This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

Comprehensive Study on MIMO-related Interference Management in WLANs

Junmei Yao, Jun Xu, Sheng Luo, Lu Wang, Chao Yang, Kaishun Wu and Wei Lou

***Abstract*—With the explosive growth of smart devices and mobile data traffic, the wireless local area network (WLAN) faces persistent challenge on increasing the network throughput to meet the customer's requirement. Multiple-Input and Multiple-Output (MIMO) is a promising technology to improve the network capacity and has been adopted to WLANs in current IEEE 802.11 standards. Interference management for MIMO-based WLANs has become a hot topic in recent years due to its huge potential in improving the network performance. This paper surveys the MIMO-related interference management mechanisms from both the single collision domain and multiple collision domain scenarios. It first provides some background information about the evolution of IEEE 802.11 standards, then introduces the MIMO-related physical layer and MAC layer mechanisms as preliminaries. After this, it investigates the current advances on MIMO-related interference management in both the single collision domain and multiple collision domain through the physical layer and MAC layer mechanism design. It also provides discussions on important findings and research challenges, and finally gives potential future research directions.**

***Index Terms*—Keywords: MIMO; Interference Management; WLANs; IEEE 802.11.**

I. Introduction

The rapid development of smart devices and mobile communications results in the explosive growth of mobile data traffic, which cannot be supported only by the cellular networks. As pointed out by the Cisco Visual Networking Index [1] (VNI), about 50% of the traffic will be offloaded to the wireless local area networks (WLANs, often called Wi-Fi networks). To combat the challenge of supporting the huge data traffic in WLANs, the IEEE task group has made great efforts on throughput improvement in WLANs through the evolution of the 802.11 standards. From 1997 when the base version of IEEE 802.11 standard which supports $1/2Mbps$ data rate was released, a series of 802.11 standards have been released to increase the physical layer data rate. 802.11b [2] and 802.11a/g [3] increase the data rate of single stream up to $11Mbps$ and $54Mbps$ respectively, through exploiting higher-order modulations. 802.11n [4] and 802.11ac [5] further increase the data rate up to $600Mbps$ and $> 6Gbps$ respectively, through exploiting the Multiple-Input and Multiple-Output (MIMO) technology. The adoption of MIMO has led the WLANs into gigabit era.

A typical WLAN contains multiple access points (APs) and a large number of stations, as shown in Fig. 1. Each AP serves a set of mobile stations, which access the Internet or the server through AP. The set of stations associated with each AP is called a basic service set (BSS). The transmission of one node (an AP or station) may interfere with the other nodes' transmissions within one BSS, which is called the *single collision domain*. Meanwhile, there are always multiple BSSs in a WLAN and the signal transmission in one BSS may interfere with that in another, we call the multiple BSSs as the *multiple collision domain*. The 802.11 standard recommends the Distributed Coordination Function (DCF), which adopts carrier sense multiple access (CSMA[1]) mechanism to avoid interference among nodes in both the single and multiple collision domains. According to this mechanism, a node can transmit a signal only when the channel is determined to be idle. In the traditional WLANs with single streams, CSMA is well-known to be inefficient due to failing to avoid interference effectively or prohibiting concurrent transmissions, leading to the everlasting research on interference management in the past twenty years [6]–[12]. When MIMO is introduced to WLANs by the 802.11 standard, the MIMO transmission is limited to the single collision domain, while nodes still adopt CSMA to access the channel.

The MIMO technology has a long history in the wireless communication area, although it was brought to the WLAN context as a part of standard only in recent years. It utilizes multiple transmit and receive antennas to increase wireless capacity through exploiting the multiple path propagation. *Beamforming*, a MIMO physical layer technique, is adopted to precode the transmit signal, so that the transmit energy can be focused towards the intended receiver to improve the signal reception [5]. Both 802.11n and 802.11ac recommend Single-User MIMO (SU-MIMO) which makes an access point (AP) communicate with one station at each time through the multiple

Junmei Yao, Jun Xu, Sheng Luo and Lu Wang are with the College of Computer Science and Software Engineering, Shenzhen University, P. R. China. Email: {yaojunmei, xujun, sluo, wanglu}@szu.edu.cn.

Chao Yang is with the School of Automation, Guangdong University of Technology, Guangzhou, and the Department of Computing, Hong Kong Polytechnic University, Hong Kong, China. E-mail: yangchaoscut@aliyun.com.

Kaishun Wu is with the College of Computer Science and Software Engineering, Shenzhen University, P. R. China, and PCL Research Center of Networks and Communications, Peng Cheng Laboratory, Shenzhen, P. R. China. Email: wu@szu.edu.cn.

Wei Lou is with the Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong. Email: csweilou@comp.polyu.edu.hk.

[1]CSMA in this paper represents CSMA/CA (carrier sense multiple access with collision avoidance).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

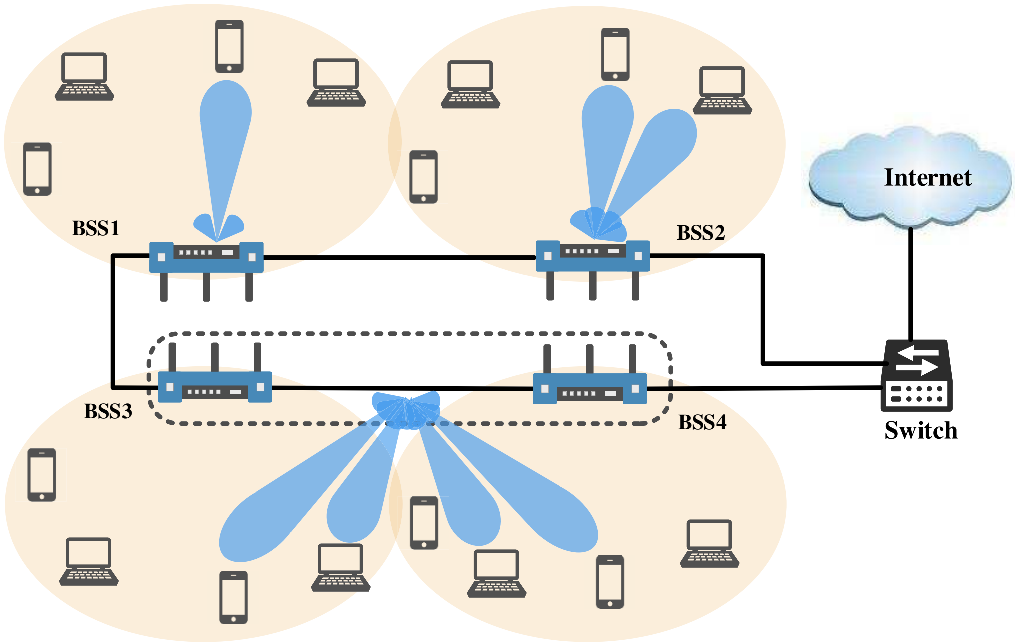


Fig. 1. An example of WLAN topology.

antennas (as shown in BSS1 of Fig. 1). 802.11ac further recommends downlink Multi-User MIMO (MU-MIMO) which makes an AP communicate with multiple stations simultaneously to improve the system capacity (as shown in BSS2 of Fig. 1); for this end, another beamforming technique called null steering is adopted to precode the transmit signal, thus to avoid the inter-user interference [5].

The implementation of MIMO transmission depends on two critical mechanisms: *channel sounding* and *user selection*. Channel sounding lets nodes obtain the accurate channel state information (CSI), which is essential for both SU-MIMO and MU-MIMO transmissions, as nodes need this information to precode the transmit signal and allow the signal to arrive at the intended receivers in designated manners. User selection plays an important role in MU-MIMO transmission since the transmitter should choose the stations with the most uncorrelated channels to be receivers, thus to minimize the inter-user interference.

The MIMO transmission relies on the interference management mechanisms in the MAC layer to achieve high network performance. As shown in Fig. 2, researchers have made great efforts on this area to optimize the performance of WLANs. In the single collision domain, previous research includes improving the efficiency of the distributed CSMA mechanism and mitigating the inter-user interference in MU-MIMO transmission, while the latter further consists of reducing the huge overhead of channel sounding and designing efficient user selection process. In the multiple collision domain, current advances focus on interference management through cross layer design: two advanced MIMO physical layer techniques, *interference alignment* and *interference nulling*, are always exploited to enable concurrent transmissions and avoid interference among BSSs; in addition, besides SU-MIMO and MU-MIMO, another MIMO application scenario, network MIMO which makes multiple APs act as a giant AP and communicate with multiple stations simultaneously (as shown in BSS3 and BSS4 of Fig. 1), is also adopted by some researchers to improve the network performance; finally, MAC layer schemes are proposed to make the physical layer mechanisms work well in the networks. This paper will study the current MIMO-related interference management mechanisms in WLANs from both collision domains.

The key point of this survey is to investigate how the MIMO physical layer techniques are exploited and what the corresponding MAC layer protocols are proposed in WLANs to improve the network performance. The main contribution of this paper is twofold. First, we conclude the evolution of IEEE 802.11 standards, and report the current MIMO-related physical and MAC layer mechanisms in WLANs, and also the MIMO application scenarios. Second, we present the literatures on MIMO-related interference management mechanisms in WLANs from both the single collision domain and the multiple collision domain. We also identify the important findings and research challenges after the literature review, and give our thoughts about possible future research directions in this area.

There are several survey papers related to this work. Authors in [13] presented the existing protocols on interference management for ad hoc wireless networks, but only focused on the single stream systems. Authors in [14] surveyed four technologies which can be applied for high efficient WLANs, including the PHY layer techniques, MAC layer strategies, spatial frequency reuse schemes and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

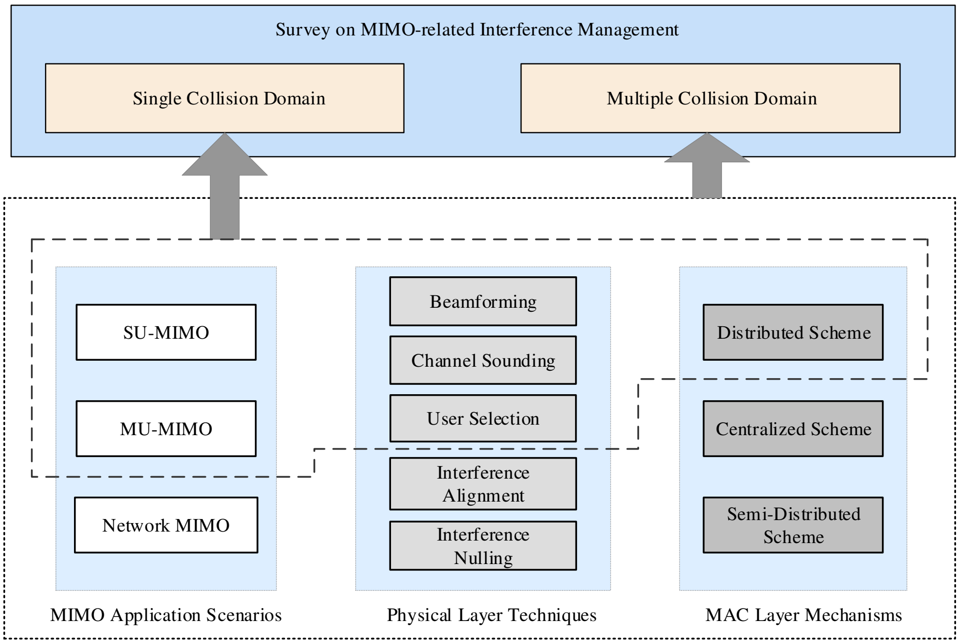


Fig. 2. The basic categorization of this survey.

TABLE I
ABBREVIATIONS.

Access Point (AP)	Acknowledgement (ACK)	Basic Service Set (BSS)
Beamforming Report Poll (BRP)	Block Acknowledgement Request (BAR)	Binary Phase Shift Keying (BPSK)
Carrier Sense Multiple Access (CSMA)	Channel State Information (CSI)	Channel State Information at Transmitter (CSIT)
Channel State Information at Receiver (CSIR)	Clear-To-Send (CTS)	DCF Inter Frame Space (DIFS)
Degree of Freedom (DoF)	Direct Sequence Spread Spectrum (DSSS)	Distributed Coordination Function (DCF)
Frequency Hopping Spread Spectrum (FHSS)	Medium Access Control (MAC)	Microsoft Research Software Radio (Sora)
Minimum Mean Square Estimation (MMSE)	Multiple Input Multiple Output (MIMO)	Multi-User MIMO (MU-MIMO)
Network Allocation Vector (NAV)	Null Data Packet (NDP)	Orthogonal Frequency Division Multiplexing (OFDM)
Orthogonal Frequency Division Multiple Access (OFDMA)	Physical Layer (PHY)	Point Coordination Function (PCF)
Quadrature Amplitude Modulation (QAM)	Quadrature Phase Shift Keying (QPSK)	Request-To-Send (RTS)
Single-User MIMO (SU-MIMO)	Short Inter Frame Space (SIFS)	Signal to Interference and Noise Ratio (SINR)
Successive Interference Cancellation (SIC)	Time Slot (TS)	Universal Software Radio Peripheral (USRP)
Wireless Local Area Networks (WLAN)	Wireless Open Access Research Platform (WARP)	Zero Forcing (ZF)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

TABLE II
IEEE 802.11 PHY EVOLUTION.

Version	Release Time	Maximum Data Rate	Frequency Band (GHz)	Bandwidth (MHz)	Modulation	MIMO
802.11-base version	Jun. 1997	2Mbps	2.4	20	BPSK, QPSK DSSS, FHSS	/
802.11b	Sep. 1999	11Mbps	2.4	20	BPSK, QPSK DSSS (CCK)	/
802.11a	Sep. 1999	54Mbps	5	20	BPSK, QPSK, 16-QAM, 64-QAM, OFDM	/
802.11g	Jun. 2003	54Mbps	2.4	20	BPSK, QPSK, 16-QAM, 64-QAM, OFDM, DSSS	/
802.11n	Oct. 2009	600Mbps	2.4 & 5	20, 40	BPSK, QPSK, 16-QAM, 64-QAM, OFDM	SU-MIMO
802.11ac	Dec. 2013	6.933Gbps	2.4 & 5	20, 40, 80, 160	BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM, OFDM	SU-MIMO Downlink MU-MIMO
802.11ax	Approx. 2019	> 10Gbps	< 6	----	BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM, 1024-QAM, OFDM, OFDMA	SU-MIMO Downlink/Uplink MU-MIMO

power saving mechanisms. Authors in [15] investigated some novel technologies in the future WLAN standards, such as machine-to-machine communication in 802.11ah; it just payed a little attention on high throughput WLAN recommended by 802.11ac. Authors in [16] discussed some key issues in large-scale MIMO (LS-MIMO), such as practical channel models application scenarios, but their investigation focused mainly on the cellular networks. In [17], authors surveyed the recent works on applying MU-MIMO in WLANs, containing the key technologies in the physical layer and the assisted protocols in the MAC layer; in [18], authors investigated the transmission methods and efficiency using MIMO in WLANs, and then introduced the MAC protocol designs in this kind of systems; however, these two works only consider the single collision domain. So far, there is no comprehensive survey focusing on MIMO-related interference management in WLANs, especially for the multiple collision domain.

The rest of this paper is organized as follows. Section II gives an overview of the IEEE 802.11 family, including the MAC and PHY layer functions. Section III introduces the MIMO application scenarios and MIMO physical layer techniques. Section IV and Section V investigate related works on improving the efficiency of MIMO transmission in the single collision domain and multiple collision domain respectively, then give some important findings. Section VI puts some future directions. Section VII concludes this paper.

