Cloud Based mmWave WLANs: Architectural Paradigms, Proposals and Perspectives

Kaijun Cheng, Xuming Fang, and Xianbin Wang

Abstract

With the explosive growth of data traffic over wireless access networks, the use of the millimeter-wave band becomes inevitable due to the spectrum scarcity in lower frequencies and the availability of large chunks of underutilized spectrum in this band. The rapid evolution of wireless technologies and co-existence of legacy and emerging wireless infrastructures are leading to complex network operation scenarios. For instance, to achieve the orchestration of co-existed networks and a large number of devices, dynamic operations of conventional distributed WLANs will become extremely challenging. Various technologies such as dense network deployment, dual-band cooperation, C-RAN, and AI, which have been developed for cellular networks, could be adopted to achieve intelligent and efficient WLAN operations. To meet the future traffic requirements of evolving WLAN, this article first outlines the architecture of future WLAN and the corresponding challenges. A centralized control architecture, named WLAN C-RAN, is proposed to achieve orchestrated coordination and improved network throughput. Critical technologies of WLAN C-RAN, for example, dual-band protocol stack and resource management schemes are also developed, and the simulation results demonstrate the improvements on system capacity. Finally, we present some potential challenges and feasible solutions for the future WLAN.

Introduction

In conventional wireless local area networks (WLANs), the low frequency (LF) transmission in the 2.4 and 5.8 GHz unlicensed bands (also called LF interchangeably) is relatively reliable, and the related technologies are well studied. However, the landscape of wireless communications is turning to higher frequency band due to the growing data traffic and the spectrum crunch in LF. High frequency (HF) millimeter-wave (mmWave) band (also called HF interchangeably) with range of 30-300 GHz, has tremendous underutilized unlicensed frequency resources and has already been adopted by IEEE 802.11ad/ay standards (802.11ad/ay for short) [1, 2]. However, high attenuation of mmWave band signal significantly limits its reliable communication distance in line-of-sight (LoS) scenarios [3]. Moreover, mmWave communication is more vulnerable to complex environments since it can be severely affected by blockages due to poor diffraction. To overcome these constraints, beamforming is adopted in mmWave band communications to provide high directional gain and compensate the signal attenuation [1, 2]. In addition, in order to solve the blockage problem of mmWave communication, the method of finding redundant links is leveraged in beam tracking, HF and LF cooperation and relay operation [1, 4]. Beam tracking allows both transmitter and receiver to measure the neighbor backup beams of the current beam by appending training units. The Fast Session Transfer (FST) mechanism allows the ongoing HF traffics between transmitter and receiver to be transferred to LF in case of blockages. The relay operation can be regarded as another remedy method in which the transmitter will find a relay station to form multi-hop transmission to the receiver.

During the last two decades, the IEEE 802.11 based WLAN has developed rapidly around the world. This low cost and highly efficient access technology has become the mainstream of indoor wireless networks, and will play a critical role in enabling smart cities. In some bandwidth-hungry scenarios, such as enterprises, factories, university campuses, stadiums and libraries, the dense deployment of WLAN is greatly needed to improve the system throughput by improving the signal quality and frequency reuse factor. In addition, high-performance products and new 802.11 standards with appealing features will be introduced to meet manifold traffic demands [5]. Given the different variation of 802.11 standards, this article only considers 802.11ad/ay and 802.11ax. Among them, 802.11ad is a groundbreaking WLAN standard to leverage the mmWave 60 GHz unlicensed band for pencil-beam directional communication and very high transmission rate [1]. While operating in the 2.4 and 5 GHz bands, 802.11ax is a great leap in achieving ubiquitous connectivity. It allows stations (STAs) to transmit concurrently by using some mechanisms such as basic service set color and multiple network allocation vectors [6]. Thus, it can improve spatial sharing and transmission efficiency in dense network (DN). Furthermore, 802.11ay is an extension of 802.11ad, intending to provide ultra-high-speed directional communications [2], and it paves the way for beamforming and beam tracking, channel bonding and aggregation, single-user/multi-user multiple-input multiple-output, spatial sharing and so on [4].

Current WLANs are operated in a de-centralized way with independent working access points (APs). The benefits of this framework include low

Digital Object Identifier: 10.1109/MWC.001.1900062

Kaijun Cheng and Xuming Fang are with Southwest Jiaotong University; Xianbin Wang is with Western University.

deployment cost and the elimination of signaling overhead for inter-AP coordination. However, the de-centralized operation cannot support interference minimization and resource scheduling among co-existing APs in a complex environment. In particular, with current de-centralized operation mechanisms, it is impossible to support mobility through coordinated APs. When an STA moves, the coordination among multiple APs close to the STA plays a crucial role in guaranteeing seamless handover. In this regard, an efficient central control and unified coordination architecture is necessary to achieve global optimization of resource allocation. Along this direction, Yang *et al*. [7] solved load balance and access problems under the Access Controller-AP architecture in WLANs to improve the aggregated throughput and user fairness; however, only the LF band operations were considered in [7]. Dely *et al*. [8] proposed a CloudMAC architecture that could be integrated with OpenFlow to achieve similar performance as normal WLANs while implementing novel services faster and easier. Dezfouli *et al.* [9] summarized the central control mechanisms operating on Software-Defined WLAN and identified some key challenges and solutions for user association and channel assignment. However, only the network layer operations were considered in these works instead of the MAC layer, and they did not consider the trends and characteristics of future WLAN, such as mmWave communication and dual-band cooperation in 802.11ad/ay.

