheduling and power allocation. Recently, Wu *et al.* [49] studied DC-enabled traffic offloading through small cells that are powered by energy harvesting. Compared to these studies, our study focuses on load balancing in a multi-operator scenario. In such a case, the amount of incoming traffic is not fully observable information as assumed in the above related studies, but a partially observable information.

A branch of recent studies has resorted to game theory to design rules that can guarantee fair and reasonable usages of the shared spectrum. Reference [50] studied a multi-operator transmit power allocation game, where each operator strategically decided the fraction of the total transmit power budget to be allocated to the shared and private bands. They showed that, if the network configurations satisfy certain conditions, the best-response dynamics always converge to the Nash equilibrium (NE). Reference [51] studied a spectrum partition game, where a common spectrum pool was partitioned into several orthogonal segments. The partitioned spectrum segments were then dynamically allocated to the operators based on the channel conditions. The operators mutually decided the allocation scheme, such that the Nash bargaining solution was reached. Reference [52] conducted an in-depth study of the sharing of unlicensed spectrum by strategic operators. Both of transmit power control and bandwidth partition were discussed in [52]. The authors formulated a repeated game and design rules for operators to share the spectrum under time-varying conditions. Most of existing game-based solutions require either the exchange of traffic information or a central controller, such as

the spectrum broker. Such entity, however, may not exist in some cases.

Chapter 3

Spatial Reuse for WLANs with Early Identification of Interfering Transmitters

3.1 Introduction

The significant growth in the number of WLANs devices in recent years [13–15] has resulted in the common occurrence of OBSSs, i.e., co-located WLANs cells operating in the same frequency channel. In dense OBSSs scenario, the throughput degradation becomes a severe problem because concurrent transmissions among OBSSs were not allowed in previous IEEE 802.11 standards, e.g., IEEE 802.11n, 11ac. In particular, an AP or an STA has to defer its channel access when it detects the transmission of any other APs or STAs.

To overcome this challenge, in this chapter, the spatial reuse of WLANs is formulated as a stochastic decision process and a learning-based spatial reuse scheme is proposed. The proposed scheme has two distinctive features. First, the proposed scheme utilizes the information in the detected frame header to identify the interferer and makes decisions accordingly. This solves the problem in Fig. 2.2 in that the proposed scheme has the freedom of deciding whether or not to transmit concurrently with a particular interferer, rather than setting a common threshold to all OBSS interferers. Second, the proposed scheme learns from past experiences of success or failure in concurrent transmissions. The main merit of

using learning-based scheme is that it does not assume any prior knowledge of the correspondence between the interference power and the packet loss probability. Besides, the proposed scheme does not need additional information report from the receiver other than acknowledgment (ACK).

Moreover, this chapter theoretically analyzes the performance gains due to identifying interferers in exploiting spatial reuse. The key idea of the analysis is to partition the state space of the original decision process. Thereby, this chapter uses the partitioned decision process to model the spatial reuse operation where the agent does not identify interferers. By calculating the deviation between the partitioned decision process and the original decision process, a theoretical bound on the gains in the value function due to identifying interferers is derived.

The novelty and contributions of this work are as follows:

- A learning-based spatial reuse scheme which utilizes the information in the detected frame header to identify the interferer and makes decisions accordingly.
- A stochastic decision process and its partition are used to model the spatial reuse operations where the transmitter does and does not identify interferers, respectively. Their deviation in Markov environment is further calculated.
- By using this deviation, a theoretical bound on the gains in value function due to identifying interferers is derived.

The reminder of this chapter is organized as follows. Stochastic decision process formulation is presented in Section 3.3. The learning algorithm is presented in Section 3.4. Section 3.5 theoretically analyzes the gains in the value function due to identifying interferers in exploiting spatial reuse. Evaluation results are presented in Section 3.6. Section 3.7 concludes this chapter.

3.2 Preliminaries: Early Identification of Interfering Transmitters

Before starting to describe the proposed scheme, this section explains how to utilize the information in the detected frame header to identify the interferer. [2]

In IEEE 802.11ax, the physical (PHY) preamble contains a mandatory high efficiency signal-A (HE-SIG-A) field. It lasts for 16 μs and provides some basic information about the frame. One important information is the BSS color. It is a 6-bit numerical identifier of BSSs. An AP and all its associated STAs share the same BSS color, while co-located co-channel APs use different BSS colors. Besides the BSS color bits, the HE-SIG-A field also contains information showing whether the frame is sent in downlink or uplink.

Note that, when an AP or an STA overhears a frame, it always first decodes the PHY layer preamble. By decoding the BSS color bits in the preamble, it can make an early identification of which BSS the frame belongs to. Furthermore, if the detected frame is sent in downlink, the transmitting AP can be uniquely identified, since there is only one AP in each BSS. Note that, this identification is done before it starts to receive the remaining part of the frame.

3.3 A Stochastic Decision Process Formulation of Spatial Reuse

In this section, the spatial reuse is formulated in WLANs as a stochastic decision process. The spatial reuse is focused from a single-agent perspective. In particular, this chapter considers an AP that is sending packets to an associated STA. All other interferers are treated as a part of the environment. Note that, although this chapter considers the AP as the transmitter and the STA as the receiver, it works in the same manner if the transmitter is an STA. Hereafter, the AP under consideration is referred to as the agent.

A stochastic decision process is defined as a four-tuple $(\Omega, \mathcal{A}, q, R)$, where Ω denotes the set of states and $\mathcal{A}(\omega[t])$ denotes the set of possible actions under state $\omega[t] \in \Omega$. When the agent selects an action $a[t] \in \mathcal{A}(\omega[t])$ at instant t, the state transits to $\omega[t+1] \in \Omega$ according to a probability distribution given by q_t. The agent receives a stochastic reward $R(\omega[t], \omega[t+1], a[t])$ at the same time.

The Markov decision process (MDP) is a special type of stochastic decision process where the state transitions satisfy the Markov property. Let $q(\omega[t], \omega[t+1], a[t])$ denote the probability that the state transits to $\omega[t+1]$ when the agent selects an action $a[t] \in \mathcal{A}(\omega[t])$ at instant t.

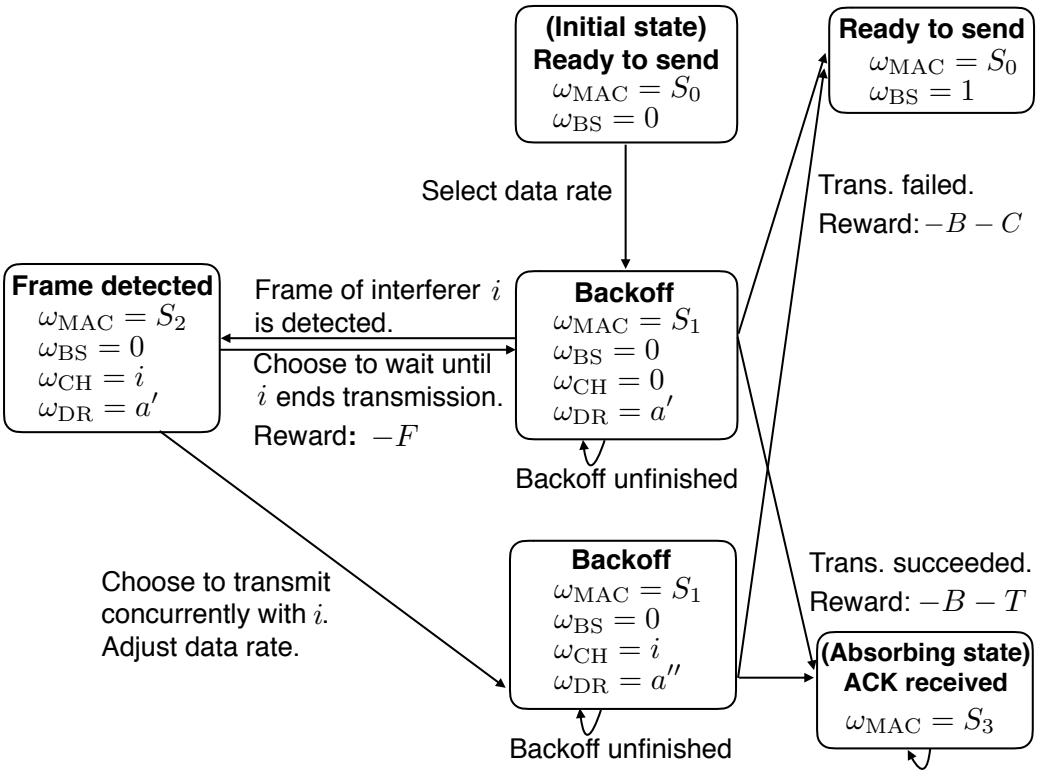


Figure 3.1: State transition diagram. ©2020 IEEE

Note that, in practice, the environment is not a stationary MDP due to many factors, e.g., topology changes, STAs arrivals or departures, the existence of multiple agents. In these cases, the probability distribution q_t is time-varying. In this work, the stationary MDP is assumed for the theoretical analysis presented in Section VI. The simulation evaluations in Section VII, however, evaluate the proposed scheme in various scenarios including time-varying topology and multiple agents scenarios.

State

As shown in Fig. 3.1, the state space Ω is defined from the union and the Cartesian product of the MAC state space Ω_{MAC}, the backoff stage state space Ω_{BS}, the channel state space Ω_{CH}, and the data rate state space Ω_{DR}.

The MAC state space $\omega_{\mathrm{MAC}}[t] \in \Omega_{\mathrm{MAC}}$ has four possible values, i.e., $\Omega_{\mathrm{MAC}} \coloneqq \{S_0, S_1, S_2, S_3\}$. State S_0 denotes that the agent is ready to contend for channel

access. State S_1 denotes the state of backoff, where the agent keeps on sensing the channel while reducing the backoff counter. State S_2 denotes the state that the agent has detected the preamble of a transmission. State S_3 is an absorbing state, which denotes that the packet has been successfully received.

The backoff stage state $\omega_{\mathrm{BS}}[t] \in \Omega_{\mathrm{BS}}$ denotes the current backoff stage, i.e., the times of consecutive transmission failures at present. In IEEE 802.11 WLANs [8], the contention window size doubles after a failed transmission and is reset after a successful transmission. Consider $\Omega_{\mathrm{BS}} \coloneqq \mathcal{J} \coloneqq \{0, 1, \ldots, J_{\max}\}$, i.e., the contention window size does not grow further after $J_{\max}$ times of consecutive transmission failures [8].

The channel state $\omega_{\mathrm{CH}}[t] \in \Omega_{\mathrm{CH}}$ denotes the index of the transmitting interferer that has been identified by the agent. Note that, the agent keeps on sensing the channel during the backoff period. If the agent detects a transmission, the agent immediately identifies the interferer through checking the information in the detected frame header. Note that, if multiple interferers are transmitting preambles at the same time, there are two possibilities. Either the preamble with the strongest received power is decoded or any preamble is unable to be successfully decoded, i.e., preamble error. This depends on whether the SINR requirement for decoding the preamble is met. Note that, the preamble is modulated using the lowest modulation and coding scheme (MCS), i.e., BPSK. [8]

The channel state space is defined as $\Omega_{\mathrm{CH}} \coloneqq \{0\} \cup \mathcal{N} \coloneqq \{0, 1, 2, \ldots, N\}$, where $\omega_{\mathrm{CH}}[t] = 0$ denotes that the channel is idle or the interferer is unable to be identified.

The data rate state $\omega_{\mathrm{DR}}[t] \in \Omega_{\mathrm{DR}}$ denotes the currently chosen data rate for transmission. Consider $\Omega_{\mathrm{DR}} = \mathcal{K} \coloneqq \{1, \ldots, K\}$, where K denotes the number of available MCS.

As shown in Fig. 3.1, the backoff stage state ω_{BS} is defined if the MAC state is S_0, S_1, or S_2. The channel state ω_{CH} and data rate state ω_{DR} are defined if the MAC state is S_1 or S_2. In summary, the entire state space Ω is defined from the union and the Cartesian product of four spaces as follows:

$$\Omega \coloneqq (\{S_0\} \times \Omega_{\mathrm{BS}}) \cup (\{S_1, S_2\} \times \Omega_{\mathrm{BS}} \times \Omega_{\mathrm{CH}} \times \Omega_{\mathrm{DR}}) \cup \{S_3\}. \tag{3.1}$$

Description of State Transitions

The stochastic decision process denotes the entire process of transmitting a packet as illustrated in Fig. 3.1. When a new packet arrives at the head of the transmission queue, the state is initialized as $\omega_{\mathrm{MAC}}[0] = S_0$, $\omega_{\mathrm{BS}}[0] = 0$. When the packet is successfully received, the state transits to S_3. State transition happens when following events occur: the agent begins contending for the channel access, the agent detects an on-going transmission, the agent makes a CCA decision, or the agent transmits.

- When the MAC state is S_0, i.e., when the agent is ready to contend for channel access, the agent selects a data rate from $\mathcal{A}(\omega_{\mathrm{MAC}} = S_0)$ for transmission. Then, the MAC state transits to S_1, the channel state transits to 0, and the data rate state transits to the chosen data rate.

- When the MAC state is S_1, i.e., during the backoff period, the agent keeps on carrier sensing while reducing the backoff counter. If the agent detects the preamble of a transmission, the MAC state transits to S_2 and the channel state transits to the index of the detected interferer.

- When the MAC state is S_2, the agent decides whether to wait until the detected transmission ends or to transmit concurrently with the identified interferer i. If the agent chooses to wait, the agent freezes the backoff counter until the detected transmission ends. After that, the MAC state transits from S_2 to S_1. The channel state ω_{CH} is reset to 0. On the other hand, if the agent chooses to transmit concurrently with interferer i, the MAC state transits from S_2 to S_1, the channel state ω_{CH} stays to be the index of the identified interferer. In this case, the agent also adjusts the data rate for concurrent transmission. The data rate state transits to the re-selected data rate.

- The agent starts transmission when the backoff counter is reduced to 0. According to the transmission results, the MAC state transits to either S_0 or S_3. If the transmission has failed, the MAC state transits to S_0 and the backoff stage $\omega_{\mathrm{BS}}[t]$ transits to $\min(\omega_{\mathrm{BS}}[t] + 1, J_{\max})$. If the packet is successfully transmitted, i.e., the agent has received the ACK from the receiver, the MAC state transits to S_3.

The MAC state S_3 is an absorbing state with zero reward. It indicates that the packet has been received successfully. The learning episode ends when the MAC state reaches S_3 and there is no further state transition. When a new packet arrives at the head of the transmission queue, another learning episode begins and the state is initialized as $\omega_{\text{MAC}}[0] = S_0$, $\omega_{\text{BS}}[0] = 0$.

Action

The set of available actions depends on the MAC state. First of all, when the MAC state is S_0, i.e., when the agent is ready to contend for channel access, the agent needs to select a data rate for transmission, i.e., $\mathcal{A}(\omega_{\text{MAC}} = S_0) = \mathcal{K}$.

Second, when the MAC state is S_2, i.e., when a transmission is detected, the agent chooses whether or not to ignore the detected transmission. If the agent chooses to ignore the detected transmission, i.e., to transmit concurrently with interferer i, the agent also adjusts the data rate for concurrent transmission. Hence, $\mathcal{A}(\omega_{\text{MAC}} = S_2) = \{0\} \cup \mathcal{K}$, where $a = 0$ means to wait until the interferer ends its transmission.

Finally, when the MAC state is S_1, i.e., when the backoff counter has not been reduced to zero, the only available action is to continue carrier sensing. Denote it as $\mathcal{A}(\omega_{\text{MAC}} = S_1) = \{0\}$.

Metric and Reward

The metric this work uses to evaluate a spatial reuse operation is the MAC layer service time [53] of a packet. Formally, the MAC service time is defined as the duration from the instant when a packet arrives at the head of the transmission queue and the agent begins contending for the channel, to the instant when the agent has received an ACK from the receiver [53]. It comprises four parts. Given that the agent has successfully transmitted a packet after J times of consecutive packet transmission failures, where $J \in \mathbb{N}_{\geq 0}$, the MAC service time of a packet includes:

- the duration of J times of failed transmissions,
- the duration of the successful transmission,

- the backoff duration before each transmission attempt,
- the duration that the agent freezes its backoff counter.

The MAC service time of a packet is formulated as follows [53]:

$$D = \underbrace{\sum_{j=0}^{J-1} C_j}_{\substack{J \text{ times of failed} \\ \text{transmissions}}} + \underbrace{T_J}_{\substack{\text{Successful} \\ \text{transmission}}} + \underbrace{\sum_{j=0}^{J} B_j}_{\substack{\text{Backoff} \\ \text{countdown}}} + \underbrace{\sum_{i=1}^{Y} F_i}_{\text{Freeze}}, \tag{3.2}$$

where C_j, T_J, B_j, Y, and F_i are stochastic values. Here, C_j denotes the duration of the unsuccessful transmission in backoff stage j, which includes the duration of transmitting the header, data, ACK timeout, and DIFS [8]. T_J denotes the duration of the successful transmission, which includes the duration of transmitting the header, data, ACK, SIFS, and DIFS [8]. B_j denotes the backoff countdown duration in backoff stage j. Y denotes the number of times that the agent has freezed its backoff counter. F_i denotes the duration that the agent freezes its backoff counter.

Note that, the MAC service time D of a packet is exactly the duration from the initial state to the absorbing state as shown in Fig. 3.1. Since this work interests in reducing the MAC service time, the reward is considered as the negative value of the MAC service time.

The agent measures the duration of each event and calculates the corresponding reward when state transition happens. In particular, the agent receives a reward $-B_j - C_j$ when the transmission failed, i.e., when the MAC state transited from S_1 to S_0. The agent receives a reward $-B_J - T_J$ when the transmission succeeded, i.e., when the MAC state transited from S_1 to S_3. The agent receives a reward $-F_i$ when it has frozen the backoff counter to wait until the detected transmission ends, i.e., when $a = 0$ and the MAC state transits from S_2 to S_1.

3.4 Learning-Based Spatial Reuse Operation

The ultimate goal of the learning algorithm presented in this section is to find the policy of making spatial reuse decisions such that the maximum discounted sum of rewards can be achieved. Intuitively speaking, it is desirable that the agent can learn to transmit concurrently with those OBSS interferers whose interference is

tolerable at the receiver. On the other hand, the agent is supposed to refrain from transmitting concurrently with those interferers whose interference is not tolerable at the receiver. Since the problem have already been formulated as a stochastic decision process, the reinforcement learning algorithms can be applied to solve this problem.

3.4.1 Learning Algorithm

A policy π is a solution concept of the stochastic decision process. It is a mapping from the state space Ω to the action space $\mathcal{A}(\omega[t])$. Given a state ω_0 and a policy π, let $V^\pi(\omega_0)$ denote the expectation of the discounted sum of reward the agent would receive within one episode, i.e.,

$$V^\pi(\omega_0) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R[t] \,\middle|\, \omega[0] = \omega_0\right], \tag{3.3}$$

where $\gamma \in [0, 1)$ is the discounted factor. Note that, in the considered stochastic decision process, the physical meaning of $V^\pi(\omega_0)$ is the negative discounted value of the MAC service time, which is supposed to be maximized. The discounted factor γ indicates how important future rewards are to the current state. Generally, a large γ yields a better outcome after convergence, but it requires longer time to converge [54]. The γ is often set to a value close to one in related studies, e.g., [54]. In this chapter, it is assumed that $\gamma = 0.99$.

It is known that the Q-learning (QL) algorithm derives the optimal policy in a stationary MDP environment [55]. The proposed scheme, however, does not directly apply the conventional QL as the learning algorithm. The reason is that the real environment is not always stationary due to many factors, e.g., changes in communication distance, new interferers arrivals, the existence of multiple agents.

To tackle the non-stationarity, modifications are applied to the QL algorithm. First, the technique of repeated update Q-learning (RUQL) [56] is incorporated to solve the policy bias problem of the conventional QL algorithm. As pointed in [56], the policy bias problem causes performance degradation in noisy non-stationary environments. The policy bias problem refers to the problem that, those optimal actions with temporal lower values are executed less often during the learning process in the conventional QL algorithm. As a result, the values of

those actions are updated less often. This leads to performance degradation since the environment may have already changed before the agent learns the optimal action.

The basic idea of RUQL is to adjust the learning rate in the conventional QL algorithm so that less-chosen actions have higher learning rate. Let $Q(\omega, a)$ denote the state-action value corresponding to state ω and action a. Given the current state $\omega[t]$, the selected action $a[t]$, the new state $\omega[t+1]$, and the associated reward $R[t]$, the RUQL updates the state-action value $Q(\omega, a)$ according to the following expression [56], i.e.,

$$Q(\omega[t], a[t]) \leftarrow (1 - z_n)Q(\omega[t], a[t]) + z_n \left[R[t] + \gamma \max_{a'} Q(\omega[t+1], a') \right], \quad (3.4)$$

where γ is the discounted factor and z_n is the learning rate at episode n. The learning rate z_n is given as follows [56]:

$$z_n = 1 - [1 - \alpha_n]^{\frac{1}{\pi_n(\omega[t], a[t])}}, \quad (3.5)$$

where α_n denotes the learning rate in the conventional QL algorithm and $\pi_n(\omega[t], a[t])$ denotes the probability of choosing action $a[t]$ at state $\omega[t]$ at episode n.

This chapter assumes the ε-greedy exploration policy, i.e.,

$$\pi_n(\omega, a) = \begin{cases} 1 - \varepsilon, & a = \arg\max_{a' \in \mathcal{A}(\omega)} Q(\omega, a'); \\ \varepsilon, & a \neq \arg\max_{a' \in \mathcal{A}(\omega)} Q(\omega, a'), \end{cases} \quad (3.6)$$

where ε is a small constant that denotes the exploration rate. Note that, ε controls the trade-off between exploration and exploitation. Given a higher ε, the agent explores the action space more aggressively. On the other hand, it cannot guarantee acceptable run-time performance since non-greedy actions will be taken frequently [54]. This chapter considers $\varepsilon = 0.1$.

Moreover, in non-stationary environment, it is also important that the agent can respond quickly to environment changes. The learning rate $0 < \alpha < 1$ plays an important role in determining the learning speed of the agent. In fact, there is a trade-off between the stability and the speed of the learning algorithm [56, 57]. If α is small, the agent cannot respond quickly to the environment changes. If α is large, the algorithm may not be robust and stable under stationary environment.

This chapter considers that the learning rate α satisfies the following two conditions: $\sum_0^\infty \alpha_n = \infty$ and $\sum_0^\infty \alpha_n^2 = 0$. Specifically, $\alpha_n = 1000/(1000 + n)$ in simulation evaluations, where n is the learning episode. It is worth mentioning that,

if this condition is satisfied, the RUQL guarantees stability, i.e., convergence, in stationary MDP environment. [56] On the other hand, the agent reset the learning episode $n = 1$, if n is large and the agent detects any arrivals of new interferers or significant changes in the topology. This can be detected by periodically measuring the received signal strength indicator (RSSI) of neighbors [58].

3.4.2 Transmit Power Restriction

This chapter considers that the transmit power of the agent is restricted when the agent starts a concurrent transmission. The reason of restricting transmit power is to protect the on-going transmission in OBSS from being corrupted by the newly issued concurrent transmission of the agent. Note that, the transmit power restriction is also considered in the OBSS_PD-based spatial reuse operation as shown in Fig. 2.1 [13–15].

When the agent decides to transmit concurrently with an on-going transmission, the transmit power of the concurrent transmission is given as follows:

$$p = \min\left(P_{\text{ref}}, \frac{P_{\text{ref}}\Theta_{\text{min}}}{I}\right), \tag{3.7}$$

where P_{ref} denotes the maximum possible transmit power of the agent, $\Theta_{\text{min}} = -82\,\text{dBm}$ [8] denotes the default CCA threshold of legacy devices, and I denotes the measured interference strength. The intuition of this rule is to adjust the transmit power inversely proportional to the detected interference strength. Note that if the agent does not detect any on-going transmissions, it transmits with its maximum possible transmit power.

3.5 Analysis of Gains Due to Identifying Interferers

As shown in Fig. 2.2, it is desirable if the agent can identify which interferer is transmitting before deciding whether or not to transmit concurrently with that interferer. In this section, the concept of state aggregation is introduced. This concept is used to analyze the gains due to identifying interferers.

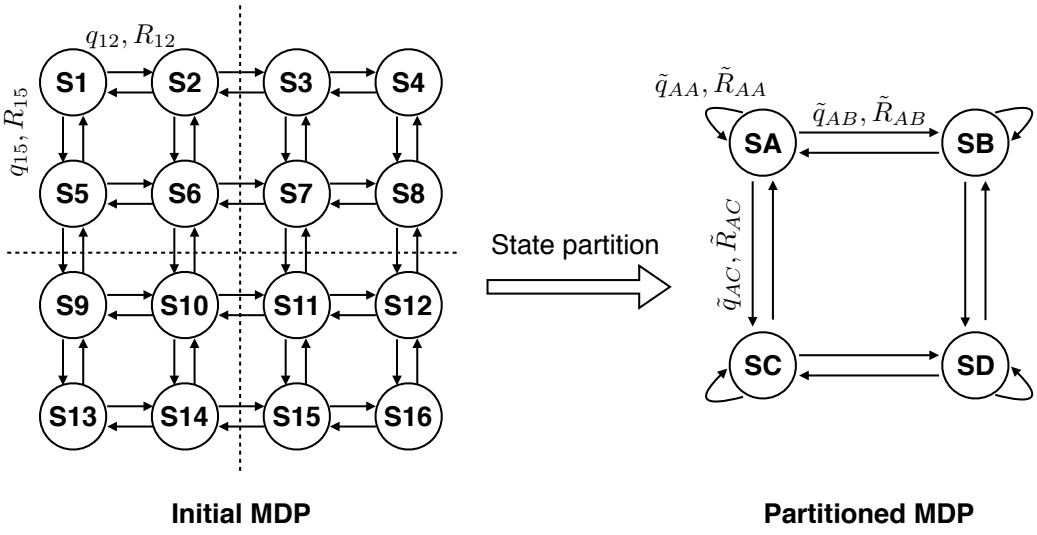


Figure 3.2: The concept of state partition. ©2020 IEEE

3.5.1 Preliminary: State Partition

First the concept of state partition (or called state aggregation) is introduced. The state partition is originated as an approximate method to solve the MDP whose state space is large [59, 60]. The intuition of the state partition is to aggregate multiple similar states into meta-states to create smaller state space. An illustration of the state partition is given in Fig. 3.2, where 16 states are aggregated into 4 states. Thereby, the optimal policy derived on the partitioned MDP can be considered as an approximate solution of the initial MDP. Since the state space is smaller, the derivation is generally easier.

The definition of state partition is given as follows:

Definition 3. *[60] An MDP $\tilde{\Psi} = (\tilde{\Omega}, \mathcal{A}, \tilde{q}, \tilde{R})$ is an δ-homogenous partition of MDP $\Psi = (\Omega, \mathcal{A}, q, R)$ if there exists a mapping $\phi\colon \Omega \to \tilde{\Omega}$, such that ϕ is surjective and for all $\omega \in \Omega$ and $a \in \mathcal{A}$ the following conditions hold:*

$$\left| \sum_{\omega'' \in \tilde{\Omega}} \left| \sum_{\omega' : \phi(\omega') = \omega''} q(\omega, \omega', a) - \tilde{q}(\phi(\omega), \omega'', a) \right| \right| \leq \delta, \tag{3.8}$$

$$\max_{\omega' : \phi(\omega') = \omega} \frac{\left| R(\omega', a) - \tilde{R}(\omega, a) \right|}{|R|_{\max}} \leq \delta, \tag{3.9}$$

where $|R|_{\max}$ denotes the maximum achievable absolute value of reward.

The metric $\delta \in [0, 1]$ describes how much the partitioned MDP $\tilde{\Psi}$ deviates from the original Ψ. In this chapter, δ is referred to as the *deviation* between two MDPs. The constant $|R|_{\max}$ is a normalization factor such that $\frac{|R|}{|R|_{\max}} \in [0, 1]$.

Let $\tilde{\pi}^*$ denote the optimal policy derived on the partitioned MDP $\tilde{\Psi}$. It induces a policy on the original MDP Ψ as follows:

$$\hat{\pi}^*(\omega) \coloneqq \tilde{\pi}^*(\phi(\omega)). \tag{3.10}$$

The induced policy $\hat{\pi}^*$ is an approximate solution of the real optimal policy π^* on the original MDP Ψ. It has been shown in [60] that the value functions of $\hat{\pi}^*$ and π^* satisfies the following property.