As a great number of abbreviations and technical terms will be used in this paper, Table I summarizes some important terms in alphabetical order for the ease of understanding.

II. OVERVIEW OF THE 802.11 STANDARD

This section briefly introduces the standards related to the throughput improvement in the IEEE 802.11 family, from the PHY and MAC layer functions.

A. IEEE 802.11 PHY

As shown in Table II, the 802.11 base version was released in 1997, which specified two data rates of $1/2Mbps$ working in the $2.4GHz$ frequency band, and using the Direct Sequence Spread Spectrum (DSSS) and Frequency Hopping Spread Spectrum (FHSS) as the modulation types. After that, they released a series of 802.11 standards to increase the physical layer data rate [3]. In 1999, 802.11a, which operates in the 5GHz frequency band and uses the Orthogonal Frequency Division Multiplexing (OFDM) as the modulation type, was introduced to enhance the data rate to $54Mbps$. In the same year, 802.11b, which operates in the same frequency band and uses the same modulation technique with the base version, was also introduced to support up to $11Mbps$ data rate. Due to its dramatic throughput increase and the similar PHY technique compared with the base version, the 802.11b became the definitive WLAN technology. In 2003, the IEEE task group released 802.11g, which uses OFDM as its modulation technique (the same as 802.11a), and finally achieves up to $54Mbps$ data rate in the 2.4 GHz band. It is regarded as an extension of 802.11b.

Utilizing the MIMO technology substantially improves the spectral efficiency, and finally improves the physical layer data rate of WLANs. In 2009, 802.11n, the first standard that supports MIMO was ratified to achieve up to $600Mbps$ data rate through four spatial streams and 40MHz wider bandwidth; it recommends only Single-User

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

MIMO (SU-MIMO) which makes an AP to communicate with one station at each time through the multiple antennas and allows up to 4×4 SU-MIMO transmission. The latest released version is 802.11ac [5], which increases the data rate to $> 6Gbps$ through 160MHz wider bandwidth, up to 8×8 SU-MIMO transmission, and higher-order 256-QAM modulation; it also recommends downlink Multi-User MIMO (MU-MIMO) which makes one AP communicate with up to four stations simultaneously in the downlink. The upcoming 802.11ax, which is currently in an early stage of development, is predicted to achieve $> 10Gbps$ data rate, through further introducing the Orthogonal Frequency Division Multiple Access (OFDMA) technique and the higher order 1024-QAM modulation.

B. IEEE 802.11 MAC

Although the 802.11 PHY has been updated to higher order modulations, wider bandwidth and MIMO to increase the physical layer data rate, the MAC layer process, which was first proposed in the 802.11 base version [2] to manage interference in wireless networks, only had a few modifications accordingly.

The basic channel access process recommended by 802.11 MAC has barely changed during the update of the 802.11 standard. The 802.11 recommends two kinds of coordination functions in the MAC layer to manage interference: a distributed mechanism called Distributed Coordination Function (DCF) and a centralized mechanism called Point Coordination Function (PCF). PCF lets the AP schedule both the downlink and uplink transmissions in a centralized way. It is not implemented in most devices due to its application limit in the multiple collision domain, and is even not part of the Wi-Fi Alliance's standard. DCF further has two mechanisms, the carrier sense multiple access (CSMA) mechanism which allows a transmitter to transmit its signal only when it senses the channel is idle, and RTS/CTS mechanism which is based on CSMA but further makes the transmitter and receiver exchange the Request-To-Send (RTS) and Clear-To-Send (CTS) frames to reserve the channel for the following data transmission. CSMA is the widely deployed MAC protocol in current WLANs, while RTS/CTS is disabled in most cases due to its high transmission overhead induced by the RTS and CTS transmissions.

Fig. 3 gives the detailed processes of CSMA. Before transmitting, a sender first senses the medium to determine whether the channel is available; if the channel is determined to be busy when a nearby node is transmitting a signal, the sender should defer its transmission until the channel is available again to avoid interference. When the channel is idle for a DIFS (distributed interframe space) period, the node will wait for an additional backoff period before transmitting. The backoff period is randomly generated and is decremented along with the time when the channel stays idle and is suspended otherwise. Only when the backoff period expires, the sender will initiate the data transmission. When SU-MIMO and MU-MIMO

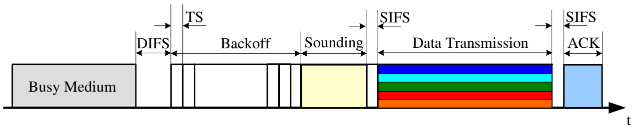


Fig. 3. The channel access processes for MIMO transmission recommended by the 802.11n/ac standard.

are introduced to WLANs through 802.11n/ac, nodes still adopt CSMA to avoid interference among the multiple transmissions; the channel sounding period is added to obtain CSI for the following data transmission through MIMO. Once receiving the data packet correctly, each receiver should reply an acknowledgement (ACK), so that the sender can confirm the successful data transmission; otherwise, the sender will retransmit the packets. Although multiple data packets can be transmitted simultaneously through MU-MIMO in the downlink direction, the ACKs must be transmitted individually in the uplink direction to avoid interference [5].

III. PHY Mechanisms for MIMO Transmissions

In this part, we first present three typical MIMO application scenarios, then give some common MIMO precoding and decoding processes as preliminary. After that, we introduce both the basic MIMO PHY techniques adopted by the 802.11 standard, and the advanced MIMO PHY techniques proposed by researchers for further interference management.

A. MIMO Application Scenarios

According to whether the communication is point to point, point to multi-point, or multi-point to multi-point, the MIMO technology has three basic application scenarios in WLANs: SU-MIMO, MU-MIMO, and the network MIMO. Both SU-MIMO and downlink MU-MIMO have been adopted in the 802.11 standard, while uplink MU-MIMO and network MIMO are now open research issues in WLANs. We will give brief description of each scenario through examples.

1) SU-MIMO: SU-MIMO makes an AP communicate with only one station at each time through multiple antennas, which can be used for two purposes. Under the high SNR situations, the antennas are always used to transmit different data packets to increase the data rate, that is to improve the *spatial multiplexing*. Under the low SNR situations, they are always used to transmit the same packet to increase the SNR at the receiver side, that is to improve the *receiver diversity*.

The performance of SU-MIMO is limited by the number of antennas at both the transmitter and receiver sides. In practical WLANs, the AP often has much more antennas than smart devices due to the device's constraints on portability concern, energy consumption and so on, resulting in poor performance due to under utilization of the AP antennas. This problem impels the 802.11ac task group to bring into MU-MIMO.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

TABLE III
THE MIMO PRECODING/DECODING SCHEMES WITH KNOWN CSI.

Scheme	Implementation	Scenario	Main Feature	Complexity
Singular Vector Decomposition (SVD)	Combination of precoding and decoding	SU-MIMO	Linear scheme, achieve theoretical capacity	High
Zero Forcing (ZF)	Precoding or Decoding	SU-MIMO Downlink/Uplink MU-MIMO	Linear scheme, cancel the inter-stream interference	Low
Minimum Mean Square Error (MMSE)	Precoding or Decoding	SU-MIMO Downlink/Uplink MU-MIMO	Linear scheme, cancel the inter-stream interference while consider noise	Medium
Successive Interference Cancellation (SIC)	Decoding	SU-MIMO Uplink MU-MIMO	Nonlinear scheme, use together with ZF and MMSE	High
Maximum Likelihood (ML)	Decoding	SU-MIMO Uplink MU-MIMO	Nonlinear scheme, exhaustive searching	Very high

2) MU-MIMO: MU-MIMO allows an AP to communicate with multiple stations at each time, so as to fully utilize AP's antennas. The main challenge in the MU-MIMO transmission is the inter-user interference, which occurs because the receivers of the MU-MIMO transmissions are not sufficiently separated. Therefore, in the downlink the AP should select a group of users (stations) as the receivers and precode the transmit signal to null the inter-user interference at the transmitter side, in the uplink the AP should decode and separate the signals transmitted simultaneously from stations for demodulation.

3) Network MIMO: Network MIMO is also called virtual MIMO, distributed MIMO or cooperative MIMO. Its basic idea is to group multiple individual nodes into an antenna array for MIMO transmission. It has already had thorough study in the cellular network scenario [19]. In the WLAN context, network MIMO always represents treating multiple APs as a virtual AP and exploiting all their antennas to serve multiple stations simultaneously. It is just like an extension of MU-MIMO to achieve the communication between a giant AP and the stations. The network MIMO is mostly regarded as a theoretical solution due to some serious challenges in the implementation of this system, such as the ability to achieve tight phase and timing synchronization among multiple APs to make joint precoding, and also its huge computational complexity. However, in recent years, some pioneers have already been prepared in combating these challenges and implementing the network MIMO systems, which will be discussed later in Section V.

B. MIMO Precoding/Decoding Processes

As shown in Fig. 4, MIMO transmission relies on precoding at the transmitter side and decoding at the receiver side to achieve the performance improvement. Here we would like to give some well-known precoding and decoding schemes, including their advantages and limitations, as included in Table III. Although there are substantial research works on these schemes [20]–[22], we do not want to give deep analysis as they are out of the scope of this paper.

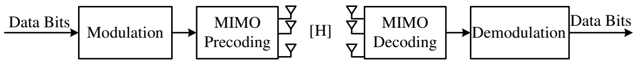


Fig. 4. The simplified PHY layer processes with MIMO in wireless communication systems.

We consider the complex baseband model for MIMO transmission:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{H}$ is the $N_r \times N_t$ channel matrix [2], also called the channel state information (CSI). N_t and N_r represent the number of transmit and receive antennas, respectively. $\mathbf{x} = [x_1, x_2, \cdots, x_M]^T$ is a vector of the M data streams that will be transmitted from the N_t antennas, $\mathbf{y} = [y_1, y_2, \cdots, y_{N_r}]^T$ is a vector of the received signal from the N_r antennas, $\mathbf{n}$ is the background noise. Please note that the vectors $\mathbf{x}$ and $\mathbf{y}$ may include signals from multiple nodes in the MU-MIMO and network MIMO scenarios.

1) Singular Vector Decomposition (SVD): SVD is a method to decompose a MIMO channel into multiple parallel subchannels. Mathematically, SVD decomposes the channel matrix $\mathbf{H}$ as:

$$\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H, \tag{2}$$

[2]In WLANs where OFDM is adopted, the channel is highly frequency-dependent and $\mathbf{H}$ may have different values among the subcarriers. Here we just use a simplified model.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices with $N_r \times N_r$ and $N_t \times N_t$ sizes, $\mathbf{V}^H$ is the conjugate transpose of $\mathbf{V}$, Σ is the $N_r \times N_t$ rectangular diagonal matrix with positive numbers in descending order on the diagonal.

At the transmitter side, the transmitter left-multiplies $\mathbf{x}$ by the matrix $\mathbf{V}$, then the received signal is represented as:

$$\mathbf{y} = \mathbf{U}\Sigma\mathbf{V}^H \cdot \mathbf{V}\mathbf{x} + \mathbf{n} = \mathbf{U}\Sigma\mathbf{x} + \mathbf{n}. \tag{3}$$

At the receive side, the receiver left-multiplies $\mathbf{y}$ by $\mathbf{U}^H$ to decode the signal, that is:

$$\mathbf{U}^H\mathbf{y} = \mathbf{U}^H(\mathbf{U}\Sigma\mathbf{x} + \mathbf{n}) = \Sigma\mathbf{x} + \mathbf{U}^H\mathbf{n} \approx \Sigma\mathbf{x}. \tag{4}$$

The number of parallel data streams M should be no more than $min\{N_t, N_r\}$, and the first M columns of $\mathbf{V}$ are selected as the steering matrix.

SVD is not suitable for MU-MIMO or network MIMO scenarios, as in these situations either the transmitter or receiver could not acquire the whole information of $\mathbf{H}$. In the SU-MIMO scenario, SVD is well-known to achieve theoretical channel capacity together with water filling for power allocation [23]. However, this scheme depends highly on the accuracy of CSI, even a small error in the CSI estimation may deteriorate the channel performance significantly. Thus, SVD is rarely deployed in practical systems and is always used to derive the theoretical capacity for the linear schemes.

2) Zero-Forcing (ZF): ZF is a simple method to eliminate the interference among multiple concurrent data streams through making them mutually orthogonal [24].

Consider the communication system $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$. When ZF is performed at the receiver side to allow no interference among the streams, it left-multiplies $\mathbf{y}$ with a matrix $\mathbf{V}$, that is:

$$\mathbf{V}\mathbf{y} = \mathbf{V}\mathbf{H}\mathbf{x} + \mathbf{V}\mathbf{n} \approx \mathbf{x}, \tag{5}$$

here $\mathbf{V}$ is the pseudo-inverse of $\mathbf{H}$ and $\mathbf{V} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$.

When ZF is performed at the transmitter side through precoding, the steering matrix $\mathbf{V} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$ is left-multiplied with the transmit signal $\mathbf{x}$. Then the received signal is:

$$\mathbf{y} = \mathbf{H}\mathbf{V}\mathbf{x} + \mathbf{n} \approx \mathbf{x}. \tag{6}$$

The ZF scheme is easy to implement and can be utilized in all the MIMO application scenarios to enable concurrent data stream transmissions. Especially, 802.11ac adopts ZF to cancel the inter-user interference for the MU-MIMO transmission in WLANs. However, its major problem is that it only works well under high SNR situations, as the least square solution obtained through Eq. (5) or Eq. (6) would reduce the signal strength of each component [3]; meanwhile, the background noise $\mathbf{n}$ may be amplified when the channel matrix $\mathbf{H}$ is ill-conditioned. The performance of ZF will be very poor when SNR is low.

[3]This point is clearly illustrated in Fig. 6(b), where the received signal strength of x_2 after ZF is $|\mathbf{H}_{proj}\mathbf{H}_2 x_2|$, which is much smaller than that without ZF ($|\mathbf{H}_2 x_2|$). The x_2's received SNR is reduced in this way.

3) Minimum mean square error (MMSE): Distinct from ZF which does not consider the noise and simply eliminate the inter-user interference, the MMSE scheme aims to minimize the overall error rate through adjusting the matrix $\mathbf{V}$ with the noise. That is, $\mathbf{V} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I})^{-1}$. This matrix is similar with that of ZF but is adjusted by an identify matrix $\mathbf{I}$ weighted with the factor α that is determined by the noise $\mathbf{n}$.

With the steering matrix $\mathbf{V}$, the implementation of MMSE for both the MIMO precoding and decoding has the similar processes as ZF, referring to Eq. (5) and Eq. (6). It is well-known that MMSE can achieve better performance than ZF under low SNR environments and can approach the performance of ZF under high SNR environments, but with higher complexity due to the noise estimation [24].

