Hopping on Spectrum: Measuring and Boosting a Large-scale Dual-band Wireless Network

Haibo Wang*†, Tao Gao*†, Weizhen Dang*†, JingAn Xue*†, Jiahao Cao*†, Fenghua Li*† and Jilong Wang*†
* Institute for Network Sciences and Cyberspace, Tsinghua University
† Beijing National Research Center for Information Science and Technology
Email: {hbwang1994, jinganxue, caojh15}@gmail.com,{gt18, dangwz16}@mails.tsinghua.edu.cn, {lifh, wjl}@cernet.edu.cn

***Abstract*—In recent years, more and more wireless networks support both 2.4GHz and 5GHz bands. However, in large-scale dual-band wireless networks, lack of understanding on the behavior and performance makes the network diagnosis and optimization extremely challenging. In this paper, we conduct a comprehensive measurement to characterize the behavior and performance in a large-scale dual-band wireless network (TD_WLAN). We make several meaningful observations. (1) Although the 5GHz band outperforms the 2.4GHz band, 60% of devices tend to be associated with the 2.4GHz band. The device association behavior has a large impact on the performance. (2) Rogue and non-WiFi devices are prevalent, wherein hidden terminal interference increases the average loss rate by 8%, carrier sense interference increases the average WiFi latency by 45%, and RF interference further aggravates both packet loss and channel contention. (3) The dynamic channel assignment strategy is not always effective. On this basis, we propose a novel and easy-to-implement strategy to improve the wireless performance by intelligent band navigation and heuristic channel optimization. The actual deployment in TD_WLAN shows the packet loss reduces by 40% on average and the WiFi latency for more than 60% of devices is below 5ms.**

***Index Terms*—WiFi, measurement, management, performance**

I. Introduction

WiFi has become one of the most popular ways to access the Internet [1]. In recent years, mobile devices experience explosive growth [2]. The statistical report indicates that more than 22 billion devices are going to be connected via wireless networks in 2021 [3]. To meet the ever growing demand, many organizations (*e.g.*, university) deploy large-scale wireless networks to provide high-quality access service. Moreover, with the development of WiFi technologies, more and more wireless networks support both 2.4GHz and 5GHz bands [4], [5]. In such large-scale dual-band wireless environments, network administrators face two typical challenges.

On the one hand, around a large-scale wireless network, there may exist a lot of neighbor 802.11 networks. For example, many users prefer to use self-built access points (APs) to get stronger received signal strength indication (RSSI). The configurations of self-built APs (noted as rogue APs hereafter) generally take no account of the surrounding environment, which greatly increases the probability of channel contention and packet collision. In addition to the interference across WiFi enabled devices, there may exist the RF interference caused by many non-WiFi devices, *e.g.*, Bluetooth devices. Different types of interference increase the difficulty of the spectrum resource management. On the other hand, the band characteristics and the association behavior of user devices are highly different in the 2.4GHz and 5GHz bands, which could affect the access performance. Without understanding the behavior and performance in large-scale dual-band wireless networks, it is difficult for network administrators to make accurate decisions for performance optimization.

There are extensive studies that characterize the behavior and performance of different wireless networks [4], [6]–[19]. However, many of them are limited by the scale of measurements [6]–[10], [17]–[19]. Although a few studies conduct measurements on large-scale wireless networks [4], [11]–[16], most of them only focus on the 2.4GHz band [11], [13]–[15]. Recent studies have begun to explore the characteristics of 802.11ac networks [4], [12], but they still lack a comprehensive comparison between 2.4GHz and 5GHz bands. Besides, little research explores to improve the performance in large-scale dual-band wireless networks.

In this paper, we conduct a comprehensive measurement to understand the wireless behavior of a large-scale dual-band wireless network TD_WLAN. It has more than 7000 APs deployed in 54 dormitory buildings of T university. We further evaluate the effect of different wireless behavior on performance in two bands. Based on the measurement study, we propose a novel strategy to improve the wireless performance by optimizing the spectrum resources in TD_WLAN.