To the best of our knowledge, the combination of cloud-based centralized architecture in mmWave WLAN has not been well investigated. In this article, the WLAN architecture and the state-of-the-art technologies corresponding to the latest 802.11 standards are summarized. The related challenges based on future traffic requirements and evolution trends of WLAN are also discussed. Meanwhile, a central control framework named *WLAN cloud radio access network (W-CRAN)* is proposed, considering both the reliability of LF and the high rate directional transmission of HF. In *W-CRAN*, the control functionalities of AP are transferred to the *central control units (CCUs)*, which can get massive information (e.g., link quality, channel state information, and AP load condition) from APs and STAs to allocate wireless resources optimally. Then, the beam management and AP selection schemes are proposed, and its performance simulation is given. Furthermore, some future perspectives are presented. Finally, the article is concluded.

Key Issues for Future WLANs Architecture

This section gives an overview of the state-of-the-art technologies in current WLAN architecture, and introduces some potential technologies for future WLAN design.

Dense Network and AP Coordination

In supporting the ever-growing data traffic in WLAN, multiple APs will be deployed in a hotspot to form a DN. It can improve the system capacity through spatial sharing, realize a blanket signal coverage, and meet the forthcoming trends of massive wireless communication devices. However, densely distributed APs in the network makes the distance between AP and STA even smaller, thus it will increase the interference of STAs since there are more APs transmitting in the downlink concurrently. Meanwhile, more operating APs means more energy consumption of the network [10]. Therefore, in achieving effective interference coordination and flexible resource scheduling, full coordination among neighboring APs with the help of a central controller becomes necessary. The IEEE 802.11ay draft and the related proposals have taken this coordination as a solution, aiming to improve the wireless resources utilization and reduce interference among devices, which are still in their infancy. Apparently, it is imperative to design an efficient AP coordination or centralized control mechanism in DN to provide much larger coverage under certain interference level.

Dual-Band Cooperation

Due to the limited coverage of mmWave signal and its instability, it is generally appropriate for short distances and LoS environments. However, the LF band has the benefits of wider coverage and stronger signal diffraction capability. The dual-band integration of HF and LF, which possesses both high capacity and high robustness, has been widely studied in cellular networks [11]. Nevertheless, it has not been investigated well in WLAN. Although 802.11ad defines FST operation [1], enabling the device to switch between HF and LF, it can only allow the device to work on the same frequency band at a time. Therefore, in the future WLAN, it is necessary to consider that the device can concurrently work on two or more frequency bands, which not only ensures the transmission stability but also improves the transmission rate. Meanwhile, the control information will be transmitted only once in an omnidirectional pattern through LF, so it can avoid being transmitted to all beam directions in a directional pattern through HF. In this way, the substantial signaling overhead and delay can be saved. Such a cooperation mechanism can be leveraged in WLAN to implement beam training, beam tracking, and data transmission. For the matter of handover, the LF can ensure the seamless switching of the STA between the source AP and target AP by maintaining control information unbroken. Consequently, how to improve the STAs' quality of experience (QoE) by using dual-band cooperation remains a promising research topic.

Due to the limited coverage of mmWave signal and its instability, it is generally appropriate for short distances and LoS environments. However, the LF band has the benefits of wider coverage and stronger signal diffraction capability. The dual-band integration of HF and LF, which possesses both high capacity and high robustness, has been widely studied in cellular networks. Nevertheless, it has not been investigated well in WLAN.

Cloud Radio Access Network

Figure 1 shows the evolution of wireless network architecture. In a cellular network, C-RAN is a typical centralized control architecture which transfers the building baseband unit (BBU) of each base station (BS) into a central office [12]. These BBUs connect with remote radio heads (RRHs) via optical fibers. The RRHs only have simple physical radio frequency (RF) and are commonly deployed in hotspots to provide high data rate services. In this way, it can reduce the cost of network deployment, energy consumption, and system complexity. Furthermore, user mobility management becomes simple and effective under the control of BBUs. However, C-RAN has limited fronthaul capacity, which is detrimental for the total throughput of the system.

To address this limitation, heterogenous C-RAN (H-CRAN) is proposed [13], which takes advantage of heterogeneous networks (HetNets) and C-RAN.

While AI is a better alternative which can get the historical data from wireless network and map it to the agent. Then, through continuous training in the offline stage, it can simulate the relationship between the input and desired output of the agent in the online stage. Thus, it can adjust the resource allocation intelligently, according to the changes of wireless network environment, eventually to achieve real-time processing with simple computations.