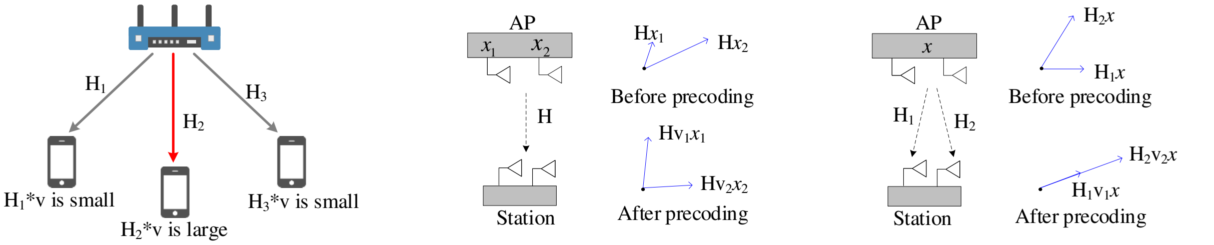
4) Successive Interference Cancellation (SIC): The concept of SIC is to make a receiver separate parallel signals through exploiting the power characteristics. The receiver first detects the signal with highest power, then subtracts it from the received mixed signals to recover the signal with inferior power. SIC could be used together with both ZF and MMSE to improve the MIMO system capacity. For example, according to Eq. 5, the receiver obtains the estimated transmitted signal $\hat{\mathbf{x}} = [x_1, x_2, \cdots, x_M]^T$, it subtracts the effect of x_i with the highest power from $\mathbf{y}$, then re-estimates the inferior signal x_j from the remaining signal through maximum ratio combining (MRC). The process is conducted iteratively until the signal with the lowest power is detected. The main problem of this scheme is the error propagation in the iterative decoding processes [25].

5) Maximum Likelihood (ML): ML is a kind of nonlinear MIMO decoding scheme that obtains the transmitted signals through exhaustive searching. With the channel matrix $\mathbf{H}$, the receiver feeds all possible transmitted data streams $\hat{\mathbf{x}}$ into this channel, then compares the output signals $\hat{\mathbf{y}} = \mathbf{H}\hat{\mathbf{x}}$ with the actual received signal $\mathbf{y}$ to determine the most likely one. This scheme has a high computational complexity which is exponential with the number of data streams, making it hard to be deployed in practical systems. Sphere decoding is a faster algorithm compared to the natural ML decoding [26], it reduces the complexity of ML through limiting the search space. These nonlinear schemes can achieve optimal performance compared with the linear ones, but at the cost of much higher complexity.

C. Basic 802.11 MIMO PHY Techniques

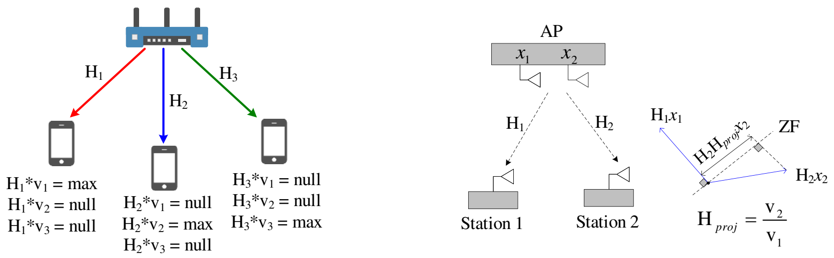
The 802.11n/ac recommend SU-MIMO transmission, while 802.11ac further supports downlink MU-MIMO transmission. Both of them adopt *beamforming* to improve the signal reception at the receiver. In this part, we first describe how beamforming can achieve the SU-MIMO and MU-MIMO transmissions in WLANs, then introduce two important mechanisms to implement beamforming: channel sounding to acquire the channel matrix $\mathbf{H}$, and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials



(a) Effects of beamforming in SU-MIMO (b) Precoding to achieve spatial multiplexing (c) Precoding to increase receive diversity

Fig. 5. An example of beamforming in the SU-MIMO scenario.



(a) Effects of beamforming in MU-MIMO (b) Precoding to cancel inter-user interference

Fig. 6. An example of beamforming in the MU-MIMO scenario.

the user selection to mitigate inter-user interference in the MU-MIMO scenario.

1) Beamforming: Beamforming refers in particular to a technique in which the transmitter exploits the previously known CSI to precode its transmit signal, so that the transmitter can focus its transmit energy towards the intended receiver to improve the signal reception. With beamforming, the transmitter left-multiplies the transmit signal $\mathbf{x}$ with a steering matrix $\mathbf{V}$, so that the received signal $\mathbf{y} = \mathbf{HVx}+\mathbf{n}$ can achieve an enhanced performance. As almost all the surveyed papers use vectors to illustrate their works, in this part we also utilize vector to describe how beamforming can achieve the SU-MIMO and MU-MIMO transmissions.

i) Single-User Beamforming

Single-user beamforming makes the transmitted signals from each antenna reenforce at the preferred receiver to become stronger, as depicted in Fig. 5(a). Fig. 5(b) illustrates an example on how beamforming can improve the spatial multiplexing when the parallel data streams x_1 and x_2 are transmitted by AP simultaneously. Actually, without precoding, the receiver is also able to detect x_1 and x_2 through solving the linear equations $\mathbf{y} = \mathbf{Hx}$. However, the performance in this situation may be very poor due to the unexpected receiving signals, e.g., their intersection angle could be relatively small or the SNR is poor. Precoding improves the system performance through letting the two signals easier to be detected.

Fig. 5(c) demonstrates how beamforming can enhance the receiver diversity when a single signal x is transmitted by the AP through the two antennas. We let $\mathbf{H}_i$ indicate the channel matrix from AP to the ith antenna of the station. Without precoding, the station gets a signal $\mathbf{H}_1 x + \mathbf{H}_2 x$ which may not be strong enough for detection. As shown in Fig. 5(c), after precoding, the two received signals $\mathbf{H}_1\mathbf{v}_1 x$ and $\mathbf{H}_2\mathbf{v}_2 x$ constructively add up towards the same direction ideally. This condition can be satisfied, for example, through making $(\mathbf{v}_1, \mathbf{v}_2) = (\mathbf{H}_2^*, \mathbf{H}_1^*)$. The signal strength now is obviously much stronger than that before precoding.

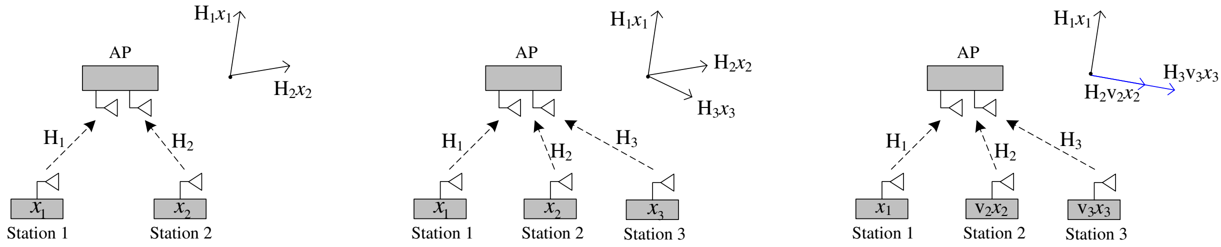
ii) Multi-User Beamforming

Multi-User Beamforming allows an AP to transmit data packets to multiple receivers simultaneously. The main challenge of applying MU-MIMO is to solve the problem of inter-user interference, which occurs because two receivers of the MU-MIMO transmissions are not sufficiently separated. To avoid this interference, ZF, also called *null steering* in 802.11ac [5], is exploited to derive the steering matrix for the AP, thus produces a high signal at this station but nulls the signal at other stations. As shown in Fig. 6(a), in the ideal scenario, a set of steering vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ will make the transmit signal x_i arrive at station i with high power, while that at other stations with no power. However, in real network scenarios, the unintended signal cannot be completely nulled out and it is always left with small power.

Fig. 6(b) shows an example on how null steering can avoid the inter-user interference between two stations through the vector diagram. The vector of H_2x_2 (the signal will be transmitted to station 2), is projected onto the space orthogonal to its interferer $\mathbf{H}_1 x_1$ (the signal will be transmitted to station 1), while the projection is represented as $\mathbf{H}_2\mathbf{H}_{proj}x_2$. As the two vectors $\mathbf{H}_1 x_1$ and $\mathbf{H}_2\mathbf{H}_{proj}x_2$ are orthogonal to each other, they can be transmitted concurrently without interference.

The implementation of beamforming depends upon CSI which is obtained through the so-called channel sounding

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials



(a) x_1 can be decoded through solving the two linear equations

(b) No signal can be decoded when client 3 joins in

(c) x_1 can be decoded when applying interference alignment

Fig. 7. An example of interference alignment, where $\mathbf{H}_i$ indicates the channel matrix from station i to AP. When Station 1 transmits a signal x_1 to AP, this signal can be decoded if there is one interferer (Station 2) through the standard MIMO decoding process as AP only has two antennas. Interference alignment makes its decoding successfully even if multiple interferers join in.

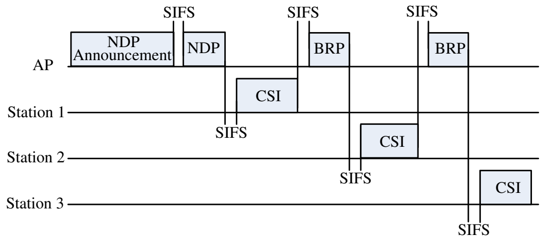
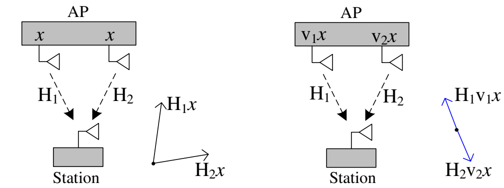


Fig. 8. The explicit channel sounding process recommended by the 802.11ac standard.



(a) Before precoding

(b) After interference nulling

Fig. 9. An example of interference nulling, signals from the AP are canceled at the station in (b).

process in WLANs. Meanwhile, how to select the best set of users for MU-MIMO transmission is also very important to achieve the optimal performance. In the following parts we introduce some detailed descriptions of these two mechanisms according to the 802.11 standard.

2) Channel Sounding: The 802.11n/ac recommend channel sounding to obtain CSI for the transmitter, the CSI will be used for both the SU-MIMO and MU-MIMO precoding in data transmission and the user selection in MU-MIMO.

In general, there are two channel sounding methods for MIMO transmission, the implicit method which relies on channel reciprocity feature and uses the CSI at the transmitter to estimate that at the receiver, and the explicit method which requires the receiver to estimate the channel and feedback the CSI to the transmitter in time. The implicit method has channel estimation error as the channel condition at the receiver is always different from that at the transmitter, leading to non-effective interference management [27], [28]. The explicit method provides more accurate channel information to the transmitter, but may induce high transmission overhead. While 802.11n supports both kinds of methods, the 802.11ac only recommends the explicit method as the imperfect channel estimation is less tolerable in MU-MIMO than in the SU-MIMO transmission, due to the induced harmful inter-user interference.

The explicit channel sounding process relies on the MAC layer design to achieve the frame exchange between the transmitter and receivers. As depicted in Fig. 8, at the beginning of the process, AP transmits an NDP (Null Data Packet) Announcement frame, which contains the information of stations that need to feedback their CSI. It then transmits an NDP frame, which will be used by the receiver to estimate the channel. In the following time, for SU-MIMO, only one station needs to report its CSI; for MU-MIMO, the stations will report sequentially according to the Beamforming Report Poll (BRP) frame.

3) User Selection: The user selection process plays a very important role in the MU-MIMO transmissions, although it is not specified by the 802.11ac standard. As described in Fig. 6(b), if the angle between the two vectors is very small, the projection will become very small as well, making the SINR of x_2 decreased dramatically. In this situation, it is essential to select the users with the most uncorrelated channels to increase the performance of MU-MIMO.

Intuitively, the optimal users can be determined through exhaustive searching over all the possible users and antenna sets. However, its computational complexity increases exponentially with the number of users and antennas, making it infeasible to be deployed in practical networks. A well-known practical user selection scheme is the greedy algorithm [29], which lets an AP first select one user with the best channel condition, then select another user which results in the highest aggregate capacity. The users are selected until the highest sum capacity is achieved. This algorithm is easy to implement and with low complexity, but has suboptimal performance, e.g., in some situations only 50 $\sim$ 60% optimal capacity can be achieved [30].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

D. Advanced MIMO PHY Techniques

Besides beamforming recommended by the 802.11 standards, here we introduce two advanced MIMO PHY techniques which have been adopted in many current research works for interference management: *interference alignment* and *interference nulling*.

1) Interference Alignment: The basic idea of interference alignment is to make all the interfering signals aligned along a specific direction at the receiver, thus allowing the signal from the intended transmitter to be interference free for decoding [31]. It already has many applications in wireless networks [32]–[36], while this paper focuses on the WLAN context.

Consider the scenario shown in Fig. 7(a), the channel matrix from station i to AP is denoted as $\mathbf{H}_i$. When a signal x_1 is transmitted concurrently with an interfering signal x_2, both signals can be decoded successfully as the AP has two antennas which correspond to two received signals, while the decoding process is to solve the two linear equations. AP just discards x_2 and gets its desired signal x_1. This process obviously cannot work when another interference x_3 joins in, as shown in Fig. 7(b). Through interference alignment shown in Fig. 7(c), both station 2 and station 3 precode their transmitted signals as $\mathbf{v}_2 x_2$ and $\mathbf{v}_3 x_3$, making the two signals aligned along the same direction at AP. AP then treats $\mathbf{v}_2 x_2 + \mathbf{v}_3 x_3$ as one signal and decodes x_1 successfully. This mechanism can certainly be extended to the scenario with more than two interferers.

2) Interference Nulling: Interference nulling is to make a transmitter cancel its signal completely at a particular receiver. As shown in Fig. 9, when the AP intends to null its signal at this station, it precodes the transmitted signals from each antenna with $\mathbf{v}_1$ and $\mathbf{v}_2$, and finally makes $\mathbf{H}_1\mathbf{v}_1 x + \mathbf{H}_2\mathbf{v}_2 x = 0$ at the station. This condition can be easily satisfied through making $\mathbf{H}_1\mathbf{v}_1 = -\mathbf{H}_2\mathbf{v}_2$. We note that interference nulling is different from the aforementioned null steering, as the former is an interference management technique which only cares about whether the signal is canceled out at the intended receiver, while the latter is a kind of beamforming technique which tries to focus each transmit signal on the intended receiver but null the signal at other receivers, thus to achieve multiple concurrent transmissions.

IV. Survey on MIMO-related Interference Management in Single Collision Domain

Some research works focus on the single collision domain where there are only one AP and a group of associated stations. These works use the 802.11 standard as the basis, try to investigate and solve the problems in current MIMO transmission, including channel access, channel sounding and user selection. Nearly all of the surveyed works adopt distributed MAC layer mechanisms. In this part we conclude the advances on these efficient MIMO-related interference management mechanism design.

A. Channel Access

Some researchers find that the CSMA mechanism in the MIMO system is even less efficient than that in the single stream system and try to optimize it.

CHRoME [41] finds out that the retransmission scheme in the 802.11 MAC is inefficient due to the re-sounding process. After a failed transmission, the AP should wait for an incremented backoff period to retransmit the data packets, and in which situation it should re-sound the channel, as the CSI after this long period has a large probability to be changed. It proposes a beamformed probe for channel measurements and selects the optimal bit rate accordingly. acPad [42] proposes to increase the concurrent transmission opportunities in MU-MIMO where the streams have heterogeneous transmission durations. It adds additional packets to fill up the idle channel time, while the users are identified not to harm all the ongoing streams. Authors in [43] study the cross-layer performance of MU-MIMO for the TCP traffic in the single collision domain and indicate that, even with dominant downlink traffic, the downlink MU-MIMO would result in severe low performance when it is coupled with single-user uplink transmission, due to the intrinsic limitations in the random channel contention. They also provide some potential solutions to improve the network performance, such as cumulative ACK and uplink MU-MIMO. These works, however, mainly focus on the single collision domain, the problems need further study in the multiple collision domain, in which situation they would lead to worse performance.