The main contributions can be summarized as follows:

(1) Data Collection. We design an efficient data collection method to get wireless metrics without any dedicated measurement hardware. Compared to the Simple Network Management Protocol (SNMP) based polling method, our method reduces the average CPU usage of wireless access controllers by about 25% and improves the data collection efficiency by 16 times.

(2) Measurement. We conduct a large-scale measurement on the wireless behavior and present a systematic analysis on the wireless performance in TD_WLAN. We find that although the 5GHz band significantly outperforms the 2.4GHz band in both loss rate and WiFi latency, 60% of devices tend to be associated with the 2.4GHz band in dual-band settings. We reveal the prevalence of rogue and non-WiFi devices, most of which work in the 2.4GHz band, and quantify

978-1-6654-4131-5/21/$31.00 ©2021 IEEE

the effect of three types of interference. Hidden terminal interference increases the average loss rate by 8%, carrier sense interference increases the average WiFi latency by 45%, and RF interference further aggravates both channel contention and packet loss. In large-scale dual-band wireless networks, the dynamic channel assignment (DCA) strategy is not always effective to mitigate the effect of interference.

(3) Performance Optimization. We propose a novel strategy from the perspectives of band selection and channel configuration to improve the overall wireless performance, which could be implemented by easy-to-operate parameter settings. First, the strategy could achieve the intelligent band navigation by learning the directed acyclic graph (DAG) for causal inference. Second, the strategy could achieve the channel optimization by heuristic configurations based on the measurement results. The actual deployment in TD_WLAN shows that the packet loss reduces by 40% on average and the WiFi latency is below 5ms for more than 60% of devices.

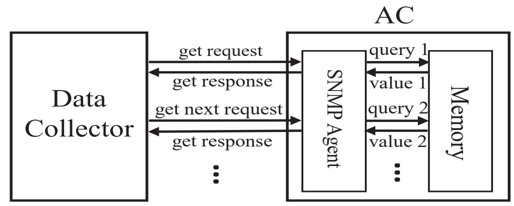
II. Dataset and Methodology

TD_WLAN is one of the largest dual-band wireless networks. There are more than 7000 homogeneous APs deployed indoors in 54 dormitory buildings of T university, which provide network access for more than 45,000 individual users. The APs are generic commercial devices (Huawei 4030DN) and support dual-band with management provided by 13 separate access controllers (ACs). In TD_WLAN, there are 3 available orthogonal channels in the 2.4GHz band and 9 available orthogonal channels in the 5GHz band. TD_WLAN is deployed in dense mode with one AP serving two or three rooms. The placement of APs is based on the signal coverage of the service area in both bands. In this paper, we mainly collect three weeks of data for measurement analysis and evaluation. The data from April 13th, 2019 to April 19th, 2019 is used to characterize the behavior and performance of TD_WLAN in Section III and Section IV. Because the collected data in TD_WLAN presents a weekly pattern, one week of data is representative for the measurement study. The data from April 20th, 2019 to April 26th, 2019 is used to evaluate the effectiveness of the DCA strategy in Section IV-B. The data from May 11th, 2019 to May 17th, 2019 is used to evaluate the proposed optimization strategy in Section V-E. Besides, to achieve the intelligent band navigation in Section V-B, we collect data of one month before April 13th, 2019 for model training. Inquiries about data sharing may be directed to the authors of the paper.

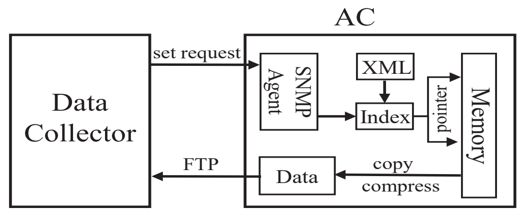
Ethical issues: The data collection is authorized by *Network Operation Center* in T university and we have signed a confidentiality agreement to ensure no private data leakage. We anonymize all privacy information such as IP and MAC addresses. User identifiable information cannot be traced back.