Theorem 3.1. *[60] Let $\tilde{\Psi}$ be a δ-homogeneous partition of MDP Ψ, then the optimal policy in $\tilde{\Psi}$ induces an $\frac{2\delta|V|_{\max}}{1-\gamma}$-optimal policy in Ψ, i.e., $\forall\omega \in \Omega$*

$$\left|V^{\hat{\pi}^*}(\omega) - V^{\pi^*}(\omega)\right| \leq \frac{2\delta|V|_{\max}}{1-\gamma}, \tag{3.11}$$

where $|V|_{\max}$ is the maximal achievable absolute value of the value function.

3.5.2 Analysis of Gains Due to Identifying Interferers

The main idea of the analysis is that an MDP $\Psi = (\Omega, \mathcal{A}, q, R)$ is used to represent the decision process where the agent identifies interferers and treats different OBSS interferers differently. On the other hand, its partition $\tilde{\Psi} = (\tilde{\Omega}, \mathcal{A}, \tilde{q}, \tilde{R})$ is used to represent the decision process where the agent does not identify interferers and only distinguishes whether the interference power of the detected OBSS interferer is above or below the OBSS_PD. Note that, the environment is assumed to be stationary MDP only in this section. The environment is not assumed to be stationary MDP in the learning algorithm and in the simulation evaluations. The illustrations of Ψ and $\tilde{\Psi}$ are shown in Figs. 3.3 and 3.4. For notational simplicity, a vector $(\omega_{\text{MAC}}, \omega_{\text{BS}}, \omega_{\text{CH}}, \omega_{\text{DR}})$ is used to represent the state. Some states are omitted for illustrative simplicity.

As shown in Figs. 3.3 and 3.4, the original channel state $\Omega_{\text{CH}} = \{0, 1, \ldots, N\}$ is partitioned into $\tilde{\Omega}_{\text{CH}} = \{0, \text{L}, \text{H}\}$, where $\tilde{\omega} = 0$ denotes that the channel is idle or the preamble of the detected frame is unable to be decoded, $\tilde{\omega} = \text{H}$ denotes that the agent has detected the transmission of an OBSS interferer whose interference

power is above the OBSS_PD, and $\tilde{\omega} = \mathrm{L}$ denotes that the interference is below the OBSS_PD. Hence, the partitioned state space $\tilde{\Omega}$ is defined as follows:

$$\tilde{\Omega} \coloneqq (\{S_0\} \times \Omega_{\mathrm{BS}}) \left(\{S_1, S_2\} \times \Omega_{\mathrm{BS}} \times \tilde{\Omega}_{\mathrm{CH}} \times \Omega_{\mathrm{DR}} \right) \cup \{S_3\}. \tag{3.12}$$

Let $\mathcal{N}_{\mathrm{L}}$ and $\mathcal{N}_{\mathrm{H}}$ denote the set of interferers whose average interference power to the agent is below or above the OBSS_PD, respectively. Note that, $\mathcal{N}_{\mathrm{L}}$ and $\mathcal{N}_{\mathrm{H}}$ are two disjoint subsets of $\mathcal{N}$, where $\mathcal{N} = \mathcal{N}_{\mathrm{L}} \cup \mathcal{N}_{\mathrm{H}}$ and $\mathcal{N}_{\mathrm{L}} \cap \mathcal{N}_{\mathrm{H}} = \emptyset$. Hence, the mapping $\phi_{\theta,\mathrm{CH}} : \Omega_{\mathrm{CH}} \to \tilde{\Omega}_{\mathrm{CH}}$ that maps the original channel state space to the partitioned channel state space is given as follows:

$$\phi_{\theta,\mathrm{CH}}(\omega_{\mathrm{CH}}) = \begin{cases} 0, & \omega_{\mathrm{CH}} = 0; \\ \mathrm{L}, & \omega_{\mathrm{CH}} \in \mathcal{N}_L; \\ \mathrm{H}, & \omega_{\mathrm{CH}} \in \mathcal{N}_H, \end{cases} \tag{3.13}$$

where θ denotes the OBSS_PD. Therefore, the mapping $\phi_\theta : \Omega \to \tilde{\Omega}$ that maps the original state space to the partitioned state space is given as follows:

$$\phi_\theta(\omega) = \begin{cases} (\omega_{\mathrm{MAC}}, \omega_{\mathrm{BS}}, \phi_{\theta,\mathrm{CH}}(\omega_{\mathrm{CH}}), \omega_{\mathrm{DR}}), & \omega_{\mathrm{MAC}} \in \{S_1, S_2\}; \\ \omega, & \omega_{\mathrm{MAC}} \in \{S_0, S_3\}. \end{cases}$$

Next, a series of parameters are introduced. They are used to analyze how much the partitioned MDP $\tilde{\Psi}$ deviates from the original MDP Ψ, i.e., the deviation δ.

Let $\tau_{0j} \in [0, 1]$ denote the probability that the backoff counter has been reduced to zero and the agent starts transmission in an idle slot time in backoff stage j. Given the current contention windows size CW_j, this probability can be approximately calculated as follows [61]:

$$\tau_{0j} = \frac{2}{1 + \mathrm{CW}_j}. \tag{3.14}$$

Let $p_{ik} \in [0, 1]$ denote the expected packet error probability when the agent transmits concurrently with interferer $i \in \mathcal{N}$ using data rate $k \in \mathcal{K}$. Specifically, let $p_{0k} \in [0, 1]$ denote the expected packet error probability when the agent transmits with data rate $k \in \mathcal{K}$ when channel is idle or the preamble of the detected frame is unable to be decoded. Hence, the probability that the agent fails a transmission at backoff stage j is given as follows:

$$q((S_1, j, i, k), (S_0, \min(j + 1, J_{\max})), 0) = \tau_{0j} p_{ik}, \tag{3.15}$$

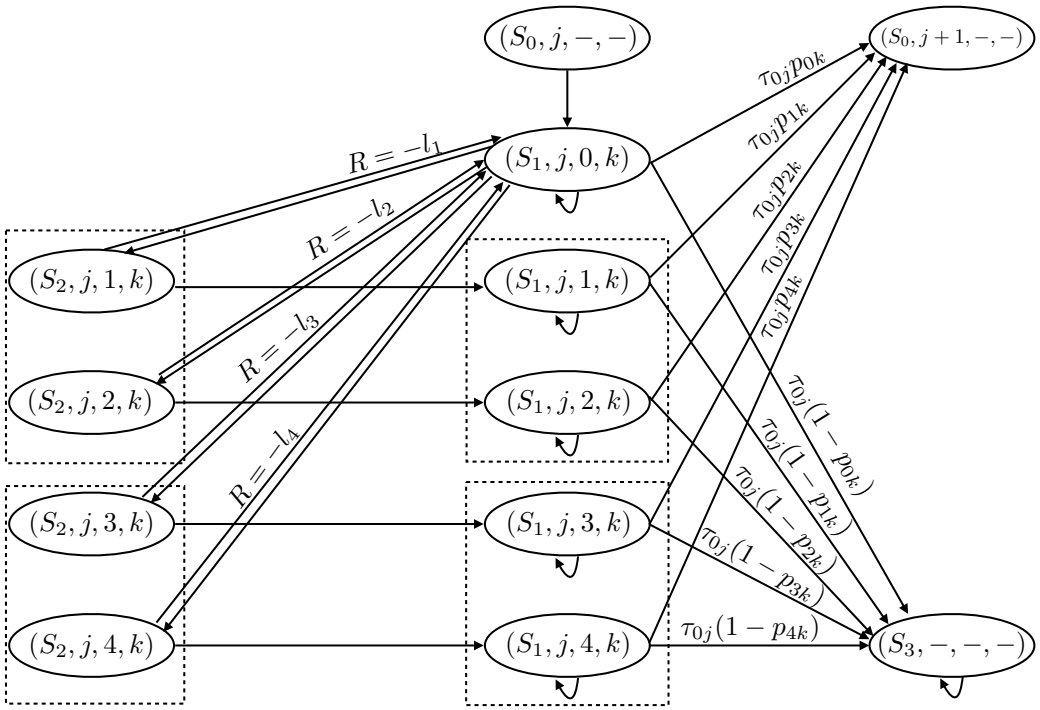


Figure 3.3: The MDP Ψ that represents the spatial reuse operation where agent identifies interferers and treats different OBSS interferers differently. ©2020 IEEE

where $i \in \{0, 1, \ldots, N\}$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$. On the contrary, the probability that the agent succeeds in transmitting the packet and reaches the absorbing state is given as follows:

$$q((S_1, j, i, k), S_3, 0) = \tau_{0j}(1 - p_{ik}), \tag{3.16}$$

where $i \in \{0, 1, \ldots, N\}$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$. Let l_i denote the expected duration of the transmission of interferer $i \in \mathcal{N}$. Hence, the reward that the agent receives when it chooses to wait until the transmission of interferer i ends is given as follows:

$$R((S_2, j, i, k), (S_1, j, 0, k), 0) = -l_i, \tag{3.17}$$

where $i \in \mathcal{N}$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$.

On the other hand, in the partitioned MDP, the probability that the agent fails

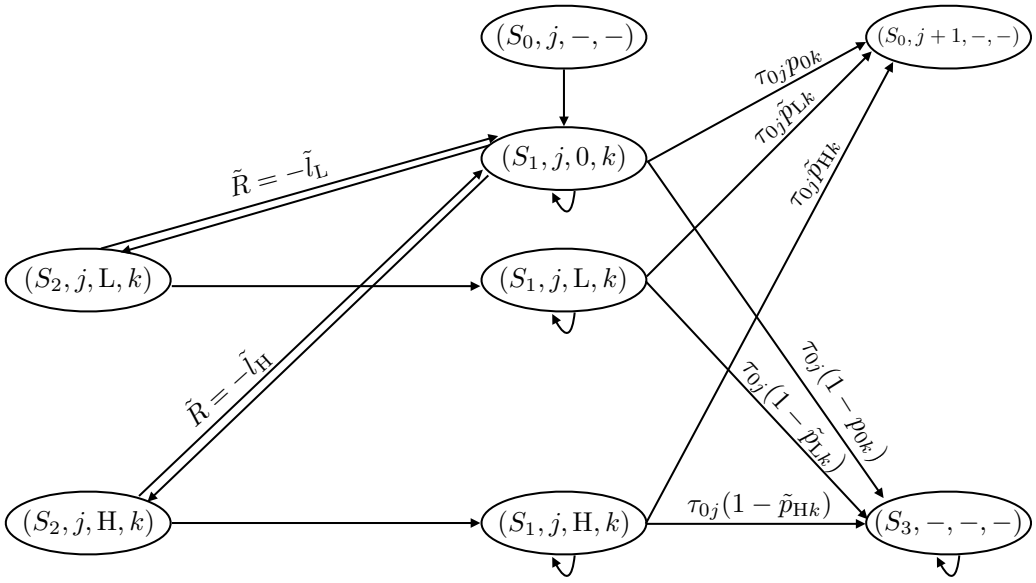


Figure 3.4: This partitioned MDP $\tilde{\Psi}$ represents the spatial reuse operation where agent does not identify interferers and only distinguishes whether the interference power of the detected OBSS interferer is above or below the OBSS_PD. ©2020 IEEE

or succeeds in a transmission are given as follows:

$$\tilde{q}_\theta((S_1, j, i, k), (S_0, \min(j+1, J_{\max})), 0) = \tau_{0j} \tilde{p}_{ik}, \tag{3.18}$$

$$\tilde{q}_\theta((S_1, j, i, k), S_3, 0) = \tau_{0j}(1 - \tilde{p}_{ik}), \tag{3.19}$$

respectively, where $i \in \{0, \mathrm{L}, \mathrm{H}\}$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$. Here, $\tilde{p}_{0k} = p_{0k}$ while $\tilde{p}_{\mathrm{L}k}$ and $\tilde{p}_{\mathrm{H}k}$ denote the expected packet error probability when the agent transmits concurrently with an interferer whose interference power is below or above the OBSS_PD, respectively. The reward that the agent receives when it chooses to wait until the detected transmission ends is given as follows:

$$\tilde{R}_\theta((S_2, j, i, k), (S_1, j, 0, k), 0) = -\tilde{l}_i, \tag{3.20}$$

where $i \in \{\mathrm{L}, \mathrm{H}\}$. Here, $\tilde{l}_{\mathrm{L}}$ and $\tilde{l}_{\mathrm{H}}$ denote the expected durations of the transmission of an interferer whose interference power is below or above the OBSS_PD, respectively.

The following theorem derives the deviation of the considered two MDPs.

Theorem 3.2. *$\tilde{\Psi}$ is a δ_θ-homogenous partition of Ψ, where δ_θ is given as follows:*

$$\delta_\theta = \max\left(\max_{(i,k)}\left(2\tau_{00}\left|p_{ik}-\tilde{p}_{i'k}\right|\right), \max_{i}\left(\frac{\left|l_l-\tilde{l}_{i'}\right|}{|R|_{\max}}\right)\right), \tag{3.21}$$

where $i \in \mathcal{N}$, $i' = \phi_{\theta,\mathrm{CH}}(i)$, $k \in \mathcal{K}$, and $|R|_{\max}$ denotes the maximum achievable absolute value of reward.

Proof. Let us calculate δ_θ based on (3.8) and (3.9) for every $\omega \in \Omega$.

First of all, when $\omega_{\mathrm{MAC}} = S_0$, the agent selects a data rate and ω_{MAC} transits from S_0 to S_1. This state transition is deterministic and does not depend on the channel state. The reward associated with this state transition is zero. Hence, there is no deviation between Ψ and $\tilde{\Psi}$ when $\omega_{\mathrm{MAC}} = S_0$.

Secondly, when $\omega_{\mathrm{MAC}} = S_1$ and the agent fails a transmission, ω_{MAC} transits from S_1 to S_0. The reward associated with the state transition does not depends on the channel state. The state transition probability, however, has a deviation between Ψ and $\tilde{\Psi}$ as follows:

$$\begin{aligned}&\left|q((S_1,j,i,k),(S_0,\min(j+1,J_{\max})),0)\right.\\&\qquad\left.-\tilde{q}_\theta((S_1,j,i',k),(S_0,\min(j+1,J_{\max})),0)\right| = \tau_{0j}\left|p_{ik}-\bar{p}_{i'k}\right|,\end{aligned} \tag{3.22}$$

where $i \in \{0,1,\ldots,N\}$, $i' = \phi_{\theta,\mathrm{CH}}(i)$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$. Similarly, when the agent succeeds in a transmission, the deviation between Ψ and $\tilde{\Psi}$ is given as follows:

$$\begin{aligned}&\left|q((S_1,j,i,k),S_3,0)-\tilde{q}_\theta((S_1,j,i',k),S_3,0)\right|\\&\qquad=\tau_{0j}\left|(1-p_{ik})-(1-\tilde{p}_{i'k})\right| = \tau_{0j}\left|p_{ik}-\tilde{p}_{i'k}\right|,\end{aligned} \tag{3.23}$$

where $i \in \{0,1,\ldots,N\}$, $i' = \phi_{\theta,\mathrm{CH}}(i)$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$. When $\omega_{\mathrm{MAC}} = S_1$ and the agent detects a transmission of interferer i, ω_{MAC} transits from S_1 to S_2. The reward associated with this state transition is zero. The deviation of the state transition probability between Ψ and $\tilde{\Psi}$ is calculated as follows:

$$\left|\sum_{i:\phi_{\mathrm{CH}}(i)=i'} q((S_1,j,0,k),(S_2,j,i,k),0)-\tilde{q}_\theta((S_1,j,0,k),(S_2,j,i',k),0)\right|, \tag{3.24}$$

where $i \in \mathcal{N}$, $i' = \phi_{\theta,\mathrm{CH}}(i)$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$. Since the events of detecting the transmission of interferers are mutually exclusive, the first term is equal to the second term. Hence, for $\omega_{\mathrm{MAC}} = S_1$, the deviation between Ψ and $\tilde{\Psi}$ is $2\tau_{0j}\left|p_{ik}-\tilde{p}_{i'k}\right|$, where $i \in \{0,1,\ldots,N\}$, $i' = \phi_{\theta,\mathrm{CH}}(i)$, $j \in \mathcal{J}$ and $k \in \mathcal{K}$.

Thirdly, when $\omega_{\text{MAC}} = S_2$ and the agent chooses to wait, ω_{MAC} transits from S_2 to S_1. The state transition probability is deterministic and does not depends on the channel state. The associated reward, however, has a deviation as follows:

$$\left|R((S_2, j, i, k), (S_1, j, 0, k), 0) - \tilde{R}_\theta((S_2, j, i', k), (S_1, j, 0, k), 0)\right| = \left|l_i - \tilde{l}_{i'}\right|, \quad (3.25)$$

where $i \in \mathcal{N}$, $i' = \phi_{\theta,\text{CH}}(i)$, $j \in \mathcal{J}$ and $k \in \mathcal{K}$.

Finally, note that $\tau_{00} = \max_{j\in\mathcal{J}} \tau_{0j}$ and $p_{0k} = \tilde{p}_{0k}$, the deviation between Ψ and $\tilde{\Psi}$ for every state $\omega \in \Omega$ and action $a \in \mathcal{A}$ is calculated as (3.21). □

Plugging (3.21) into Theorem 2 and noticing the fact that $|V|_{\max}$ is bounded by $\frac{|R|_{\max}}{1-\gamma}$, the following lemma is derived.

Lemma 1. *Let V^{π^*} denote the optimal value function where the agent identifies the interferer and treats different OBSS interferers differently. Let $V^{\hat{\pi}^*}$ denote the optimal value function where the agent does not identify the interferer and only distinguishes whether the interference power is above or below θ. Then, $\forall\omega \in \Omega$,*

$$\left|V^{\hat{\pi}^*}(\omega) - V^{\pi^*}(\omega)\right| \leq \frac{2\delta_\theta |R|_{\max}}{(1-\gamma)^2}, \quad (3.26)$$

where $|R|_{\max}$ denotes the maximum achievable absolute value of reward.

It can be seen from this theorem that, identifying interferes is more attractable if δ_θ is large. This is more likely to happen if the agent and its associated STA are apart from each other as shown in Fig. 2.2, where p_{ik} may deviate largely from $\tilde{p}_{i'k}$.

3.6 Numerical Evaluation

3.6.1 Evaluation Settings

The proposed learning-based spatial reuse scheme is proposed through MATLAB based simulations. The simulation evaluations are dividend into three scenarios: single agent and static topology, single agent and time-varying topology, and multiple agents scenarios. Note that, the AP under consideration is referred to as the agent.

Firstly, the simulation to be shown in Figs. 8 to 11 and Fig. 13 evaluate the proposed scheme in a single agent and static topology scenario. In these simulations, the agent and other OBSS transmitters are randomly placed in a square region with side length 100 m as shown in Fig. 3.5. The agent and each OBSS transmitter have one associated receiver. The distance between them is called communication distance. The agent and each OBSS transmitter are assumed to have saturated downlink traffic. The number of OBSS transmitters and the communication distance of the agent are stated in each evaluation. Unless otherwise stated, each OBSS transmitter adopts a fixed OBSS_PD of −82 dBm and the communication distance of each OBSS transmitter is 1 m. In these simulations, 100 different random patterns of locations are generated. Each pattern of locations for 10 s is evaluated. The average of the evaluation results is taken.

Secondly, the simulation to be shown in Fig. 12 evaluate the proposed scheme in a single agent and time-varying topology scenario. The evaluation conditions are basically the same as the first scenario, whereas the locations of each OBSS transmitter changes once a second randomly. This simulation runs for 30 s.

Finally, the simulation to be shown in Figs. 14 and 15 evaluate the proposed scheme in a multiple agents scenario. The evaluation conditions are basically the same as the first scenario, whereas all the transmitters are non-cooperative and independent agents. Same as the first scenario, 100 different random patterns of locations are generated. Each pattern of locations is evaluated for 10 s. The average of the evaluation results is taken.

The simulation program to evaluate the proposed scheme is an event-driven simulator written in MATLAB. The program mainly attempts to simulate the MAC layer operation of APs and STAs, including random back-off procedure. Note that this kind of simulation model is widely used in related studies such as [61]. Evaluations parameters are as shown in Table II and III [2, 62, 63]. The distance-based path loss model in [64] is considered, where the center frequency f_c is 5200 MHz and the path loss coefficient N_{PL} is 30 for indoor residential scenario, i.e.,

$$PL_{\mathrm{dB}}(d) = 20\log_{10}(f_c) - 28 + N_{\mathrm{PL}}\log_{10}(\min(d, 1)). \tag{3.27}$$

It is assumed that a packet will be successfully decoded if the required SINR is met, where the payload is 4096 B [2].

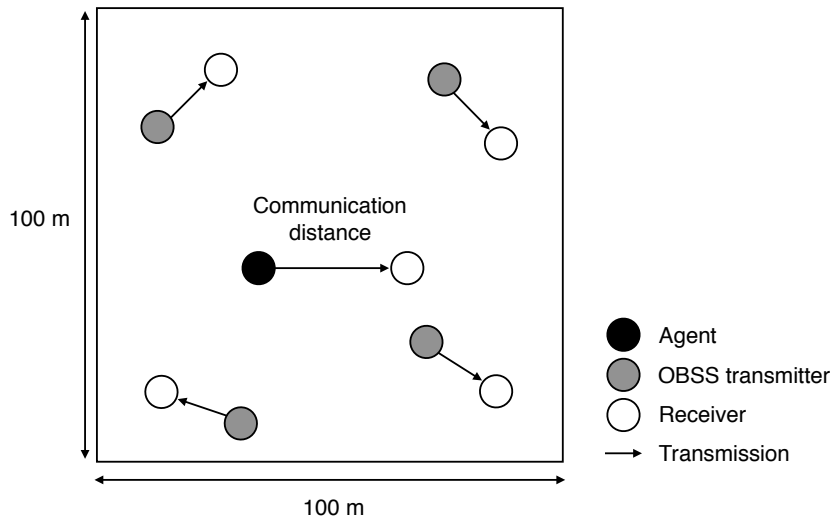


Figure 3.5: Evaluation topology. Here, the AP under consideration is referred to as the agent. ©2020 IEEE

The contention window size updates according to the binary exponential back-off (BEB) algorithm in the IEEE 802.11 [8]. The contention window size of after j times of consecutive packet loss is given as follows:

$$\mathrm{CW}_j = \begin{cases} 2^j(\mathrm{CW}_{\min}+1)-1, & j = 0, 1, \ldots, 6; \\ 2^6(\mathrm{CW}_{\min}+1)-1, & j > 6, \end{cases} \tag{3.28}$$

where $\mathrm{CW}_{\min} = 15$ is the minimum contention window size.

Several performance comparison benchmarks in the following evaluations. The benchmark *optimal OBSS_PD* represents the highest performance under an exhaustive search of all the integer values of OBSS_PD within $[-82\,\mathrm{dBm}, -62\,\mathrm{dBm}]$. In other words, the optimal OBSS_PD can be seen as the performance upper bound of threshold-based CCA policies. In the comparison benchmarks, the rate selection scheme auto-rate feedback (ARF) [65] is used, which tunes up the data rate after two consecutive successful transmissions and tunes it down after one unsuccessful transmission.

Table 3.1: Evaluation parameters.

Parameters	Value
Slot time	9 μs
DIFS	34 μs
SIFS	16 μs
ACK	44 μs
ACK timeout	60 μs
Maximum transmit power	21 dBm
Noise power	-101 dBm
Discount factor γ	0.99
Exploration factor ε	0.1

Table 3.2: Required SINR and transmission time for IEEE 802.11ax.

Data rate	Required SINR	Transmission time (header+ data)
8.6 Mbit/s	1 dB	3844 μs
17.2 Mbit/s	4 dB	1937 μs
25.8 Mbit/s	6 dB	1302 μs
34.4 Mbit/s	9 dB	984 μs
51.6 Mbit/s	13 dB	666 μs
68.8 Mbit/s	17 dB	508 μs
77.4 Mbit/s	18 dB	455 μs
86 Mbit/s	19 dB	412 μs
103.2 Mbit/s	24 dB	349 μs
114.7 Mbit/s	26 dB	317 μs
129 Mbit/s	29 dB	285 μs
143.4 Mbit/s	31 dB	260 μs

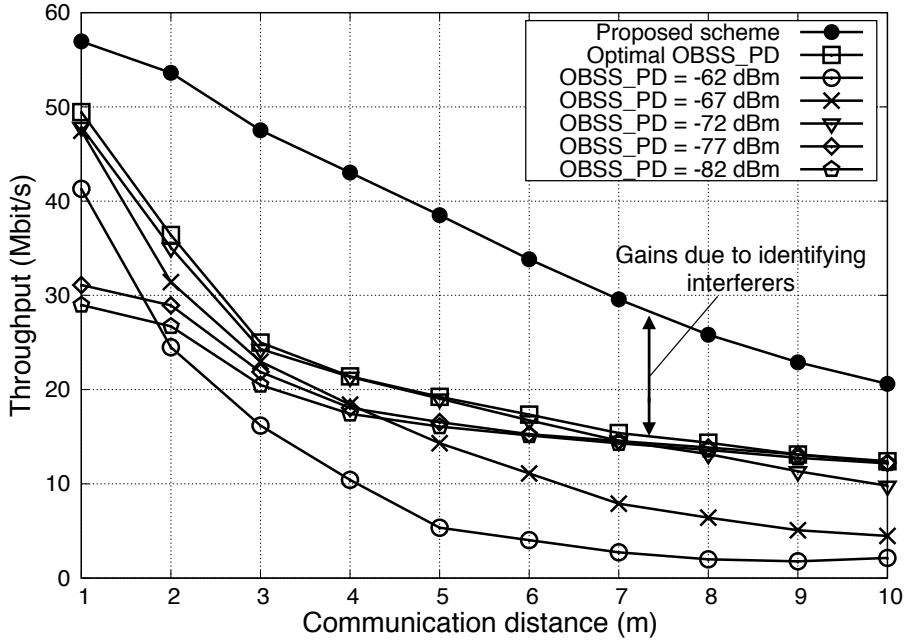


Figure 3.6: Throughput of the agent with the communication distance. ©2020 IEEE

3.6.2 Evaluations Results

Throughput

In Figs. 3.6 and 3.7, the throughput of the proposed scheme is evaluated under different conditions of communication distance and the number of OBSS transmitters.

From Fig. 3.6, it is clear that the comparison benchmark optimal OBSS_PD outperforms other comparison benchmarks of fixed OBSS_PD under different conditions of communication distance. The throughput gain of the optimal OBSS_PD compared to fixed OBSS_PD schemes comes from the ability to adjust the OBSS_PD according to communication distance. When the communication distance is short, a high value of OBSS_PD achieves good performance. This is because the received signal strength is high and conservative concurrent transmission policy decreases transmission opportunities. When the communication distance increases,

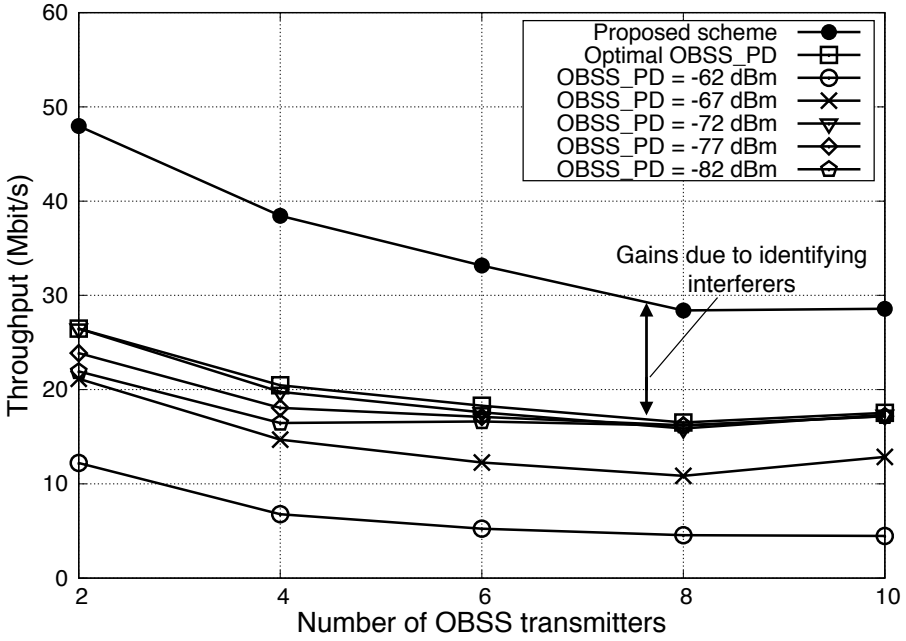


Figure 3.7: Throughput of the agent with the number of OBSS transmitters. ©2020 IEEE

the OBSS_PD value that achieves the highest throughput decreases. This is because the received signal strength is low and aggressive concurrent transmission policy increases the possibility of transmission failures.