We note that the research works are evaluated based on the testbed WARP (Wireless Open Access Research Platform) [44] or USRP (Universal Software Radio Peripheral) [45]. Actually, many current researchers in wireless networks try to investigate the advantages of their mechanisms through system testbeds, which are commonly constructed based on some typical software defined radio platforms, such as WARP, USRP and Sora (Microsoft Research Software Radio) [46]. Experimental results from the testbeds are more convincing than theoretical ones as they have considered the complex wireless environments. Nearly all the surveyed works in this paper have been evaluated through testbeds.

B. Channel Sounding

One of the main problems of the explicit channel sounding process is its huge transmission overhead. For example, as demonstrated by MUTE [37] through experimental results, the overhead required by a four-user transmission can reach 30% of the total transmission time in a 20MHz channel, and 60% in an 80*MHz* one. When the number of users or the channel bandwidth increases, the capacity gain from MIMO transmission may not even offset the transmission time induced by the channel sounding frame exchange. Thus, the 802.11ac lets the station compress its CSI to reduce this overhead from three aspects: sending

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

TABLE IV
RESEARCH ON EXPLICIT CHANNEL SOUNDING FOR DOWNLINK MU-MIMO TRANSMISSION.

Mechanism	Main features	Evaluation Tool
MUTE [37]	Quantify the overhead induced by explicit channel sounding, and reduce the overhead based on the channel stability	WARP
AFC [38]	Optimize the CSI compression level according to the information loss	WARP
CUiC [39]	Allow the CSI report at the same slot and propose MMSE-SIC to detect all the CSI	WARP
Gabriel [40]	Let the CSI be reported only when it is different from the history one	USRP

the channel matrix less frequently, allowing adjacent subcarriers to share the same matrix, and quantizing the matrix values into a small number of bits. However, a tradeoff should be made between the overhead and the matrix's accuracy.

There are many research works striving to reduce this overhead [38], [39]. AFC [38] indicates that the CSI transmission may still significantly degrade the link capacity despite CSI compression; it then proposes to model the information loss due to compression, and finally identifies the appropriate compression level that optimizes the overall network throughput, considering both the capacity loss and the overhead reduction. Different from 802.11n/ac which make each user report its CSI in different time slot, as shown in Fig. 3(b), CUiC [39] lets all the uses report their CSI at the same slots and proposes the MMSE-SIC algorithm to decode all the frames, thus significantly reduces the channel sounding time. Gabriel [40] compares the history CSI to a new estimated one based on current training frame, and lets a user report its CSI only when the verification fails.

As summarized in Table IV, all the research on improving the efficiency of explicit channel sounding focuses on the downlink MU-MIMO scenario in the single collision domain, and it is obvious that the induced overhead will increase dramatically in the multiple collision domain, especially in the network MIMO scenarios [47], [48].

C. User Selection

The user selection process requires the AP to first obtain the CSI at all the stations, and then select the best group of users for beamforming. It faces two main challenges when applied in real networks: the channel sounding overhead and the computational complexity.

Some researchers try to reduce the overhead induced by CSI acquisition for user selection. OPUS [49] makes each user evaluate its potential contribution to the downlink capacity when grouped with those already selected ones, only the one which will lead to maximum aggregate capacity should send its CSI to AP in each round, so as to reduce the CSI transmission overhead. It is a greedy-like algorithm, but without acquiring all stations' CSI at AP. [4] PUMA [50] proposes a mode and user selection algorithm without the channel sounding process. It first predicts the expected per-user data rate only using the number of transmit antennas M and that of receive antennas K; it then estimates the throughput of MU-MIMO with a set of potential user groups, mainly based on the priori information, including M at AP and K at each client, each user's link state (the ommi-directional SNR), and the per-user expected data rate; it finally identifies a user group with the highest throughput for the following MU-MIMO transmission. It still needs channel sounding before the actual data transmission as CSI is required for precoding. MUSE [51] captures the inter-user interference by computing the correlation of the compressed CSI feedback matrix among users to improve the efficiency of user selection; it also considers the diversity of user's bandwidth configurations, and selects the bandwidth combination with the highest throughput for each group to improve the MU-MIMO capacity. Instead of performing user selection process for each MU-MIMO transmission, MAPS [52] proposes to identify clients with uncorrelated channels and associate them to the same AP. It adopts an implicit channel sounding method through obtaining CSI at the APs to measure the correlation among clients' channels. DiFuse [53] is a distributed user selection protocol for the downlink MU-MIMO transmission. It makes each user cleverly compute the expected sum-capacity gain by overhearing the CSI from other users, then feedbacks the result concurrently in the frequency domain to reduce the overhead. Based on all the feedbacks from users, the AP will select and add the user with optimal sum-capacity into the current user group. BUSH [54] proposes to make user selection in large-scale MIMO systems. It selects users according to the obtained angle-of-arrival (AoA) and power information, users with large angle distances are grouped and should feedback CSI for precoding. While the above advances focus on the downlink MU-MIMO transmission, MIMOMate [55] tries to improve the efficiency of uplink

[4] As discussed in Section III-C3, the greedy algorithm requires all the channel conditions (CSI) for selecting users, thus this algorithm always implies high channel sounding overhead.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

TABLE V
RESEARCH ON USER SELECTION.

Mechanism	Transmission Direction	Main features	Channel Sounding Type & Overhead	Beamforming Capacity	Evaluation Tool
OPUS [49]	Downlink	Reduce the channel sounding overhead: let only one user feedback its CSI in each round	Explicit & Low	$\approx C_{greedy}$	WARP
PUMA [50]	Downlink	Reduce the channel sounding overhead: identify the user group based on priori information	No need & None	$< C_{greedy}$	WARP
MUSE [51]	Downlink	Reduce the channel sounding overhead: exploit the correlation of the compressed CSI feedback matrix to improve user selection	Explicit & Medium	$\approx C_{greedy}$	WARP
MAPS [52]	Downlink	Reduce the channel sounding overhead: associate the stations with uncorrelated channels to the same AP	Implicit & Low	$\approx C_{greedy}$	Commodity testbed
DiFuse [53]	Downlink	Reduce the channel sounding overhead: identify the user with optimal expected capacity according to the capacity feedback from users	Explicit & Low	$> C_{greedy}$	USRP
BUSH [54]	Downlink	Reduce the channel sounding overhead and optimize the user selection: select users according to AoA and power	Explicit & Low	$> C_{greedy}$	WARP
MIMOMate [55]	Uplink	Reduce the channel sounding overhead: group users based on the updated CSI	Implicit & Low	$\approx C_{greedy}$	USRP
SIEVE [30]	Downlink	Balance between computational complexity and capacity: achieve optimal capacity in stable networks while suboptimal capacity in dynamic networks	Implicit & Low	$> C_{greedy}$	Sora
Signpost [56]	Uplink	Reduce both the channel sounding and computational complexity: design a set of orthogonal signpost directions for transmission, with low complexity and suboptimal capacity	Implicit & Low	$> C_{greedy}$	WARP
SAMU [57]	Downlink	Solve the frequency-selective fading problem	Explicit & High	$> C_{greedy}$	WARPLab

user selection. It groups stations as concurrent transmitters (MIMO-Mates) based on the channel characteristics to maximize the network performance; the MIMO-Mates will be rescheduled only when the channels of certain clients change due to channel variation or user mobility.

The computational overhead in the user selection process is induced by the large searching space, which increases exponentially with both the number of antennas at AP and the number of mobile stations. For example, to select a best beamforming group for an 8-antenna AP with 30 active users, it takes 269 seconds even for a modern 16-core server [30]. It is obvious that the computational cost is now unacceptable in practical, and this problem will be more serious when the uplink MU-MIMO is adopted in the upcoming 802.11ax standard, as the station has much more limited resources than AP. To solve the problem in the downlink transmission, SIEVE [30] keeps only a small set of good candidates and refines this candidate set according to the channel variations. It further utilizes a central database to track the channel and the coherence time for each mobile user, and proposes a progressive update strategy to avoid unnecessary computing. Signpost [56] achieves distributed user selection in the uplink direction. It predefines a set of Signpost uplink directions, which are the orthogonal channel vectors, and selects the users which are best aligned with each Signpost direction for uplink MU-MIMO transmission.

Some researchers consider other problems in user selection. For example, the WLAN's wide bandwidth has been increased from $20MHz$ to $160MHz$ with the 802.11 update, making the frequency-selective fading a dominant channel effect which degrades the system capacity. To solve this problem, SAMU [57] proposes a selective-aware MU-MIMO transmission, which divides the transmission channel into subchannels, and selects a set of users for each subchannel that is considered frequency-flat.

We summarize the current research on user selection in Table V. The performance of MU-MIMO transmission is related to both the beamforming capacity and the channel sounding overhead. The beamforming capacity is affected by the computational complexity [30], and the works with higher capacity normally have higher complexity. No work in Table V achieves the optimal capacity due to the impractically high complexity for exhaustive searching. Here we use the well-known suboptimal greedy user selection algorithm as a baseline, which achieves the beamforming

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

capacity C_{greedy}, to demonstrate the capacity of these research works. Please note that the mechanisms with low beamforming capacity in Table V would not result in low performance, as the channel sounding overhead has been reduced significantly. In addition, we should mention that nearly all the schemes are designed based on the single-AP scenarios, it is worthy to study whether they can be extended to the multiple-AP scenarios and what the performance will be.

V. Survey on MIMO-related Interference Management in Multiple Collision Domain

In the multiple collision domain scenarios, the interference management mechanisms need to coordinate among multiple MIMO transmissions, and the corresponding channel access schemes fall into three categories: centralized, distributed and semi-distributed. In this part, we will investigate the related works from the three categories, as shown in Table VI. For each category, the mechanisms will be studied from the three MIMO application scenarios, SU-MIMO, MU-MIMO and network MIMO. In addition, as the MIMO-related interference management in the multiple collision domain always contains both the PHY and MAC layer design, we will also study how the advanced MIMO physical layer techniques are exploited and combined with the MAC layer protocols to improve the network performance.

TABLE VI
Categorization of the survey for MIMO-related interference management.

	Centralized	**Distributed**	**Semi-Distributed**
SU-MIMO	Section V-A1	Section V-B1	—
MU-MIMO	Section V-A2	Section V-B2	Section V-C1
Network MIMO	Section V-A3	—	Section V-C2
Discussion	Section V-A4	Section V-B3	Section V-C3

A. The Centralized Schemes

In this kind of schemes, a central node (e.g., a leader AP or a central server) is responsible to coordinate the data transmissions and apply the MIMO physical layer techniques to improve the network performance, through exploiting the AP wired backbone network, which makes APs exchange high speed information through Ethernet. This kind of design extends the PCF mechanism in 802.11 to achieve centralized coordination in the multiple collision domain scenarios.

1) SU-MIMO: Based on the SU-MIMO transmission supported by 802.11n and 802.11ac, researchers further utilize the interference alignment and interference nulling in the multiple collision domain to exploit more concurrent streams, so as to improve the network capacity.

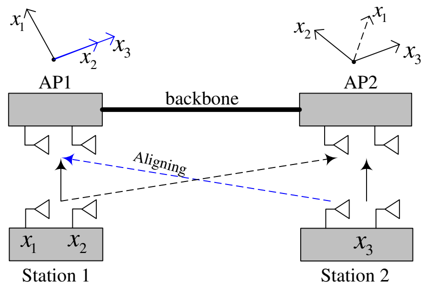


Fig. 10. A scenario of IAC. Three streams can be transmitted concurrently when the AP only has two antennas.

Considering that the throughput of MIMO is limited by the number of antennas per AP, IAC [58] is proposed to overcome this limit through exploiting both the interference alignment and interference cancellation mechanism. An example of the IAC decoding process is shown in Fig. 10. When station 1 transmits two packets x_1 and x_2 to AP1, AP1 can obtain both packets through the normal MIMO decoding process. In this situation, IAC further permits station 2 to transit its packet x_3 to AP2, without interfering with AP1's data reception. It makes use of interference alignment to align x_3 with x_2 at AP1, then allows AP1 to decode x_1 locally; the decoded x_1 will be informed to AP2 through the backbone network, making AP2 cancel it from the received signal to decode the remaining x_2 and x_3. We note that this is a SU-MIMO scenario as each station only delivers its data packets to one particular AP, while the signals received by other APs are used for interference cancellation. IAC designates one AP as the leader, which acts as the coordinator to poll the data transmissions of all the clients and APs. It adopts implicit channel sounding for downlink channel estimation.

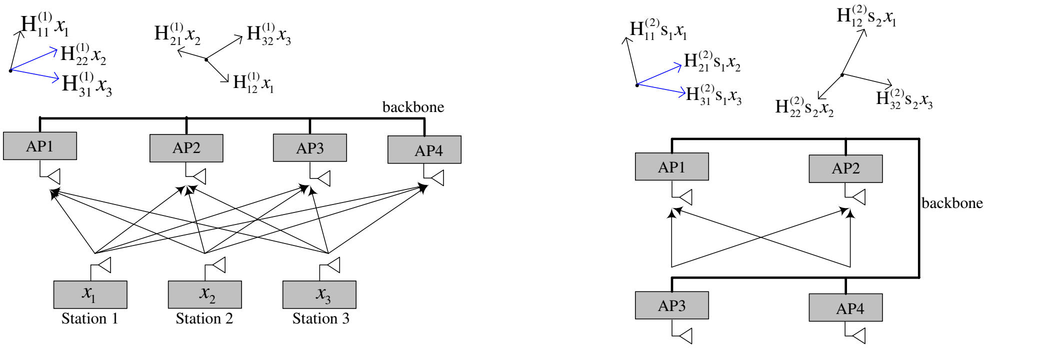
MoMIMO [59] extends IAC to perform interference alignment or interference nulling at the receiver side through sliding one of its antennas slightly. It also proposes a Stochastic Hill Climbing algorithm to seek the optimal position of the sliding antenna, owe to the continuous channel characteristics over the space. This design acquires CSI at the receiver side (CSIR) to achieve the interference alignment or nulling process, thus avoids the overhead induced by acquiring CSI at the transmitter side (CSIT).

2) MU-MIMO: There are a large number of research works falling into the MU-MIMO scenario. As the mechanisms designed for the uplink and downlink directions are different to a great extent, here we survey them from the transmission directions.

i) Downlink Direction

The downlink MU-MIMO is comparatively easier to be implemented as the AP can precode all the signals centrally to ensure their decoding at the stations. Although with high theoretical performance, the MU-MIMO faces many challenges when applying to the real networks, which attract some researchers to implement MU-MIMO

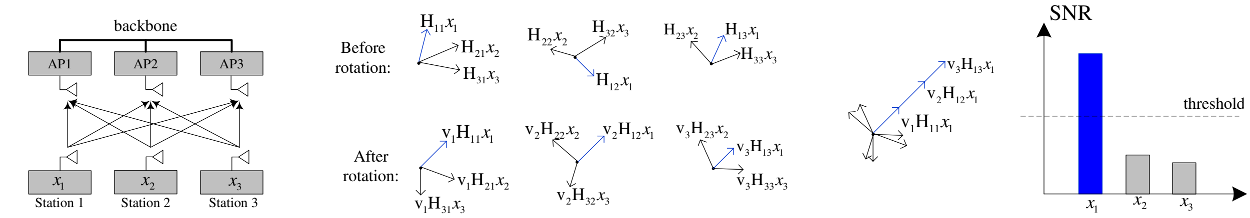
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials



(a) In the first slot, the three stations transmit x_1, x_2 and x_3, respectively. All the APs can receive the three signals.

(b) In the second slot. AP3 and AP4 retransmit the received signal to AP1 and AP2. They make x_2 and x_3 at AP1 aligned with their received signal in the first slot through precoding.