A. Efficient Data Collection

SNMP is widely used to monitor the network device status. For a given object identifier (OID), we can initiate SNMP requests and get object values in the form of key-value pairs



(a)



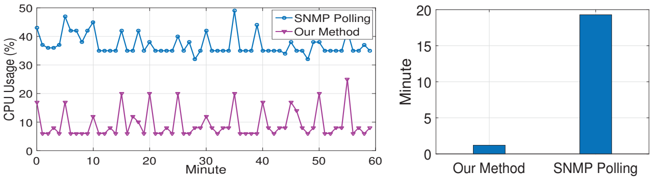
(b)

Fig. 1. (a). The packet interaction process of SNMP based polling method for obtaining all key-values pairs of **each object**. (b). The sketch of our proposed data collection method.

[20]. Commercial ACs generally support SNMP and would organize the monitored wireless objects in the form of SNMP with the device MAC address or the AP MAC address as the key value. The wireless monitor information measured by commercial ACs is relatively accurate and valuable for characterizing the wireless behavior and performance.

To collect these objects, one possible method is to implement an SNMP based polling tool [11], [21], [22]. However, it does not work well in large scale wireless networks such as TD_WLAN. First, the SNMP service has a low priority in ACs. We can only serially get all required objects in each round of polling, which leads to low data collection efficiency. Besides, the data of each object needs to be transmitted by a large amount of SNMP packet interactions. Fig. 1(a) shows the packet interaction process for obtaining all key-values pairs of each object. At the beginning, the collector initiates an SNMP *GET* request. The SNMP agent in AC will query the first key-value pair and transmit it back by the SNMP response. After that, the collector will continue to initiate a series of *GET-NEXT* requests until all key-value pairs of the object are obtained. Frequent data requests and responses may lead to excessive CPU usage for ACs, which could affect other important services (*e.g.*, data forwarding). Second, at the rush hour, each object includes a large amount of data. The data acquisition for each object may consume huge time, which will cause different objects in the same round of polling to be out of sync. Even worse, sometimes the next round of polling begins while the current round of polling has not been finished. It will not only cause some requests to be unresponsive, but also greatly limit the polling frequency.

To solve the problems, we design an efficient and lightweight data collection method, which can be implemented as extensions of the SNMP service on commercial ACs. The sketch of the method is shown in Fig. 1(b). We organize the OIDs of all required objects into the form of XML file and transfer it to each AC in advance. Once the XML file is



(a) (b)

Fig. 2. (a). Comparison of CPU usage between the SNMP based polling method and our method. (b). The average time of data collection in each round for the SNMP based polling method and our method.

received, AC will parse it and build the index structure ($e.g.$, pointers) to help quickly find the storage location of different objects in memory, which could avoid the query overhead in Fig. 1(a). Note that although we organize the requested objects into XML format for the convenience of parsing, other types of data organization formats can also be adopted here. In each round of polling, instead of initiating a large amount of SNMP *GET* requests to obtain all key-value pairs for each object, we initiate an SNMP *SET* request to set a predefined state parameter to 1, which would trigger the SNMP agent to directly fetch all required objects from the memory at the same time with the help of the index structure without SNMP packet interactions in Fig. 1(a). After that, the state parameter will be set to 0 until the SNMP *SET* request of the next round of polling is received. After the data for all objects has been prepared, ACs will compress it and send it to the collector by File Transfer Protocol (FTP).

To demonstrate the superiority, we respectively deploy the SNMP based polling method and our proposed method to collect the same wireless objects in two different days. The collection period is set to 5 minutes and we record the CPU usage every 1 minute. Fig. 2(a) shows the CPU usage of the two methods in the same hour. We can see that the CPU usage of our method is almost below 20% and it is only high within the one or two minutes of requests being initiated, while the CPU usage of the SNMP based polling method is always high during the hour. We further compare the average time of data collection in each round for the two methods in Fig. 2(b). We observe that the average time is about 19.3 minutes for the SNMP based polling method, which far exceeds the collection period. It is difficult to synchronize across different objects. For our method, the average time is about 1.2 minutes, which increases the data collection efficiency by 16 times. Considering that SNMP objects in ACs are updated at regular intervals, our method is very efficient and can ensure the data synchronization to a large extent. In our measurement study, we set the the data collection period as 5 minutes.