In Fig. 3.6, the proposed scheme outperforms the comparison benchmark optimal OBSS_PD. This demonstrates the performance gains of identifying interferers in learning concurrent transmissions under different conditions of communication distance. Although the optimal OBSS_PD allows to adjust the OBSS_PD to communication distance, it treats different OBSS interferers indifferently.

This work has also conducted one-tail paired *t*-test to evaluate the statistical accuracy of this result. The *t*-test is a statistic method that is commonly used to determine if the means of two sets of data are significantly different from each other [66]. Consider the null hypothesis that the proposed scheme does not achieve higher throughput than the benchmark optimal OBSS_PD. Given the simulation data generated from 100 different random patterns of locations, the null

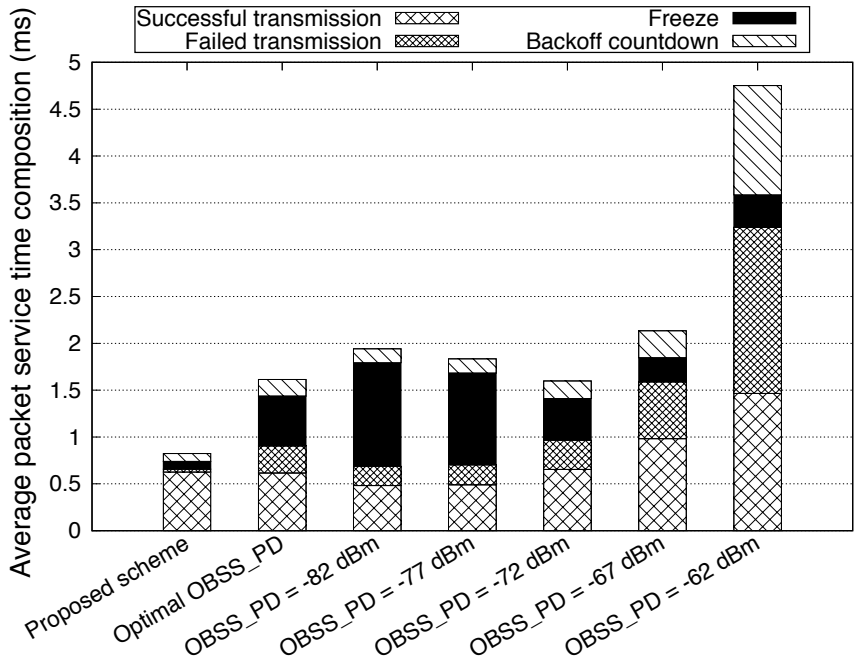


Figure 3.8: MAC service time composition. ©2020 IEEE

hypothesis is rejected at a p-value of 0.05. In other words, the result that the proposed scheme achieves higher throughput than the optimal OBSS PD is at least 95% confident.

It can also be confirmed in Fig. 3.7 that the optimal OBSS_PD outperforms other comparison benchmarks of fixed OBSS_PD and the proposed scheme outperforms the optimal OBSS_PD. This work has also conducted one-tail paired t-test where the null hypothesis is that the proposed scheme does not achieve higher throughput than the benchmark optimal OBSS_PD. The t-test shows that the null hypothesis is rejected at a p-value of 0.05. This demonstrates the performance gains of interferer identification in facilitating concurrent transmissions under different number of OBSS transmitters.

MAC Service Time Composition

The composition of MAC service time is evaluated in Fig. 3.8. Remember that the MAC service time is the duration from the instant when the packet arrives

at the head of the transmission queue to the instant when the agent has received the ACK from the receiver. As shown in (3.2), the MAC service time comprises four components: the duration of counting down backoff counter, freezing backoff counter, failed transmissions, and successful transmission.

The evaluation results in Fig. 3.8 reveal that the proposed scheme reduces the time of freezing backoff counter compared to the fixed OBSS_PD of -82 dBm while keeping the percentage of failed transmission low. The proposed scheme achieves the shortest packet service time of less than 0.85 ms. In the comparison benchmark with -82 dBm OBSS_PD, the average packet service time is larger than 1.9 ms where about 57% of the MAC service time is freezing backoff timer. This indicates that the wireless channel is under high contention.

From Fig. 3.8, the negative effects of increasing the OBSS_PD for facilitating concurrent transmissions is confirmed. Note that increasing the OBSS_PD indeed reduces the time of freezing backoff counter; however, the other three components also increase. This is because, as the OBSS_PD increases, infeasible concurrent transmissions are more likely to occur. Transmission failures increase the contention window size and hence increase the backoff countdown duration. Transmission failures result in the agent choosing a slower data rate. Hence, the time of successful transmission also increases.

Performance Gains Due to Identifying Interferers

Figure 3.9 evaluates the performance gains due to identifying interferers from both the simulation and the theoretical perspective.

From the simulation perspective, the difference between the packet service time of the proposed scheme and the optimal OBSS_PD is evaluated. Note that, this difference corresponds to the gaps in Figs. 3.6 and 3.7.

From the theoretical perspective, the deviation between the partitioned MDP and the original MDP is calculated. Here is how δ is calculated based on (3.21). Note that, there is no difference in the expected transmission durations of each interferer according to our considered simulation setting. Hence, δ_θ as $\delta_\theta = \max_{ik} \tau_{00} |p_{ik} - \tilde{p}_{i'k}|$ is calculated. Furthermore, in the optimal OBSS_PD scheme, the agent searches for the optimal value of OBSS_PD that maximizes throughput.

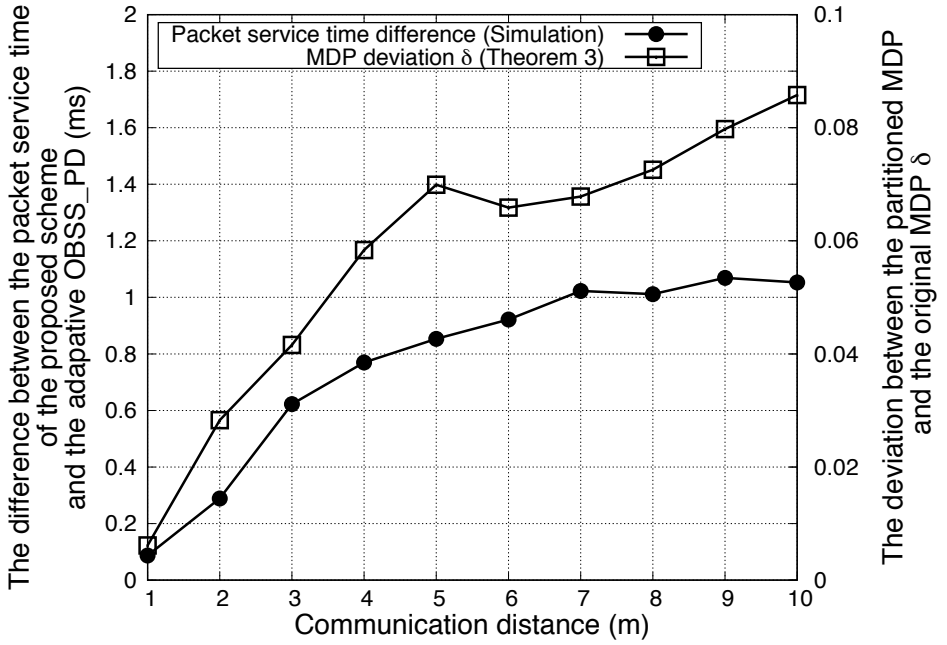


Figure 3.9: Performance gains due to identifying interferers. ©2020 IEEE

Hence, δ is calculated as follows:

$$\delta = \min_{\theta} \delta_{\theta} = \min_{\theta} \max_{ik} \tau_{00} |p_{ik} - \tilde{p}_{i'k}|. \tag{3.29}$$

Remember that p_{ik} denotes the packet error probability when the agent transmits concurrently with i by using data rate k. The value of p_{ik} can be calculated from the locations and parameters in Tables 3.1 and 3.2.

First, notice from Fig. 3.9 that the deviation δ increases as the distance increases. This is because when the communication distance increases, the interference power measured by the agent is more likely to deviate largely from that measured by its receiver. As a consequence, comparing the interference power with the OBSS_PD has limited predictive value in determining the success or failure of transmissions. Also notice that the theoretical deviation δ and the packet service time difference evaluated through simulations have similar shapes. This shows that, given node locations and necessary parameters, calculating δ provides a quantitative method of analyzing the gains due to identifying interferers.

Time-varying Topology

All abovementioned evaluations evaluate the proposed scheme in the static topology environment. Figure 3.10 presents the evaluation result of the proposed scheme under time-varying topology. In this evaluation, the locations of each OBSS transmitter change once a second. The agent does not reset the learning algorithm within the evaluation period. Simulation result reveals that the proposed scheme achieves less throughput in the time-varying topology than that in the static topology in Figs. 3.6 and 3.7. The reason is that the relative distance between the interferer and the agent may have already changed significantly before the agent learns the optimal action. Note that, the performance degradation due to the time-varying environment is in a sense inevitable as long as the learning-based scheme is utilized.

On the other hand, simulation results also confirm that the proposed scheme with RUQL algorithm outperforms the conventional QL algorithm in the time-varying topology. In the RUQL, the value functions of those optimal actions with temporal lower values are updated more frequently than that in the conventional QL algorithm. Hence, it can be inferred that the agent learns the optimal action faster by using RUQL. As a result, the RUQL outperforms the conventional QL in the time-varying topology.

Impact to Legacy Transmitters

All abovementioned simulations focus on the evaluations of the performance gains of the proposed scheme. Notice that, there is a concern if the concurrent transmission of the agent can corrupt on-going transmissions in OBSSs. This evaluation evaluates the percentage of packets transmitted by the OBSS transmitters that are corrupted by the transmission of the agent. These packets are defined as failed transmissions which would be received successfully if the agent is not transmitting. Note that all the OBSS transmitters are assumed to be legacy devices with a fixed OBSS_PD of -82 dBm. The communication distance of each AP is randomly distributed from 1 m to 10 m.

The simulation result in Fig. 3.11 shows that the proposed scheme corrupts approximately 4% of the packets transmitted by legacy devices. This indicates that, by applying the transmit power restriction rule (3.7), the agent does not cause seri-

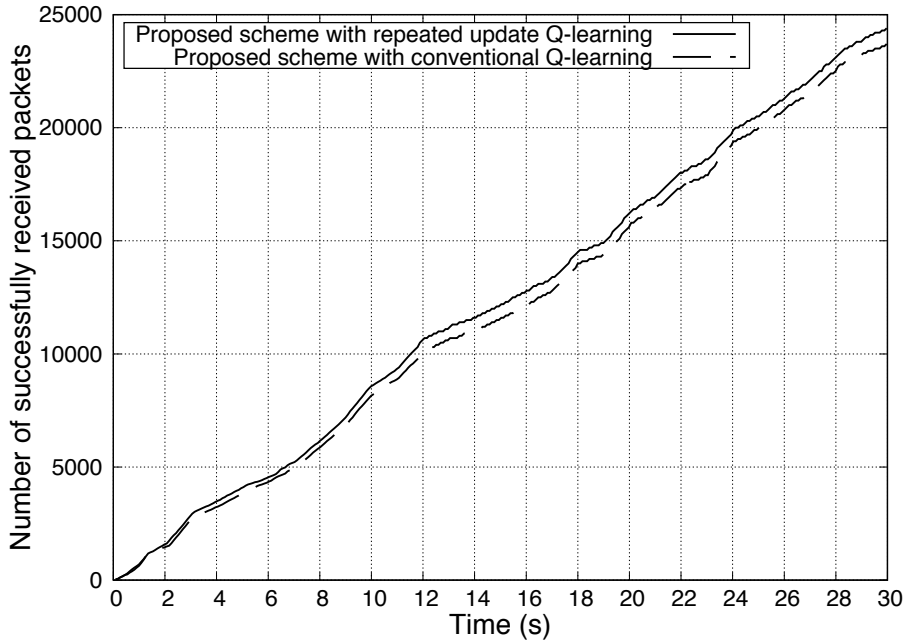


Figure 3.10: Number of successfully transmitted packets. ©2020 IEEE

ous corruptions to legacy devices. Although the comparison benchmark of a fixed OBSS_PD of −62 dBm causes fewer corruptions to legacy devices, it restricts concurrent transmission power by 20 dB and causes low throughput performance.

Multiple Agents

All abovementioned evaluations evaluate the proposed scheme from a single agent perspective. Figures. 3.12 and 3.13 evaluate the multiple agents case where there are multiple agents using the proposed scheme and each agent wishes to maximize its own value function. The communication distance of each agent is randomly distributed from 1 m to 10 m.

It is clear that the proposed scheme achieves better performance than comparison benchmarks in the multiple agents case. Fig. 3.12 evaluates the system throughput in multiple agents case, where all the transmitters adopt the same scheme as shown in the legend. The case where all the APs use the proposed scheme achieves the highest area throughput. Fig. 3.13 evaluates the packet ser-

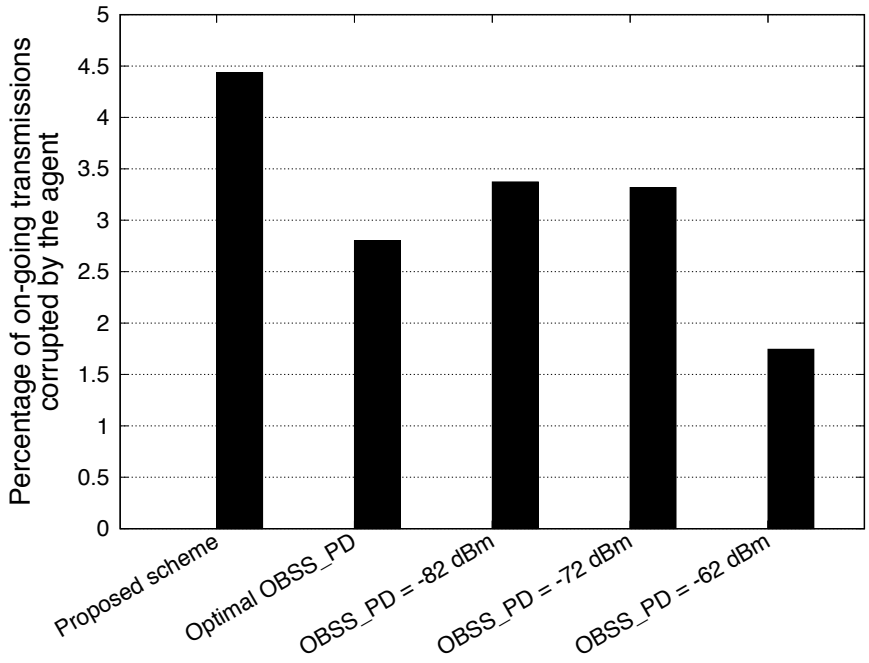


Figure 3.11: Percentage of on-going transmissions corrupted by the agent. ©2020 IEEE

vice time composition in multiple agents case. The case where all the APs use the proposed scheme achieves the shortest packet service time. The gain of identifying interfering transmitters is also confirmed in the multiple agents case. Figure 14 has conducted one-tail paired t-tests. The null hypothesis is that the proposed scheme does not achieve higher throughput than the benchmark optimal OBSS_PD. The t-tests show that the null hypothesis is rejected at a p-value of 0.05.

Note that there is no obvious performance degradation in the multiple-agent scenario compared to that in the single-agent case. One reason is that the transmit power restriction rule (3.7) has alleviated the interactions among agents. When one agent decides to transmit concurrently with another on-going transmission, it does not generate strong interference to that on-going transmissions.

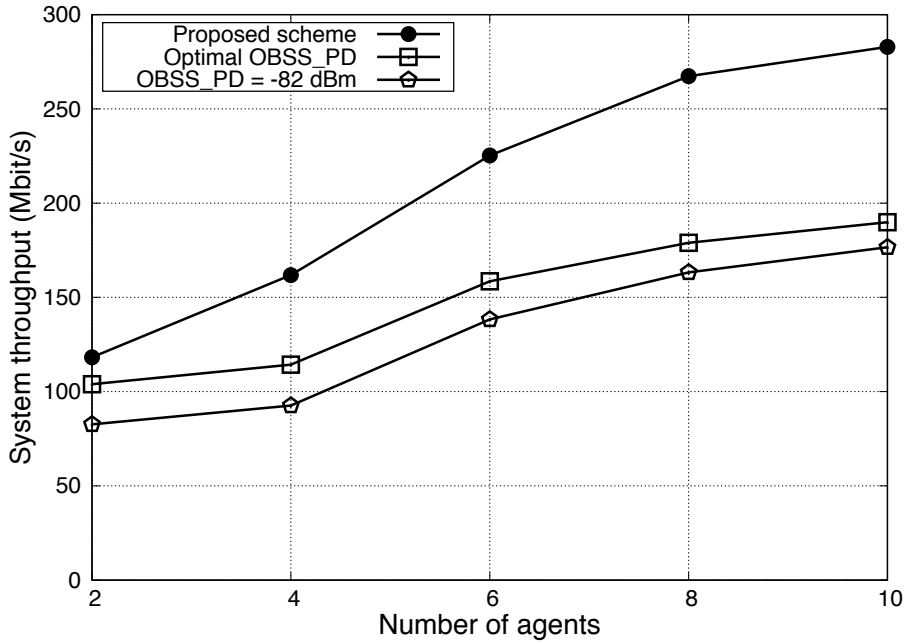


Figure 3.12: System throughput in the multiple agents case. ©2020 IEEE

3.6.3 Pratical Implications

The above presented simulation evaluations provide some insights into designing spatial reuse operation of WLANs. First of all, OBSSs should be avoided as much as possible, especially when the communication between AP and STA is large. Since if the communication distance is large, AP does not have an accurate understanding of the interference at a remote STA by performing carrier sensing. Second, if the existence of OBSSs is inevitable and the communication distance between AP and STA is large, it is unreasonable to set a common CCA threshold to all the OBSS interferers. It is desirable that the AP can treat different interferers differently for deciding whether or not to exploit spatial reuse.

3.7 Conclusion

In this work, a reinforcement learning-based spatial reuse scheme is proposed. When the agent overhears an on-going transmission, it utilizes the information

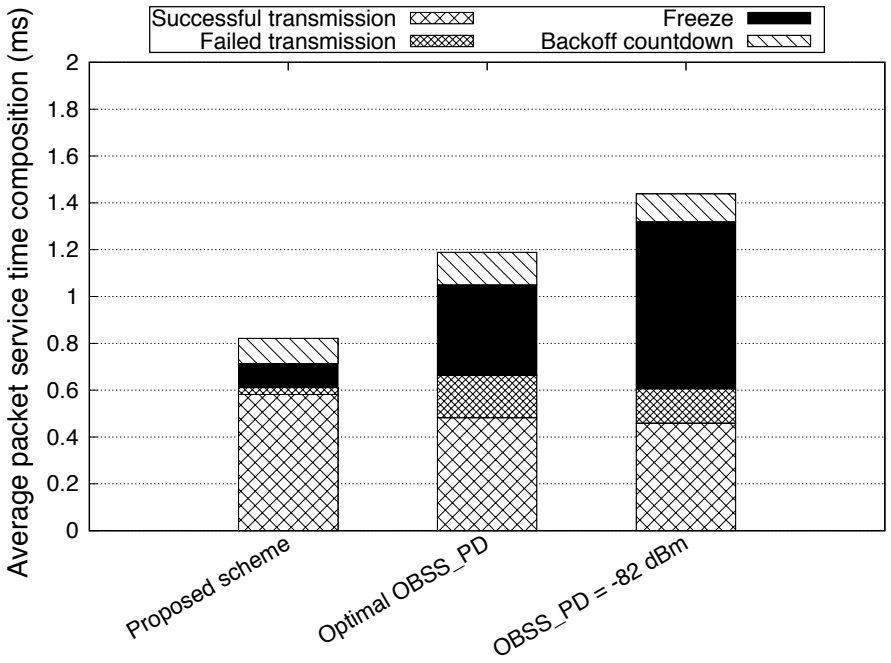


Figure 3.13: Packet service time composition in the multiple agents case. ©2020 IEEE

in the detected frame header to identify the interferer and decides whether or not to freeze the backoff counter accordingly. The proposed scheme is evaluated under various scenarios through simulations. Specifically, the composition of MAC layer service time is analyzed. Results show that the proposed scheme reduces the time of freezing the backoff counter while keeping the number of failed transmissions low. This confirms that, on the one hand, the agent learns to transmit concurrently with those OBSS interferers whose interference is tolerable at the receiver. On the other hand, the agent learns to refrain from transmitting concurrently with those interferers whose interference is not tolerable at the receiver. Moreover, the concept of state partition in MDP is utilized to study the performance gains due to making non-binary identifications of interferers on exploiting spatial reuse in WLANs. A theoretical bound on the gains in value function due to identifying interferers is obtained.

Chapter 4

Potential Game-Based Channel Selection for Mitigating FIM Starvation

4.1 Introduction

Providing an efficient way for wireless internet access, IEEE 802.11 [8] WLAN APs have been widely deployed by mobile operators as commercial hotspots and by individuals for residential usages. The deployment of public Wi-Fi® hotspots has been constantly growing since the 2000s [9], and this trend is predicted to continue. It is predicted that, globally, total public Wi-Fi® hotspots will grow sevenfold from 64.2 million in 2015 to 432.5 million by 2020 [10]. Mobile offload traffic will increase from 3.9 exabytes/month to 38.1 exabytes/month in the meanwhile [10].

The densification of WLANs and the growth of traffic will no doubt make the contention among co-located WLAN cells a critical issue. The IEEE 802.11 MAC supports shared access to the wireless medium through a carrier sense multiple access with collision avoidance (CSMA/CA) technique. CSMA/CA is known to suffer from unfairness issue under heavy traffic conditions, i.e., the equity of media access among neighboring transmitters is not always guaranteed [11, 67, 68]. Certain transmitters even suffer from extremely low transmission opportunity (i.e., throughput starvation) due to unfair carrier sensing [11, 69].

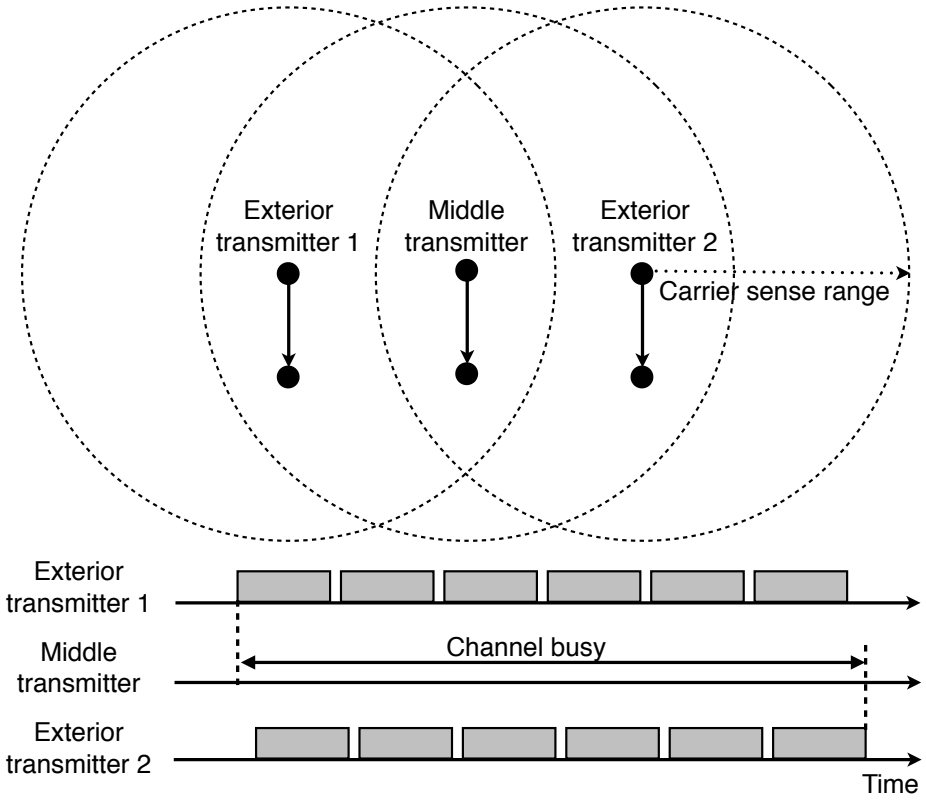


Figure 4.1: Flow-in-the-middle (FIM) starvation. ©2017 IEICE

This chapter focuses on the flow-in-the-middle (FIM) starvation problem defined in [11]. Consider three co-channel transmitters as shown in Fig. 4.1. The middle transmitter is within the carrier sense range of two exterior transmitters while the two exterior transmitters are not within the carrier range of each other. Under heavy traffic conditions, the middle transmitter detects the channel being occupied for a prolonged time and suffers from extremely low transmission opportunities, because the transmissions of exterior transmitters might overlap. Similar phenomenon occurs even if the middle transmitter is operating in a wider channel bandwidth of 40 or 80 MHz [70].

The contribution of this work is to propose distributed channel selection schemes to mitigate the FIM starvation problem. The intuition is to reduce the three-node-chain topologies in the same channel. The proposed channel selection schemes are formulated in the framework of strategic form game [71], in which each AP is supposed to distributedly select a channel based on its payoff function.

The payoff function is designed so that the games are proved to be potential games, in which the existence of a Nash equilibrium is guaranteed, and thus unilateral improvement dynamics is guaranteed to converge in a finite number of steps [72]. Two similar payoff functions with guaranteed convergence are presented. The physical meaning of the first payoff function is the minus number of three-node-chain topologies within two-hop neighbors. The other payoff function has a slightly different physical meaning while it requires less information to be conducted in a distributed manner.

4.2 Proposed Channel Selection Scheme

4.2.1 Contention Graph

The contention graph [73] is used to model the considered network topology. Details about mapping the network topology to the contention graph can be found in several previous works, e.g., [9, 18, 73–75].

In our contention graph $(\mathcal{N}, \mathcal{E})$, $\mathcal{N}$ represents the set of APs. An undirected edge $ij \in \mathcal{E}$ denotes that AP i detects the transmission from AP j and vice versa. Note that, here the following assumptions or approximations are made in this model:

- If AP i detects the transmission from AP j, AP j detects the transmission from AP i. Note that, this condition is not always guaranteed when APs have heterogenous transmission power or carrier sense thresholds. Such case is out of the scope of this work but is studied in [35, 76]. The authors proved this condition is always guaranteed if the carrier sense threshold and transmission power are jointly adjusted in an inversely proportional manner.

- This work focuses on the the carrier sense between APs. In other words, this work focuses on the FIM starvation in the downlink traffic scenario.

Fig. 4.2(a) gives an example of contention graph, and Fig. 4.2(b) shows the subgraph induced by nodes in channel 1. Note that the induced subgraph is constructed by selecting all of the nodes operating in channel 1 and the edges connecting them.

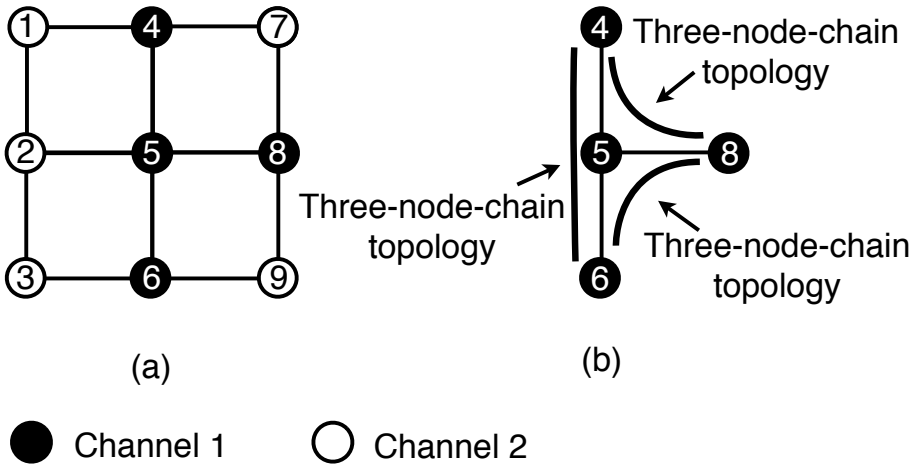


Figure 4.2: An example of contention graph and the subgraph induced by nodes in channel 1. ©2017 IEICE

Furthermore, $\mathcal{N}_i \coloneqq \{\, j \in \mathcal{N} \setminus \{i\} \mid ij \in \mathcal{E} \,\}$ denotes the neighbors of AP i on the graph, a_i denotes the frequency channel that AP i is operating on, and $\mathcal{A}_i$ denotes the set of available channels for AP i. Define $\mathcal{A}_{\mathcal{N}_i} \coloneqq \prod_{j \in \mathcal{N}_i} \mathcal{A}_i$ and $\mathcal{A} \coloneqq \prod_{i \in \mathcal{N}} \mathcal{A}_i$. Let N denote the number of APs, and $\boldsymbol{a} = (a_1, a_2, \ldots, a_N) \in \mathcal{A}$ denote a channel assignment.