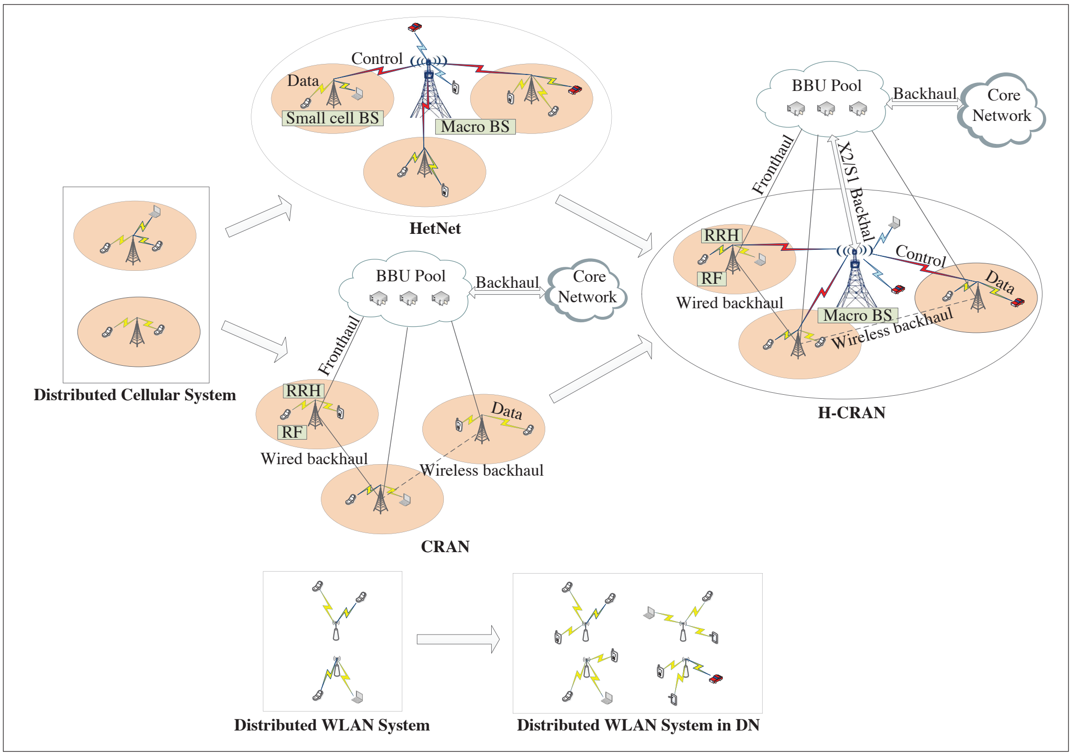


FIGURE 1. The system architecture evolution of cellular network and WLAN.

The control plane and user plane in H-CRAN are decoupled where the control information and system broadcasting information are transferred on the backhaul between BBU Pool and macro BS. Some users can have access to macro BS to alleviate the deficiency of limited fronthaul capacity and delay. Meanwhile, it inherits the merits of C-RAN, which has less network energy consumption, less deploying cost, and improves the spectral efficiency.

Although the existing WLAN standard has been developed from 802.11a/b/n/ac to current 802.11ad/ay/ax, its architecture paradigm is always distributed. That is, STAs are served by the AP whose coverage can reach them. There is no centralized control node to manage the resource of the entire network, and the optimum wireless resource allocation cannot be obtained. In contrast, C-RAN not only saves baseband processing unit resources and improves the efficiency of resource utilization, but also enhances the flexibility of network updating, and reduces capital expenditure and operating expenditure. Therefore, it is necessary to make a foray into cloud-based architecture in WLAN to improve network performance. Thus, the existing protocols and functionalities need to be adjusted on the AP side and transferred to the remote cloud for the centralized control.

Artificial Intelligence Enabled Network Management

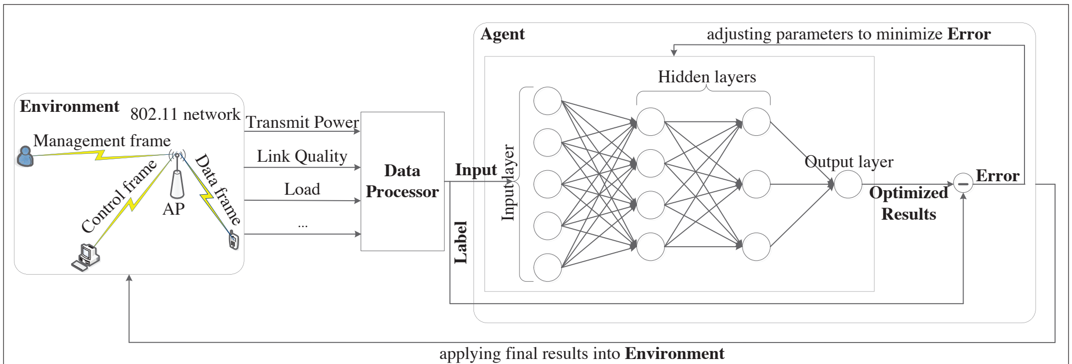
In the era of data deluge, artificial intelligence (AI) is a powerful technique to manage massive information. Using optimization theory to solve the problem of wireless resource allocation requires many complex and nonlinear computations, which results in disparities of theoretical researches and practical applications, and thus it can be hard to achieve real-time processing [14]. While AI is a better alternative which can get the historical data from wireless network and map it to the agent. Then, through continuous training in the offline stage, it can simulate the relationship between the input and desired output of the agent in the online stage. Thus, it can adjust the resource allocation intelligently, according to the changes of wireless network environment, eventually to achieve real-time processing with simple computations.

From the review of the current literature, Sun *et al*. [14] solved an NP hard power allocation problem through a deep neural network (DNN) in a low complexity but high accuracy way. Zhou *et al.* [15] also used DNN to solve the resource allocation for mmWave 802.11ay in terms of STA scheduling, power allocation, and beamwidth. Both simulations proved the superiority of AI over traditional algorithms in terms of computational complexity.

Figure 2 presents a model of AI in WLAN, where DNN is used as an example. In 802.11 networks, if an AP has AI capability, it can use the information (e.g., transmission power, link quality, and load condition) collected from the interactions with STAs, and after data processing, the information can be infused into DNN. The optimized output is obtained through the forward network and the error between this optimized output and the data label obtained from 802.11 networks can be calculated. It is necessary to adjust the parameters in DNN iteratively to reduce this error. After a series of AI training, it can simulate the changes of wireless network in real time and allocate the wireless resources intelligently. Figure 2 also enumerates some problems to be solved with AI in WLAN.

Cloud Based WLAN Architecture

This section introduces a cloud-based WLAN architecture as an infrastructure network and proposes a dual-band protocol stack to meet the demand of high capacity, massive data traffic, and large-scale connections.