B. Objects for APs and User Devices

Based on the data collection method introduced in Section II-A, we collect two types of wireless objects. Objects in the first type are related to APs and user devices in TD_WLAN, which can be directly measured by commodity WiFi hardware working in the normal data forwarding mode. Objects in the second type are related to rogue and non-WiFi devices, which will be introduced in Section II-C. In this section, we mainly describe 7 important objects in the first type.

(1) Channel Utilization: It represents the time proportion occupied by all WiFi and non-WiFi devices for the current working channel of a given AP. **(2) Interference Utilization:** It represents the time proportion of the interference from other 802.11 networks working on the same channel with the current AP. **(3) Transmitted Frames:** It represents the number of successfully transmitted 802.11 frames from APs during the period of data collection. It means APs receive layer-2 (MAC) acknowledgements (ACK) for these frames. The object is abbreviated as $Trans_{num}$. **(4) Retransmitted Frames:** It represents the number of retransmitted frames due to the packet loss or packet error in the MAC layer during the period of data collection. The object is abbreviated as $Retry_{num}$. **(5) Failed Frames:** It represents the number of frames that finally fail to transmit after several retransmission attempts (The maximum retransmission threshold in TD_WLAN is set to 7). The object is abbreviated as $Fail_{num}$. **(6) RSSI:** It represents the average packet RSSI (received signal strength indication) observed in APs. **(7) SNR:** It represents the signal and noise strength ratio of frame reception.

C. Objects for Rogue and Non-WiFi Devices

Because APs in TD_WLAN lack extra radios to measure the information related to rogue and non-WiFi devices, we configure them to alternately work in the data forwarding mode and the monitor mode in different periods. When APs work in the monitor mode, they can scan the channels by the 2 radios (2.4GHz and 5GHz) but cannot provide the network access service. Therefore, to mitigate the effect as much as possible, the period in the monitor mode should be far less than that in the data forwarding mode. In practice, the period of the monitor mode is set to 60ms and the period of the data forwarding mode is set to 10s. Because TD WLAN is deployed in dense mode and all APs in the monitor mode would scan all channels of 2.4GHz and 5GHz bands in turn, rogue and non-WiFi devices can be well detected.

For rogue devices, we mainly collect the following objects: **(1) MAC address**, **(2) Rogue RSSI** and **(3) Working Channel**. For non-WiFi devices, we mainly collect the following objects: **(1) Device Type**, **(2) Non-WiFi RSSI**, **(3) Occupied Channels** and **(4) Detected Time**. The method for type identification of non-WiFi devices is consistent with the previous work [23]. In brief, according to the spectral samples generated by APs in the monitor mode, a set of fingerprinting features ($e.g.$, frequency, bandwidth, duty cycle, pulse distribution, $etc.$) will be extracted to capture the properties of non-WiFi devices. Based on specific rules and classification methods (more details can refer to the work [23]), different types of non-WiFi devices can be accurately distinguished.

D. Performance Metrics

In this section, we formally define two performance metrics, which will be used to evaluate the wireless performance.

(1) Loss Rate in MAC Layer ($loss_{mac}$)

TABLE I
THE DISTRIBUTION OF USER DEVICES ASSOCIATED WITH DIFFERENT BANDS OF THE TWO SSIDS AT THE RUSH HOUR

SSID	T-University		T-University-5G
Band	2.4GHz	5GHz	5GHz
Devices (2.4GHz only) (#1829)	1829	0	0
Devices (5GHz only) (#1420)	0	201	1219
Devices (dual-band) (#17986)	6972	5494	5520
Total Devices (#21235)	8801	5695	6739

It is defined as follows:

$$loss_{mac} = \frac{Retry_{num} + Fail_{num}}{Trans_{num} + Retry_{num} + Fail_{num}} \quad (1)$$

where $Trans_{num}$, $Retry_{num}$, and $Fail_{num}$ are defined in Section II-B. When the packet loss occurs in the wireless MAC layer, the corresponding packets will be retransmitted until they are successfully transmitted or the number of retransmissions reaches the threshold. In the rest of the paper, loss rate in MAC layer is also called "loss rate" for short.