Fig. 11. An example of BBN which utilizes the dense deployment of APs to enable concurrent transmissions. It divides the transmissions into two slots, and combines the received signals at AP1 and AP2 in the two slots to decode x_1 at first.



(a) An example of POLYPHONY where three stations transmit signals to their APs

(b) The received signals at each AP before and after rotation

(c) The combined signal

Fig. 12. An example of POLYPHONY. The three APs rotate their received signals to make x_1 aligned at the same direction, the x_1 will have a high SNR for decoding after combining the rotated signals at three APs.

in the WLAN testbeds to investigate its efficiency [60], [61] and demonstrate its feasibility to deploy in real networks [62].

The first experimental evaluation of downlink MU-MIMO was conducted by [60] in 2010, based on the WARP platform. The authors reveal that, in the one AP indoor environment, two stations can achieve approximately the maximum throughput when their distance is over a quarter of a wavelength. Explicit channel sounding is adopted in the experiments to get CSIT, whose required update rate is related with the user mobility, environment variation and so on. They also demonstrate that ZF is promising to reduce interference at the intended stations. As for the multiple collision domain scenario, they let the APs jointly beamform signals just like the process in the network MIMO scenario.

Kardia [63] utilizes the degree of freedom (DoF) [64] to analyze how maximum performance can be achieved in MU-MIMO scenarios. It points out that the MU-MIMO throughput does not increase linearly with more added stations due to the correlated user channels, especially in the multiple collision domain. To solve this problem, each AP uses part of its DoF to serve the selected users with the most orthogonal channels for downlink data transmission, and exploits the remaining DoFs to coordinate the neighboring APs' transmissions through nulling mutual interference. Kardia designs a lightweight multi-cell orthogonality probing (MOP) mechanism to reduce the CSI acquisition overhead and facilitate user selection, it also utilizes a centralized coordinator to allocate APs carefully to enable concurrent downlink MU-MIMO transmissions.

Although existing standards only support up to eight MU-MIMO stations in WLANs, some researchers implement large scale MU-MIMO system to investigate its performance, due to its theoretical high potential to improve the spatial reuse, such as Argos [62], BigStation [65], Hekaton [66] and FlexCore [67]. These works have mentioned few MAC layer mechanism.

ii) Uplink Direction

The uplink MU-MIMO transmission is much different from its downlink counterpart, as the stations are uncoordinated and they cannot perform precoding efficiently to ensure the decoding at the receiver.

RobinHood [68] and BBN [69] utilize the dense deployment of APs to enable concurrent transmissions from multiple stations to multiple nearby APs through exploiting interference alignment and nulling. An example is shown in Fig. 11, where the three stations transmit their signals to APs. BBN divides the transmissions into two slots. In the first slot, the three stations transmit their signals x_1, x_2 and x_3 concurrently, making each of the four APs receive a combination of the three signals. In the second slot, AP3 and AP4 will retransmit their received signals after proper precoding, so that both x_2 and x_3 at AP1 aligned along the same direction with that in the first slot. x_1 can be decoded at first through nulling both x_2 and x_3 at AP1, then it will be passed to AP2 through the backbone network. AP2 will subtract x_1 from the received signals at both slots, and decode x_2 and x_3 through solving the linear equations. To make BBN work in the multi-collision domain, the authors de-composite the APs into groups through a central server, such that all APs in one group can hear each other to accomplish the BBN detection process like that in Fig. 11. They then propose centralized scheduling mechanism to make neighboring groups not transmit simultaneously to avoid interference. Meanwhile, they also use a centralized method to schedule the pilot signal transmissions of clients and APs, so as to obtain the required CSI for performing interference alignment and nulling.

POLYPHONY [70] exploits the idea of interference alignment for enabling more concurrent transmissions in the uplink through exploiting the AP backbone network. It designs a "Shadow Preamble" mechanism to estimate the CSI between each station and APs; based on the obtained CSIR, it makes one packet from a station aligned along the same direction at all the APs, so that the accumulated SNR for this packet is high enough for decoding; the same process can be applied to decode the remaining packets. Fig. 12 shows an example of decoding x_1 through POLYPHONY. The three APs first rotate their received signals to let x_1 aligned at the same direction, making x_1 have a high SNR for decoding after combining the rotated signals at three APs. The process is repeated to decode the other signals x_2 and x_3. This mechanism certainly needs a central server to schedule the uplink transmissions and decode the received signals.

iii) Both Directions

FlexRadio [71] exploits both the interference alignment and full-duplex [72] technologies [5] to make the design fully flexible, thus to increase concurrent streams to improve the system capacity. An example of FlexRadio is shown in Fig. 13, AP1 and AP2 cannot listen to each other, but station 1 and station 2 can listen to each other and both APs. According to the traditional MIMO communications, only two packets can be transmitted concurrently, e.g., the two packets from AP1 to station 1 in the SU-MIMO scenario, or the two packets from AP1 to the two stations in the MU-MIMO scenario. However, an additional concurrent stream can be transmitted with FlexRadio. In Fig. 13, AP1 sends x_1 and x_2 to station 1 and station 2 respectively; since AP1 has two antennas, it further nulls x_2 at station 1, making station 1 able to receive x_1 through only one antenna and decode it correctly; station 1 can transmit another signal x_3 to AP2

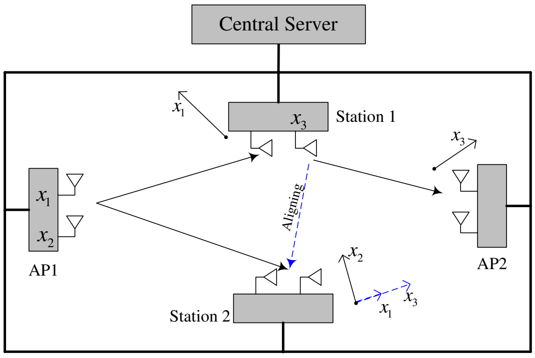


Fig. 13. An example of FlexRadio.

concurrently through the second antenna, it also aligns x_3 with x_1 at station 2 to make station 2 decode x_2 correctly; in this situation AP2 can decode x_3 obviously. We note that although the transmit antenna of station 1 will interfere with its receive antenna, this interference can be canceled out through the full-duplex technology. Implicit channel sounding is adopted to obtain CSIT for precoding. This work has a strong assumption that a central server knows the RF resources of all nodes in the network and their respective traffic demands, in order to configure the FlexRadio communications.

Under full-duplex operation, a node activates equal number of RF chains for transmission as it does for simultaneous reception. Thus, when a node has N RF chains, under full duplex, N/2 RF chains are active transmitting RF chains while the remaining N/2 RF chains are receiving RF chains.

3) Network MIMO: Network MIMO is like an extension of MU-MIMO to allow the communications between multiple APs and the stations. It has the main challenge of achieving tight phase and timing synchronization among multiple APs, so that the multiple APs can work as one giant AP to make joint precoding for the MIMO transmission [75]–[79]. The wired backbone among APs is fully utilized to achieve this goal, as shown in Fig. 14, where the central server may be virtual or real in a particular mechanism.

MegaMIMO [75] is the first work that implements network MIMO in the WLAN testbed based on both USRP and off-the-shelf devices. It makes the transmitted data packets shared by all the APs through the wired backbone network, and adopts the precoding process similar with that in the MU-MIMO transmission. It achieves phase synchronization among independent APs through electing one AP as a leader and making the other APs change the phase of their signals to maintain a required alignment with the leader AP. MegaMIMO adopts explicit channel sounding before the actual data transmissions, and the process is similar as that in the MU-MIMO transmission recommended by 802.11ac. During the channel sounding process, APs transmit their SYNC header and get the clients' CSI feedback one after another, while the induced

[5] Although some research works exploit additional antenna to cancel the self interference for full-duplex transmission [73], current advances have no such limitation [72], [74]. FlexRadio [71] has the following assumption for full-duplex: a node with N RF (Radio Frequency) chains can activate $N/2$ RF chains for transmission while the remaining RF chains for reception.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

overhead increases proportionally to the number of both APs and stations, making network MIMO unapplicable to real networks when the network size is relatively large. To address the challenge of real-time CSI update in the explicit channel sounding, the authors extend MegaMIMO's design and implementation to MegaMIMO 2.0 [79], which proposes a new technique for extending reciprocity to the network MIMO scenario, thus improves the accuracy of implicit channel sounding. The main challenge of accomplishing it is to estimate the calibration factor for reciprocity correctly.

Authors in [76] implement a network MIMO system which is also based on the backbone network, but on the WARP platform. The server is responsible for jointly precoding the transmitted signals and passing the final frequency domain symbols to APs, thus assures all the APs are phase synchronized. Based on this system, they implement zero-force precoding and Tomlinson-Harashima precoding [80] to achieve the network MIMO transmissions, where CSI is assumed to be available through exploiting the channel reciprocity. Both precoding mechanisms demonstrate the near-theoretical performance. Under the condition when CSI cannot be reliably obtained, the authors exploit blind interference alignment [81] to increase the data rate. They also provide some centralized scheduling algorithms [82] in the MAC layer to convert the physical layer gains into the transmission rate gains.

AirSync [83] utilizes the central server for joint signal processing, including the joint precoding in the downlink transmission and the joint decoding in the uplink transmission, while the CSI used for precoding is obtained from explicit channel sounding as that in the 802.11ac. It makes a reference AP broadcast a pilot signal, and makes the other APs closely track the phase drift of this signal through the Kalman filter and de-rotate their baseband signals, finally achieves tight phase synchronization among all the APs. Experimental evaluation based on WARP testbed also indicates that AirSync can achieve the theoretical optimal multiplexing gain. AirShare [78], [84] proposes to enable synchronization among APs through a reference clock, which is shared through the wireless medium and fed into the wireless nodes, so that all the APs have the same clock to achieve tight time and phase synchronization. No MAC scheme is mentioned in this work.

4) Discussions about Centralized Schemes: Table VII summarizes the main characteristics of the surveyed centralized MIMO-related interference management schemes, from which we get the following key points:

- The centralized schemes rely highly on the AP wired backbone network to make APs exchange information with each other. This design is reasonable for the enterprise WLANs where the Ethernet is often created among the APs and can satisfy the requirements for low delay and high speed information exchange. However, in the other kinds of WLANs where Ethernet does not exist, such as the home WLANs where APs are connected through Internet, these schemes may not be suitable due to the inefficient information exchange among the APs.
- The advanced MIMO physical layer techniques, interference alignment and interference nulling, are widely exploited in the SU-MIMO and MU-MIMO scenarios, so as to increase the concurrent transmission streams to improve the network performance, while the research works on the network MIMO scenario only adopt beamforming and null steering, they currently only strive to extend the MU-MIMO transmission to the multiple-AP scenario.
- The introduction of interference alignment and interference nulling can increase concurrent streams, however, the negative effect is that they require more CSI for precoding. For example, as shown in Fig. 10, IAC needs the extra CSI of the links station1⟶AP2 and station2⟶AP1, while the traditional SU-MIMO transmission only needs CSI of the data transmission links station1⟶AP1 and station2⟶AP2. The extra CSI requirements will certainly induce more overhead to the system.
- Although the 802.11ac standard only recommends explicit channel sounding, its high transmission overhead impedes its wide usage. Many researchers adopt the implicit channel sounding process in their design to achieve high performance. Meanwhile, for the uplink transmissions, some research works [59], [69], [70] propose to perform the MIMO physical layer techniques at the receiver (AP side), where CSIR is adopted and the channel sounding process is no longer required.
- As a key feature in the MU-MIMO transmission, user selection is also considered by researchers when designing mechanisms in the multiple collision domain scenarios. However, most of them adopt the greedy algorithm which obtains the best user candidates through exhaustive searching, and are not suitable for large scale networks. Designing proper user selection mechanisms for the multiple collision domain scenario is worthy for thorough study.

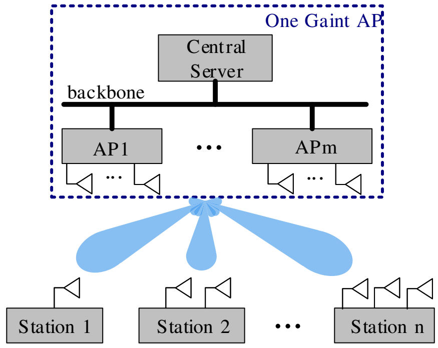


Fig. 14. The illustration of network MIMO transmission.

TABLE VII
MIMO-RELATED INTERFERENCE MANAGEMENT MECHANISMS WHICH ADOPT CENTRALIZED MAC SCHEMES.

Mechanism	Scenario	Transmission Coordinator	Transmission Direction	Channel Sounding	User Selection	Advanced MIMO PHY Techniques	Evaluation Tool
IAC [58]	SU-MIMO	Leader AP	Uplink	Implicit CSIT	—-	Interference Alignment, Interference Nulling	USRP
MoMIMO [59]	SU-MIMO	Leader AP	Uplink	CSIR	—-	Interference Alignment, Interference Nulling	USRP
MU-MIMO testbed [60]	MU-MIMO	Central Server	Downlink	Explicit CSIT	Greedy Algorithm	—-	USRP
Kardia [63]	MU-MIMO	Central Server	Downlink	Explicit CSIT	New Schemes	Interference Nulling	WARP
BBN [69]	MU-MIMO	Central Server	Uplink	CSIR	Not Mentioned	Interference Alignment, Interference Nulling	Off-the-shelf
POLYPHONY [70]	MU-MIMO	Central Server	Uplink	CSIR	Not Mentioned	Interference Alignment	Off-the-shelf
FlexRadio [71]	MU-MIMO	Central Server	Downlink & Uplink	Implicit CSIT	Not Mentioned	Interference Alignment	FPGA
Network MIMO testbed [76]	Network MIMO	Central Server	Downlink	Implicit CSIT	Greedy Algorithm	—-	WARP
MegaMIMO [75]	Network MIMO	Central Server	Downlink	Explicit CSIT	Greedy Algorithm	—-	USRP, Off-the-shelf
MegaMIMO 2.0 [79]	Network MIMO	Central Server	Downlink	Implicit CSIT	Greedy Algorithm	—-	FPGA
AirSync [83]	Network MIMO	Central Server	Downlink	Explicit CSIT	Greedy Algorithm	—-	WARP

B. The Distributed Schemes

In these schemes, nodes inherit the CSMA mechanism for random channel access in MIMO networks.

1) SU-MIMO: 802.11n$^+$ [85] is a distributed random access protocol for MIMO networks. It allows the nodes to carrier sense the channel, and contend not only for the transmission time, but also for the DoFs. According to this design, when there is an ongoing transmission link, another node with more antennas than the ongoing transmitter may transmit concurrently, through utilizing both the interference nulling and interference alignment technology, so as to cancel its interference at the ongoing receivers and maximize its signal at the intended receivers. Fig. 15 shows an example of this mechanism. The TX or RX node can be either an AP or a station. When there is an ongoing link $TX1 \longrightarrow RX1$, $TX2$ can utilize its two antennas to transmit concurrently through nulling its signal at RX1. $TX3$ can further utilize its three antennas to transmit concurrently through both nulling its signal at RX1, and aligning its signal with x_1 at RX2. Thus, all the signals x_1, x_2 and x_3 can be decoded successfully by their intended receivers. 802.11n$^+$ requires at least k+1 antennas for enabling new concurrent transmissions when there are k ongoing transmission links. It adopts a light-weight RTS/CTS mechanism [86] for random channel access. The node pair that wins the contention exchanges a light-weight RTS/CTS which includes the used DoF for the data transmission. Other nodes which have more antennas can learn this information from the prior RTS/CTS messages and contend for the unused DoF for concurrent transmission. Implicit channel sounding is adopted as nodes utilize the received RTS/CTS messages to estimate CSI.