Resource allocation	STA association; load balancing; interference mitigation; dual-band cooperation; energy harvesting
Prediction	beamforming and beam tracking; handover; caching; STA location; data traffic; STA behavior
Physical layer	channel estimation; anomaly detection; fault detection; power control

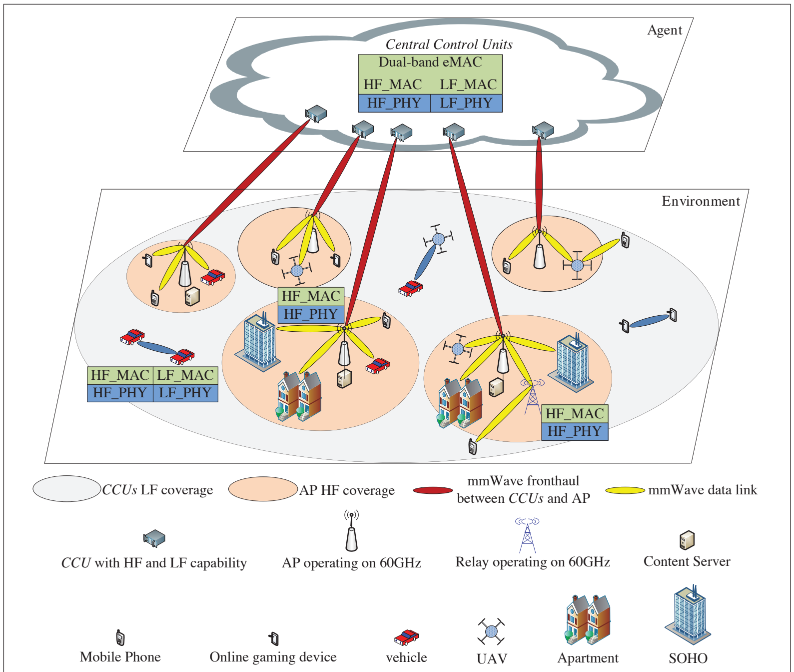
FIGURE 2. An example when AI meets WLAN.

System Architecture

Figure 3a shows the *W-CRAN* architecture. *W-CRAN* is densely deployed with *q* APs which are denoted as {AP 1, AP 2, ⋯, AP *j*, ⋯, AP *q*}, and *n* STAs which are denoted as {STA 1, STA 2, ⋯, STA *i*, ⋯, STA *n*}. Each AP is configured with a content server, which can cache some popular content during the off-peak period to alleviate the transmission delay and save signaling overheads. Each STA supports HF and LF protocols, and they can be mobile phones, cars, unmanned aerial vehicles (UAVs), residential or office hubs, depending on the specific scenarios. In addition, *CCUs* can manage the wireless resources for AP and STA, and its protocol stack is also combined with HF and LF. The *CCUs* can provide reliable connections with STAs through LF links. AP can perform association, authentication, beamforming, and data transmission.

The communication in *W-CRAN* is similar to the control/user plane decoupled architecture in [11]. That is, HF is used to transmit data flows, improving transmission rate, while LF is used to transmit system scheduling information, ensuring the reliability of communication. To reap the full benefits of HF and LF in *W-CRAN*, STAs not only can communicate with *CCUs* through LF (gray shadow in Fig. 3a), which is mainly used for scheduling information transmission to guarantee the QoE without sharp decline, but also can communicate with the serving AP through directional links on the mmWave band (yellow beams in Fig. 3a), which can improve transmission rate and reduce interference. The fronthaul refers to the high-speed mmWave links (red beams in Fig. 3a) between APs and *CCUs*, which carries high rate directional communications. With a certain height of *CCUs* and APs, the directional links between them can be regarded as LoS, which naturally avoid the blockages. Thus, the scheduling information can be transmitted by HF. Figure 3b presents the main functions of each entity in *W-CRAN*.

Specifically, the physical layer (PHY) of *CCUs* includes HF_PHY and LF_PHY, where HF_PHY is used to communicate with APs via mmWave beams, and LF_PHY is used to transmit the scheduling information and some burst data flows to STAs. Similarly, the enhanced medium access control (eMAC) layer is also divided into HF_MAC and LF_MAC. It not only has the conventional MAC functions, but also has unified scheduling functions such as load balance, AP selection, and beam management, to coordinate each AP and STA. These functions are achieved by *dual-band station management entity (DB-SME)* in MAC layer of *CCUs*. *CCUs* can efficiently manage the wireless resources of the entire network via the information from each AP and each STA. The contents cached on servers are determined by *CCUs* based on the STA distribution and traffic demands. The interference



(a)

Capability / Entity	Authentication	Association	Beam Management	Mobility Management	Load Balancing
CCUs	√	√	manage the beam pairs between AP and STA	master the handover decision	master the load balancing decision
AP	√	√	coordinated by *CCUs*		coordinated by *CCUs*
Relay	√	√	receive and forward the data between AP and STA		
STA	√	√	coordinated by *CCUs*	coordinated by *CCUs*	

(b)

FIGURE 3. The *W-CRAN* architecture and its functionalities: a) architecture of *W-CRAN* which is composed of *CCUs*, AP w or w/o content server, and different terminals; b) functions of entities in *W-CRAN*.

mitigation schemes can also be obtained by *CCUs* through the information collected from the network.