4.2.2 Proposed Channel Selection Scheme

To mitigate the FIM starvation problem, this work proposes to reduce the number of three-node-chain topologies on the contention graph. A three-node-chain topology is a topology as shown in Fig. 4.2. Notice that, it is the simplest topology where the FIM starvation occurs.

Since the middle transmitter is being starved in a three-node-chain topology, AP i can select the channel to reduce three-node-chain topologies, in which node i is the middle transmitter. The payoff function is given as follows.

$$u0_i(\boldsymbol{a}) = -f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}). \tag{4.1}$$

Here, a function $f_i : \mathcal{A}_i \times \mathcal{A}_{\mathcal{N}_i} \to \mathbb{R}$ is defined to denote the number of three-node-chain topologies, where node i is the middle transmitter. For any node $i \in \mathcal{N}$, given its channel a_i and the channels of its neighbors $\boldsymbol{a}_{\mathcal{N}_i} \coloneqq \{\, a_j \mid j \in \mathcal{N}_i \,\}$, the

$f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i})$ is defined as

$$f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}) \coloneqq \sum_{\{j,k\}\subseteq\mathcal{N}_i} \mathbb{1}_{\{jk\notin\mathcal{E}\}}\, \mathbb{1}_{\{a_i=a_j\}}\, \mathbb{1}_{\{a_i=a_k\}}, \tag{4.2}$$

where $\mathbb{1}_{\{\text{condition}\}}$ is equal to 1 if the condition is true, and is equal to 0 otherwise.

4.2.3 Design of the Payoff Functions for Guaranteed Convergence

Based on the proposed payoff function $u0_i(\boldsymbol{a})$, two payoff functions $u1_i(\boldsymbol{a})$ and $u2_i(\boldsymbol{a})$ are presented so that the convergence of the channel selection scheme is guaranteed. Here, $u1_i(\boldsymbol{a})$ and $u2_i(\boldsymbol{a})$ are referred to as the proposed payoff functions 1 and 2, and $u0_i(\boldsymbol{a})$ as the non-potential game (non-PG) payoff. The potential game formulations of schemes with $u1_i(\boldsymbol{a})$ and $u2_i(\boldsymbol{a})$ are given later in Section 5.4.

Payoff function $u1 : \mathcal{A} \to \mathbb{R}$ is proposed as follows.

$$u1_i(\boldsymbol{a}) \coloneqq -f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}) - \sum_{j\in\mathcal{N}_i} f_j(a_j, \boldsymbol{a}_{\mathcal{N}_j}). \tag{4.3}$$

Note that $u1_i(\boldsymbol{a})$ consists of two parts: f_i, i.e, the number of three-node-chain topologies whose middle transmitter is node i, and the sum of f_j among the neighbors of i. When an AP updates its channel, it not only reduces three-node-chain topologies whose middle transmitter is itself, but also those whose middle transmitters are its neighbors. In other words, the physical meaning of $u1_i(\boldsymbol{a})$ corresponds to the minus number of three-node-chain topologies within two-hop neighbors node i.

In addition to $f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i})$, define $g_i : \mathcal{A}_i \times \mathcal{A}_{\mathcal{N}_i^{(2)}} \to \mathbb{R}$ to denote the number of three-node-chain topologies on the contention graph where node i is an exterior transmitter. Here, $\mathcal{N}_i^{(2)} \coloneqq \bigcup_{j\in\mathcal{N}_i} \mathcal{N}_j \setminus \{i\}$ denotes the set of nodes that are within two hops of node i, and $\mathcal{A}_{\mathcal{N}_i^{(2)}} \coloneqq \prod_{j\in\mathcal{N}_i^{(2)}} \mathcal{A}_i$.

$$g_i(a_i, \boldsymbol{a}_{\mathcal{N}_i^{(2)}}) \coloneqq \sum_{j\in\mathcal{N}_i} \sum_{k\in\mathcal{N}_j\setminus\{i\}} \mathbb{1}_{\{ik\notin\mathcal{E}\}}\, \mathbb{1}_{\{a_i=a_j\}}\, \mathbb{1}_{\{a_i=a_k\}}\,. \tag{4.4}$$

Payoff function $u2 : \mathcal{A} \to \mathbb{R}$ is proposed as follows.

$$u2_i(\boldsymbol{a}) \coloneqq -f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}) - g_i(a_i, \boldsymbol{a}_{\mathcal{N}_i^{(2)}}). \tag{4.5}$$

The physical meaning of $u2_i(\boldsymbol{a})$ is slightly different from that of the $u1_i(\boldsymbol{a})$. It corresponds to the minus number of three-node-chain topologies where node i is either a middle transmitter or an exterior transmitter.

For example, given the contention graph and the channel as shown in the Fig. 4.2 (b), the proposed payoff functions 1 and 2 have different values on Node 8, i.e., $u1_8(\boldsymbol{a}) \neq u2_8(\boldsymbol{a})$. Since $f_5(\boldsymbol{a}) = 3$, and $f_4(\boldsymbol{a}) = f_6(\boldsymbol{a}) = f_8(\boldsymbol{a}) = 0$, $u1_8(\boldsymbol{a}) = -3$. On the other hand, $g_8(\boldsymbol{a}) = 2$, thus $u2_8(\boldsymbol{a}) = -2$. Note that the three-node-chain topology consisted of nodes 4, 5 and 6 are not considered in the $u2_8(\boldsymbol{a})$.

4.3 Potential Game Formulation

4.3.1 Basic Framework

As a branch of applied mathematics, game theory has been used in a wide variety of study fields, including wireless communications [71]. Potential game is a special class of game and has been applied to channel allocation problem, referring to [77] and the references therein. Potential game gives the benefit that the convergence is always guaranteed when unilateral improvement dynamics are adopted [72].

In this work, the channel selection is modeled as a strategic form game, $\mathcal{G} \coloneqq (\mathcal{N}, (\mathcal{A}_i)_{i\in\mathcal{N}}, (u_i)_{i\in\mathcal{N}})$, given simply as $(\mathcal{N}, (\mathcal{A}_i), (u_i))$. $\mathcal{N}$ is a finite set of players and $\mathcal{A}_i$ is the set of strategies for player $i \in \mathcal{N}$. Player set $\mathcal{N}$ and strategy set $(\mathcal{A}_i)$ are identical to the set of AP indexes and channels, respectively. $u_i : \mathcal{A} \to \mathbb{R}$ denotes the payoff function of player $i \in \mathcal{N}$. Define $\boldsymbol{a}_{-i} \in \mathcal{A}_{-i}$, where $\mathcal{A}_{-i} \coloneqq \prod_{j\in\mathcal{N}\setminus\{i\}} \mathcal{A}_j$.

The Nash equilibrium (NE) is a solution concept of strategic form game, in which no player can increase payoff from changing strategies, assuming other players remain their strategies.

Definition 4 (Nash equilibrium)**.** *A strategy profile $\boldsymbol{a}^* = (a_1^*, a_2^*, \ldots, a_N^*) = (a_i^*, \boldsymbol{a}_{-i}^*)$ is a Nash equilibrium if and only if it satisfies the following inequality for any $i \in \mathcal{N}$, and $a_i \in \mathcal{A}_i$.*

$$u_i(a_i^*, \boldsymbol{a}_{-i}^*) \geq u_i(a_i, \boldsymbol{a}_{-i}^*). \tag{4.6}$$

Generally, the existence of a NE is not guaranteed in a strategic form game. The existence of a NE is guaranteed if a strategic form game is a potential game with a finite strategy space [72].

Definition 5 (Potential game [72])**.** *A strategic form game is a potential game if there exists a potential function $V : \mathcal{A} \to \mathbb{R}$ satisfying*

$$u_i(a_i', \boldsymbol{a}_{-i}) - u_i(a_i, \boldsymbol{a}_{-i}) = V(a_i', \boldsymbol{a}_{-i}) - V(a_i, \boldsymbol{a}_{-i}), \tag{4.7}$$

for any $i \in \mathcal{N}, a_i, a_i' \in \mathcal{A}$, and $\boldsymbol{a}_{-i} \in \mathcal{A}_{-i}$.

Game with Payoff Function 0

Game $\mathcal{G}0 \coloneqq (\mathcal{N}, (\mathcal{A}_i), (u0_i))$ is not a potential game. This can be shown by presenting a counterexample. Consider a *path*, i.e., a sequence of channel assignments as shown in Fig. 4.3. It starts from the left-top and moves counterclockwise.

Nodes 1, 2, and 3, form a three-node-chain topology on the contention graph. Each of them has two available channels to choose from. Hence, $\mathcal{N} = \{1, 2, 3\}$ and $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A}_3 = \{1, 2\}$. A potential function $V(a_1, a_2, a_3)$ satisfying Definition 5 does not exist. Define a potential function V that correctly captures the change in payoff function $u0_i(\boldsymbol{a})$. Hence, $V(1,1,1) < V(1,2,1) = V(2,2,1) = V(2,2,2) < V(2,1,2) = V(2,1,1) = V(1,1,1)$, which is contradictory.

Game with Payoff Function 1

In this section, the game with the payoff function $u1_i$ is proved to be a potential game. This game is similar to the local altruistic game discussed in [74].

Theorem 1. *Game $\mathcal{G}1 \coloneqq (\mathcal{N}, (\mathcal{A}_i), (u1_i))$ is a potential game with the following potential function,*

$$V1(\boldsymbol{a}) = -\sum_{i \in \mathcal{N}} f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}). \tag{4.8}$$

Proof. Assume that player i updates its strategy from a_i' to a_i. The increment of

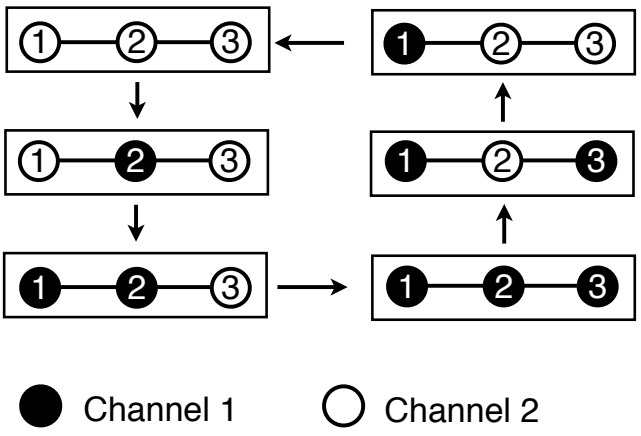


Figure 4.3: A counterexample that shows $\mathcal{G}0 := (\mathcal{N}, (\mathcal{A}_i), (u0_i))$ is not a potential game. ©2017 IEICE

payoff function $u1_i$ is,

$$\begin{aligned} &u1_i(a_i, \boldsymbol{a}_{-i}) - u1_i(a'_i, \boldsymbol{a}_{-i}) \\ &\quad = -f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}) - \sum_{j \in \mathcal{N}_i} f_j(a_j, a_i, \boldsymbol{a}_{\mathcal{N}_j \setminus \{i\}}) + f_i(a'_i, \boldsymbol{a}_{\mathcal{N}_i}) + \sum_{j \in \mathcal{N}_i} f_j(a_j, a'_i, \boldsymbol{a}_{\mathcal{N}_j \setminus \{i\}}). \end{aligned}$$

On the other hand,

$$\begin{aligned} V1(a_i, \boldsymbol{a}_{-i}) &= -\sum_{j \in \mathcal{N}} f_j(a_j, \boldsymbol{a}_{\mathcal{N}_j}) \\ &= -f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}) - \sum_{j \in \mathcal{N}_i} f_j(a_j, a_i, \boldsymbol{a}_{\mathcal{N}_j \setminus \{i\}}) - \sum_{j \in \mathcal{N} \setminus (\mathcal{N}_i \cup \{i\})} f_j(a_j, \boldsymbol{a}_{\mathcal{N}_j}). \end{aligned}$$

The third term does not depend on a_i. Therefore,

$$\begin{aligned} &V1(a_i, \boldsymbol{a}_{-i}) - V1(a_i, \boldsymbol{a}'_{-i}) \\ &\quad = -f_i(a_i, \boldsymbol{a}_{\mathcal{N}_i}) + f_i(a'_i, \boldsymbol{a}_{\mathcal{N}_i}) - \sum_{j \in \mathcal{N}_i} f_j(a_j, a_i, \boldsymbol{a}_{\mathcal{N}_j \setminus \{i\}}) + \sum_{j \in \mathcal{N}_i} f_j(a_j, a'_i, \boldsymbol{a}_{\mathcal{N}_j \setminus \{i\}}) \\ &\quad = u1_i(a_i, \boldsymbol{a}_{-i}) - u1_i(a'_i, \boldsymbol{a}_{-i}). \end{aligned}$$

According to Definition 5, game $\mathcal{G}1 = (\mathcal{N}, (\mathcal{A}_i), (u1_i))$ is a potential game with a potential function $V1(\boldsymbol{a})$. □

Game with Payoff Function 2

Before the potential game formulation of the game with payoff function $u2_i$, define the following notation for simplicity. It means that nodes i, j, and k is a three-node-chain topology, where node j is the middle transmitter.

$$\mathbb{1}_{\{i\text{-}j\text{-}k\}} \coloneqq \mathbb{1}_{\{ij\in\mathcal{E}\}}\,\mathbb{1}_{\{jk\in\mathcal{E}\}}\,\mathbb{1}_{\{ik\notin\mathcal{E}\}}\,\mathbb{1}_{\{a_i=a_j\}}\,\mathbb{1}_{\{a_i=a_k\}}\,.$$

A necessary and sufficient condition for a strategic game to be a potential game is the existence of an interaction potential [78].

Theorem 2 (Interaction potential [78])**.** *A strategic form game is a potential game if and only if, for every subset $\mathcal{S}\subseteq\mathcal{N}$, there exists an iteration potential $\Phi_{\mathcal{S}}:\prod_{i\in\mathcal{S}}\mathcal{A}_i\to\mathbb{R}$ such that*

$$u_i(\boldsymbol{a}) = \sum_{\mathcal{S}\subseteq\mathcal{N}:\, i\in\mathcal{S}} \Phi_{\mathcal{S}}(\boldsymbol{a}_{\mathcal{S}}). \tag{4.9}$$

The potential function is given as

$$V(\boldsymbol{a}) = \sum_{\mathcal{S}\subseteq\mathcal{N}} \Phi_{\mathcal{S}}(\boldsymbol{a}_{\mathcal{S}}). \tag{4.10}$$

Theorem 3. *Game $\mathcal{G}2 \coloneqq (\mathcal{N},(\mathcal{A}_i),(u2_i))$ is a potential game.*

Proof. According to (4.2) and (4.4), $u2_i(\boldsymbol{a})$ can be transformed as follows.

$$\begin{aligned} u2_i(\boldsymbol{a}) &= -f_i(a_i,\boldsymbol{a}_{\mathcal{N}_i}) - g_i(a_i,\boldsymbol{a}_{\mathcal{N}_i^{(2)}}) \\ &= -\sum_{\{j,k\}\subseteq\mathcal{N}_i} \mathbb{1}_{\{jk\notin\mathcal{E}\}}\,\mathbb{1}_{\{a_i=a_j\}}\,\mathbb{1}_{\{a_i=a_k\}} - \sum_{j\in\mathcal{N}_i}\sum_{k\in\mathcal{N}_j\setminus\{i\}} \mathbb{1}_{\{ik\notin\mathcal{E}\}}\,\mathbb{1}_{\{a_i=a_j\}}\,\mathbb{1}_{\{a_i=a_k\}}\,. \end{aligned}$$

Using the notation defined above, it is transformed as follows.

$$\begin{aligned} u2_i(\boldsymbol{a}) &= -\sum_{\{j,k\}\subseteq\mathcal{N}} \mathbb{1}_{\{j\text{-}i\text{-}k\}} - \sum_{\{j,k\}\subseteq\mathcal{N}} (\mathbb{1}_{\{i\text{-}j\text{-}k\}} + \mathbb{1}_{\{i\text{-}k\text{-}j\}}) \\ &= -\sum_{\{j,k\}\subseteq\mathcal{N}} (\mathbb{1}_{\{j\text{-}i\text{-}k\}} + \mathbb{1}_{\{i\text{-}j\text{-}k\}} + \mathbb{1}_{\{i\text{-}k\text{-}j\}}) \\ &= -\sum_{\mathcal{S}\subseteq\mathcal{N},\, i\in\mathcal{S}} \sum_{\{i,j,k\}\subseteq\mathcal{S}} \mathbb{1}_{\{\#\{\mathcal{S}\}=3\}}\,(\mathbb{1}_{\{j\text{-}i\text{-}k\}} + \mathbb{1}_{\{i\text{-}j\text{-}k\}} + \mathbb{1}_{\{i\text{-}k\text{-}j\}}) \\ &= \sum_{\mathcal{S}\subseteq\mathcal{N},\, i\in\mathcal{S}} \Phi_{\mathcal{S}}(\boldsymbol{a}_{\mathcal{S}}). \end{aligned}$$

Here, define $\Phi_{\mathcal{S}}(\boldsymbol{a}_{\mathcal{S}})$ as follows.

$$\Phi_{\mathcal{S}}(\boldsymbol{a}_{\mathcal{S}}) = -\sum_{\{i,j,k\}\subseteq\mathcal{S}} \mathbb{1}_{\{\#\{\mathcal{S}\}=3\}}(\mathbb{1}_{\{j\text{-}i\text{-}k\}} + \mathbb{1}_{\{i\text{-}j\text{-}k\}} + \mathbb{1}_{\{i\text{-}k\text{-}j\}}).$$

According to Theorem 2, $\mathcal{G}2 \coloneqq (\mathcal{N}, (\mathcal{A}_i), (u2_i))$ is a potential game with interaction potential $\Phi_{\mathcal{S}}(\boldsymbol{a}_{\mathcal{S}})$. □

4.4 Distributed Algorithm

This work adopts spatial adaptive play (SAP) algorithm [79] for each player to update strategies in a distributed manner. Each player is randomly assigned a channel at the beginning of the algorithm. In the kth iteration ($k \in \mathbb{N}^+$), player i is randomly selected to update its strategy according to the following distribution,

$$x_i(\, a_i \mid \boldsymbol{a}_{-i}[k]\,) = \frac{\exp[\beta u_i(a_i, \boldsymbol{a}_{-i}[k])]}{\sum_{a_i'\in\mathcal{A}_i}\exp[\beta u_i(a_i', \boldsymbol{a}_{-i}[k])]}. \tag{4.11}$$

The learning parameter β ($0 < \beta < \infty$) can be viewed as the inverse of temperature in the Boltzmann-Gibbs distribution. Notice that, in the situation $\beta \to \infty$, the SAP approaches to the best response (BR) dynamics [74] where one player i is randomly selected in the kth iteration to update its strategy according to

$$a_i[k+1] \in \arg\max_{a_i\in\mathcal{A}_i} u_i(a_i, \boldsymbol{a}_{-i}[k]). \tag{4.12}$$

For player $i \in \mathcal{N}$, to calculate the value of payoff function $u1_i(\boldsymbol{a})$, it only requires limited knowledge of the topology.

- The contention graph within two hops, i.e., the nodes within two hops and the edges connecting them.
- The operating channels of nodes within two hops.

In such case, the information can be collected in a distributed manner through message passing. The process is summarized in Algorithm 5.1.

Calculating the proposed payoff functions $u0_i(\boldsymbol{a})$ and $u2_i(\boldsymbol{a})$ requires less information than the $u1_i(\boldsymbol{a})$. Specifically,

- The edges that connecting two-hop nodes are not required.

In such case, Step 3 in Algorithm 5.1 in not required.

Algorithm 4.1 Spatial adaptive play (SAP) algorithm

1: Let each player $i \in \mathcal{N}$ randomly select a channel $a_i \in \mathcal{A}_i$ with equal probability.
2: Let each player $i \in \mathcal{N}$ broadcast a message $\mathcal{M}1_i$ over all its neighbors $j \in \mathcal{N}_i$. The message $\mathcal{M}1_i$ contains a list of neighbors of i and the channels they are operating on, i.e., $\mathcal{M}1_i = (\mathcal{N}_i, \{ a_j \mid j \in \mathcal{N}_i \})$.
3: Let each player $i \in \mathcal{N}$ broadcast a message $\mathcal{M}2_i$ over all its neighbors $j \in \mathcal{N}_i$. The message $\mathcal{M}2_i$ contains a set of messages $\{ \mathcal{M}1_j \mid j \in \mathcal{N}_i \}$.
4: A player i is randomly selected from $\mathcal{N}$.
5: Let selected player i calculate the payoff function over all its available actions, i.e., $u_i(a_i, \boldsymbol{a}_{-i}), \forall a_i \in \mathcal{A}_i$.
6: Let player i update its strategy according to the (4.11) while all the other players remain their strategies.
7: Stop if the maximum number of iterations is reached; else, go to Step 2.

4.5 Simulation

4.5.1 Compared Scheme

Channel assignment schemes in WLANs have been extensively studied under various approaches [80], including potential game approaches [75]. To enhance successful access probability in collision channels, Chen and Huang [75] discussed a game whose payoff function is given by the probability that transmitter i acquires channel access. It is expressed as the probability of getting the minimal back-off timer among all the contending nodes, specifically,

$$u3_i(\boldsymbol{a}) := \Pr\Big\{ \lambda_i < \min_{j \in \mathcal{N}_i :\, a_j = a_i} \{\lambda_j\} \Big\}. \tag{4.13}$$

Here, λ_i denotes the back-off timer of transmitter i, a_i denotes the channel of transmitter i, and $\mathcal{N}_i$ denotes the neighbors of node i on the graph.

It can easily be proven that when the contention window size goes to infinity, i.e., the upper limit of λ goes to infinity, (4.13) can be approximated as follows.

$$u4_i(\boldsymbol{a}) := \frac{1}{1 + \sum_{j \in \mathcal{N}_i} \mathbb{1}_{\{a_j = a_i\}}}. \tag{4.14}$$

Here, this compared scheme is referred to as the Least Overlapping Basic Service set (Least OBSS) scheme, since selecting the channel of least neighbors gets the highest payoff.

4.5.2 Graph-based Throughput Model

To evaluate the throughput distribution in the network, this work uses the graph-based back-of-the-envelope (BoE) model proposed in [18]. Previous studies have used this model to calculate the throughput in dense deployments of APs, e.g., Ref. [9]. BoE model ignores packet collisions. It also assumes a saturated traffic condition, that is, each transmitter always has packets to send.

Due to the CSMA mechanism, the set of simultaneous transmitting nodes would be an *independent set* [18] on the contention graph. An independent set is a set of nodes no two of which are connected by an edge.

The key of BoE model is to neglect the probability of all the other states except the *maximum independent sets* [18] (MISs) states. A maximum independent set is an independent set with the maximum numbers of nodes. The normalized throughput of node i is given by the ratio of the number of MISs in which node i is included to the number of MISs on the entire contention graph as shown in (2.1) in Chapter II.

4.5.3 Graph-based Simulations

The graph-based simulation considers a densely deployed WLANs environment distributed on a 1200 m $\times$ 1200 m plane. Any two transmitters whose distance is closer than 240 m are assumed to be connected by an edge on the contention graph. There are 4 orthogonal frequency channels for each AP to choose from.

Fig. 4.4 shows the potential functions along with the iterations. Notice that, both potential functions $V1(\boldsymbol{a})$ and $V2(\boldsymbol{a})$ correspond to the minus number of three-node-chain topologies on the entire graph. Consider 200 APs randomly distributed on the region. In the case of SAP, the temperature is set to decrease exponentially along with the iteration. As shown in the simulation results, in both payoff functions 1 and 2, BR dynamics converges faster than the SAP, while the SAP achieves a higher value of the potential function, i.e., there is a speed-optimality trade-off between BR and SAP.

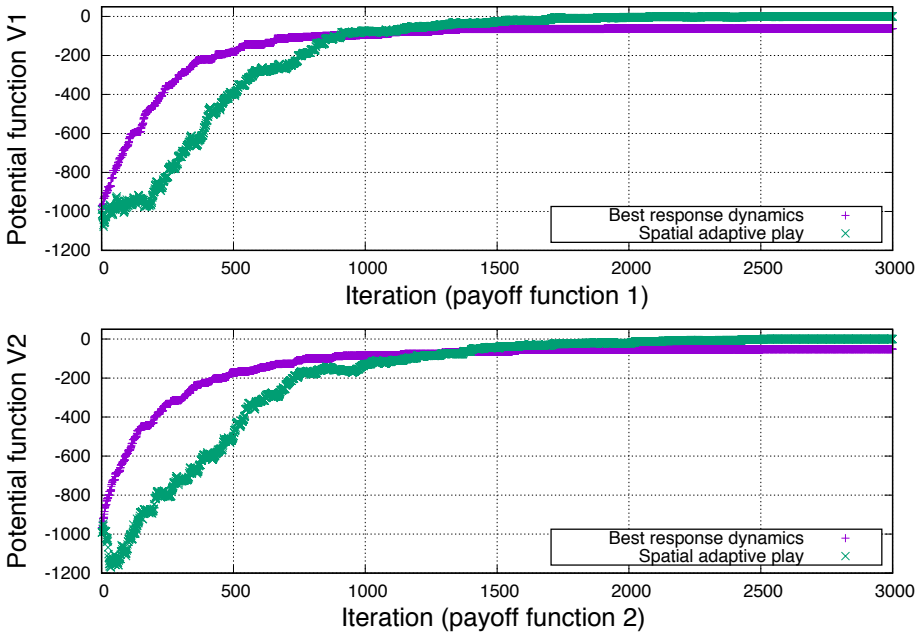


Figure 4.4: Potential function v.s. iteration. ©2017 IEICE

To convey the intuition of the proposed channel selection schemes in reducing the number of three-node-chain topologies, Fig. 4.5 is plotted, i.e., the subgraph of contention graph induced by nodes in each channel when the proposed scheme with payoff function 1 converges. Here, BR dynamics is adopted and only the proposed scheme with payoff function 1 is plotted for simplicity. Similar results are obtained in the payoff function 2. In Fig. 4.5, the nodes represent the APs and the colors represent the frequency channels.

As shown, the three-node-chain topologies are completely eliminated, and only isolated *cliques* [18] remain on the graph. A clique is a set of the nodes such that every two distinct nodes are adjacent. In this case, the potential function (4.8), whose physical meaning is the minus number of three-node-chain topologies, is maximized to zero. In Fig. 4.6, the subgraph after the convergence of compared scheme (4.14), there are three-node-chain topologies remained.

Fig. 4.7 through Fig. 4.9 make a comparison between the random channel allocation, proposed schemes, and the compared scheme i.e., Least OBSS, whose

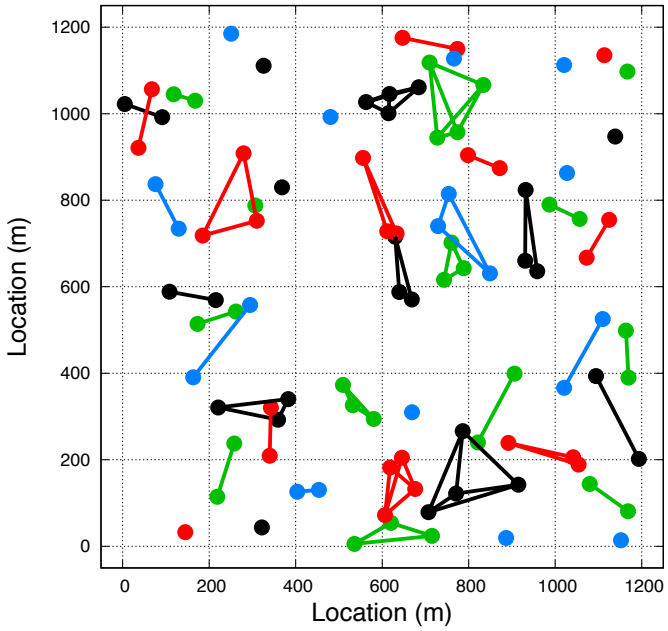


Figure 4.5: Subgraph induced by nodes in each channel when the proposed scheme with payoff function 1 converges. ©2017 IEICE

payoff function is given in (4.14). In these simulations, the learning parameter is set as $\beta \to \infty$ for simplicity, i.e., best response dynamics is adopted, because the payoff functions of these schemes have different order of magnitudes and the optimal value of β in SAP varies from schemes to schemes. Here the average of 100 realization of random deployment is taken.