802.11mc [87] extends 802.11n$^+$ [85] to increase more concurrent transmissions and reduce the coordination overhead. It also utilizes the light-weight RTS/CTS mechanism for channel access, and adopts the implicit channel sounding through exploiting the RTS/CTS messages. The main difference is that, it designs a new RTS frame with a postamble at the end of RTS, allowing the receiver to obtain the CSI when the RTS collisions occur, so as to reduce the number of control message transmissions.

MDMA [88] proposes to maximize the concurrent transmissions in the WLANs in a distributed way, and meanwhile obtain more receiver diversity gains. Interference alignment is exploited to achieve this goal. Fig. 16 shows an example of MDMA's design where three APs contend for transmission. Suppose AP1 wins the contention and firstly transmits its packet to station 1 through beamforming to gain the receive diversity, AP2 can then join this

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

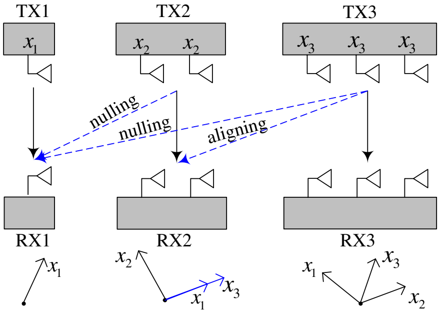


Fig. 15. An example of 802.11n$^+$. When accessing channel, nodes contend not only for the transmission time, but also for the degrees of freedom provided by the multiple antennas.

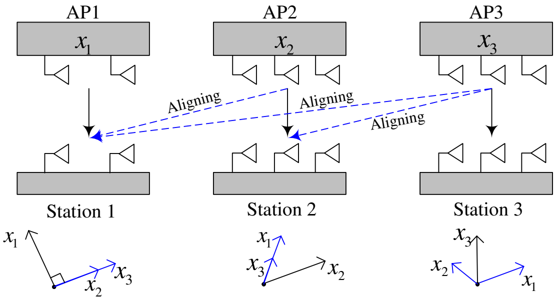


Fig. 16. A scenario of MDMA. Interference alignment is exploited to enable more concurrent transmissions as well as achieve high receive diversity.

transmission if its precoding makes its signal x_2 aligned orthogonal to x_1 at station 1; AP3 can further join if its precoding makes the signal x_3 aligned along the same direction with x_2 at station 1, and with x_1 at station 2. According to this design, one node can join the kth concurrent transmission if it does not interfere with the $k-1$ previously joined transmissions. MADA adopts frequency domain contention [89] to allow as many nodes as possible to win the contention. It lets each contending transmitter send an M-RTS message simultaneously on a randomly selected OFDM subchannel. The M-RTS carries the information of DoF, each receiver picks the node with smallest DoF as its transmitter and feedbacks an M-CTS, according to which a selected lead transmitter allocates the link pairs, and each pair then joins the transmissions one by one through a preceded light-weight RTS/CTS handshake. Similar with 802.11n, it adopts implicit channel sounding to obtain CSI.

2) MU-MIMO: The study for the MU-MIMO scenario is conducted from the transmission directions.

i) Downlink Direction

MIDAS [90] spatially separates the antennas of an AP through RF or optical cabling to achieve a distributed antenna system (DAS), which has the potential to improve the WLAN performance due to the following reasons: 1) reducing the path loss as one client may find an AP antenna closer to it, 2) naturally mitigating the inter-user interference as one client will get the strongest signal from a specific antenna, and 3) enabling more concurrent transmissions as each antenna has different channel states (busy or idle). It then designs a new power-balancing zero-force precoding mechanism for this system to achieve MU-MIMO transmission. In the MAC layer, it makes each antenna carrier sense the channel independently to get the available antennas, and selects clients for MIMO transmission according to the signal strength from different antenna to each client. The acquisition of CSI for precoding adopts the same channel sounding process in the 802.11 standard.

ii) Uplink Direction

SAM [91] proposes a chain-decoding technique to decode the uplink MU-MIMO signals, through performing the interference nulling at the receiver side. An example of the SAM chain-decoding process is shown in Fig. 17. When an AP receives two signals x_1 and x_2 through the two antennas and gets y_1 and y_2, it first finds a proper linear transform of each received signal to make x_1 at both y_1 and y_2 aligned at the same direction (shown in Fig. 17(c)), the signal x_1 can be easily nulled out through making substraction between y_1 and y_2; the AP can then decode x_2 correctly. The decoded x_2 is in turn canceled from both y_1 and y_1 to recover x_1. This process can be extended to the scenario with more than two stations. SAM requires only CSIR as decoding is performed at the receiver side. It further extends CSMA to Carrier Counting Multiple Access (CCMA), which makes multiple stations coordinate their transmissions according to the chain-decoding requirement.

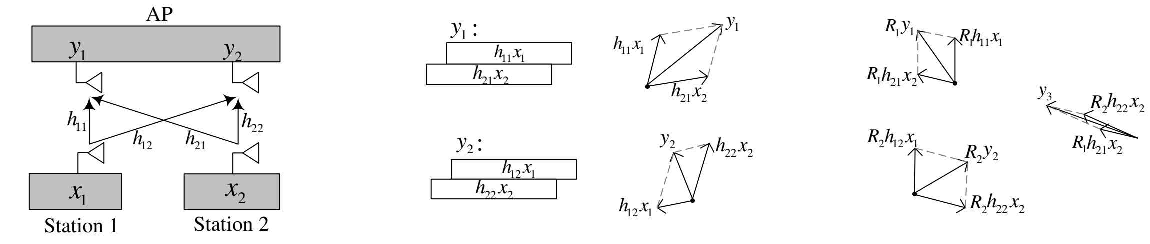
MUSE [92] also achieves the uplink MU-MIMO transmission through decoding at the receiver side. It makes AP obtain clean channel estimation from all the clients based on the preamble. Instead of user selection according to the obtained channel matrix, this work exploits the cyclic shift delays (CSD) introduced by 802.11n to decorrelate the uplink channels. In the MAC layer, it proposes a single medium access contention mechanism for the users' channel access: when the first user wins the medium after a random backoff, it sends a triggering message to all the selected users to grant a transmission opportunity; the users then transmit the uplink packets simultaneously after triggered by the AP.

iii) Both Directions

iBeam [93] implements the MU-MIMO from the station's point of view. It allows each station to associate with multiple APs and make it communicate with these APs simultaneously at both the uplink and downlink directions. In the uplink direction, each station can transmit different data packets to the APs through using the multi-user beamforming, just like the downlink MU-MIMO. In the downlink direction, the station can decode different packets transmitted simultaneously from APs, through utilizing interference nulling and cancellation mechanism. It adopts implicit channel sounding to avoid the CSI transmission overhead, and obeys CSMA for channel access.

3) Discussions about Distributed Schemes: Table VIII summarizes the main characteristics of the surveyed dis-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials



(a) An example of SAM where two stations transmit signals to AP

(b) The original received signals and the vectors at AP

(c) The signals after interference nullifying

Fig. 17. An example of the chain-decoding process in SAM. AP finds a proper linear transform of the received signals y_1 and y_2 to make x_1 aligned along the same direction at both signals, so as to nullify it to decode x_2.

TABLE VIII
MIMO-RELATED INTERFERENCE MANAGEMENT MECHANISMS WHICH ADOPT DISTRIBUTED MAC SCHEMES.

Mechanism	Scenario	Channel Access	Transmission Direction	Channel Sounding	User Selection	Advanced MIMO PHY Techniques	Evaluation Tool
802.11n+ [85]	SU-MIMO	RTS/CTS	Uplink & Downlink	Implicit CSIT	—-	Interference Alignment, Interference Nulling	USRP
802.11mc [87]	MU-MIMO	RTS/CTS	Uplink & Downlink	Implicit CSIT	—-	Interference Alignment, Interference Nulling	WARP
MDMA [88]	SU-MIMO	RTS/CTS	Uplink & Downlink	Implicit CSIT	—-	Interference Alignment	USRP
MIDAS [90]	MU-MIMO	CSMA	Downlink	Explicit CSIT	New Schemes	—-	WARP
SAM [91]	MU-MIMO	CCMA	Uplink	CSIR	Not mentioned	Interference Nulling	SORA
MUSE [92]	MU-MIMO	CSMA	Uplink	CSIR	Greedy Algorithm	—-	WARP
iBeam [93]	MU-MIMO	CSMA	Uplink & Downlink	Implicit CSIT	Greedy Algorithm	—-	WARP

tributed MIMO-related interference management schemes, from which we get the following key points:

- Explicit channel sounding is harder to be implemented efficiently in a distributed manner, and nearly all the schemes adopt the implicit channel sounding method to obtain the CSI. This is reasonable as, besides the high transmission overhead, the control packets have a high probability of collision, which may significantly reduce the accuracy of the estimated CSI even when explicit channel sounding is adopted.
- The distributed random access schemes always include RTS/CTS packet exchange prior to the actual data transmission, which process has two advantages: 1) with RTS and CTS, the transmitter and receiver can negotiate a proper DoF for data transmission, while other neighboring nodes can learn this information and contend for the unused DoF for concurrent transmission; 2) the RTS and CTS transmission can be exploited for channel calibration in the implicit channel sounding process, so as to maintain accurate CSI for MIMO transmission.
- One problem in the distributed schemes is how to make the independent backoff timers of the concurrent transmitters (such as TX1, TX2 and TX3 in Fig. 15) expire nearly at the same time to perform MIMO transmission. Frequency domain contention seems to be a promising way to solve this problem meanwhile reduce the coordination overhead induced by backoff; it has been adopted in some recent advances such as MDMA [88] and FlexNEMO [94].
- Most of the research works also adopt the greedy algorithm to search for the best user candidates in the MU-MIMO transmission. More efficient user selection mechanisms are expected in this scenario.

C. Semi-distributed Schemes

This kind of schemes try to benefit from both the random channel access and the centralized coordination.

1) MU-MIMO: CoaCo [95] is an effective extension of 802.11n+ in the MU-MIMO scenario to combat the inter-cell interference, that is, interference among multiple BSSs. It uses one antenna for either transmitting streams,

TABLE IX
MIMO-RELATED INTERFERENCE MANAGEMENT MECHANISMS WHICH ADOPT SEMI-DISTRIBUTED MAC SCHEMES.

Mechanism	MIMO Application	Random Channel Access	Transmission Direction	Channel Sounding	User Selection	Advanced MIMO PHY Techniques	Evaluation Tool
OpenRF [97]	MU-MIMO	CSMA	Downlink	Implicit CSIT	Not Mentioned	Interference Alignment, Interference Nulling	Off-the-shelf devices
CoaCo [95]	MU-MIMO	CSMA	Uplink & Downlink	Explicit CSIT	New Schemes	Interference Alignment, Interference Nulling	WARP
NEMOx [98]	Network MIMO	CSMA	Downlink	Explicit CSIT	New Schemes	—-	WARP
NURA [99]	Network MIMO	CSMA	Uplink	Implicit CSIT	New Schemes	—-	WARP
FlexNEMO [94]	Network MIMO	CSMA	Uplink	Explicit CSIT	New Schemes	—-	USRP

or receiving streams, or canceling inter-cell interference. Based on this concept, each AP or station optimizes its antenna usage to increase concurrent transmitting streams, then optimizes the beamforming weights to cancel the inter-cell interference. CoaCo groups all cells into clusters, each of which has up to three cells. It employs CSMA/CA for the channel access between clusters to avoid inter-cluster interference. Within each cluster, it proposes inter-leaved channel sounding to obtain the required CSI among nodes, through making APs send their NDP sequentially; it then adopts CHAIN technique [96] to let clients respond their CSI, according to which APs perform the downlink MU-MIMO transmissions finally.

2) Network MIMO: As the giant MIMO cannot scale well owing to the global synchronization requirement and the overhead in sharing data packet among all the APs, some researchers intend to divide the large scale network into small practical-size clusters.

OpenRF [97] can be regarded as this kind of system. It deploys the architecture of software defined network (SDN) [100] to the WLANs with backbone networks, as shown in Fig. 14 where the central server is the control plan of SDN called OpenRF controller. It represents the OpenRF controller to automatically manage the network streams in a centralized way, enabling APs to automatically beamform their signals to their own stations but null their interference at the other stations. It also exploits interference alignment to further increase concurrent streams. This work exploits implicit channel sounding to obtain the CSIT for precoding.

NEMOx [98] first implements network MIMO in the downlink direction based on the WARP platform. In NEMOx, each cluster contains one master AP (mAP) that coordinates the other distributed APs (dAPs) centrally for the downlink transmission, and the dAPs utilize the decentralized CSMA to avoid inter-cluster interference, so as to make a balance between the spatial reuse and the dAP's cooperative gain. NEMOx also optimizes dAP's power budget and serves the set of stations opportunistically to ensure fairness of stations. It exploits the similar explicit channel sounding process as that in 802.11ac, making stations feedback their CSI sequentially for precoding. In addition, it proposes iterative user selection algorithm that selects stations which can achieve the optimal downlink transmission capacity.

NURA [99] further proposes a channel access mechanism for the uplink network MIMO transmission. It allows stations to access the channel randomly and utilizes a semi-synchronization protocol to guarantee that the new coming station will not interrupt with the ongoing uplink transmissions. It employs a request-permission mechanism to identify the stations whose channels are most uncorrelated with ongoing stations, thus optimizes the user selection in the network MIMO scenario.

FlexNEMO [94], [101] also implements the clustering of network MIMO in the uplink transmission, while it is based on the USRP platform. It contains a user-centric clustering mechanism which adapts the virtual AP (nAP) structure dynamically to ensure that every station with traffic locates at the center of an nAP, so as to reduce the inter-cluster interference, it also contains a priority-based MAC which makes the center nodes have higher transmission opportunities. Frequency-domain contention is adopted to facilitate the channel access more efficiently.

3) Discussions about Semi-Distributed Schemes: Table IX lists the surveyed semi-distributed MIMO-related interference management schemes, which divide the network into clusters, then adopt centralized schemes within each cluster while adopt distributed schemes for inter-cluster. The semi-distributed schemes seem to be more promising in the network MIMO scenario, as it can both benefit network MIMO's high capacity and solve its scalable problem well. In addition, this kind of schemes surely faces the same challenge as those in both the centralized and distributed schemes, such as the CSI acquisition and user selection within the cluster, which need better study.

VI. DISCUSSIONS AND FUTURE DIRECTIONS

The aim of the surveyed papers is to investigate efficient MIMO-related interference management mechanisms in WLANs to improve the network throughput, from both the single collision domain and the multiple collision domain

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

perspectives. Here we further give some discussions and future directions that fall into this area.

A. Channel Access

As the key component in MIMO-related interference management, the channel access process has attracted much research interest. A few researchers focus on the single collision domain scenario, where they investigate the drawbacks of the 802.11 CSMA and design effective schemes, through either increasing concurrent transmissions [42], [43] or improving the retransmission efficiency [41]. By contrast, significant efforts concentrate on the multiple collision domain, which is the more typical scenario in WLANs. Several MIMO PHY techniques are exploited to better manage interference among the multiple BSSs, such as interference alignment, interference nulling and network MIMO. There are three categories of channel access mechanisms: centralized, distributed and semi-distributed schemes. Discussions for each category has been shown in Section V. Here we further give our thoughts on potential solutions in this area.