Besides, *CCUs* have AI functionality. All the APs and STAs can be regarded as an *Environment* where massive information comes out, while *CCUs* can be regarded as an *Agent*. The real-time optimum management of wireless resources in

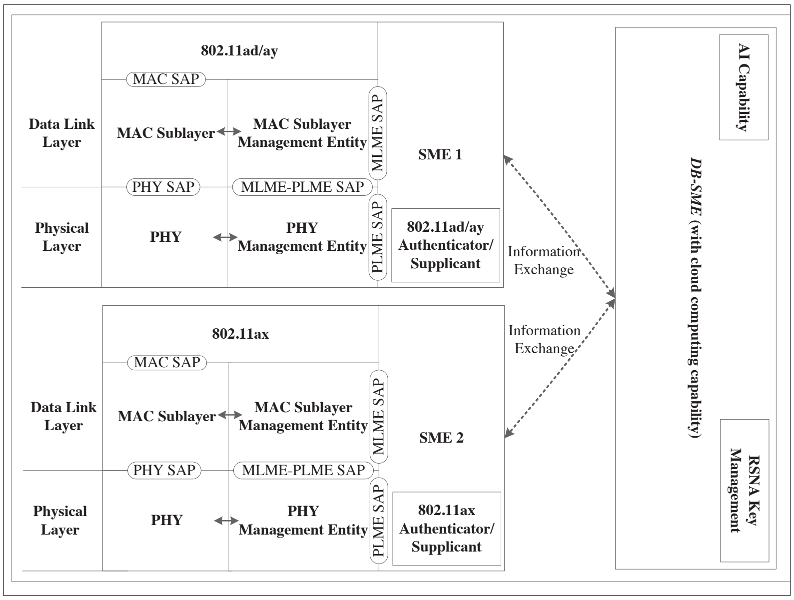


FIGURE 4. Dual-band protocol stack framework.

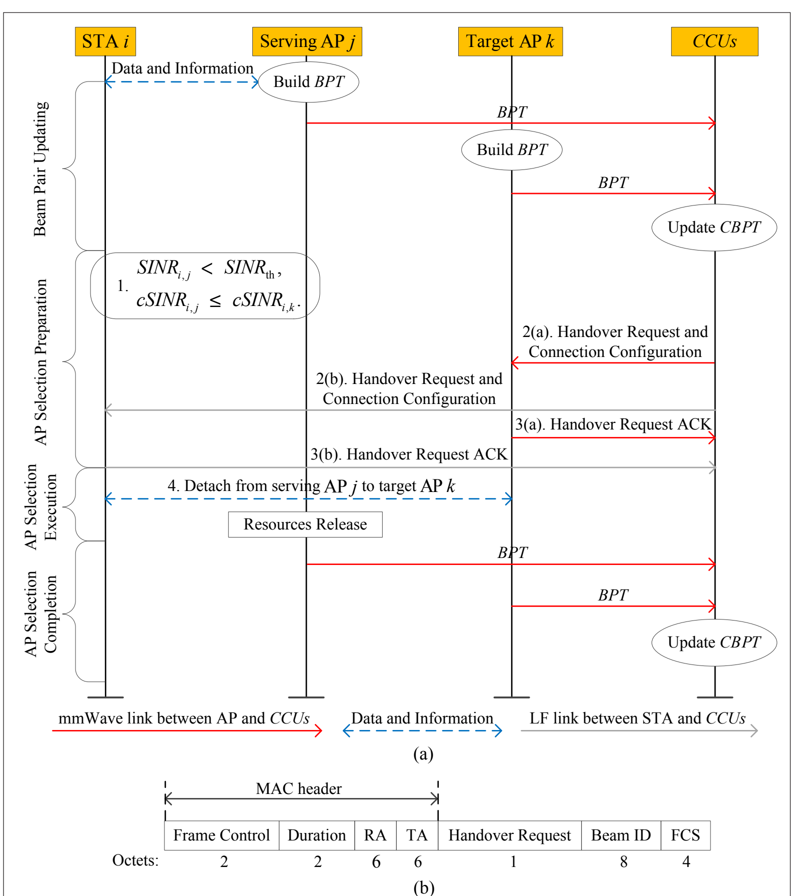


FIGURE 5. The AP selection in *W-CRAN*: a) AP selection process; b) handover request frame.

W-CRAN can be implemented by AI methods such as reinforcement learning and DNN. This will be discussed later.

Other characteristics of *W-CRAN* are summarized as follows:

- The function of AP is degenerated by transferring the coordination function to *CCUs*. AP can report the collected information to *CCUs*, which will dispatch the scheduling decisions to APs and STAs through HF and LF, respectively. Thus, *CCUs* can transmit directly to STAs through LF, which alleviates the limited capacity of the fronthaul.
- There is a fixed beam pair between *CCUs* and each AP because of their fixed positions. Namely, there are *m* beams doing beamforming training in 802.11ad, while the AP in *W-CRAN* only uses $m - 1$ beams to participate in the training. Meanwhile, the use of wireless fronthauls other than optical fibers, can improve the flexibility of network deployment and upgrade.
- *CCUs* has centralized coordination and access capabilities that can transmit scheduling information and burst data flows to STAs through LF to better support user mobility and relieve the problem of link interruption when HF link is unavailable. Therefore, it can significantly reduce the link interruptions stemming from hard handover happening in conventional WLAN systems and reduce latency caused by blockages.
- *W-CRAN* is upgraded from the original WLAN which only provides single service like data transmission, to a comprehensive infrastructure network which supports the combination of communication, computing, and caching.

However, the *W-CRAN* architecture presents the following challenges:

- Since STAs maintain stable connections with *CCUs* through LF, and only carry out HF handover between APs, a new handover process should be re-designed to achieve seamless handover. In addition, for mobile STAs with relatively fixed movement tracks, the application of AI makes *CCUs* capable of predicting handover incidents proactively, thus improving autonomy and timeliness of the network decision-making.
- It is complicated to implement STA discovery and initial access in conventional WLAN. The existing STA access criterion is generally determined by the maximum signal-to-interference-plus-noise ratio (max-SINR), which will cause load imbalance of some APs. Therefore, it is a challenge to consider dual-band cooperation to accelerate the STA discovery and access process, and relieve the load imbalance problem in the environment of large-scale devices.
- How to improve the throughput of edge STAs on the condition that *CCUs* can well organize APs is urgent to be settled. On the one hand, joint transmission can be implemented by multiple APs; however, it is necessary to consider the interference mitigation of inter-AP. To this end, efficient joint transmission process should be designed. On the other hand, AI can be used to select several APs to serve edge STAs at different time according to their movements and communication environment, so as to maximize the throughput of edge STAs.