Fig. 4.7 shows the empirical cumulative distribution function (ECDF) of the normalized throughput of all the nodes. In the case of random channel allocation, the cumulative probability of zero normalized throughput is 20%, which indicates that there are 20% of the nodes being starved. Similarly, there are 8% of the nodes being starved in the case of compared scheme, whereas in proposed schemes with payoff functions 1 and 2, the cumulative probability of zero normalized throughput is less than 1%. Note that, payoff functions 1 and 2 also outperform the non-PG payoff function.

Fig. 4.8 shows the 5th percentile throughput. Fig. 4.9 shows the average of the

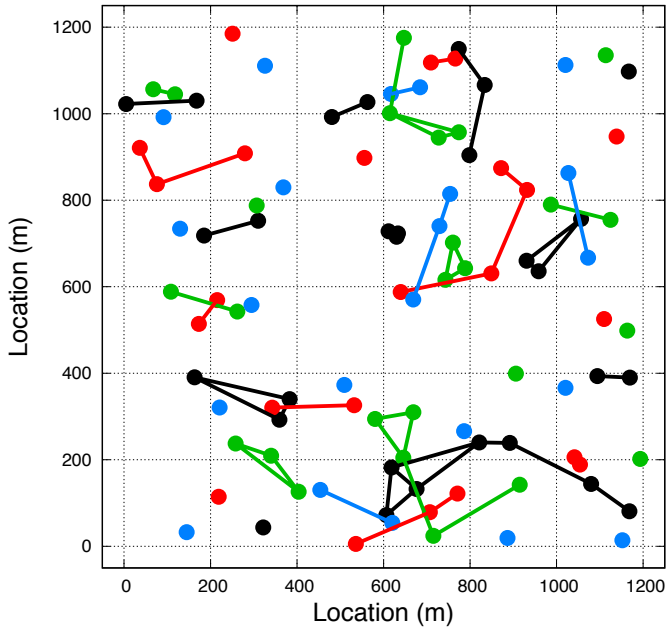


Figure 4.6: Subgraph induced by nodes in each channel when compared scheme converges.©2017 IEICE

throughput against the number of APs. The proposed channel selection schemes result in lower average normalized throughput than the compared scheme but higher 5th percentile throughput, i.e., there is a trade-off between starvation mitigation and average throughput. Based on these results, it is clear that the proposed schemes reduce the number of nodes whose throughput are unfairly low due to their neighbors, at a slight loss of average throughput.

4.6 Conclusion

To mitigate the FIM starvation in densely deployed WLANs, distributed channel selection schemes are proposed to reduce the number of three-node-chain topologies on the contention graph. The proposed channel selection schemes are conducted in a distributed manner, requiring limited knowledge of the topology. Moreover, the schemes are formulated in the framework of strategic form game.

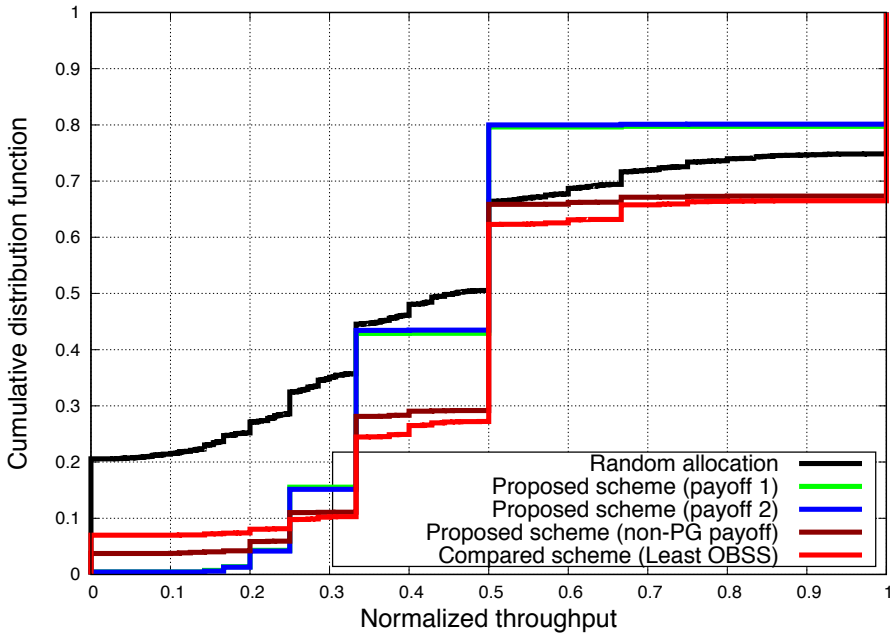


Figure 4.7: ECDF of normalized throughput. ©2017 IEICE

Payoff functions are designed so that the games are proved to be potential games. Simulations confirmed the convergence of the proposed scheme with payoff functions 1 and 2. Simulation results also proved the effectiveness of the proposed schemes in reducing the number of three-node-chain topologies and in enhancing the percentile throughput.

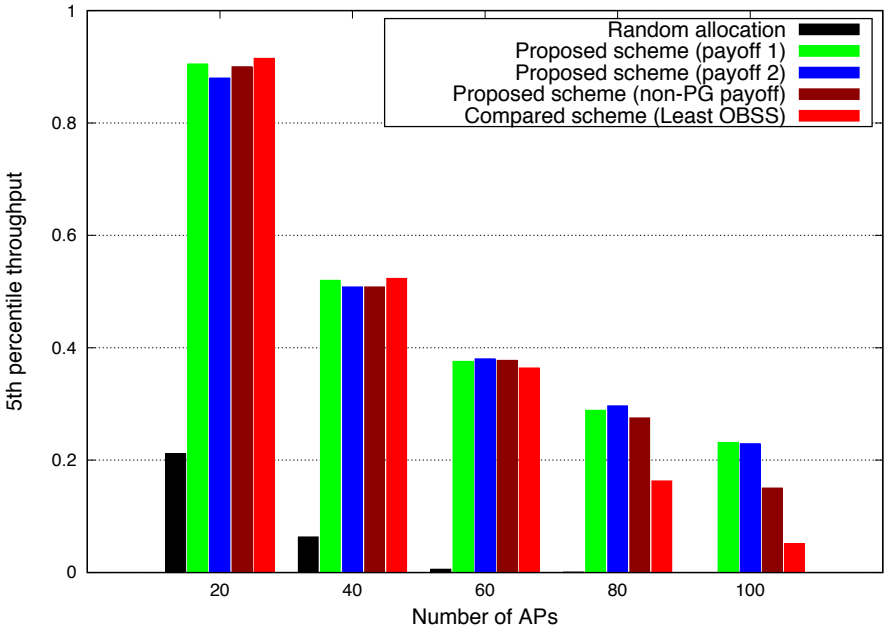


Figure 4.8: 5th percentile throughput. ©2017 IEICE

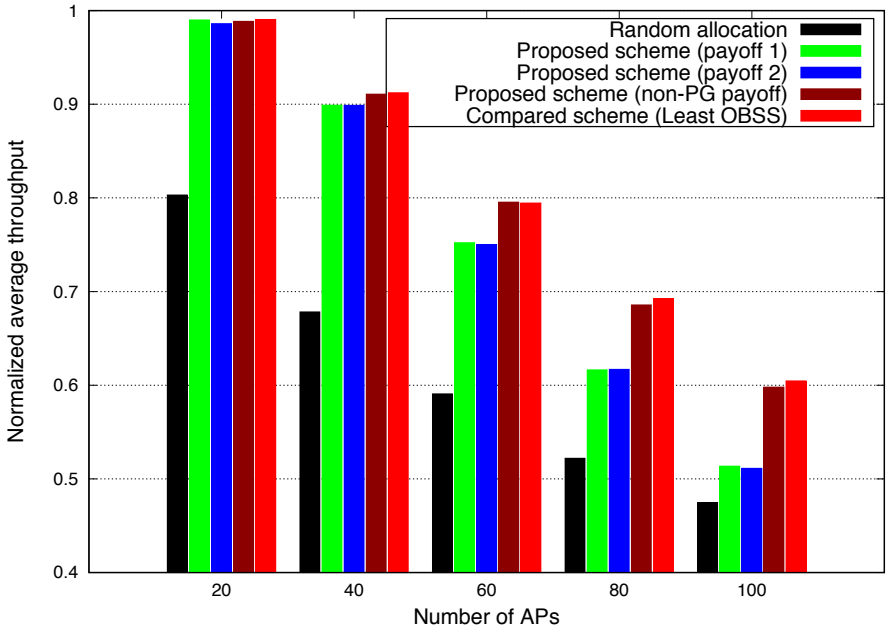


Figure 4.9: Normalized average throughput. ©2017 IEICE

Chapter 5

Load Balancing Between Exclusive and Multi-Operator Shared RANs

5.1 Introduction

The fundamental problem studied in this chapter is as follows. From an operator's perspective, how to split the user traffic over the shared RAN and non-shared RANs to accomplish offloading? This problem is studied from a load-balancing perspective. Specifically, the *cell load* [12, 81] is utilized as an indicator of the network load conditions. It is a more reliable indicator compared to other metrics, e.g., the number of users. It is defined as the fraction of airtime resources that are used in a BS. By considering the cell load, it is possible to study the interactions of offloading policies among multiple operators. Moreover, it is desirable to minimize the cell load from a quality of service (QoS) guaranteeing perspective. This is because UEs experience higher scheduling delay if the cell load is heavy.

5.1.1 Motivation

There are two aspects that have been less well studied in the previous works.

Analysis Incorporating Cell Load Equation

The cell load equation has been investigated in [12, 81, 82]. The image behind this equation is the causal chain as shown in Fig. 5.1. It is a fixed-point equation

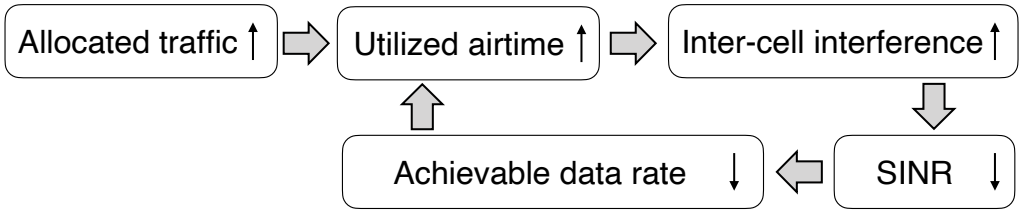


Figure 5.1: Causal chain of the load-interference coupling.

where the cell load appears on both sides of the equation. Existing literatures only solve the cell load equation numerically, i.e., iteratively searching for the cell load satisfying this equation. How to incorporate this equation into theoretical analysis is a topic worth studying.

Load Balancing without Exchanging Information

In a multi-operator shared RAN, it is desirable that the usage of shared resources adapts to the temporal fluctuations in the traffic. Nevertheless, there may exist situations where operators do not hope to share private information for confidential reasons such as the amount of traffic. Moreover, there may be a lack of information-exchanging entity among the operators, where the operators do not have any observations of any private information about the other operators, e.g., traffic demands and number of users.

5.1.2 Main Contributions

Corresponding to the above two motivations, the novelty and contributions are as follows.

Load Analysis in a Multi-Operator Setting

In order to provide guidelines for multi-operator RAN sharing and dynamic offloading, this work perform stochastic geometric analysis that incorporates the cell load equation. This wokr theoretically derives the conditions on multi-operator traffic density such that the networks do not overload. Moreover, this work theoretically proves that the proportionally fair (PF) load balancing can accommodate

more traffics than the round robin (RR) load balancing. The analysis also indicates that the PF balancing scheme impairs inter-operator resource isolation, i.e., the decisions made by one operator may increase the cell load in the exclusive RANs of other operators.

This work theoretically analyze two types of typical load balancing schemes in a multi-operator setting, i.e., the round robin (RR) and proportionally fair (PF) load balancing schemes. Our main finding is that the PF-like load balancing scheme impairs inter-operator resource isolation but can accommodate more traffic such that the RAN does not overload. To the best of our knowledge, this is the first study indicating that the PF load balancing impairs inter-operator resource isolation. This is because this work has taken into account the cell load [12, 81] and the load-interference coupling effect. This analysis would not be possible without considering this effect.

Load Balancing Without Exchanging Information

A scenario is considered where operators adapt to temporal fluctuations in network traffic but do not want to reveal their traffic load information or BS densities to their sharing counterparts for confidentiality reasons. The dynamic load balancing game is formulated as a stochastic game with individual states. Owing to the existence of potential function, a branch of reinforcement learning algorithms is guaranteed to guide multiple operators to an NE under incomplete information. Despite the information is incomplete, simulation results show that the reinforcement learning-based solutions achieve as good performance as the complete information scenario, if the traffic demand is highly correlated among different operators temporally.

The rest of this chapter is organized as follows. The system model and preliminary knowledge are presented in Section 5.2. The analysis of load balancing schemes in a multi-operator setting is presented in Section 5.3. The game formulation, the existence of NE and a fully distributed learning algorithm are discussed in Section 5.4. Section 5.5 presents the results of numerical evaluations. Section 5.6 concludes this chapter.

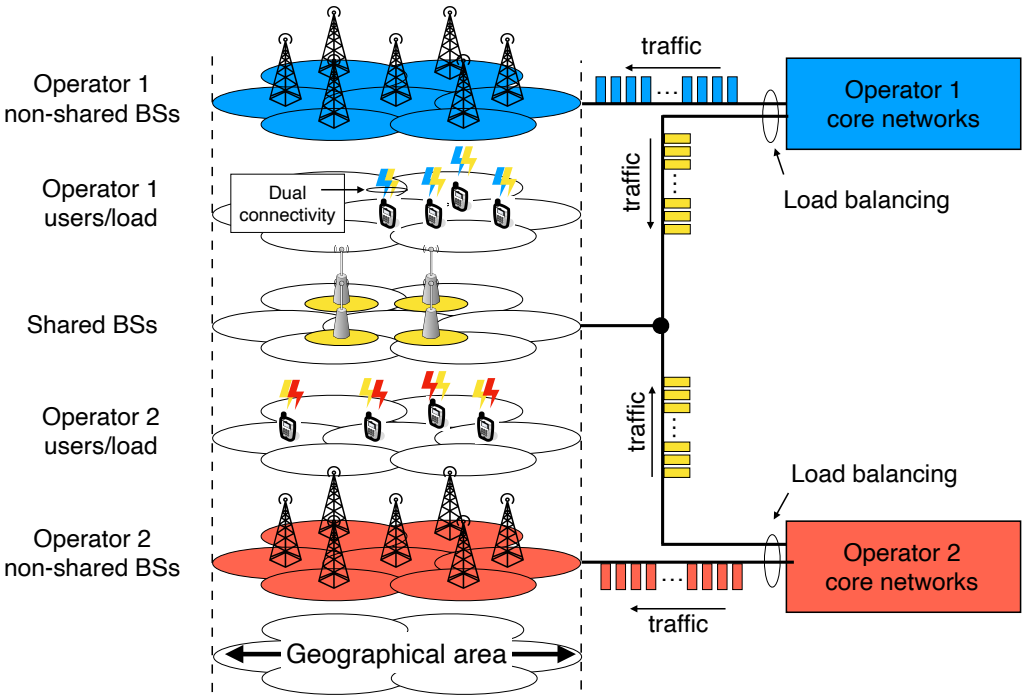


Figure 5.2: A multi-operator cellular network with dual-connectivity capabilities. The locations of shared BSs and non-shared BSs follow independent Poisson point processes (PPPs).

5.2 System Model

5.2.1 Network Deployment and Traffic Model

Consider a two-operator-three-tier RAN sharing model as illustrated in Fig. 5.2. For cost reduction purposes, two operators deploy and share the use of a group of shared BSs for the open access from users of the other operator, e.g., the MOCN architecture [25]. The shared BSs are connected to the core networks of both operators. Meanwhile, each operator also independently deploys non-shared BSs for the exclusive use of its own users. The exclusive BSs operate in an exclusive licensed spectrum, while the shared BSs operate in a commonly shared spectrum. Denote the index set of operators as $\mathcal{M} := \{1, 2\}$. Note that, the analysis and the algorithm used in this chapter can also be extended to the scenario of an arbitrary number of operators. Furthermore, it is assumed that each UE has dual-connectivity capabilities, i.e., each UE is associated with the nearest shared BS

while maintaining a connection with the nearest exclusive BS of its own operator. For the shared BSs, UEs from different operators are treated equally without prioritizing about any specific operator. The traffics allocated to the shared BS from any operators are transmitted with full bandwidth.

The locations of BSs are modeled as independent homogeneous Poisson point processes (PPPs). Note that, the PPP is assumed to obtain tractable theoretical insights as it is a tractable model accounting for the randomness of BSs deployment [21, 28, 43]. The learning algorithm to be presented in Section IV, however, does not require the PPP assumption. Denote three tiers of BSs as Φ_{E1}, Φ_{E2}, and Φ_{S}, with intensities λ_{E1}, λ_{E2}, and λ_{S}, respectively. Here, E represents *exclusive use* and S represents *shared use*. Denote the bandwidths of the exclusive spectrum and the shared spectrum as W_{E1}, W_{E2}, and W_{S}, respectively.

This chapter considers a dynamic traffic scenario. Depending on the number of subscribers and users activities, the traffic is a time-varying stochastic process that follows certain periodic patterns [83, 84]. For example, it is high in certain hours of the day and much lower in some others. The traffic is a *private* information. An operator can only observe its own traffic but does not hope to reveal its traffic information to its sharing counterparts for confidentiality reasons. Let a stochastic process $\omega_i[t]$ denote the downlink network traffic density of operator $i \in \mathcal{M}$ at time stage t. This is interpreted as the downlink traffic per area, whose unit is $\mathrm{bit/s/m^2}$. As the time variations are not considered in Sections II and III, $\omega_i[t]$ is denoted as ω_i for notational simplicity.

5.2.2 Downlink SINR and Cell Load

Consider a UE that is associated with BS $k \in \Phi_j$ of tier $j \in \{\mathrm{E1}, \mathrm{E2}, \mathrm{S}\}$. Let P_j denote the transmit power of BSs of tier j. Let β_j and α_j denote the path-loss coefficient and the path-loss exponent, respectively. Let x_{jk} and h_{jk} denote the distance from BS $k \in \Phi_j$ to the considered UE and the Rayleigh fading gain. The Rayleigh fading gain is an exponentially distributed variable with unit mean. The received signal power is formulated as:

$$P_j \beta_j h_{jk} x_{jk}^{-\alpha_j}.$$

For the interference power, this chapter mainly takes into account the inter-cell interference rather than the intra-cell interference [21, 43]. In our model, BSs do

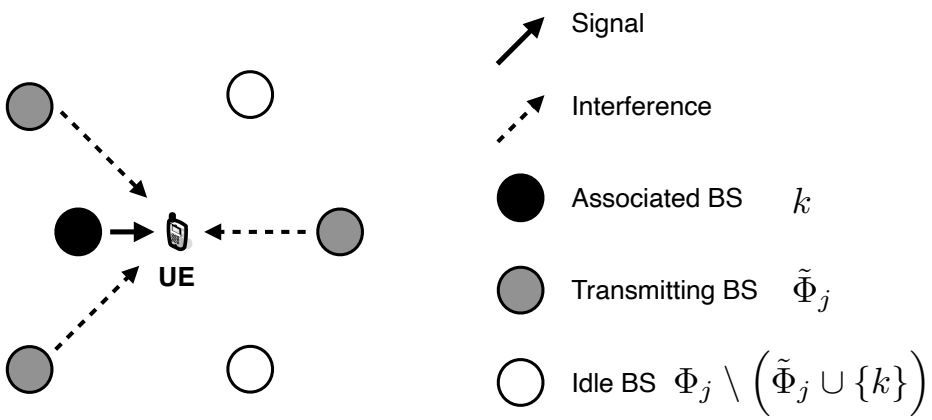


Figure 5.3: The interference structure considered in this chapter.

not continuously emit interference power. The probability that a BS transmits interference depends on its load conditions [81,82]. This is because during the peak hours of traffic, BSs consume more airtime resources. Hence, from a probabilistic perspective, a BS has a higher probability to be transmitting at a given instant during peak hours. To model this fact, the *cell load* concept [12,81,82] is introduced to characterize the network load condition. The cell load is denotes as $\rho_j \in [0,1]$, where $j \in \{\text{E1}, \text{E2}, \text{S}\}$. It interprets the average probability that an arbitrary BS of tier j is transmitting at a certain instant. A higher value of ρ_j indicates a heavier load condition. For example, $\rho_j = 0.3$ indicates that, on average 30% of BSs of tier j are transmitting at an instant. Meanwhile, followed from the ergodicity of PPP, it also indicates that a BS is transmitting with a probability of 0.3 at a certain instant.

The considered interference model is illustrated in Fig. 5.3. Denote the set of BSs that are transmitting interference at an instant as $\tilde{\Phi}_j$ where $j \in \{\text{E1}, \text{E2}, \text{S}\}$. Hence, the set of idle BSs is $\Phi_j \setminus \left(\tilde{\Phi}_j \cup \{k\}\right)$. The aggregated interference power is characterized as:

$$\sum_{k' \in \tilde{\Phi}_j} P_j \beta_j h_{jk'} x_{jk'}^{-\alpha_j},$$

where the interfering BSs $\tilde{\Phi}_j$ can be modeled as a thinned PPP with density $\tilde{\lambda}_j = \lambda_j \rho_j$ by using the cell load ρ_j [81, 82]. Let $W_j N_0$ denote the noise power. The

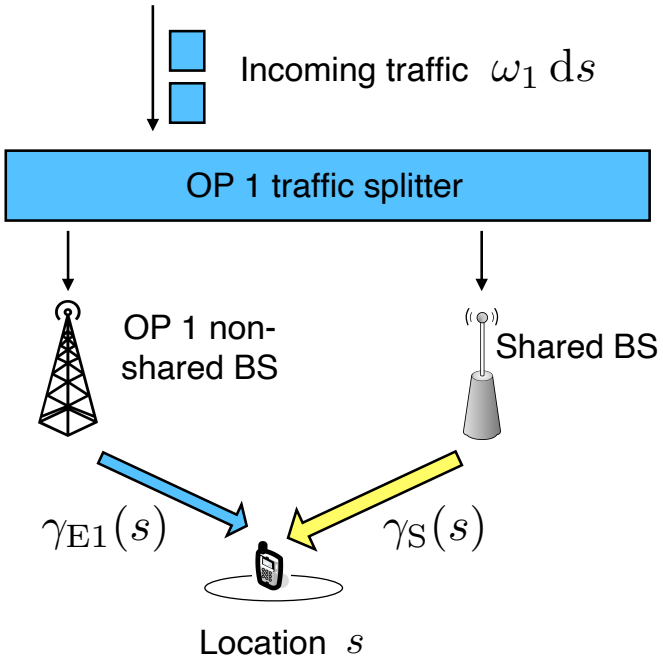


Figure 5.4: Th load balancing model considered in this chapter.

downlink SINR is characterized as follows:

$$\gamma_j = \frac{P_j \beta_j h_{jk} x_{jk}^{-\alpha_j}}{\sum_{k' \in \tilde{\Phi}_j} P_j \beta_j h_{jk'} x_{jk'}^{-\alpha_j} + W_j N_0}, \tag{5.1}$$

where $j \in \{\mathrm{E1}, \mathrm{E2}, \mathrm{S}\}$.

5.2.3 Load Balancing Model

Load balancing is an important task from a QoS-guaranteeing perspective. The fundamental problem studied in this chapter is: how to split user traffic over the shared RAN and the non-shared exclusive RAN to accomplish load balancing? To study this problem from a tractable approach, a basic load balancing model is present here.

An illustration of the traffic splitting model is shown in Fig. 5.4. Consider a small area at location s. The incoming traffic raised from UEs of operator i in this area is given as $\omega_i\,\mathrm{d}s$. Operators use traffic splitters to split the incoming traffic over the shared BSs and the non-shared exclusive BSs. The incoming traffic

Table 5.1: List of load balancing schemes.

Scheme	$\xi_{\mathrm{E1}}(s)$	$\xi_{\mathrm{E2}}(s)$	$\xi_{\mathrm{S}}(s)$
RR	$\frac{\omega_1}{1+b_1}$	$\frac{\omega_2}{1+b_2}$	$\frac{b_1\omega_1}{1+b_1} + \frac{b_2\omega_2}{1+b_2}$
PF	$\frac{R_{\mathrm{E1}}(s)\omega_1}{R_{\mathrm{E1}}(s)+b_1R_{\mathrm{S}}(s)}$	$\frac{R_{\mathrm{E2}}(s)\omega_2}{R_{\mathrm{E2}}(s)+b_2R_{\mathrm{S}}(s)}$	$\frac{R_{\mathrm{S}}(s)b_1\omega_1}{R_{\mathrm{E1}}(s)+b_1R_{\mathrm{S}}(s)} + \frac{R_{\mathrm{S}}(s)b_2\omega_2}{R_{\mathrm{E2}}(s)+b_2R_{\mathrm{S}}(s)}$

is divided into two streams and aggregated at the UE through dual connections. Based on this model, the original traffic densities ω_1 and ω_2 are splitted into three tiers. Given the downlink traffic densities ω_1 and ω_2, operator 1 splits ω_1 into tiers E1 and S while operator 2 splits ω_2 into tiers E2 and S. Specifically, let $\xi_{\mathrm{E1}}(s)$ and $\xi_{\mathrm{E2}}(s)$ denote the traffic densities allocated to the exclusive tier by operators 1 and 2 at location s, respectively. Let $\xi_{\mathrm{S}}(s)$ denote the traffic density allocated to the shared tier jointly by operators 1 and 2 at location s. For any location s, the following equation holds:

$$\xi_{\mathrm{E1}}(s) + \xi_{\mathrm{E2}}(s) + \xi_{\mathrm{S}}(s) = \omega_1 + \omega_2. \tag{5.2}$$

Depending to whether or not the channel state information (CSI) is utilized, two typical load balancing schemes are considered:

- (Biased) Round Robin (RR)
- (Biased) Proportionally Fair (PF)

The reason of considering these two schemes is that they are typical examples of two types of load balancing schemes. Their fundamental difference is that: the PF scheme is rate-dependent, while the RR scheme is rate-independent. Both of these two schemes use the control variable b_i to accomplish load balancing, where $i \in \{1, 2\}$. Hereafter, the control variable b_i is called *offload bias*. The amounts of traffic allocated to each tier are summarized in Table 5.1.

In the RR scheme, the traffic is split into shared BS and non-shared exclusive BS according to a constant ratio. The ratio is fixed regardless of the location or temporal achievable data rate. In other words, the traffic splitter does not require any instantaneous channel state information from the UE. Consider UEs of operator i located in a small area at location s, whose total incoming traffic is $\omega_i \, \mathrm{d}s$.

The amounts of traffic split to the shared BS and the non-shared exclusive BS are $b_i\omega_i\,\mathrm{d}s/(1+b_i)$ and $\omega_i\,\mathrm{d}s/(1+b_i)$, respectively. In this scheme, the operators may adjust the offload bias b_i to adapt to the time-varying incoming traffic density ω_i. As operator i increases the offload bias b_i, it splits more traffic into the shared RANs.

In the PF scheme, the traffic is splitted according to the instantaneous downlink data rate. As the UE has dual-connectivity with two BSs, it is reasonable to allocate more traffic to the link with a higher achievable data rate. Consider UEs of operator i located in a small area at location s, whose total incoming traffic is $\omega_i\,\mathrm{d}s$. Denote the instantaneous downlink data rate as $R_{\mathrm{S}}(s)$ and $R_{\mathrm{E1}}(s)$ or $R_{\mathrm{E2}}(s)$. The traffic splitter allocates traffic to the exclusive BS and the non-shared BS proportionally to the achievable data rate as summarized in Table 5.1. In this scheme, the operators may also adjust the offload bias b_i to adapt to the time-varying incoming traffic density ω_i. Note that, the allocated traffic $\xi_j(s)$ is rate-dependent in the PF scheme.