1) Semi-Distributed Schemes: The centralized channel access schemes have obvious advantages to improve the efficiency of interference management, but rely highly on prompt information exchange among APs, thus have the deployment and scalable problems in real networks. The distributed schemes are more flexible to be deployed but with high transmission overhead. The semi-distributed schemes which adopt centralized schemes within the cluster while adopt distributed schemes among clusters, may benefit from both kinds of schemes, and seem to be a promising way for MIMO-related channel access. We have already seen a few advances on this kind of design, but there are still many practical problems remained. For example, how to group clusters dynamically and flexibly according to the current network environments? How to determine the optimal cluster size? What's the optimal channel access schemes among clusters? The solutions to all these problems are critical for application in real networks.

2) Centralized Coordination: Centralized interference management schemes are more efficient due to their low coordination overhead. However, they incline to utilizes the wired AP backbone to share the transmitted or received signals among APs, so as to apply proper MIMO PHY techniques to achieve high throughput. These designs are only suitable for the enterprise WLANs, which naturally have the wired backbone (the Ethernet) for AP management. For the other kind of WLANs, such as the home WLANs where APs are connected through Internet, these backbone-based interference management mechanisms may not be suitable due to the inefficient information exchange among the APs, such as the high delay and low bandwidth.

Actually, there are already some advances on centralized coordination in home WLANs. For example, COAP [102] proposes a cloud-based centralized framework for home AP coordination and management, using an open API implemented by commodity APs. It is implemented based on the architecture of software defined network (SDN) and designs a cloud-based SDN controller to coordinate among the individual APs. It would be a choice to exploit this kind of architecture for MIMO-related interference management in the non-enterprise WLANs, but the performance of the cloud-based information exchange needs to be evaluated carefully.

3) Utilization of MIMO PHY Techniques: The two advanced MIMO physical layer techniques, interference alignment and interference nulling, are widely exploited to increase the concurrent transmission streams and avoid interference among BSSs. Current advances have demonstrated significant throughput improvement through exploiting these techniques, no matter adopting the centralized or distributed channel access schemes. The major concern about these schemes is their ability to be deployed in real networks. The following two considerations need further study. At first, overhead induced by the extra CSI requirements should be reduced, especially in the mobility scenarios; at second, distinct methods should be provided to select the optimal links and schedule them to transmit concurrently, so as to achieve interference alignment and nulling at the receivers.

4) Optimization of Heterogeneous Transmissions: Current channel access scheme is inefficient when heterogeneous transmission are dominant in the network. As discussed in acPad [42], data streams always have heterogeneous transmission duration which is determined by the packet length and the transmission rate. If they are grouped into one MU-MIMO transmission, channel resources are wasted due to the idle time after part of the streams. acPad proposes to solve this problem through filling up the idle channel time with additional packets. Since it only considers the single collision domain, the problem in the multiple collision domain needs to be investigated.

Besides the transmission duration, data streams often occupy different channel bandwidth. As shown in Table II, the maximum bandwidth has been increased from 20MHz in 802.11a/b/g to 40MHz in 802.11n, and even 160MHz in 802.11ac. The 802.11ac recommends a dynamic bandwidth operation, which makes the 160MHz divided into a set of narrow bandwidth (20MHz, 40MHz, and 80MHz), and utilizes the exchange of RTS and CTS to reserve a specific bandwidth for each data transmission. However, this operation is only suitable to the simple scenario where RTS/CTS is rare to be collided and their transmission overhead is not serious. Designing efficient mechanisms to solve the heterogeneous bandwidth problem in the complex real networks, especially the multiple collision domain scenarios, is also worthy further study.

5) Full Duplex Transmission: Nowadays, the wireless full duplex transmission, which makes a station transmit one signal and receive another signal simultaneously through two antennas at the same frequency band, is becoming a promising communication technology [103]

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

and may totally change the traditional channel access methods [104]–[106]. Some advances already exploit this technology when designing MIMO-related interference management mechanisms. For example, FlexRadio [71] makes the antennas of a node fully flexible to transmit or receive signals, thus increases concurrent streams, iBeam [93] makes a station transmit an ACK packet while receiving packets from APs. Actually, the full duplex transmission can be applied to many other scenarios to improve the network throughput, such as providing the instant CSI feedback for precoding [107], obtaining prompt information from the data receiver for efficient interference management, etc. We consider the introduction of wireless full duplex transmission will lead to more efficient interference management in WLANs.

B. Channel Sounding

One of the key problems in the MIMO transmission is CSI acquisition, which will be used for both precoding and user selection. There is a clear dilemma between the CSI accuracy and its overhead, as the accurate CSI needs frequent information exchange, which induces more transmission overhead. Especially, the adoption of interference alignment and interference nulling may bring more CSI acquisition requirements, which will further increase the overhead.

This problem may be even worse in the multiple collision domain scenarios. As shown in Fig. 3(b), explicit channel sounding is a typical point coordination process, as one AP should coordinate all the stations' CSI transmissions. This design can definitely work in the scenario where there is only one BSS and one AP in the network. However, once it is extended to the multiple collision domain, interference among BSSs should be considered as it can affect the accuracy of CSI estimation. Centralized research works [75] make the explicit channel sounding in each BSS interleaved to avoid interference, they would obviously induce huge transmission overhead, especially in large scale networks. Distributed schemes mainly adopt implicit channel sounding due to the problem in its explicit counterpart. Channel sounding will be an everlasting topic in the MIMO-related interference management.

C. User Selection

User selection, which selects the users with the most uncorrelated channels for concurrent transmissions based on the obtained CSI, plays a very important role in both MU-MIMO and network MIMO scenarios. The optimal users can be selected through exhaustive searching, which is impractical due to the high computational complexity. The greedy algorithm is easy to implement but has sub-optimal performance.

Current advances on optimizing the user selection process mainly focus on the single collision domain, through reducing either the channel sounding overhead or the computational complexity. However, most of the MIMO-related interference management mechanisms designed for the multiple collision domain still adopt the greedy algorithm for user selection. It is worthy to study efficient user selection algorithms in this scenario.

VII. Conclusion

In this paper, we investigate the MIMO-related interference management mechanisms from both the single collision domain and multiple collision domain perspective. After giving some background information about the evolution of IEEE 802.11 standards, we first present the PHY mechanisms for MIMO transmissions, including two basic MIMO PHY techniques recommended in the 802.11n/ac standards and two advanced techniques adopted in current research; we also introduce three MIMO application scenarios and present how they can be achieved through the MIMO PHY techniques. Based on that, we discuss three MAC components for achieving high performance through MIMO, including channel access, channel sounding and user selection. After that, we study the MIMO-related interference management mechanisms in the single collision domain from optimizing the three MAC components. We then investigate the MIMO-related interference management mechanisms in the multiple collision domain, which contain both the PHY and MAC layer design and are mainly categorized into centralized, distributed and semi-distributed schemes according to the channel access process, we also present some discussions for each kind of schemes. Finally, we clarify some discussions and future directions related to both efficiently exploiting the MIMO PHY techniques and designing effective MIMO-related interference management mechanisms. We hope this survey would help the readers summarize the current research progress and inspire their future work.

Acknowledgement

This work was supported in part by grants from China NSFC (61702343, 61872248, U1736207, 61603099), Hong Kong RGC (PolyU-521312), Hong Kong PolyU (4-BCB6, M-N020,G-YBXY), Guangdong NSF 2017A030312008, Shenzhen Science and Technology Foundation (No. JCYJ20170302140946299, JCYJ20170412110753954), Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China(Grant No.161064), Guangdong Talent Project 2015TX01X111, PCL Future Regional Network Facilities for Large-scale Experiments and Applications (PCL2018KP001) and GDUPS (2015). This work is partially supported by Tianjin Key Laboratory of Advanced Networking (TANK), School of Computer Science and Technology, Tianjin University, Tianjin China, 300350. Kaishun Wu is the corresponding author.

References

[1] Cisco, "Cisco Visual Networking Index: global mobile data traffic forecast update, 2016 – 2021 white paper," 2017.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

[2] IEEE Computer Society. 802.11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," 1999.

[3] IEEE Computer Society. 802.11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," 2007.

[4] IEEE Computer Society. 802.11, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 5: Enhancements for higher throughput," 2009.

[5] M. S. Gast, "802.11ac: A survival guide," *O'Reilly Media*, 2013.

[6] M. Vutukuru, K. Jamieson, and H. Balakrishnan, "Harnessing exposed terminals in wireless networks," in *Proc. of ACM NSDI*, 2008.

[7] T. Xiong, J. Zhang, J. Yao, and W. Lou, "Symbol-level detection: A new approach to silencing hidden terminals," in *Proc. of IEEE ICNP*, 2012.

[8] J. Yao, T. Xiong, J. Zhang, and W. Lou, "On eliminating the exposed terminal problem using signature detection," *IEEE Transactions on Mobile Computing*, vol. 15, no. 8, pp. 2034 – 2047, Aug. 2016.

[9] J. Lee, H. Lee, Y. Yi, S. Chong, E. W. Knightly, and M. Chiang, "Making 802.11 DCF near-optimal: design, implementation, and evaluation," *IEEE/ACM Transactions on Networking*, vol. 24, no. 3, pp. 1745 – 1758, Jun. 2016.

[10] H. Qiu, K. Psounis, G. Caire, K. M. Chugg, and K. Wang, "High-rate WiFi broadcasting in crowded scenarios via lightweight coordination of multiple access points," in *Proc. of ACM MOBIHOC*, 2016.

[11] A. Baiocchi, I. Tinnirello, D. Garlisi, and A. L. Valvo, "Random access with repeated contentions for emerging wireless technologies," in *Proc. of IEEE INFOCOM*, 2017.

[12] R. Bhardwaj, K. Chintalapudi, and R. Ramjee, "Skip-correlation for multi-power wireless carrier sensing," in *Proc. of NSDI*, 2017.

[13] S. Kumar, V. S. Raghavan, and J. Deng, "Medium access control protocols for ad hoc wireless networks: a survey," *Elsevier on Ad Hoc Networks*, vol. 4, no. 3, pp. 326 – 358, May 2006.

[14] 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 Communications Surveys* & *Tutorials*, vol. 18, no. 4, pp. 2315 – 2344, Fourthquarter 2016.

[15] B. Bellalta, L. Bononi, R. Bruno, and A. Kassler, "Next generation IEEE802.11 wireless local area networks: current status, future directions and open challenges," *Elsvier on Computer Communications*, vol. 75, no. 1, pp. 1–25, Feb. 2016.

[16] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, "Survey of Large-Scale MIMO Systems," *IEEE Communications Surveys* & *Tutorials*, vol. 17, no. 3, pp. 1738 – 1760, Apr. 2015.

[17] R. Liao, B. Bellalta, M. Oliver, and Z. Niu, "MU-MIMO MAC protocols for wireless local area networks: a survey," *IEEE Communications Surveys* & *Tutorials*, vol. 18, no. 1, pp. 162 – 183, Firstquarter 2016.

[18] T. Hiraguri and K. Nishimori, "Survey of transmission methods and efficiency using MIMO technologies for wireless LAN systems," *IEICE Transactions on Communications*, vol. E98-B, no. 7, pp. 1250 – 1267, Jul. 2015.

[19] S. Fu, H. Wen, J. Wu, and B. Wu, "Energy-efficient precoded coordinated multi-point transmission with pricing power game mechanism," *IEEE Systems Journal*, vol. 11, no. 2, pp. 578 – 587, Jun. 2017.

[20] K. Yang, N. Yang, C. Xing, and J. Wu, "Relay antenna selection in MIMO two-way relay networks over Nakagami-*m* fading channels," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2349 – 2362, Jun. 2014.

[21] G. Li, J. Wu, Z. Chen, X. Luo, T. Tang, and Z. Xu, "Performance analysis and evaluation for active antenna arrays under three-dimensional wireless channel model," *IEEE Access*, vol. 6, no. 1, pp. 19 131 – 19 139, Jan. 2018.

[22] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2841 – 2853, May 2015.

[23] A. Goldsmith, *Wireless Communications*, 2004.

[24] T. Brown, E. D. Carvalho, and P. Kyritsi, "Practical guide to the mimo radio channel," *Willy*, 2012.

[25] P. Xiao, J. Wu, and C. F. N. Cowan, "MIMO detection schemes with interference and noise estimation enhancement," *IEEE Transactions on Communications*, vol. 59, no. 1, pp. 26 – 32, Jun. 2011.

[26] K. Nikitopoulos, J. Zhou, B. Congdon, and K. Jamieson, "Geosphere: consistently turning MIMO capacity into throughput," in *Proc. of ACM SIGCOMM*, 2014.

[27] H. Lou, M. Ghosh, P. Xia, and R. Olesen, "A Comparison of implicit and explicit channel feedback methods for MU-MIMO WLAN systems," in *Proc. of IEEE PIMRC*, 2013.

[28] F. Mehran, K. Nikitopoulos, P. Xiao, and Q. Chen, "Eliminating channel feedback in next-generation celluar networks," in *Proc. of ACM SIGCOMM*, 2016.

[29] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3857 – 3868, Oct. 2005.

[30] W.-L. Shen, K. C.-J. Lin, M.-S. Chen, and K. Tan, "SIEVE: scalable user grouping for large MU-MIMO systems," in *Proc. of IEEE INFOCOM*, 2015.

[31] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425 – 3441, Aug. 2008.

[32] N. Zhao, F. R. Yu, M. Jin, Q. Yan, and V. C. M. Leung, "Interference alignment and its applications: a survey, research issues, and challenges," *IEEE Communications Surveys* & *Tutorials*, vol. 18, no. 3, pp. 1779 – 1803, Third Quarter 2016.

[33] B. Koo and D. Part, "Interference alignment with cooperative primary receiver in cognitive networks," *IEEE Communications Letter*, vol. 16, no. 7, pp. 1072 – 1075, Jul. 2012.

[34] X. Li, N. Zhao, Y. Sun, and F. R. Yu, "Interference alignment based on antenna selection with imperfect channel state information in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5497 – 5511, Jul. 2016.

[35] H. Zeng, Y. T. Hou, Y. Shi, and W. Lou, "Shark-IA: an interference alignment algorithm for multi-hop underwater acoustic networks with large propagation delays," in *Proc. of IEEE INFOCOM*, 2013.

[36] H. Zeng, F. Tian, Y. T. Hou, W. Lou, and S. F. Midkiff, "Interference alignment for multihop wireless networks: challenges and research directions," *IEEE Network*, vol. 30, no. 2, pp. 74 – 80, Mar. 2016.

[37] O. Bejarano, E. Magistretti, O. Gurewitz, and E. W. Knightly, "MUTE: sounding inhibition for MU-MIMO WLANs," in *Proc. of IEEE SECON*, 2014.

[38] X. Xie, X. Zhang, and K. Sundaresan, "Adaptive feedback compression for MIMO networks," in *Proc. of ACM MOBICOM*, 2013.

[39] O. Bejarano, S. Quadri, O. Gurewitz, and E. W. Knightly, "Scaling multi-user MIMO WLANs: the case for concurrent uplink control messages," in *Proc. of IEEE SECON*, 2015.

[40] Y. Cai, W. Xi, Z. Wang, K. Zhao, J. Han, C. Qian, H. Ding, and J. Zhao, "Poster: CSI feedback reduction by checking its validity period," in *Proc. of ACM MOBICOM*, 2016.

[41] O. Bejarano, R. P. F. Hoefel, and E. W. Knightly, "Resilient multi-user beamforming WLANs: mobility, interference, and imperfect CSI," in *Proc. of IEEE INFOCOM*, 2016.