Dual-Band Protocol Stack Framework

The MAC layer and PHY layer in IEEE 802.11 both have corresponding entities, and they serve the upper layer via service access point (SAP) [1]. MAC SAP refers to the MAC service access point, which provides services to the logical link control layer. There is a MAC sublayer management entity (MLME), which interfaces with MLME SAP. All the services (e.g., power management, synchronization, authentication, association and channel measurement) provided by MLME must pass through MLME SAP before being sent to the station management entity (SME) for further processing. The PHY sublayer management entity (PLME) SAP is PLME SAP, which provides services to the MAC layer. The PLME performs management of the local PHY functions in conjunction with MLME. To facilitate the unified operation and management of the MAC layer and PHY layer, the SAP in MAC and PHY are controlled by SME.

In view of the characteristics of HF and LF, it is necessary to support both protocols simultaneously in the future WLAN. Figure 4 shows the protocol stack framework of *DB-SME* which makes *CCUs* transmit and receive in HF and LF concurrently. The HF and LF SMEs are controlled by *DB-SME* (dotted arrows in Fig. 4), which is responsible for setup, configuration, teardown, and transfer of sessions from one band to another. Each MAC sublayer has a separate MAC SAP and MLME SAP, and different SMEs have different RFs and PHYs. The two SMEs share the same set of robust security network association (RSNA) key. Once the STA connects to any device (HF AP or LF *CCUs*) in *W-CRAN*, its information can be shared with them. It indeed saves the procedure of second access. The two frequency bands of *CCUs* can report the information (e.g., beam information and link quality status) to *DB-SME* through their respective SMEs, and then the resource allocation results are returned to each SME, and they are finally transmitted to the corresponding AP and STA.

Resource Management Schemes

Based on the *W-CRAN* discussed above, the corresponding beam pairing and updating, and AP selection process will be presented in this section to promote the network performance.

Beam Pairing and Updating

It is assumed that the paths between *CCUs* and APs are basically LoS and fixed in *W-CRAN* because they are relatively high. It is also assumed that STAs are already connected with *CCUs* through LF before the transmission with APs. Thus, the beam pairing and updating process is only considered among APs and STAs.

As discussed before, the load imbalance of APs will degrade the network performance in terms of the unavailable services in overload APs and the redundant wireless resources in underload APs. Therefore, we take the load balance into consideration in the beam pairing and updating process, and define a commuted SINR (cSINR): $cSINR_{i,j} = SINR_{i,j}/N_j$, where $SINR_{i,j}$ is the SINR value between AP j, and STA i, N_j is the number of STAs in AP j. Both the link quality and the number of STAs served by the AP are considered in cSINR. Thus, STAs can be assigned to an AP which has higher quality of service (QoS) and lower load condition. The specific process is divided into two steps:

- Each AP conducts its beamforming training in turn with each STA, which is in line with 802.11ad [1]. Then both AP and STA get their list of the best transmitting and receiving beam pairs. Thus, the *beam pair table (BPT)* at each AP can be obtained, which contains the highest SINR values of the beam pairs between APs and STAs.
- Each AP reports its *BPT* to *CCUs*, then $cSINR_{i,j}$ is calculated by *CCUs* according to *BPT*. Finally, the *complete BPT (CBPT)* which includes the *cSINR* values of n STAs and q APs is established.

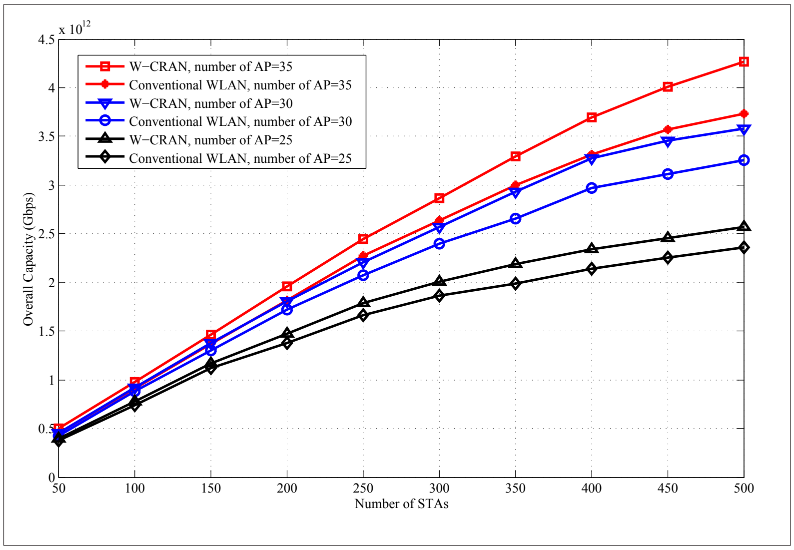


FIGURE 6. Overall capacity comparison between the conventional WLAN and *W-CRAN* in terms of the number of STAs.