5.3 Analysis of Load Balancing Schemes in a Multi-Operator Setting

In this section, the cell load equation is introduced. It is used to perform theoretical analysis on load balancing in multi-operator shared RANs. By using this equation, it is able to demonstrate that the PF load balancing scheme has the ability to admit more traffic than the RR scheme. On the other hand, the RR load balancing scheme has certain property that the PF load balancing scheme does not have.

The cell load equation is a fix point equation that the cell load is supposed to satisfy, which has been studied in previous works, e.g., [12, 81, 82]. Some slight changes are made to fit it to our system model.

Definition 6. *Consider tier* $j \in \{\mathrm{E1}, \mathrm{E2}, \mathrm{S}\}$. *Let* ξ_j *denote the downlink traffic density splitted to tier* j *at an arbitrary location, whose unit is* $\mathrm{bit}/\mathrm{s}/\mathrm{m}^2$. *Let* R_j *denote the achievable downlink data rate at this location. The cell load* ρ_j *satisfies*

the following equation [12, 81, 82]:

$$\rho_j \coloneqq \mathbb{E}\left[\frac{\xi_j}{R_j \lambda_j}\right], \tag{5.3}$$

where $R_j = W_j \log_2(1+\gamma_j)$*,* λ_j *denotes BSs density, and* W_j *denotes the spectrum bandwidth*

Note that, equation (5.3) is a fixed point equation on the cell load ρ_j, as the SINR γ_j is depends on the cell load ρ_j. The intuition behind this equation is as follows. The (normalized) airtime that is required to transmit traffics at the considered location is ξ_j/R_j. The expected cell area of a BS in tier j is $1/\lambda_j$. Hence, the expectation evaluates the expected (normalized) airtime that is required to transmit all traffics demand within a cell.

To provide guidelines for operators to deploy sufficient numbers of shared BSs, this chapter also defines the *maximum allowed traffic density region*. It is defined as the traffic density (ω_1, ω_2) such that there exists offload bias b_1 and b_2 where the network does not overload, i.e.,

$$O = \left\{(\omega_1, \omega_2) \,\middle|\, {}^{\exists} 0 \le b_1, b_2 < \infty, \text{ s.t. } 0 \le \rho_{\mathrm{S}}, \rho_{\mathrm{E1}}, \rho_{\mathrm{E2}} < 1\right\}.$$

Notations O_{RR} and O_{PF} are used to distinguish between the regions of the RR and PF schemes.

5.3.1 Round Robin Load Balancing

Due to the randomness of location s and channel fading gains, the SINR $\gamma(s)$ is a stochastic variable. The complementary cumulative distribution function (CCDF) of the SINR, however, can be characterized under proper assumptions. Recall that, the Rayleigh fading channel, nearest BS association, and that BSs follow PPPs are assumed. The CCDF can be derived by following the stochastic geometry approach [85].

Lemma 2. *The CCDF of* γ_j *conditioned on* ρ_j *is derived as (5.4), where* $j \in \{\mathrm{E1}, \mathrm{E2}, \mathrm{S}\}$ *and* ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ *is the Gauss hypergeometric function. Furthermore, if the path-loss exponent* $\alpha_{\mathrm{E1}} = \alpha_{\mathrm{E2}} = \alpha_{\mathrm{S}} = 4$*, a simplified expression (5.5) can be obtained.*

$$\begin{aligned}\mathbb{P}\left[\gamma_j > \theta \,\middle|\, \rho_j\right] &= \int_0^{\infty} 2\pi x \lambda_j \exp\Bigg(-\pi x^2 \lambda_j - \frac{\theta N_0 W_j x^{\alpha_j}}{P_j \beta_j} \\ &\qquad -\frac{2\pi\rho_j\lambda_j\theta x^2}{\alpha_j - 2} {}_2F_1\left(1, 1 - 2/\alpha_j; 2 - 2/\alpha_j; -\theta\right)\Bigg) \mathrm{d}x.\end{aligned} \tag{5.4}$$

$$\begin{aligned}\mathbb{P}\left[\gamma_j > \theta \,\middle|\, \rho_j\right] &= \frac{\pi\lambda_j}{2}\sqrt{\frac{\pi P_j \beta_j}{\theta N_0 W_j}} \exp\left(\frac{\pi^2 \lambda_j^2 P_j \beta_j}{4\theta N_0 W_j}\left(1 + \rho_j \sqrt{\theta}\arctan\sqrt{\theta}\right)^2\right) \\ &\qquad \operatorname{erfc}\left(\frac{\pi\lambda_j}{2}\sqrt{\frac{P_j\beta_j}{\theta N_0 W_j}}\left(1 + \rho_j\sqrt{\theta}\arctan\sqrt{\theta}\right)\right).\end{aligned} \tag{5.5}$$

Proof. See Appendix A. □

When the RR load balancing scheme is applied, it is clear from Table 5.1 that the allocated traffic ξ_j is location-independent. Consequently, the cell load equation can be transformed as follows:

$$\rho_j = \frac{\xi_j}{\lambda_j}\mathbb{E}\left[{R_j}^{-1} \mid \rho_j\right], \tag{5.6}$$

where the expectation of the inverse Shannon capacity needs to be derived. As the CCDF of the SINR has been derived in Lemma 1, the cell load equation (5.3) can be derived based on Lemma 1.

Lemma 3. *Given the density of BSs λ_j and the bandwidth W_j, the CCDF of γ_i as shown in (5.5), the cell load equation is transformed as follows for $j \in \{\mathrm{E1}, \mathrm{E2}, \mathrm{S}\}$:*

$$\rho_j = \frac{\xi_j}{\lambda_j}\mathbb{E}_s\left[{R_j}^{-1}\middle|\rho_j\right] = \frac{\xi_j}{\lambda_j}\int_{\theta_{\min}}^{\infty} \frac{\ln 2\left(1 - \mathbb{P}[\gamma_j > \theta \,|\, \rho_j]\right)}{W_j(1+\theta)[\ln(1+\theta)]^2}\,\mathrm{d}\theta, \tag{5.7}$$

where $\theta_{\min}$ is the minimum allowable SINR for association.

Proof. See Appendix B. □

The parameter $\theta_{\min}$ denotes the minimum SINR for cell association [81, 86]. Note that, the selection of $\theta_{\min}$ determines the trade-off between coverage probability and cell load. Since the selection of $\theta_{\min}$ is beyond our scope, $\theta_{\min}$ is considered as a constant close to zero.

Substituting the parameters listed in Table 5.1 into Definition 1, obtain the following cell load equations.

$$\begin{aligned}
\rho_{\mathrm{S}} &= \left(\frac{b_1\omega_1}{1+b_1} + \frac{b_2\omega_2}{1+b_2}\right) \cdot \frac{\mathbb{E}\left[{R_{\mathrm{S}}}^{-1} \middle| \rho_{\mathrm{S}}\right]}{\lambda_{\mathrm{S}}}, \\
\rho_{\mathrm{E1}} &= \frac{\omega_1}{1+b_1} \cdot \frac{\mathbb{E}\left[{R_{\mathrm{E1}}}^{-1} \middle| \rho_{\mathrm{E1}}\right]}{\lambda_{\mathrm{E1}}}, \\
\rho_{\mathrm{E2}} &= \frac{\omega_2}{1+b_2} \cdot \frac{\mathbb{E}\left[{R_{\mathrm{E2}}}^{-1} \middle| \rho_{\mathrm{E2}}\right]}{\lambda_{\mathrm{E2}}},
\end{aligned} \tag{5.8}$$

The derived cell load equations are a set of fixed point equations of ρ_{S}, ρ_{E1}, and ρ_{E2}. From the equations, it is clear that the RR load balancing scheme guarantees *inter-operator resource isolation*. Precisely, the offload decisions made by one operator does not affect the cell load of the exclusive RAN of the other operator. Specifically, ρ_{E1} does not depend on b_2 and ρ_{E2} does not depend on b_1.

Deriving the explicit expressions of the cell loads can be difficult, if not impossible. Fortunately, it is possible to derive the maximum allowed traffic density such that the network does not overload, i.e., $\rho_j < 1$. It is derived as follows.

Proposition 1. *The closed form of the maximum allowed traffic density region O_{RR} is given as the following inequalities when the RR load balancing scheme is applied.*

$$\begin{aligned}
\omega_1 &< \frac{\lambda_{\mathrm{S}}}{\mathbb{E}\left[{R_{\mathrm{S}}}^{-1} \middle| \rho_{\mathrm{S}} = 1\right]} + \frac{\lambda_{\mathrm{E1}}}{\mathbb{E}\left[{R_{\mathrm{E1}}}^{-1} \middle| \rho_{\mathrm{E1}} = 1\right]}, \\
\omega_2 &< \frac{\lambda_{\mathrm{S}}}{\mathbb{E}\left[{R_{\mathrm{S}}}^{-1} \middle| \rho_{\mathrm{S}} = 1\right]} + \frac{\lambda_{\mathrm{E2}}}{\mathbb{E}\left[{R_{\mathrm{E2}}}^{-1} \middle| \rho_{\mathrm{E2}} = 1\right]}, \\
\omega_1 + \omega_2 &< \frac{\lambda_{\mathrm{S}}}{\mathbb{E}\left[{R_{\mathrm{S}}}^{-1} \middle| \rho_{\mathrm{S}} = 1\right]} + \frac{\lambda_{\mathrm{E1}}}{\mathbb{E}\left[{R_{\mathrm{E1}}}^{-1} \middle| \rho_{\mathrm{E1}} = 1\right]} \\
&\quad + \frac{\lambda_{\mathrm{E2}}}{\mathbb{E}\left[{R_{\mathrm{E2}}}^{-1} \middle| \rho_{\mathrm{E2}} = 1\right]}.
\end{aligned} \tag{5.9}$$

This is a sufficient condition for ω_1 and ω_2 such that there exists offload bias b_1 and b_2, where the network does not overload, i.e., $\rho_{\mathrm{E1}} < 1$, $\rho_{\mathrm{E2}} < 1$, and $\rho_{\mathrm{S}} < 1$.

Proof. Consider the extreme case in which all tiers have been overloaded. Setting the cell loads $\rho_{\mathrm{S}} = \rho_{\mathrm{E1}} = \rho_{\mathrm{E2}} = 1$ in (5.8), the result can be obtained. □

5.3.2 Proportionally Fair Load Balancing

To analyze the maximum allowable traffic in the PF load balancing case, this section begins with characterizing the cell load equation in three tiers. Following the definition of cell load in (5.3), the PF scheme defined in Table 5.1, the cell load equation is expressed as follows:

$$\rho_{\mathrm{E}i} = \mathbb{E}\left[\frac{\xi_{\mathrm{E}i}}{R_{\mathrm{E}i}\lambda_{\mathrm{E}i}}\middle|\rho_{\mathrm{E}i},\rho_{\mathrm{S}}\right] = \frac{1}{\lambda_{\mathrm{E}i}}\mathbb{E}\left[\frac{\omega_i}{R_{\mathrm{E}i}+b_iR_{\mathrm{S}}}\middle|\rho_{\mathrm{E}i},\rho_{\mathrm{S}}\right]. \tag{5.10}$$

where $i \in \mathcal{M}$. The cell load equation of tier S is determined by the traffic jointly allocated to tier S by two operators.

$$\rho_{\mathrm{S}} = \frac{1}{\lambda_{\mathrm{S}}}\sum_{i\in\mathcal{M}}\mathbb{E}\left[\frac{b_i\omega_i}{R_{\mathrm{E}i}+b_iR_{\mathrm{S}}}\middle|\rho_{\mathrm{E}i},\rho_{\mathrm{S}}\right]. \tag{5.11}$$

Given the CCDF of γ_i as shown in (5.5), the expected value of inverse Shannon rate conditioned on the cell load is given as follows for $j = \{\mathrm{E1},\mathrm{E2}\}$, $i = 1, 2$:

$$\begin{aligned}&\mathbb{E}\left[(R_{\mathrm{E}i}+b_iR_{\mathrm{S}})^{-1}\middle|\rho_{\mathrm{E}i},\rho_{\mathrm{S}}\right]\\&=\iint_{\theta_{\min}}^{\infty}\frac{\mathrm{d}\mathbb{P}[\gamma_{\mathrm{E}i}>\theta\,|\,\rho_{\mathrm{E}i}]\,\mathrm{d}\mathbb{P}[\gamma_{\mathrm{S}}>\varphi\,|\,\rho_{\mathrm{S}}]}{W_i\log_2(1+\theta)+b_iW_{\mathrm{S}}\log_2(1+\varphi)},\end{aligned} \tag{5.12}$$

where $\theta_{\min}$ is the minimum allowable SINR for association.

From the derived cell load equations from (5.10) to (5.11), it is clear that the PF load balancing scheme no longer guarantees inter-operator resource isolation. Precisely, the cell load ρ_{E1} depends on the offload bias b_2 and the cell load ρ_{E2} depends on the offload bias b_1. This is because the PF load balancing scheme depends on the instantaneous data rate R_j, which depends on the cell load ρ_j. This is the load-interference coupling effect as illustrated in Fig. 5.1. Consequently, the offload decisions made by one operator would influence the cell load of the exclusive RANs of the other operator.

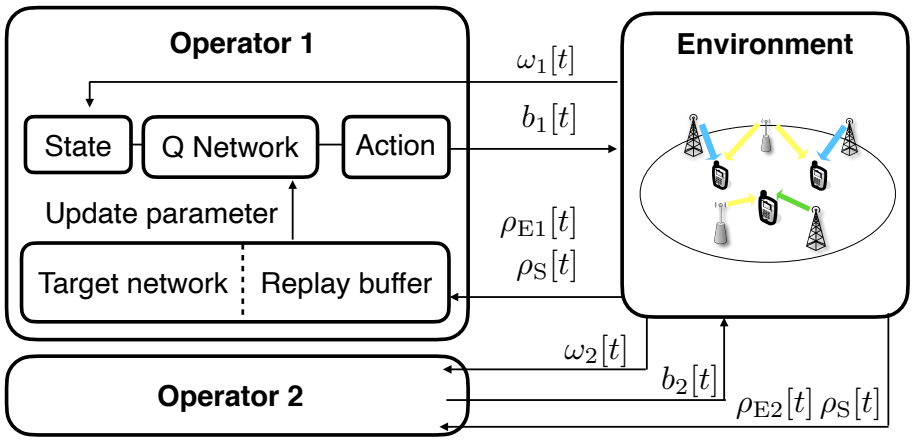


Figure 5.5: Dynamic load balancing game framework.

The explicit expressions of cell load equations from (5.10) to (5.11) are difficult to derive. The following proposition, however, can be theoretically proved. It indicates that the PF load balancing scheme has the ability to admit more traffic than the RR load balancing scheme.

Proposition 2. *The maximum allowed traffic density region of the PF load balancing scheme is a super set of that of the RR load balancing scheme.*

$$\mathcal{O}_{\mathrm{RR}} \subseteq \mathcal{O}_{\mathrm{PF}}. \tag{5.13}$$

Proof. See Appendix C. □

This theorem shows that exploiting a PF load balancing scheme expands the maximum allowed traffic density region. Given a traffic density, the PF load balancing scheme would not overload the RANs if the RR load balancing scheme does not. On the other hand, the RR load balancing scheme guarantees inter-operator resource isolation. Moreover, it is easy to implement as the temporal achievable data rate information is not required.

5.4 Dynamic Load Balancing Game Formulation: Stochastic Game with Individual States

The previous section has derived the cell load ρ_j as a function of offload bias b_i and traffic density ω_i. Based on this result, this section considers the multi-operator load balancing game in a dynamic settings where the traffic density state ω_i is a time-varying stochastic variable. Operators dynamically adjust their offload bias b_i to the variation in ω_i without exchanging any information. To model the incompleteness of the information, a repeated stochastic game with individual states is formulated. Moreover, both of the RR and the PF load balancing schemes are discussed under the same framework.

5.4.1 Game Formulation

The repeated stochastic game is built from a one-stage game. The one stage game is denoted as $\mathcal{G} \coloneqq (\mathcal{M}, \{\Omega_i\}_{i\in\mathcal{M}}, \{\mathcal{A}_i\}_{i\in\mathcal{M}}, \{u_i\}_{i\in\mathcal{M}})$, where $\mathcal{M}$, Ω_i, $\mathcal{A}_i$ and u_i denote the index set of players, state space of player i, action space of player i and the utility function of player i, respectively. The players of this game are the operators $\mathcal{M} \coloneqq \{1, 2\}$. The state of a player is the network traffic density $\omega_i \in \Omega_i$. The action of a player is the offload bias $b_i \in \mathcal{A}_i$. For notation simplicity, denote $\boldsymbol{b} \coloneqq (b_1, b_2)$ and $\boldsymbol{\omega} \coloneqq (\omega_1, \omega_2)$. Note that, both of the traffic density ω_i and the offload bias b_i are continuous variables. To study this problem in a game-theoretic framework, this section considers a uniform quantized finite state space $\Omega_i \coloneqq \{0, \Delta\omega, 2\Delta\omega, \ldots, \omega_{\max}\}$ and a uniform quantized finite space, $\mathcal{A}_i = \{0, \Delta b, 2\Delta b, \ldots, b_{\max}\}$.

The utility function is a function defined as $u_i \colon \prod_{i\in\mathcal{M}} \Omega_i \times \prod_{i\in\mathcal{M}} \mathcal{A}_i \to \mathbb{R}$. For the utility function, this section considers the scenario where operators are willing to minimize the cell load. This is because minimizing the cell load reduces the scheduling delay and is important from a QoS-guaranteeing perspective. Note that, the cell load ρ_j depends on the traffic density $\boldsymbol{\omega}$ and offload bias $\boldsymbol{b}$ as have been studied in the previous section. The considered utility function is given as follows:

$$u_1(\boldsymbol{\omega}, \boldsymbol{b}) \coloneqq -\left(C_1\rho_{\mathrm{E1}} + \rho_{\mathrm{S}}\right), \tag{5.14}$$

where $C_i \in \mathbb{R}_{\geq 0}$ is a weighting constant determined by an operator. It reflects how importantly an operator judges its exclusive shared band. Similarly, the utility function of operator 2 is defined as:

$$u_2(\boldsymbol{\omega}, \boldsymbol{b}) \coloneqq -\left(C_2 \rho_{\mathrm{E}2} + \rho_{\mathrm{S}}\right). \tag{5.15}$$

The repeated stochastic game $\mathcal{G}^{\infty} = (\mathcal{M}, \{\mathcal{T}_i\}_{i\in\mathcal{M}}, \{\bar{u}_i\}_{i\in\mathcal{M}})$ is built from the one-stage game $\mathcal{G}$ defined above. Precisely, $\mathcal{G}^{\infty}$ is built by playing $\mathcal{G}$ an infinite number of stages, where $\mathcal{T}_i$ denotes the sequence of actions played by player i and $\bar{u}_i$ denotes the time average utility for $i \in \mathcal{M}$. In the repeated stochastic game, a player cares about what he gains over the whole game duration, i.e.,:

$$\bar{u}_i = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} u_i(\boldsymbol{\omega}[t], \boldsymbol{b}[t]), \tag{5.16}$$

where $\boldsymbol{\omega}[t]$ and $\boldsymbol{b}[t]$ denote the traffic density state and the offload bias in stage t, respectively.

The same game-theoretical framework can be used to discuss both of the RR and the PF load balancing schemes. The entire framework is as shown in Fig. 5.5. At the beginning of stage t, each player observes its own traffic state $\omega_i[t]$ and determines the offload bias $b_i[t]$ based on a certain algorithm. Within the time interval between stages t and $t+1$, it is assumed that $\omega_i[t]$ and $b_i[t]$ do not change. Recall that the physical meaning of the cell load $\rho_j[t]$ is the fraction of the consumed airtime resource. The cell load can be measured at the end of each stage. In the case of overload, i.e., $\rho_j[t] \geq 1$, the cell load can also be measured indirectly from the scheduling delay of associated UEs, as the cell load determines the average scheduling delay of UEs [81]. Operators use the measured cell load to calculate their utility functions and to update the parameters in their learning automata. Note that in this framework, each operator is only aware of its own traffic density and its own cell load. An operator cannot observe or memorize that of the other operators. For instance, operator 1 is aware of $\omega_1[t]$, $b_1[t]$, $\rho_{\mathrm{E}1}[t]$, and $\rho_{\mathrm{S}}[t]$, but not $\omega_2[t]$, $b_2[t]$, and $\rho_{\mathrm{E}2}[t]$.

5.4.2 Equilibrium Analysis

The NE is a solution concept of the game. It is a set of action profiles where a single player has no incentive to deviate unilaterally. The characterization of

NEs in a dynamic game can be a formidable task if the non-stationary strategy is considered. To this end, this work focuses on studying the NE of the *stationary behavioral strategy* [71]. In other words, this work assumes that the strategy of players has reached stationary distribution in the equilibrium state. A stationary distribution over strategy is a conditional probability mass function defined as:

$$g_i(b_i|\omega_i) \coloneqq \mathbb{P}\left[\text{Player } i \text{ acts } b_i \,\middle|\, \text{Player } i \text{ observes } \omega_i\right]. \tag{5.17}$$

From the ergodicity of the stationary strategy g_i, the time average of the utility function $\bar{u}_i$ is expressed as its expectation as a function of g_1 and g_2. For $i \in \mathcal{M}$:

$$\bar{u}_i = \bar{u}_i(g_1, g_2) = \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{\boldsymbol{b}\in\mathcal{A}} g_1(b_1|\omega_1) g_2(b_2|\omega_2) u_i(\boldsymbol{b}, \boldsymbol{\omega})\right], \tag{5.18}$$

where $\mathcal{A} \coloneqq \mathcal{A}_1 \times \mathcal{A}_2$. The definition of NE is hence given as follows:

Definition 7 (NE)**.** *[71, 87] A profile of probability mass functions* (g_1^*, g_2^*) *is an NE for the considered game* $\mathcal{G}^\infty$ *if it satisfies the following constraint:* $\forall i, i' \in \mathcal{M}$, $i \neq i'$, $\forall \omega_i \in \Omega_i$, $\forall b' \in \mathcal{A}_i$,

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{\boldsymbol{b}\in\mathcal{A}} g_i^*(b_i|\omega_i) g_{-i}^*(b_{-i}|\omega_{-i}) u_i(\boldsymbol{\omega}, \boldsymbol{b})\right] \\ &\qquad \geq \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{\boldsymbol{b}\in\mathcal{A}} g_i^*(b'|\omega_i) g_{-i}^*(b_{-i}|\omega_{-i}) u_i(\boldsymbol{\omega}, b', b_{-i})\right].\end{aligned}$$

From the theoretical analysis in the previous section, it is clear that RR guarantees inter-operator resource isolation. In the RR load balancing scheme, the offload decisions made by one operator does not affect the cell load of the exclusive RANs of the other operators. Owing to this fact, it is able to first prove that the considered one-stage game $\mathcal{G}$ is a potential game at least when the RR load balancing scheme is applied.

Lemma 4. *When RR load balancing is applied, the considered one-stage game* $\mathcal{G}$ *is a potential game [71, 72], i.e.,* $\forall i, i' \in \mathcal{M}$, $i \neq i'$, $\forall \omega_i \in \Omega_i$, $\forall b_i, b' \in \mathcal{A}_i$, *and* $b_{-i} \in \mathcal{A}_{-i}$ *there exists a potential function* $\phi\colon \prod_{i\in\mathcal{M}} \mathcal{A}_i \to \mathbb{R}$ *satisfying:*

$$u_i(\boldsymbol{\omega}, b_i', b_{-i}) - u_i(\boldsymbol{\omega}, b_i, b_{-i}) = \phi(\boldsymbol{\omega}, b_i', b_{-i}) - \phi(\boldsymbol{\omega}, b_i, b_{-i}). \tag{5.19}$$

The potential function $\phi(\boldsymbol{\omega}, \boldsymbol{b})$ *is given as follows:*

$$\phi(\boldsymbol{\omega}, \boldsymbol{b}) = -C_1\rho_{\mathrm{E1}}(\boldsymbol{\omega}, \boldsymbol{b}) - C_2\rho_{\mathrm{E2}}(\boldsymbol{\omega}, \boldsymbol{b}) - \rho_{\mathrm{S}}(\boldsymbol{\omega}, \boldsymbol{b}). \tag{5.20}$$

Proof. If the RR load balancing scheme is applied, the deviation of a player does not change the cell load of the exclusive band of the other operator. Taking $i = 1$ for instance, $\rho_{\mathrm{E2}}(\boldsymbol{\omega}, b_1, b_2) = \rho_{\mathrm{E2}}(\boldsymbol{\omega}, b', b_2)$. Consequently, $u_1(\boldsymbol{\omega}, b', b_2) - u_1(\boldsymbol{\omega}, b_1, b_2) = -(C_1\rho_{\mathrm{E1}}(\boldsymbol{\omega}, b', b_2)+\rho_{\mathrm{S}}(\boldsymbol{\omega}, b', b_2))+(C_1\rho_{\mathrm{E1}}(\boldsymbol{\omega}, b_1, b_2)+\rho_{\mathrm{S}}(\boldsymbol{\omega}, b_1, b_2)) = \phi(\boldsymbol{\omega}, b', b_2) - \phi(\boldsymbol{\omega}, b_1, b_2)$. This proves the lemma. □

In our considered stochastic game $\mathcal{G}^{\infty}$, the network traffic state $\boldsymbol{\omega}[t]$ is an external factor that does not depend on the actions chosen by the players. Thus, it is possible extend the existence of the potential function to the considered stochastic game $\mathcal{G}^{\infty}$.

Proposition 3. *When the RR load balancing scheme is applied, the considered load balancing game $\mathcal{G}^{\infty}$ is a potential game, i.e.,*

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{\boldsymbol{b}\in\mathcal{A}} g_i^*(b_i|\omega_i)g_{-i}^*(b_{-i}|\omega_{-i})u_i(\boldsymbol{\omega}, \boldsymbol{b})\right] \\
&- \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{\boldsymbol{b}\in\mathcal{A}} g_i^*(b'|\omega_i)g_{-i}^*(b_{-i}|\omega_{-i})u_i(\boldsymbol{\omega}, b', b_{-i})\right] \\
&= \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{\boldsymbol{b}\in\mathcal{A}} g_i^*(b_i|\omega_i)g_{-i}^*(b_{-i}|\omega_{-i})\phi(\boldsymbol{\omega}, \boldsymbol{b})\right] \\
&- \mathbb{E}_{\boldsymbol{\omega}}\left[\sum_{\boldsymbol{b}\in\mathcal{A}} g_i^*(b'|\omega_i)g_{-i}^*(b_{-i}|\omega_{-i})\phi(\boldsymbol{\omega}, b', b_{-i})\right].
\end{aligned} \tag{5.21}$$

Proof. This follows from Lemma 3 by taking the expectation over the traffic density $\boldsymbol{\omega}$ and the chosen action $\boldsymbol{b}$. □

Proposition 3 provides assurance for the convergence property of the considered multi-operator load balancing game even under incomplete information. The existence of potential function is guaranteed if the RR load balancing scheme is applied and the distribution of traffic density $\boldsymbol{\omega}$ is stationary. Owing to this fact, the incentive of all players to change their strategy can be expressed using a single global function ϕ. A branch of fully distributed algorithms is guaranteed to guide players to NE upon convergence [71].

5.4.3 Load Balancing Algorithm

A player cannot directly calculate the NE as given in (5.19), because the knowledge of traffic density $\boldsymbol{\omega}$ and the stationary behavioral g_{-i} of the other operators are unknown. To this end, a fully distributed algorithm is required to enable players to learn the NE under incomplete information. Owing to the existence of the potential function, there is a group of algorithms that can guide multiple players to the NE under incomplete information, e.g., Boltzmann Gibbs sampling, trial-and-error learning [71], and deep Q-learning algorithm [88]. These algorithm works in a similar manner: each player presumes that the opponents are playing stationary strategies and then iteratively improves the strategy based on empirical rewards. Consequently, the incentive of all players to change their strategy can be expressed using the potential function. Among these, the deep Q-learning algorithm [88] is a recently proposed algorithm, which is suitable for learning policy like $g_i(b_i|\omega_i)$, i.e., executing an action in a specific state. It also has good scalability and is easy to implement.