[42] C.-H. Lin, Y.-T. Chen, K. C.-J. Lin, and W.-T. Chen, "acPad: enhancing channel utilization for 802.11ac using packet padding," in *Proc. of IEEE INFOCOM*, 2017.

[43] P. Nayak, M. Garetto, and E. W. Knightly, "Multi-user downlink with single-user uplink can starve TCP," in *Proc. of IEEE INFOCOM*, 2017.

[44] "WARP: Wireless Open Access Research Platform," in *https://warpproject.org/trac*.

[45] Ettus Inc, "Universal software radio peripheral." [Online]. Available: http://ettus.com

[46] "Microsoft Research Software Radio (Sora)," in *https://www.microsoft.com/en-us/research/project/microsoft-research-software-radio-sora/*, 2008.

[47] M. Kobayashi, B. Ng, and W. Seah, "TARC: throughput-aware random scalable clustering for network MIMO," in *Proc. of IEEE GLOBCOM*, 2015.

[48] R. E. Guerra, N. Anand, C. Shepard, and E. W. Knightly, "Opportunistic channel estimation for implicit 802.11af MU-MIMO," in *Proc. of ITC*, 2016.

[49] X. Xie and X. Zhang, "Scalable user selection for MU-MIMO networks," in *Proc. of IEEE INFOCOM*, 2014.

[50] N. Anand, J. Lee, S.-J. Lee, and E. W. Knightly, "Mode and User Selection for Multi-User MIMO WLANs without CSI," in *Proc. of IEEE INFOCOM*, 2015.

[51] S. Sur, I. Pefkianakis, X. Zhang, and K.-H. Kim, "Practical MU-MIMO user selection on 802.11ac commodity networks," in *Proc. of ACM MOBICOM*, 2016.

[52] Y. Zeng, I. Pefkianakis, K.-H. Kim, and P. Mohapatra, "MU-MIMO-Aware AP selection for 802.11ac networks," in *Proc. of ACM MOBIHOC*, 2017.

[53] K. Lee, J. Yoo, and C. Kim, "DiFuse: distributed frequency domain user selection for multi-user MIMO networks," *Springer on Wireless Networks*, vol. 1, no. 1, pp. 1 – 18, Aug. 2016.

[54] Z. Chen, X. Zhang, S. Wang, Y. Xu, J. Xiong, and X. Wang, "BUSH: empowering large-scale MU-MIMO in WLANs with hybrid beamforming," in *Proc. of IEEE INFOCOM*, 2017.

[55] T.-W. Kuo, K.-C. Lee, K. C.-J. Lin, and M.-J. Tsai, "Leader-contention-based user matching for 802.11 multiuser MIMO networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4389 – 4400, Aug. 2014.

[56] A. Zhou, T. Wei, X. Zhang, M. Liu, and Z. Li, "Signpost: scalable MU-MIMO signaling with zero CSI feedback," in *Proc. of ACM MOBHOC*, 2015.

[57] Y. Du, E. Aryafar, P. Cui, J. Camp, and M. Chiang, "SAMU: design and implementation of selectivity-aware MU-MIMO for wideband WiFi," in *Proc. of IEEE SECON*, 2015.

[58] S. Gollakota, S. D. Perli, and D. Katabi, "Interference alignment and cancellation," in *Proc. of ACM SIGCOMM*, 2009.

[59] F. Adib, S. Kumar, O. Aryan, S. Gollakota, and D. Katabi, "Interference Alignment by Motion," in *Proc. of ACM SIGCOMM*, 2013.

[60] E. Aryafar, N. Anand, T. Salonidis, and E. W. Knightly, "Design and experimental evaluation of multi-user beamforming in wireless LANs," in *Proc. of ACM MOBICOM*, 2010.

[61] M. Esslaoui, F. R. Palou, and G. Femenias, "A fair MU-MIMO scheme for IEEE 802.11ac," in *Proc. of ISWCS*, 2012.

[62] C. Shepard, H. Yu, N. Anand, L. E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: practical many-antenna base stations," in *Proc. of ACM MOBICOM*, 2012.

[63] X. Xie, X. Zhang, and E. Chai, "Cross-cell DoF distribution: combating channel hardening effect in multi-cell MU-MIMO networks," in *Proc. of ACM HotMobile*, 2015.

[64] Y. Shi, J. Liu, C. Jiang, C. Gao, and Y. T. Hou, "A DoF-based link layer model for multi-hop MIMO networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1395 – 1408, Jul. 2014.

[65] Q. Yang, X. Li, H. Yao, J. Fang, K. Tan, W. Hu, J. Zhang, and Y. Zhang, "BigStation: enabling scalable real-time signal processing in large MU-MIMO systems," in *Proc. of ACM MOBICOM*, 2013.

[66] X. Xie, E. Chai, and X. Zhang, "Hekaton: efficient and practical large-scale MIMO," in *Proc. of ACM MOBICOM*, 2015.

[67] C. Husmann, G.Georgis, K. Nikitopoulos, and K. Jamieson, "Flex-Core: massively parallel and flexible processing for large MIMO access points," in *Proc. of NSDI*, 2017.

[68] T. Bansal, W. Zhou, K. Srinivasan, and P. Sinha, "RobinHood: sharing the happiness in a wireless jungle," in *Proc. of ACM HotMobile*, 2014.

[69] W. Zhou, T. Bansal, P. Sinha, and K. Srinivasan, "BBN: throughput scaling in dense enterprise WLANs with bind beamforming and nulling," in *Proc. of ACM MOBICOM*, 2014.

[70] P. Yang, Y. Yan, X. Li, and Y. Zhang, "POLYPHONY: scheduling-free cooperative signal recovery in enterprise wireless networks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 9, pp. 2599 – 2610, Sep. 2017.

[71] B. Chen, V. Yenamandra, and K. Srinivasan, "FlexRadio: fully flexible radios and networks," in *Proc. of NSDI*, 2015.

[72] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. of ACM SIGCOMM*, 2013.

[73] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proc. of ACM MOBICOM*, 2010.

[74] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. of ACM MOBICOM*, 2011.

[75] H. Rahul, S. Kumar, and D. Katabi, "JMB: scaling wireless capacity with user demands," in *Proc. of ACM SIGCOMM*, 2012.

[76] H. V. Balan, R. Rogalin, A. Michaloliakos, K. Psounis, and G. Caire, "Achieving high data rates in a distributed MIMO system," in *Proc. of ACM MOBICOM*, 2012.

[77] S. Yun, L. Qiu, and A. Bhartia, "Multi-point to multi-point MIMO in wireless LANs," in *Proc. of IEEE INFOCOM*, 2013.

[78] F. Adib, S. Kumar, O. Aryan, S. Gollakota, and D. Katabi, "AirShare: distributed coherent transmission made seamless," in *Proc. of IEEE INFOCOM*, 2015.

[79] E. Hamed, H. Rahul, M. A. Abdelghany, and D. Katabi, "Real-time Distributed MIMO Systems," in *Proc. of ACM SIGCOMM*, 2016.

[80] C. Windpassinger, R. Fischer, T. Vencel, and J. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Transactions on Wireless Communications*, vol. 3, no. 4, pp. 1305 – 1316, Jul. 2004.

[81] T. Gou, C. Wang, and S. A. Jafar, "Aiming perfectly in the dark - blind interference alignment through staggered antenna switching," in *Proc. of IEEE GLOBCOM*, 2010.

[82] A. Michaloliakos, R. Rogalin, V. Balan, L. Psounis, and G. Caire, "Efficient MAC for distributed multiuser MIMO systems," in *Proc. of WONS*, 2013.

[83] H. V. Balan, R. Rogalin, A. Michaloliakos, K. Psounis, and G. Caire, "AirSync: enabling distributed multiuser MIMO with full spatial multiplexing," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1681 – 1695, Dec. 2013.

[84] F. Adib, S. Kumar, O. Aryan, S. Gollakota, and D. Katabi, "Poster: clock synchronization for distributed wireless protocols at the physical layer," in *Proc. of ACM MOBICOM*, 2014.

[85] K. Lin, S. Gollakota, and D. Katabi, "Random access heterogeneous mimo networks," in *Proc. of ACM SIGCOMM*, 2011.

[86] K. Lin, Y.-J. Chuang, and D. Katabi, "A light-weight wireless handshake," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 2, pp. 28–34, Apr. 2012.

[87] K. Lee, J. Yoo, Y. Kang, and C.-K. Kim, "802.11mc: using packet collision as an opportunity in heterogeneous MIMO-based Wi-Fi networks," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 1, pp. 287 – 302, Jan. 2015.

[88] B.-S. Chen, K. C.-J. Lin, S.-L. Chiu, R. Lee, and H.-Y. Wei, "Multiplexing-diversity medium access for multi-user MIMO networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1211 – 1223, May. 2016.

[89] S. Sen and R. R. Choudhury and S. Nelakuditi, "No time to countdown: migration backoff to frequency domain," in *Proc. of ACM MOBICOM*, 2011.

[90] J. Xiong, K. Sundaresan, K. Jamieson, M. A. Khojastepour, and S. Rangarajan, "MIDAS: empowering 802.11ac networks with multiple-input distributed antenna systems," in *Proc. of ACM CoNEXT*, 2014.

[91] K. Tan, H. Liu, J. Fang, W. Wang, J. Zhang, M. Chen, and G. M. Voelker, "SAM: enabling practical spatial multiple access in wireless LAN," in *Proc. of ACM MOBICOM*, 2009.

[92] A. B. Flores, S. Quadri, and E. W. Knightly, "A scalable multi-user uplink for Wi-Fi," in *Proc. of NSDI*, 2016.

[93] Y. Du, E. Aryafar, J. Camp, and M. Chiang, "iBeam: intelligent client-side multi-user beamforming in wireless networks," in *Proc. of IEEE INFOCOM*, 2014.

[94] K. C.-J. Lin, W.-L. Shen, M.-S. Chen, and K. Tan, "User-Centric Network MIMO With Dynamic Clustering," *IEEE/ACM Transactions on Networking*, vol. 25, no. 3, pp. 1910 – 1923, Jun. 2017.

[95] H. Yu, O. Bejarano, and L. Zhong, "Combating inter-cell interference in 802.11ac-based multi-user MIMO networks," in *Proc. of ACM MOBICOM*, 2014.

[96] Z. Zeng, Y. Gao, K. Tan, and P. R. Kumar, "CHAIN: introducing minimum controlled coordination into random access MAC," in *Proc. of IEEE INFOCOM*, 2011.

[97] S. Kumar, D. Cifuentes, S. Gollakota, and D. Katabi, "Bringing cross-layer MIMO to todays wireless LANs," in *Proc. of ACM SIGCOMM*, 2013.

[98] X. Zhang, K. Sundaresan, and K. G. Shin, "NEMOx: scalable network MIMO for wireless networks," in *Proc. of ACM MOBICOM*, 2013.

[99] T. Wei and X. Zhang, "Random access signaling for network MIMO uplink," in *Proc. of IEEE INFOCOM*, 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2019.2894160, IEEE Communications Surveys & Tutorials

[100] N. Feamster, J. Rexford, and E. Zegura, "The road to SDN: an intellectual history of programmable networks," *ACM Magazine*, vol. 11, no. 12, Dec. 2013.

[101] W.-L. Shen, K. C.-J. Lin, M.-S. Chen, and K. Tan, "Client as a first-class citizen: practical user-centric network MIMO clustering," in *Proc. of IEEE INFOCOM*, 2016.

[102] A. Patro and S. Banerjee, "Outsourcing coordination and management of home wireless access points through an open API," in *Proc. of IEEE INFOCOM*, 2015.

[103] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: from the perspective of PHY and MAC layers," *IEEE Communications Surveys* & *Tutorials*, vol. 17, no. 4, pp. 2017 – 2046, Fourth Quarter 2015.

[104] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks?" in *Proc. of IEEE INFOCOM*, 2014.

[105] S. Wang, V. Venkateswaran, and X. Zhang, "Exploring full-duplex gains in multi-cell wireless networks: a spatial stochastic framework," in *Proc. of IEEE INFOCOM*, 2015.

[106] L. Chen, F. Wu, J. Xu, K. Srinivasan, and N. Shroff, "BiPass: enabling end-to-end full duplex," in *Proc. of ACM MOBICOM*, 2017.

[107] Z. Qian, F. Wu, Z. Zheng, K. Srinivasan, and N. B. Shroff, "Concurrent channel probing and data transmission in full-duplex MIMO systems," in *Proc. of ACM MOBIHOC*, 2017.

Junmei Yao received the B.E. degree in Communication Engineering from Harbin Institute of Technology, China in 2003, the M.E. degree in Communication and Information System from Harbin Institute of Technology, China in 2005, and the Ph.D. degree in Computer Science from the The Hong Kong Polytechnic University, Hong Kong, 2016. She is currently an assistant professor in the College of Computer Science and Software Engineering, Shenzhen University, China. Her research interests include wireless networks, wireless communications and RFID systems.

Jun Xu received the B.E. degree in Computer Science from Yanshan University, China in 2002, the M. E and Ph.D degree in Computer Architecture from the Institute of Computing Technology, Chinese Academy of Sciences in 2005 and 2011 respectively. He is currently a Research Associate in the College of Computer Science and Software Engineering, Shenzhen University, China. His research interests include wireless communication, computer architecture and mobile computing.

Sheng Luo received the B.Eng. and M.Eng. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2009 and 2012, respectively, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2017. Since December 2017, he has been with Shenzhen University where he is currently an Assistant Professor in the College of Computer Science and Software Engineering. His current research interests are in the areas of cooperative communication, buffer-aided relaying and wireless information and power transfer, mmWave communication and spatial modulation.

Lu Wang is currently an assistant professor in College of Computer Science and Software Engineering, Shenzhen University. She received the B.S. degree in communication engineering from Nankai University in 2009,Ph.D. degree in computer science and engineering from Hong Kong University of Science and Technology in 2013. Her research interests focus on wireless communications and mobile computing.

Chao Yang received the Ph.D. degree in Signal and Information Processing from South China University of Technology, Guangzhou, China, 2013. He currently work in the School of Automation, Guangdong University of Technology. From July, 2014 to July, 2016, he was a Research Associate in the Department of Computing, The Hong Kong Polytechnic University. His research interest focuses on VANETs, smart grid.

Kaishun Wu received his Ph.D. degree in computer science and engineering from HKUST in 2011. After that, he worked as a research assistant professor at HKUST. In 2013, he joined SZU as a distinguished professor. He has co-authored 2 books and published over 90 high quality research papers in international leading journals and primer conferences, like IEEE TMC, IEEE TPDS, ACM MobiCom,IEEE INFOCOM. He is the inventor of 6 US and over 80 Chinese pending patent. He received 2012 Hong Kong Young Scientist Award, 2014 Hong Kong ICT Awards: Best Innovation and 2014 IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award. He is an IET Fellow.

Wei Lou is currently an associate professor in the Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China. He received the B.E. degree in Electrical Engineering from Tsinghua University, China in 1995, the M.E. degree in Telecommunications from Beijing University of Posts and Telecommunications, China in 1998, and the Ph.D. degree in Computer Engineering from Florida Atlantic University, USA in 2004. Dr. Lou's current research interests are in the areas of wireless networking, mobile ad hoc and sensor networks, peer-to-peer networks, and mobile cloud computing. He has worked intensively on designing, analyzing and evaluating practical algorithms with the theoretical basis, as well as building prototype systems. His research work is supported by several Hong Kong GRF grants and Hong Kong Polytechnic University ICRG grants.