AP Selection

For the purpose of finding an AP that can provide higher QoS for STAs, the load of AP is critical in addition to the SINR. This can be implemented effectively in *W-CRAN* because *CCUs* has global information and both the load of AP and SINR can be taken into account. Specifically, as shown in Fig. 5a, each AP periodically reports its *BPT* to *CCUs*. When $SINR_{i,j} < SINR_{\text{th}}$ because of the interference or blockage, STA i should reselect an AP to continue its communication. We assume that AP k is the AP which has the highest *cSINR* with STA i, which means $cSINR_{i,j} \le cSINR_{i,k}$. Then, *CCUs* transmits a Handover Request frame (Handover Request field = 1) through HF as shown in Fig. 5b, to inform the target AP k of the information of STA i and reserving resources. Meanwhile, *CCUs* transmits a Handover Request frame (Handover Request field = 1) through LF, to inform STA i of the information of the target AP k. Wherein RA and TA fields are the receiving and transmitting address respectively, Beam ID field indicates which beam to be switched. If AP k cannot serve any more STAs because of its load condition, then the second highest $cSINR_{i,k'}$ will be selected. Repeat the above searching process until an underload AP is selected. Then, STA

In the future WLAN, AP can be endowed with certain resource management, computing, and storage functionalities in delay-sensitive scenarios. Namely, C-RAN and F-RAN can be combined together. Some computations are offloaded from the central controller to APs and diversified services can be designed according to different needs of STAs.

i hands over to the target AP *k*, and AP *j* releases resources for STA *i*. The authentication and association are omitted because the target AP *k* and STA *i* have already gotten the information of each other through the help of *CCUs*. Furthermore, during the AP selection process, the LF connection with *CCUs* is always valid, which improves the stability of the system.

In our simulation, the transmission power of LF *CCUs* is 23dBm and Binary Phase Shift Keying is selected as the modulation mode. For analytical tractability, we adopt the ideal sectored antenna model in our simulations, which considers the antenna gain to be a constant for angles in the main lobe, and to be a smaller constant in the side lobe [15]. According to [6], *CCUs* can provide 400 meters' signal coverage. There are 500 STAs randomly distributed in *CCUs*. The positions of APs are fixed and each AP has 15 beams available. The transmission power of AP is 43dBm. It is assumed that the transmitting beamwidth and receiving beamwidth are 30 for APs and STAs. The carrier frequency is 60 GHz and the bandwidth is 2.16 GHz, which are consistent with [1, 2]. In Fig. 6, depending on the global information collected by *CCUs*, the network capacity of the proposed scheme outperforms the conventional one which just considers SINR without AP's load condition. And the denser the network is, the wider the performance gap can achieve, which conforms to the development trends of future WLAN.

Future Perspectives

The future WLAN will be more sophisticated and ubiquitous with the appearance of a large volume of data traffic and various applications, and it is urgent to introduce some other potential technologies. We suggest the following three key issues as the future possible research directions.

AI

The target of improving system decision-making efficiency and reducing resource processing latency can be hit through AI. In the future WLAN, the central controller can collect massive information from the network and infuse it into the intelligent agent. Thus, it can respond adaptively and proactively to the network changes (e.g., the number of users and traffic requests) by continuous learning, so as to obtain network-aware resource allocation. Besides, it can make predictions for wireless resource scheduling such as beam alignment and beam tracking to increase the timeliness of the communication link.

While AI can simulate a complex algorithm through simple iterative computations and make the errors as small as possible, there may exist some challenges:

- How to translate the raw data into the AI-driven networks? Because only by establishing a good relationship between them, can we get the expected results for the wireless resource allocation.
- The scale of a DNN is depended on several factors, such as the complexity of the problem, the number of labeled samples and so on. An insufficient number of labeled samples will lead to inaccurate output of the DNN. The higher the number of them, the better the performance of DNN will have. However, when it exceeds a certain range especially in a complex network environment, it will cause overfitting and thereby increase the training time.
- In dynamic scenarios, the topology of wireless networks changes rapidly which will exacerbate the difficulties of its implementation. More advanced techniques such as reinforcement learning and deep reinforcement learning could be considered.

Undoubtedly, massive data are of utmost importance at the juncture between wireless networks and AI. How to combine them effectively under the premise of overcoming the aforementioned challenges is bound to be a nontrivial research direction.

Fog-RAN (F-RAN)

F-RAN solves the deficiency of slow response in C-RAN by sinking the centralized management and computing functionalities of the cloud to the BSs. Thus, these BSs can allocate wireless resources locally with an acceptable performance loss. In this manner, the huge information generated from the access network does not need to be transmitted to the core network, which reduces the number of transmission hops and enhances the end-to-end real-time transmission capacity.

In the future WLAN, AP can be endowed with certain resource management, computing, and storage functionalities in delay-sensitive scenarios. Namely, C-RAN and F-RAN can be combined together. Some computations are offloaded from the central controller to APs and diversified services can be designed according to different needs of STAs. For instance, APs can serve higher QoS STAs, while the central controller can serve lower QoS STAs.

Device-to-Device (D2D)

D2D capable STAs can communicate with each other directly (i.e., the blue beams shown in Fig. 3a). It can reduce latency, network loads and congestions, because the relative signaling no longer goes to the central controller. Therefore, when taking the demands of different STAs into account, this distributed strategy can be integrated into the network. When the central controller participates in network scheduling, global optimization is obtained; when the central controller does not participate in network scheduling for some STAs, D2D can meet the low latency requirements of the future WLAN.

Conclusion

In this article, we have outlined some cutting-edge technologies for future WLAN. Considering both the rapid growth in data traffic and the advantages of centralized control in C-RAN, we have proposed a novel architecture named *W-CRAN*, along with a dual-band protocol stack. In *W-CRAN*, the information can be collected by *CCUs*, and the global optimum resource allocation schemes can be obtained by optimization theory or AI. Then we have put forward the corresponding beam pairing and updating, and AP selection schemes. Simulation results have indicated that the AP selection scheme could improve the network capacity. Finally, we have provided some preliminary insights for the potential research directions to further improve the overall performance for the future WLAN.