The entire algorithm is given in Algorithm 5.1. Note that, the contribution of this wokr is to confirm its convergence property in the considered game. The standard deep Q-learning with experience replay algorithm [88] is applied. For both of the RR and the PF load balancing schemes, the algorithm works in the same way. As shown in Fig. 5.5, the learning automaton comprises a main Q network, a target network, and a replay buffer. The main network is used as a function approximation for the stationary behavioral strategy g_i. The input of the main network corresponds to the state ω_i. The input layer size of the main network is $|\Omega_i|$. One-hot encoding is used to encode the input traffic into a vector with size $|\Omega_i|$. The hidden layer size is 24. The input layer activation and hidden layer activation are ReLU functions, while the output activation layer is linear. The output layer is $|\mathcal{A}_i|$. The replay buffer and the target network are used to train the parameters in the main network. The target network has the same shape with the main network. The replay buffer stores the memory in each stage. The batch size is 144. The loss function is MSE. The optimizer is a gradient descent algorithm.

Table 5.2: Simulation parameters.

Parameters	Notation	Value
BS density	$\lambda_{\mathrm{E1}}, \lambda_{\mathrm{E2}}, \lambda_{\mathrm{S}}$	$1 \times 10^{-5}\ \mathrm{m}^{-2}$
Path loss exponent	$\alpha_{\mathrm{E1}}, \alpha_{\mathrm{E2}}, \alpha_{\mathrm{S}}$	4
Path loss gain	$\beta_{\mathrm{E1}}, \beta_{\mathrm{E2}}, \beta_{\mathrm{S}}$	$(4\pi f_c/(3 \times 10^8))^2$
Transmit power	$P_{\mathrm{E1}}, P_{\mathrm{E2}}, P_{\mathrm{S}}$	43 dBm
Bandwidth	$W_{\mathrm{E1}}, W_{\mathrm{E2}}, W_{\mathrm{S}}$	20 MHz
Noise power spectral density	N_0	-174 dBm/Hz
Minimum Association SINR	θ_{min}	-10 dB
Weighting constant	C_1, C_2	2
Time interval	Δt	10 min
Episode length		144 time intervals
Input layer (state space) size	$\lvert\Omega_i\rvert$	100
Output layer (action space) size	$\lvert\mathcal{A}_i\rvert$	100
Mini-batch size		144
Replay buffer memory		288
Exploration factor initial value	$\varepsilon_{\mathrm{DQN}}$	1
Exploration factor decay		0.995/time interval
Discount factor	γ_{DQN}	0.1

Algorithm 5.1 Load Balancing Algorithm based on Deep Q-learning [88].

For each operator $i = 1, 2$ executes the following algorithm distributedly.
Initialize replay buffer D_i, action-value function $Q_i(\omega, b, ; \theta_i)$ with random weights.
Initialize target action-value function $\hat{Q}_i(\omega, b; \theta_i^-)$ with weights $\theta_i^- = \theta_i$.
for every episode **do**
    for every time interval **do**
        Observe the network traffic density $\omega_i[t]$.
        With probability ε_{DQN} select a random offload bias $b_i[t] \in \mathcal{A}_i$. Otherwise select the offload bias $b_i[t] = \text{argmax}_{b \in \mathcal{A}_i} Q_i(\omega_i[t], b; \theta_i)$.
        Given the offload bias $b_i[t]$, split incoming traffic based on Table 5.1. Then, observe the cell load and calculate reward $r_i[t]$ based on (5.14) or (5.15).
        Store transition $(\omega_i[t], b_i[t], r_i[t], \omega_i[t+1])$ in D_i.
        Sample random minibatch of transition samples $(\omega_i[t'], b_i[t'], r_i[t'], \omega_i[t'+1])$ from D_i.
        Let $y[t'] = r_i[t'] + \gamma_{\text{DQN}} \text{max}_{a'} \hat{Q}(\omega_i[t'+1], a'; \theta^-)$.
        Perform a gradient descent step on $(y[t'] - Q(\omega_i[t'], a_i[t']; \theta))^2$ with respect to the network parameters θ_i.
    end for
    Reset $\hat{Q}_i = Q_i$ at the end of episode.
end for

5.5 Numerical Results

Numerical results are presented to validate the theoretical results and to evaluate the learning algorithm. Unless otherwise stated, the simulation parameters are listed in Table 5.2.

5.5.1 Impact of Offload Bias on Cell Load

In this subsection, the accuracy of the theoretical derivation of the cell load is demonstrated by comparing it with the Monte Carlo simulation results. This section also evaluates the impact of offload bias b_i on the cell load ρ_j. This section plots cell load ρ_{E1} and ρ_{S} as the function of the offload bias b_1. The downlink traf-

fic density is fixed as $\omega_1 = \omega_2 = 200\,\text{bit/s/m}^2$ and fix the offload bias as $b_2 = 0$. The RR and the PF load balancing schemes are evaluated in Figs. 5.6a and 5.6b, respectively. The theoretical results matches with the Monte Carlo simulation results, which demonstrates the accuracy of the theoretical derivation.

In Fig. 5.6a, as operator 1 offloads more traffic to the shared band, the cell load in the shared band ρ_S increases, while the cell load in the exclusive band ρ_{E1} decreases. Specifically, when $b_1 > 4$, the shared band approaches to the overloaded condition. On the other hand, when $b_1 < 0.25$, the exclusive band approaches to the overloaded condition. An appropriate offload bias in this case would be $0.25 < b_1 < 4$. Fig. 5.6a also increases the offload bias of b_2. In doing so, the cell load of the shared band ρ_S increases. Especially, when $b_2 > 1$, no matter how b_1 is set, there is no appropriate b_1 such that the networks do not overload.

Compared to the RR load balancing scheme, the PF load balancing scheme achieves lower cell load under the same traffic density. This is owing to the utilization of CSI for traffic splitting. As shown in Fig. 5.6b, when $b_1 > 30$, the shared band approaches to the overloaded condition. On the other hand, when $b_1 < 0.01$, the exclusive band approaches to the overloaded condition. An appropriate offload bias in this case would be $0.01 < b_1 < 30$, which is much wider than that in the RR load balancing scheme.

Figure 5.6b also demonstrates that the PF impairs the inter-operator resource isolation. Note that, in Fig. 5.6b, ρ_{E2} increases with b_1. This indicates that, the offloading action performed by operator 1 affects the cell load of the exclusive band of operator 2. The reason for this can be explained by the load-interference coupling effect in Fig. 5.1. As operator 1 increases the offload bias b_1, the cell load of the shared band increases. This will further decrease the achievable rate in the shared band, and hence, result in more traffic to be allocated to the exclusive band of operator 2. This effect does not appears in the RR load balancing scheme: As shown in Fig. 5.6a, ρ_{E1} does not change with b_2.

5.5.2 Maximum Allowed Traffic Density Region

Figures 5.7a and 5.7b evaluate the maximum allowed traffic density of the RR load balancing scheme and PF load balancing scheme, i.e., $\mathcal{O}_{RR}$ and $\mathcal{O}_{PF}$, respectively. Several conditions of BS densities are evaluated. For any point (ω_1, ω_2) in

the shaded region, there exists a setting of b_1 and b_2 such that the network does not overload. Comparing Figs. 5.7a and 5.7b, it is clear that $\mathcal{O}_{\text{PF}}$ is a superset of $\mathcal{O}_{\text{RR}}$, which validates Proposition 2. Moreover, as operator 1 deploys more exclusive BSs, the boundary shifts to the right parallelly. On the other hand, if the shared BSs are increased, the boundary expands. This theoretical analysis aims at providing guidelines for deploying sufficient shared BSs.

5.5.3 Dynamic Load Balancing Performance

This subsection evaluates the multi-operator load balancing in a time-varying traffic setting. The following evaluations use a real dataset of telecommunication traffic [83] to represent the time series of traffic densities $\omega_1[t]$ and $\omega_2[t]$. The dataset is measured in an area of 11.75 km $\times$ 11.75 km in the city of Milan at a time interval of 10 min. As the dataset only provides the traffic profile of one operator, to evaluate the multi-operator case, a week of data is used to represent the traffic profile of operator 1 and another week of data is used to represent that of operator 2. It is assumed that the time series of traffic densities $\omega_1[t]$ and $\omega_2[t]$ has a period of one week. This is reasonable as the traffic densities of different operators are generally highly correlated temporally [83, 84]. Without the loss of generality, the data is normalized such that operator 2 has a lower traffic density compared to operator 1, i.e., operator 2 has fewer subscribers than operator 1. The traffic pattern is shown in Fig. 5.8. For expression simplicity, denote 1 episode = 24 hours = 144 time intervals. Note that, the traffic density in the workdays is higher than that in the weekends.

The dynamics of the cell load after a sufficiently long time are shown in Figs. 5.9a and 5.9b. Referring to Fig. 5.8, observe that the cell load fluctuates in accordance with the traffic density temporally. Note that, during peak hours in Fig. 5.9a, the shared band overloads as operators jointly offload traffics. This is because the exclusive bands have higher priorities than the shared band as considered in the utility functions (5.14) and (5.15), where $C_1, C_2 > 1$. Moreover, it is clear that the PF load balancing achieves a much lower cell load than the RR load balancing scheme. This is owing to the utilization of achievable data rate information in the PF load balancing scheme. Figures 5.10a and 5.10b present the average cell load in one week as the learning algorithm proceeds. Results indicate that the

average cell load converges fast in both cases. Although the convergence is not rigorously proved in the PF case, the convergence is still confirmed in simulation evaluations.

The results from Figs. 5.9a to 5.10b validate that operators learned to offload without exchanging private information with other operators. Even though the traffic density information is partially observable, i.e., operator 1 (or 2) only observes $\omega_1[t]$ (or $\omega_2[t]$), the conditioned probability distribution $\mathbb{P}[\omega_2|\omega_1]$ and $\mathbb{P}[\omega_1|\omega_2]$ are rather stationary as $\omega_1[t]$ and $\omega_2[t]$ are highly correlated temporally. Hence, this provides the possibility for operators to learn to offload without exchanging private information, e,g., traffic densities, number of BSs, and number of UEs.

5.6 Conclusion

This chapter has investigated the load balancing problem in a multi-operator RAN sharing scenario. Each operator strategically decides the fraction of the traffic offloaded to the shared RAN without exchanging the private traffic demand information. Through theoretical derivations and simulation results, it is demonstrated that the PF load balancing scheme admits more traffic than the RR load balancing scheme. On the other hand, it impairs the inter-operator resource isolation. In other words, the offloading decision made by one operator affects the cell load of the exclusive band of the other operators. This chapter proved that the game is a potential game at least when the RR load balancing scheme is applied. Owing to the existence of potential function, a fully distributed reinforcement learning algorithm is guaranteed to guide multiple operators to the NE upon convergence. Numerical evaluations demonstrate that the cell loads of different RANs are well-managed without exchanging any information among operators.

5.7 Details of Proofs

5.7.1 Proof of Lemma 1

The proof is similar to the approach in [85]. Consider a randomly selected UE of tier j. The CCDF of its downlink SINR can be derived by conditioning on the

communication distance to the associated BS:

$$\mathbb{P}\left[\gamma_j > \theta \,\middle|\, \rho_j\right] = \int_0^\infty \mathbb{P}\left[\gamma_j > \theta \,\middle|\, x, \rho_j\right] f_j(x)\,\mathrm{d}x,$$

where f_j denotes the probability density function (PDF) of the communication distance. Under the assumption that the UE associates with the closest BS and the BSs follow PPP, $f_j(x)$ is given as follows [85]:

$$f_j(x) = 2\pi x \lambda_j \exp\left(-\pi x^2 \lambda_j\right).$$

Substituting (5.1), the conditioned probability is given as follows under Rayleigh fading channel:

$$\mathbb{P}\left[\gamma_j > \theta \,\middle|\, x, \rho_j\right] = \exp\left(-\frac{N_0 W_j \theta x^{\alpha_j}}{P_j \beta_j}\right) \mathcal{L}\left(\frac{\theta x^{\alpha_j}}{P_j \beta_j}\right),$$

where $\mathcal{L}(\cdot)_I$ denotes the Laplace transform of interference. By using the Gauss hypergeometric function, it can be expressed as follows: [85]:

$$\mathcal{L}\left(\frac{\theta x^{\alpha_j}}{P_j \beta_j}\right)_I = \exp\left(\frac{-2\pi\rho_j \lambda_j \theta x^2}{\alpha_j - 2} {}_2F_1\left(1, 1 - \frac{2}{\alpha_j}; 2 - \frac{2}{\alpha_j}; -\theta\right)\right),$$

where ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ is the Gauss hypergeometric function. This derives the result (5.4). Furthermore, substitute $\alpha_j = 4$, (5.5) is derived.

5.7.2 Proof of Lemma 2

Let $f_{\gamma,j}(\theta)$ denote the PDF of the SINR γ_j. The expectation over the inverse Shannon rate is given as:

$$\mathbb{E}_{\gamma_j}[R_j^{-1} | \rho_j] = \int_{\theta_{\min}}^\infty \frac{f_{\gamma,j}(\theta)\,\mathrm{d}\theta}{W_j \log_2(1+\theta)},$$

where $\theta_{\min}$ denotes minimum SINR for association, i.e., $A \coloneqq \{s \in \mathbb{R}^2 \mid \gamma(s) > \theta_{\min}\}$. The threshold $\theta_{\min}$ is a pre-determined threshold and it determines the trade-off between coverage probability and cell load. It is considered as a constant close to zero.

Given a radon variable X that has nonzero probability density for only positive values. Its expected value is given by integrating its CCDF $\mathbb{P}[X > x]$, i.e.,

$$\mathbb{E}_X[X] = \int_0^\infty x f_X(x)\,\mathrm{d}x = \int_0^\infty \mathbb{P}[X > x]\,\mathrm{d}x.$$

Notice that R_j^{-1} is a monotonically decreasing function of θ. Hence, it is derived as follows:

$$\mathbb{E}_{\gamma_j}\left[{R_j}^{-1}\middle|\rho_j\right] = \int_{\theta_{\min}}^{\infty}\Big(1-\mathbb{P}[\gamma_j > \theta\,|\,\rho_j]\Big)\,\mathrm{d}R_j^{-1}(\theta).$$

Finally, differentiate R_j and derive the result as follows:

$$\mathbb{E}_{\gamma}[{R_j}^{-1}\big|\rho_j] = \int_{\theta_{\min}}^{\infty}\frac{\ln 2\,\Big(1-\mathbb{P}[\gamma_j > \theta\,|\,\rho_j]\Big)}{W_j(1+\theta)[\ln(1+\theta)]^2}\,\mathrm{d}\theta.$$

5.7.3 Proof of Proposition 2

Proof. Our goal is to prove that the PF load balancing scheme can always achieve equal or lower cell load than the RR load balancing scheme. Formally, consider a traffic $(\omega_1, \omega_2) \in O_{\mathrm{RR}}$. Let b_1 and b_2 denote the offload bias in RR load balancing scheme. Denote the cell load achieved under RR load balancing scheme as $\rho_{\mathrm{E1}}, \rho_{\mathrm{E2}}$, and ρ_{S}. Denote the cell load achieved under PF load balancing scheme as $\rho'_{\mathrm{E1}}, \rho'_{\mathrm{E2}}, \rho'_{\mathrm{S}}$. Our goal is to prove that: there exists b'_1 and b'_2 such that:

$$\rho_{\mathrm{E1}} \geq \rho'_{\mathrm{E1}},\ \rho_{\mathrm{E2}} \geq \rho'_{\mathrm{E2}},\ \rho_{\mathrm{S}} \geq \rho'_{\mathrm{S}}.$$

This can be proved by determining b'_1 and b'_2 as follows:

$$\frac{b_1}{1+b_1} = \mathbb{E}_s\left[\frac{R_{\mathrm{S}}b'_1}{R_{\mathrm{E1}}+b'_1R_{\mathrm{S}}}\right],$$
$$\frac{b_2}{1+b_2} = \mathbb{E}_s\left[\frac{R_{\mathrm{S}}b'_2}{R_{\mathrm{E2}}+b'_2R_{\mathrm{S}}}\right].$$

Note that, such b'_1 and b'_2 always exist in $\mathbb{R}_{\geq 0}$.

First of all, stochastic variables R_{S} and R_{E1} are independent because the positions of different tiers of BSs are assumed to be independent. Hence, $1/R_{\mathrm{S}}$ and R_{E1} are independent. Owing to this fact, it is possible to focus on R_{E1} and consider R_{S} as a constant. Second, function $f(x) = 1/(1+ax)$ is concave when $a, x > 0$. Based on this two facts, according to Jensen's inequality,

$$\mathbb{E}\left[\frac{b'_1/R_{\mathrm{S}}}{b'_1+R_{\mathrm{E1}}/R_{\mathrm{S}}}\right] \leq \frac{\mathbb{E}[b'_1/R_{\mathrm{S}}]}{b'_1+\mathbb{E}[R_{\mathrm{E1}}/R_{\mathrm{S}}]} = \frac{\mathbb{E}[b'_1/R_{\mathrm{S}}]}{\mathbb{E}[b'_1+R_{\mathrm{E1}}/R_{\mathrm{S}}]}.$$

Applying Jensen's inequality once again,

$$\frac{\mathbb{E}[b_1'/R_S]}{\mathbb{E}[b_1' + R_{E1}/R_S]} \leq \mathbb{E}[1/R_S]\mathbb{E}\left[\frac{b_1'}{b_1' + R_{E1}/R_S}\right].$$

Combining the above inequalities, the following inequality holds:

$$\frac{b_1}{1+b_1}\mathbb{E}\left[\frac{1}{R_S}\right] = \mathbb{E}\left[\frac{R_S b_1'}{R_{E1} + b_1' R_S}\right]\mathbb{E}\left[\frac{1}{R_S}\right] \geq \mathbb{E}\left[\frac{b_1'}{R_{E1} + b_1' R_S}\right].$$

Hence,

$$\rho_{E1} = \frac{b_1}{1+b_1}\mathbb{E}\left[\frac{1}{R_S}\right]\frac{\omega_1}{\lambda_1} \geq \mathbb{E}\left[\frac{b_1'}{R_{E1} + b_1' R_S}\right]\frac{\omega_1}{\lambda_1} = \rho_{E1}'.$$

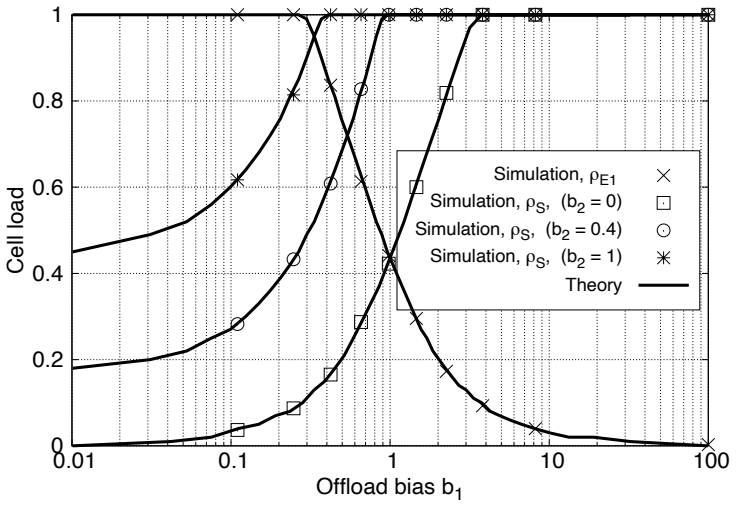
Following the same approach, it can be proved that:

$$\frac{1}{1+b_1}\mathbb{E}\left[\frac{1}{R_{E1}}\right] \geq \mathbb{E}\left[\frac{1}{R_{E1} + b_1' R_S}\right].$$

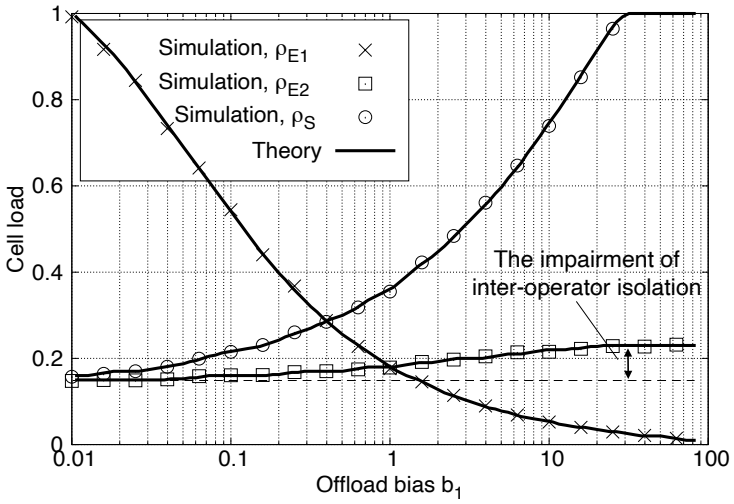
This proves the following inequalities and hence proves the proposition.

$$\rho_{E2} > \rho_{E2}',\ \rho_S \geq \rho_S'.$$

□

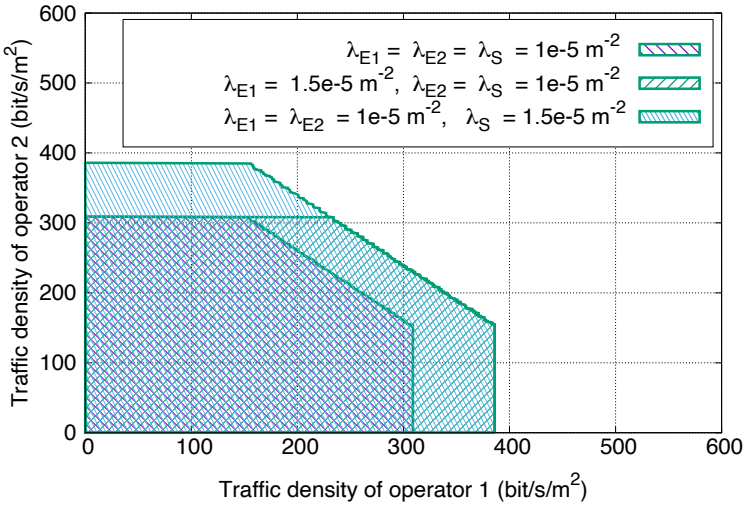


(a) RR load balancing scheme.

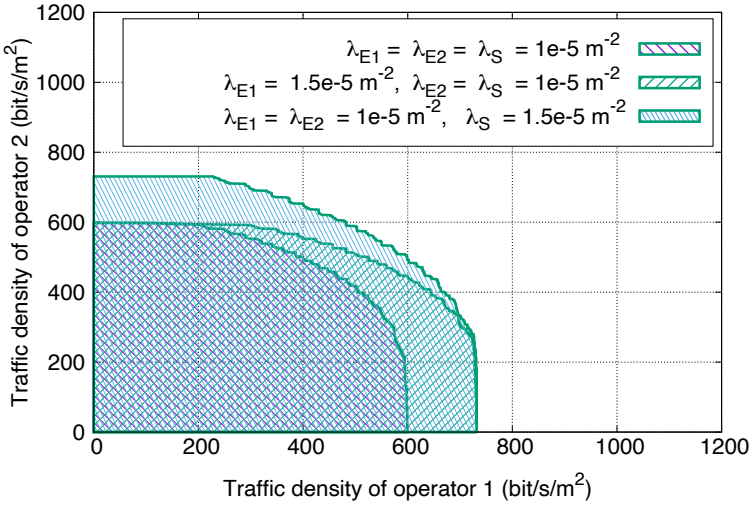


(b) PF load balancing scheme.

Figure 5.6: Cell loads when the RR or PF load balancing schemes are applied.



(a) Maximum allowed traffic density region when the RR load balancing scheme is applied, i.e., O_{RR}.



(b) Maximum allowable traffic density region when the PF load balancing scheme is applied, i.e., O_{PF}.

Figure 5.7: Maximum allowable traffic density region when the RR or PF load balancing schemes are applied.

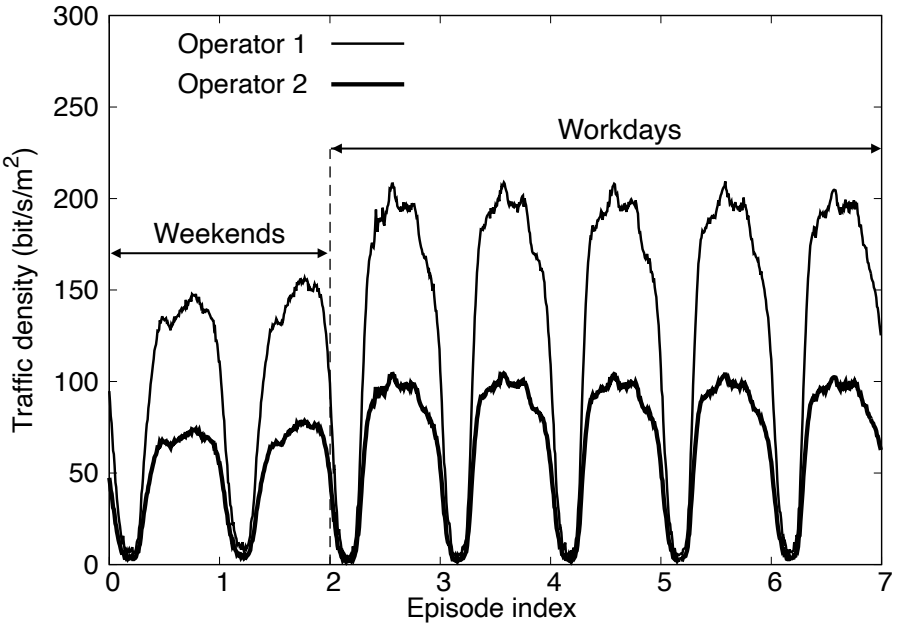
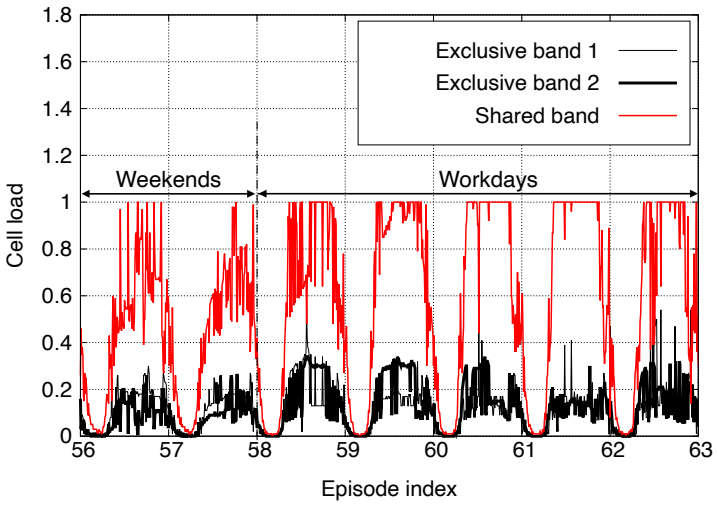
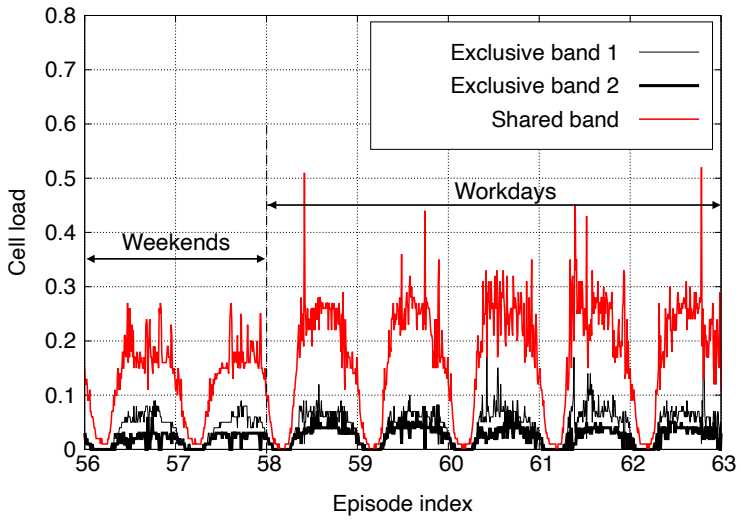


Figure 5.8: Time series of traffic densities $\omega_1[t]$ and $\omega_2[t]$ considered in the evaluations.

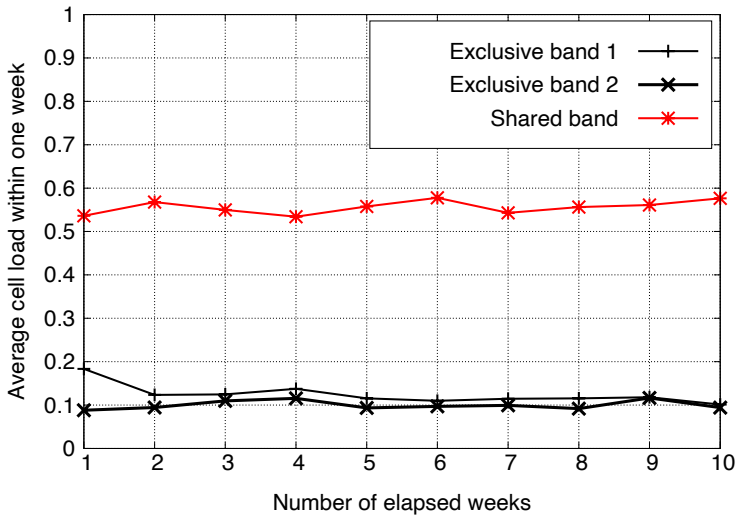


(a) RR load balancing scheme is applied.

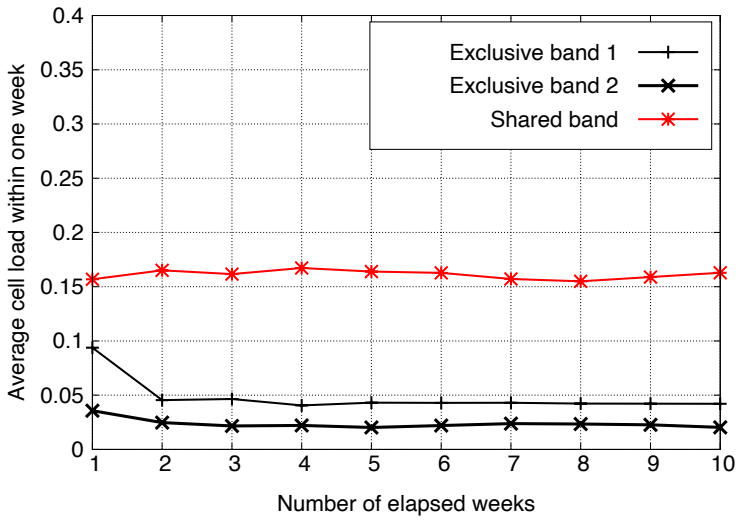


(b) PF load balancing scheme is applied.

Figure 5.9: Dynamics of cell load after after a sufficiently long time of learning.



(a) RR load balancing scheme.



(b) PF load balancing scheme.

Figure 5.10: Average cell load in one week as the learning algorithm proceeds.

Chapter 6

Conclusions

This thesis aims to reduce MAC delay in densely deployed wireless communication system by airtime resource management.

In the first topic, a reinforcement learning-based spatial reuse scheme is proposed. When the agent overhears an on-going transmission, it utilizes the information in the detected frame header to identify the interferer and decides whether or not to freeze the backoff counter accordingly. The proposed scheme is evaluated under various scenarios through simulations. Specifically, the composition of MAC layer service time is analyzed. The proposed scheme reduces the time of freezing the backoff counter while keeping the number of failed transmissions low. This confirms that, on the one hand, the agent learns to transmit concurrently with those OBSS interferers whose interference is tolerable at the receiver. On the other hand, the agent learns to refrain from transmitting concurrently with those interferers whose interference is not tolerable at the receiver. Moreover, the concept of state partition in MDP is utilized to study the performance gains due to making non-binary identifications of interferers on exploiting spatial reuse in WLANs. A theoretical bound on the gains in value function due to identifying interferers is obtained.

In the second topic, to mitigate the FIM starvation in densely deployed WLANs, distributed channel selection schemes are proposed to reduce the number of three-node-chain topologies on the contention graph. The proposed channel selection schemes are conducted in a distributed manner, requiring limited knowledge of the topology. Moreover, the schemes are formulated in the framework of strate-

gic form game. Payoff functions are designed so that the games are proved to be potential games. Simulations confirmed the convergence of the proposed scheme with payoff functions 1 and 2. Simulation results also proved the effectiveness of the proposed schemes in reducing the number of three-node-chain topologies and in enhancing the percentile throughput.

In the third topic, the load balancing problem in a multi-operator RAN sharing scenario has been investigated. Each operator strategically decides the fraction of the traffic offloaded to the shared RAN without exchanging the private traffic demand information. Through theoretical derivations and simulation results, it is demonstrated that the PF load balancing scheme admits more traffic than the RR load balancing scheme. On the other hand, it impairs the inter-operator resource isolation. In other words, the offloading decision made by one operator affects the cell load of the exclusive band of the other operators. The game is proved to be a potential game at least when the RR load balancing scheme is applied. Owing to the existence of potential function, a fully distributed reinforcement learning algorithm is guaranteed to guide multiple operators to the NE upon convergence. Numerical evaluations demonstrate that the cell loads of different RANs are well-managed without exchanging any information among operators.

It is expected that the schemes and analytical results proposed in this thesis can provided candidate technologies or guidelines for the next generation wireless communications and wireless networkings. Note that, the main topic of this thesis is focusing on reducing the average MAC delay. Few attention, however, has been paid to study the delay in the worst case. In the future works, it is indispensable to study the methods of stabilizing the MAC delay under environmental changes, such as the changes of topologies and traffic patterns. To support the real-time applications, it is important to provide deterministic guarantees and robustness to various environmental changes.

Bibliography

[1] Japanese Ministry of Internal Affairs and Communications, "Current Status of Mobile Communication Traffic in Japan (for March 2nd year of Reiwa)," Mar. 2020.

[2] 802.11 Working Group of the 802 Committee, "IEEE P802.11ax TM/D4.0 Amendment 6: Enhancements for High Efficiency WLAN," Feb. 2019.

[3] M. Bennis, M. Debbah, and H. V. Poor, "Ultra-reliable and low-latency wireless communication: Tail, risk, and scale," *Proc. of the IEEE*, vol. 106, no. 10, pp. 1834–1853, 2018.

[4] M.-T. Suer, C. Thein, H. Tchouankem, and L. Wolf, "Multi-connectivity as an enabler for reliable low latency communications-An overview," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 156–169, 2019.

[5] K. Terplan and P. A. Morreale, *The Telecommunications Handbook*. CRC Pr., 2018.

[6] R. Rom and M. Sidi, *Multiple Access Protocols: Performance and Analysis*. Springer Science & Business Media, 2012.

[7] M. Ilyas and I. Mahgoub, *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems*. CRC Pr., 2004.

[8] "IEEE Standard for Information technology - Telecommunications and information exchange between systems Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std. 802.11-2016 (Revision of IEEE Std 802.11-2012)*, Dec. 2016.

[9] A. Baid and D. Raychaudhuri, "Understanding channel selection dynamics in dense Wi-Fi networks," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 110–117, Jan. 2015.

[10] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020*. Cisco, Feb. 2016.

[11] M. Garetto, T. Salonidis, E. Knightly, and E. Edward, "Modeling per-flow throughput and capturing starvation in CSMA multi-hop wireless networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 4, pp. 864–877, Apr. 2008.

[12] T. Bonald and A. Proutière, "Wireless downlink data channels: User performance and cell dimensioning," in *Proc. ACM Annu. Int. Conf. Mobile Comput. Netw. (Mobicom), San Diego, USA*, Sept. 2003, pp. 339–352.

[13] D.-J. Deng, Y.-P. Lin, X. Yang, J. Zhu, Y.-B. Li, J. Luo, and K.-C. Chen, "IEEE 802.11ax: Highly efficient WLANs for intelligent information infrastructure," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 52–59, Dec. 2017.

[14] M. S. Afaqui, E. Villegas, and E. Aguilera, "IEEE 802.11ax: Challenges and requirements for future high efficiency WiFi," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 130–137, June 2017.

[15] H. A. Omar, K. Abboud, N. Cheng, K. R. Malekshan, A. T. Gamage, and W. Zhuang, "A survey on high efficiency wireless local area networks: Next generation WiFi," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2315–2344, Apr. 2016.

[16] T. Joshi, D. Ahuja, D. Singh, and D. P. Agrawal, "SARA: Stochastic automata rate adaptation for IEEE 802.11 networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 11, pp. 1579–1590, Oct. 2008.

[17] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris, "Link-level measurements from an 802.11b mesh network," in *Proc. ACM Special Interest Group Data Commun. (SIGCOMM) Comput. Commun. Rev., Portland, USA*, Aug. 2004, pp. 121–132.

[18] S. C. Liew, C. Kai, J. Leung, and B. Wong, "Back-of-the-envelope computation of throughput distributions in CSMA wireless networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 9, pp. 1319–1331, Sept. 2010.

[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] I. Jamil, L. Cariou, and J. Helard, "Efficient MAC protocols optimization for future high density WLANs," in *Proc. 2015 IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA, USA*, Oct. 2015, pp. 1054–1059.

[21] A. Gupta, J. Andrews, and R. Heath, "On the feasibility of sharing spectrum licenses in mmwave cellular systems," *IEEE Trans. Commun.*, vol. 64, no. 9, pp. 3981–3995, 2016.

[22] M. Rebato, M. Mezzavilla, S. Rangan, and M. Zorzi, "Resource sharing in 5G mmwave cellular networks," in *Proc. IEEE Int. Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), San Francisco, USA*, Apr. 2016, pp. 271–276.

[23] L. Cano, A. Capone, G. Carello, M. Cesana, and M. Passacantando, "Cooperative infrastructure and spectrum sharing in heterogeneous mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 10, pp. 2617–2629, 2016.

[24] R. H. Tehrani, S. Vahid, D. Triantafyllopoulou, H. Lee, and K. Moessner, "Licensed spectrum sharing schemes for mobile operators: A survey and outlook," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2591–2623, 2016.

[25] European Telecommunications Standards Institute, "Universal mobile telecommunications system (UMTS), LTE, network sharing, architecture and functional description (3GPP TS 23.251 version 15.0.0 release 15)," in *ETSI-TS 123 251*, 2018.

[26] M. M. Kassem and M. K. Marina, "Future wireless spectrum below 6 GHz: A UK perspective," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN), Stockholm, Sweden*, Sept. 2015, pp. 59–70.

[27] Samsung's 6G White Paper, "6G the next hyper-connected experience for all," July 2020.

[28] M. Rebato, F. Boccardi, M. Mezzavilla, and S. Rangan, "Hybrid spectrum sharing in mmwave cellular networks," *IEEE Trans. Cogn. Commu. and Netw.*, vol. 3, no. 2, pp. 155–168, 2017.

[29] Y. Zhu, Q. Zhang, Z. Niu, and J. Zhu, "On optimal QoS-aware physical carrier sensing for IEEE 802.11 based WLANs: Theoretical analysis and protocol design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1369–1378, Apr. 2008.

[30] H. Ma, R. Vijayakumar, S. Roy, and J. Zhu, "Optimizing 802.11 wireless mesh networks based on physical carrier sensing," *IEEE/ACM Trans. Netw.*, vol. 17, no. 5, pp. 1550–1563, Oct. 2009.

[31] L. Fu, S. C. Liew, and J. Huang, "Effective carrier sensing in CSMA networks under cumulative interference," *IEEE Trans. Mobile Comput.*, vol. 12, no. 4, pp. 748–760, Apr. 2013.

[32] J. Deng, B. Liang, and P. K. Varshney, "Tuning the carrier sensing range of IEEE 802.11 MAC," in *Proc. IEEE Global Commun. Conf. (GLOBECOM), Dallas, USA*, Dec. 2004, pp. 2987–2991.

[33] D. M. Kim and S.-L. Kim, "An iterative algorithm for optimal carrier sensing threshold in random CSMA/CA wireless networks," *IEEE Commmu. Lett.*, vol. 17, no. 11, pp. 2076–2079, Sept. 2013.

[34] H. ElSawy and E. Hossain, "A modified hard core point process for analysis of random CSMA wireless networks in general fading environments," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1520–1534, Apr. 2013.

[35] K. Yamamoto, X. Yang, T. Nishio, M. Morikura, and H. Abeysekera, "Analysis of inversely proportional carrier sense threshold and transmission power setting," in *Proc. IEEE Consum. Commun. Netw. Conf. (CCNC), Las Vegas, USA*, Jan. 2017, pp. 13–18.

[36] Y. Wen, H. Fujita, and D. Kimura, "Throughput-aware dynamic sensitivity control algorithm for next generation WLAN system," in *Proc. IEEE Symp. Pers. Indoor Mobile Radio Commun. (PIMRC), Montreal, Canada*, Feb. 2017, pp. 1–7.

[37] I. Selinis, M. Filo, S. Vahid, J. Rodriguez, and R. Tafazolli, "Evaluation of the DSC algorithm and the BSS color scheme in dense cellular-like IEEE 802.11ax deployments," in *Proc. IEEE Symp. Pers. Indoor Mobile Radio Commun. (PIMRC), Valencia, Spain*, Sept. 2016, pp. 1–7.

[38] M. S. Afaqui, E. Garcia-Villegas, E. Lopez-Aguilera, and D. Camps-Mur, "Dynamic sensitivity control of access points for IEEE 802.11ax," in *Proc. IEEE Int. Conf. Commun. (ICC), Kuala Lumpur, Malaysia*, May 2016, pp. 1–7.

[39] W. Afifi, E.-H. Rantala, E. Tuomaala, S. Choudhury, and M. Krunz, "Throughput-fairness tradeoff evaluation for next-generation WLANs with adaptive clear channel assessment," in *Proc. IEEE Int. Conf. Commun. (ICC), Kuala Lumpur, Malaysia*, May 2016, pp. 1–6.

[40] C. Thorpe and L. Murphy, "A survey of adaptive carrier sensing mechanisms for IEEE 802.11 wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1266–1293, Mar. 2014.

[41] S. Kim, S. Yoo, J. Yi, Y. Son, and S. Choi, "FACT: Fine-grained adaptation of carrier sense threshold in IEEE 802.11 WLANs," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 1886–1891, Feb. 2017.

[42] J. Zhu, B. Metzler, X. Guo, and Y. Liu, "Adaptive CSMA for scalable network capacity in high-density WLAN: A hardware prototyping approach," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM), Barcelona, Spain*, Apr. 2006, pp. 1–10.

[43] J. Kibiłda, N. Kaminski, and L. DaSilva, "Radio access network and spectrum sharing in mobile networks: A stochastic geometry perspective," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2562–2575, 2017.

[44] R. Jurdi, A. K. Gupta, J. G. Andrews, and R. W. Heath, "Modeling infrastructure sharing in mmwave networks with shared spectrum licenses," *IEEE Trans. Cognitive Commun. and Netw.*, vol. 4, no. 2, pp. 328–343, 2018.

[45] J. Park, J. G. Andrews, and R. W. Heath, "Inter-operator base station coordination in spectrum-shared millimeter wave cellular networks," *IEEE Trans. Cognitive Commun. and Netw.*, vol. 4, no. 3, pp. 513–528, 2018.

[46] M.-S. Pan, T.-M. Lin, C.-Y. Chiu, and C.-Y. Wang, "Downlink traffic scheduling for LTE-A small cell networks with dual connectivity enhancement," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 796–799, 2016.

[47] S. Singh, M. Geraseminko, S.-p. Yeh, N. Himayat, and S. Talwar, "Proportional fair traffic splitting and aggregation in heterogeneous wireless networks," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1010–1013, 2016.

[48] Y. Wu, Y. He, L. P. Qian, J. Huang, and X. Shen, "Optimal resource allocations for mobile data offloading via dual-connectivity," *IEEE Trans. Mobile Comput.*, vol. 17, no. 10, pp. 2349–2365, 2018.

[49] Y. Wu, X. Yang, L. P. Qian, H. Zhou, X. Shen, and M. K. Awad, "Optimal dual-connectivity traffic offloading in energy-harvesting small-cell networks," *IEEE Trans. Green Commun. and Netw.*, vol. 2, no. 4, pp. 1041–1058, 2018.

[50] F. Zhang, W. Zhang, and Q. Ling, "Non-cooperative game for capacity offload," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1565–1575, 2012.

[51] S. K. Joshi, K. S. Manosha, M. Codreanu, and M. Latva-aho, "Dynamic inter-operator spectrum sharing via Lyapunov optimization," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6365–6381, 2017.

[52] F. Teng, D. Guo, and M. L. Honig, "Sharing of unlicensed spectrum by strategic operators," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 3, pp. 668–679, 2017.

[53] T. Sakurai and H. L. Vu, "MAC access delay of IEEE 802.11 DCF," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1702–1710, May 2007.

[54] N. Mastronarde and M. van der Schaar, "Joint physical-layer and system-level power management for delay-sensitive wireless communications," *IEEE Trans. Mobile Comput.*, vol. 12, no. 4, pp. 694–709, Apr. 2013.

[55] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, 1st ed. Cambridge, USA: MIT Pr., 1998.

[56] S. Abdallah and M. Kaisers, "Addressing environment non-stationarity by repeating Q-learning updates," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 1582–1612, Apr. 2016.

[57] K. Levy, F. J. Vazquez-Abad, and A. Costa, "Adaptive stepsize selection for online q-learning in a non-stationary environment," in *Proc. Int. Wksh. Discrete Event Sys., Ann Arbor, USA*, July 2006, pp. 372–377.

[58] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surveys (CSUR)*, vol. 46, no. 25, pp. 25–32, Nov. 2013.

[59] R. Ortner, "Pseudometrics for state aggregation in average reward Markov decision processes," in *Proc. Int. Conf. Algorithmic Learning Theory, Sendai, Japan*, Oct. 2007, pp. 373–387.

[60] E. Even-Dar and Y. Mansour, "Approximate equivalence of Markov decision processes," in *Proc. Conf. Learning Theory (COLT), Washington DC, USA*, Aug. 2003, pp. 581–594.

[61] 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.

[62] E. Perahia and R. Stacey, *Next Generation Wireless LANs: 802.11n and 802.11ac*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Pr., 2013.

[63] Y. Inoue, K. Saitoh, T. Sakata, M. Morikura, and H. Matsue, "A study on the rate switching algorithm for IEEE 802.11 wireless LANs," *IEEJ Trans. on Electron., Inf. and Syst.*, vol. 124, no. 1, pp. 33–40, Apr. 2004.

[64] Recommendation ITU-R P.1238-10, "Propagation data and prediction methods for the planning of indoor radiocommunication systems and radio local area networks in the frequency range 300 MHz to 450 GHz," Aug. 2019.

[65] A. Kamerman and L. Monteban, "WaveLAN-II: A high-performance wireless LAN for the unlicensed band," *Bell Labs Technical Journal*, pp. 118–133, Summer 1997.

[66] D. C. Montgomery, G. C. Runger, and N. F. Hubele, *Engineering statistics*. New York: Wiley Pr., 1998.

[67] M. Durvy, O. Dousse, and P. Thiran, "On the fairness of large CSMA networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1093–1104, Sept. 2009.

[68] ——, "Self-organization properties of CSMA/CA systems and their consequences on fairness," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 931–943, Mar. 2009.

[69] C. Hua and R. Zheng, "Starvation modeling and identification in dense 802.11 wireless community networks," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM), Phoenix, USA*, Mar. 2008, pp. 1022–1030.

[70] H. Abeysekera, M. Matsui, Y. Asai, and M. Mizoguchi, "Network controlled frequency channel and bandwidth allocation scheme for IEEE 802.11a/n/ac wireless LANs: RATOP," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun. (PIMRC), Washington, USA*, Sept. 2014, pp. 1041–1045.

[71] S. Lasaulce and H. Tembine, *Game Theory and Learning for Wireless Networks: Fundamentals and Applications*. Academic Pr., 2011.

[72] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, May 1996.

[73] Z. Fang and B. Bensaou, "Fair bandwidth sharing algorithms based on game theory frameworks for wireless ad-hoc networks," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM), Hong Kong, China*, Mar. 2004, pp. 1284–1295.

[74] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y. Yao, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 2, pp. 180–194, Apr. 2012.

[75] X. Chen and J. Huang, "Database-assisted distributed spectrum sharing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 11, pp. 2349–2361, Nov. 2013.

[76] V. P. Mhatre, K. Papagiannaki, and F. Baccelli, "Interference mitigation through power control in high density 802.11 WLANs," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM), Anchorage, USA*, May 2007, pp. 535–543.

[77] K. Yamamoto, "A comprehensive survey of potential game approaches to wireless networks," *IEICE Trans. Commun.*, vol. 98, no. 9, pp. 1804–1823, Sept. 2015.

[78] T. Ui, "A Shapley value representation of potential games," *Games Econ. Behav.*, vol. 31, no. 1, pp. 121–135, April 2000.

[79] H. P. Young, *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Pr., 2001.

[80] C. Surachai, E. Hossain, and J. Diamond, "Channel assignment schemes for infrastructure-based 802.11 WLANs: A survey," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 1, pp. 124–136, Feb. 2010.

[81] G. Ghatak, A. De Domenico, and M. Coupechoux, "Coverage analysis and load balancing in HetNets with millimeter wave multi-RAT small cells," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3154–3169, 2018.

[82] C. K. Ho, D. Yuan, L. Lei, and S. Sun, "Power and load coupling in cellular networks for energy optimization," *IEEE Trans. Wireless Communs.*, vol. 14, no. 1, pp. 509–519, 2014.

[83] G. Barlacchi, M. De Nadai, R. Larcher, A. Casella, C. Chitic, G. Torrisi, F. Antonelli, A. Vespignani, A. Pentland, and B. Lepri, "A multi-source dataset of urban life in the city of milan and the province of Trentino," *Scientific data*, vol. 2, no. 1, pp. 1–15, 2015.

[84] G. P. Koudouridis and P. Soldati, "Trading off network density with frequency spectrum for resource optimization in 5G ultra-dense networks," *Technologies*, vol. 6, no. 4, p. 114, 2018.

[85] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, 2011.

[86] Y. Shen, L. Jiang, C. He, J. Ding, and Q. Xi, "Stochastic geometry based cell load analysis in heterogeneous networks," in *Proc. IEEE Globecom Workshops (GC Wkshps), Washington, DC USA*, May 2016, pp. 1–6.

[87] M. J. Neely, "A Lyapunov optimization approach to repeated stochastic games," in *Proc. IEEE Annu. Allerton Conf. Commun., Control, Comput., Monticello, USA*, Oct. 2013, pp. 1082–1089.

[88] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

Publication List

Journal Papers

1. B. Yin, S. Kamiya, K. Yamamoto, T. Nishio, M. Morikura and H. Abeysekera, "Mitigating throughput starvation in dense WLANs through potential game-based channel selection," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E100-A, no. 11, pp. 2341–2350, Nov. 2017.

2. M. Iwata, K. Yamamoto, B. Yin, T. Nishio, M. Morikura and H. Abeysekera, "Stochastic geometry analysis of individual carrier sense threshold adaptation in IEEE 802.11ax WLANs," *IEEE Access*, vol. 7, pp. 161916–161927, 2019.

3. B. Yin, K. Yamamoto, T. Nishio, M. Morikura and H. Abeysekera, "Learning-based spatial reuse for WLANs with early identification of interfering transmitters," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 151–164, March 2020.

4. B. Yin, K. Yamamoto, S.-L. Kim, T. Nishio and M. Morikura, "Learn to offload via multi-operator shared radio access networks: A cell load perspective," *(submitted to) IEEE.*

International Conference Papers

1. B. Yin, S. Kamiya, K. Yamamoto, T. Nishio, M. Morikura and H. Abeysekera, "Starvation mitigation for dense WLANs through distributed channel selection: Potential game approach," in *Proc. 2017 IEEE Annual Consumer*

Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 2017, pp. 548–553

2. B. Yin, L. Lin, K. Yamamoto, T. Nishio, M. Morikura and H. Abeysekera, "Inversely proportional carrier sense threshold and transmit power setting towards green WLANs," in *Proc. 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, Toronto, ON, Canada, 2017, pp. 1–5

3. B. Yin, K. Yamamoto, S. Kim, T. Nishio and M. Morikura, "Millimeter-wave radio access network sharing: A market-based cooperative bargaining perspective," in *Proc. 2018 IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, 2018, pp. 1–6.

4. M. Iwata, K. Yamamoto, B. Yin, T. Nishio, M. Morikura and H. Abeysekera, "Analysis of inversely proportional carrier sense threshold and transmission power setting based on received power for IEEE 802.11ax," in *Proc. 2019 IEEE 16th Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, 2019, pp. 1–6

5. H. Shimizu, B. Yin, K. Yamamoto, M. Iwata, T. Nishio, M. Morikura and H. Abeysekera, "Joint channel selection and spatial reuse for starvation mitigation in IEEE 802.11ax WLANs," in *Proc. 2019 IEEE 90th Vehicular Technology Conference (VTC-Fall)*, Honolulu, HI, USA, 2019, pp. 1–5

6. S. Kotera, B. Yin, K. Yamamoto, T. Nishio, M. Morikura and H. Abeysekera, "Latency-aware fair scheduling for spatial reuse in WLANs: A Lyapunov optimization approach," in *Proc. 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, 2021, pp. 1–6

Technical Reports and Domestic Conference Papers

1. B. Yin, S. Kamiya, K. Yamamoto and H. Abeysekera, "Potential game based distributed channel selection for starvation mitigation in WLANs," *IEICE Technical Report*, RCS2016-81, pp. 203–208, June 2016.

2. B. Yin, S. Kamiya, K. Yamamoto and H. Abeysekera, "Evaluation of potential game-based channel selection in dense CSMA networks for throughput starvation mitigation," *IEICE Technical Report*, SR2016-79, pp. 5–10, Jan. 2017.

3. B. Yin, L. Lin, K. Yamamoto and H. Abeysekera, "Improving energy efficiency in dense WLANs through inversely proportional setting of carrier sense threshold and transmit power," *IEICE Technical Report*, CQ2017-15, pp. 7–12, May 2017.

4. B. Yin, K. Yamamoto, S.-L. Kim, T. Nishio and M. Morikura, "Millimeter-wave radio access network sharing: A market-based cooperative bargaining perspective," *IEICE Technical Report*, SR2017-119, pp. 49–54, Feb. 2018.

5. B. Yin, S. Kamiya, K. Yamamoto and H. Abeysekera, "Distributed channel selection in dense WLANs for starvation mitigation," *IEICE Society Conference.*, B-11-25, Sapporo, Japan, Sept. 2016.

6. B. Yin, K. Yamamoto, T. Nishio and M. Morikura, "Stochastic geometry analysis of dynamic spectrum sharing in multi-operator shared networks," *IEICE Society Conference.*, B-17-11, Tokyo, Japan, Sept. 2017.

7. B. Yin, K. Yamamoto, T. Nishio and M. Morikura, "Coverage probability analysis of multi-operator shared millimeter-wave radio access network," *IEICE General Conference.*, B-17-6, Tokyo, Japan, Mar. 2018.

8. B. Yin, K. Yamamoto and H. Abeysekera, "An evaluation of increasing CCA threshold of WLANs supporting 40 MHz channel access," *IEICE Society Conference.*, B-5-104, Kanazawa, Japan, Sept. 2018.

9. B. Yin, K. Yamamoto and H. Abeysekera, "Q-learning based concurrent transmissions for WLANs with identification of interfering transmitters," *IEICE General Conference.*, B-5-129, Tokyo, Japan, Mar. 2019.

10. H. Shimizu, B. Yin, K. Yamamoto and H. Abeysekera "Joint channel control and spatial reuse towards starvation mitigation in WLANs," *IEICE Technical Report*, RCS2019-53, pp. 97–102, June 2019.

11. B. Yin, K. Yamamoto, T. Nishio and M. Morikura, "Inter-operator mmWave base station sharing: A user offloading game," *IEICE Society Conference.*, B-5-74, Osaka, Japan, Sept. 2019.

12. B. Yin, K. Yamamoto, T. Nishio and M. Morikura, "Mobile traffic offloading in shared spectrum: Price of incomplete traffic information," *IEICE General Conference.*, B-5-40, Higashihiroshima, Japan, Mar. 2020.

Awards

1. IEEE VTS Japan Young Researcher's Encouragement Award (2017)
2. IEEE Kansai Section Student Research Encouragement Award (2019)
3. The Telecommunications Advancement Foundation Telecom Systems Students Technical Award (2019)
4. IEEE Wireless Communications Letters Exemplary Reviewer (2020)

Patents

1. JP6579447