Acknowledgment

The work of the authors was partially supported by the NSFC Guangdong Joint Foundation under Grant U1501255; NSFC High-Speed Rail Joint Foundation under Grant U1834210; and 111 Project no.111-2-14.

References

[1] "ISO/IEC/IEEE International Standard for Information Technology — Telecommunications and Information Exchange between Systems-Local and Metropolitan Area Networks — Specific Requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: AMENDMENT 3: Enhancements for Very High Throughput in the 60 GHz Band (Adoption of IEEE Std 802.11ad-2012)," ISO/IEC/IEEE 8802-11: 2012/Amd.3: 2014 (E), Mar. 2014, pp. 1–634.

[2] "IEEE Draft Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications — Amendment 7: Enhanced Throughput for Operation in License-Exempt Bands Above 45 GHz," IEEE P802.11ay/D2.0, July. 2018 (amendment to IEEE Std. 802.11-2016), pp. 1–673.

[3] H. Shokri-Ghadikolaei *et al.*, "Millimeter Wave Cellular Networks: A MAC Layer Perspective," *IEEE Trans. Commun.*, vol. 63, no. 10, Oct. 2015, pp. 3437–58.

[4] P. Zhou *et al.*, "IEEE 802.11ay-Based mmWave WLANs: Design Challenges and Solutions," *IEEE Commun. Surveys Tutorials*, vol. 20, no. 3, Third Quarter 2018, pp. 1654–81.

[5] WiFi Alliance, "WiFi Alliance publishes 7 for '17 WiFi predictions," Jan. 2017; http://www.wi-fi.org/news-events/newsroom/wi-fi-alliance-publishes-7-for-17-wi-fi-predictions, accessed July 4, 2019.

[6] "IEEE Draft Standard for Information Technology-Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks-Specific Requirements — Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Amendment 6: Enhancements for High Efficiency WLAN," IEEE P802.11ax/D3.0, June, 2018 (amendment to IEEE Std 802.11-2016), pp. 1–682.

[7] L. Yang *et al.*, "Demand-Aware Load Balancing in Wireless LANs Using Association Control," *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Diego, CA, USA, Dec. 2015, pp. 1–6.

[8] P. Dely *et al.*, "CloudMAC-An OpenFlow Based Architecture for 802.11 MAC Layer Processing in the Cloud," *Proc. IEEE Globecom Workshops*, Anaheim, CA, USA, Dec. 2012, pp. 186–91.

[9] B. Dezfouli *et al.*, "A Review of Software-Defined WLANs: Architectures and Central Control Mechanisms," *IEEE Commun. Surveys Tutorials*, vol. 21, no. 1, First Quarter 2019, pp. 431–63.

[10] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey," *IEEE Commun. Surveys Tutorials*, vol. 18, no. 4, Fourth Quarter 2016, pp. 2522–45.

[11] L. Yan, X. Fang, and Y. Fang, "Control and Data Signaling Decoupled Architecture for Railway Wireless Networks," *IEEE Wireless Commun.*, vol. 22, no. 1, Feb. 2015, pp. 103–11.

[12] M. Peng *et al.*, "Recent Advances in Cloud Radio Access Networks: System Architectures, Key Techniques, and Open Issues," *IEEE Commun. Surveys Tutorials*, vol. 18, no. 3, Third Quarter 2016, pp. 2282–2308.

[13] M. Peng *et al.*, "Heterogeneous Cloud Radio Access Networks: A New Perspective for Enhancing Spectral and Energy Efficiencies," *IEEE Wireless Commun.*, vol. 21, no. 6, Dec. 2014, pp. 126–35.

[14] H. Sun *et al.*, "Learning to Optimize: Training Deep Neural Networks for Wireless Resource Management," *IEEE Int'l. Wksp. Signal Processing Advances in Wireless Commun. (SPAWC)*, Sapporo, Japan, July. 2017, pp. 1–6.

[15] P. Zhou *et al.*, "Deep Learning-Based Beam Management and Interference Coordination in Dense mmWave Networks," *IEEE Trans. Vehic. Tech.*, vol. 68, no. 1, Jan. 2019, pp. 592–603.

Biographies

KAIJUN CHENG (kaijuncheng@my.swjtu.edu.cn) received the B.E. degree in communication engineering from Southwest Jiaotong University, Chengdu, China, where he is currently pursuing the Ph.D. degree. His research interests include mmWave wireless communications, next generation WiFi, wireless resource optimization and management, and network architecture.

XUMING FANG [M'00, SM'16] (xmfang@swjtu.edu.cn) is a professor at Southwest Jiaotong University, China. He held visiting positions with the Technical University Berlin, Berlin, Germany, and with the University of Texas at Dallas, Richardson, TX, USA. He has published approximately 200 high-quality research papers in journals and conference publications. His research interests include wireless broadband access control, radio resource management, multi-hop relay networks, and broadband wireless access for high-speed railways.

XIANBIN WANG [S'98, M'99, SM'06, F'17] (xianbin.wang@uwo.ca) is a professor and Canada Research Chair at Western University, Canada. His research interests include 5G technologies, communications security, Internet-of-Things, signal processing for communications, and locationing technologies. He has published over 300 peer-reviewed journal and conference papers. He serves as an editor/associate editor for several IEEE journals. He is currently an IEEE Distinguished Lecturer, a Fellow of IEEE, and a Fellow of the Canadian Academy of Engineering.

When taking the demands of different STAs into account, this distributed strategy can be integrated into the network. When the central controller participates in network scheduling, global optimization is obtained; when the central controller does not participate in network scheduling for some STAs, D2D can meet the low latency requirements of the future WLAN.