(2) WiFi Latency (T_{wifi})

It represents the last-hop WiFi latency for a 802.11 frame, which can be estimated by the time difference between the transmission of packets from APs and the reception of corresponding MAC ACKs. However, because of the energy management policies of user devices [22], [24], [25], the WiFi latency could be overestimated. When user devices are in power saving mode, APs will buffer the sending data temporarily. After that, APs will notify these user devices that there is buffered data to be retrieved by *beacon* [26]. Only when user devices are waked up and send specific control frames to APs for retrieving data, these data can be transmitted. The process additionally adds a lot of delay. To solve the problem, if user devices are in power saving mode, we further record two timestamps ($i.e.$, T_{ack} and T_{wake}) in APs. T_{ack} represents the time when APs receive the MAC ACKs and T_{wake} represents the wake-up time of user devices ($i.e.$, time when APs receive the specific control frames from user devices). On this basis, T_{wifi} can be calculated by $T_{ack} - T_{wake}$. In our measurement study, during each data collection period, we record the average WiFi latency across many frames for each user device.

III. WIRELESS BEHAVIOR MEASUREMENT

In this section, based on TD_WLAN, we conduct one of the largest known measurements to understand the wireless behavior in dual-band wireless networks.

A. Characteristics for User Devices and APs

In TD_WLAN, network administrators configure two SSIDs (Service Set Identifiers), namely *T-University-5G* and *T-University*. *T-University-5G* only supports the 5GHz band and *T-University* supports both 2.4GHz and 5GHz bands. In general, when users try to connect to *T-University*, dual-band devices will proactively choose a working band and initiate association requests.

Table I shows the statistics of user devices associated with different bands of *T-University* and *T-University-5G* at the rush hour (Here the rush hour refers to the time when there are peak number of user devices associated with APs.). We observe that about 84.7% (17986/21235) of user devices support dual-band, which is a lot more than the values reported by previous studies [4], [12]. For the dual-band SSID ($i.e.$, *T-University*), about 60% of devices tend to be associated with the 2.4GHz band, which generally depends on the implementation of different device hardware. For example, some devices may prefer to choose the band with higher RSSI. Besides, we find that although users in T university have been informed that *T-University-5G* supports the 5GHz band, only 31.7% (6739/21235) of devices are associated with it, which means that users are still used to connect to *T-University*.

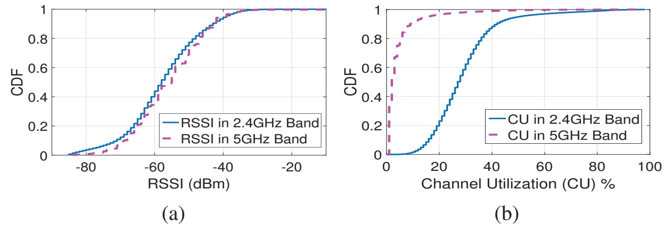


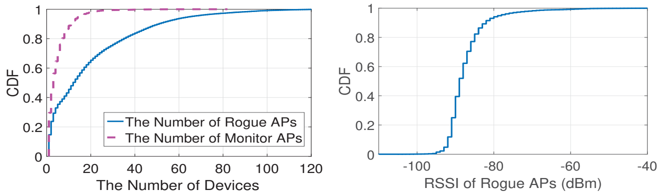
Fig. 3. (a). CDF of the packet RSSI for 2.4GHz and 5GHz bands. (b). CDF of the channel utilization for 2.4GHz and 5GHz bands.

We further study the wireless characteristics related to APs in TD_WLAN. Fig. 3(a) shows the CDF of the packet RSSI observed by APs. We can see that the overall RSSI is satisfactory. The average RSSI in the 5GHz band is slightly higher than that in the 2.4GHz band, which is counter-intuitive since the signal in the 5GHz band attenuates faster. This phenomenon can be explained by two reasons. First, when working in the 5GHz band, devices generally use a larger transmission power to prevent signal attenuation. Second, the dual-band settings may make devices work in the band with better signal strength, which means that if RSSI in the 5GHz band is very low, devices could tend to be associated with the 2.4GHz band. We also compare the channel utilization in Fig. 3(b). The average channel utilization for the 5GHz band is about 2% while the value for the 2.4GHz band is up to 27%. One reason for the large difference is that the transmission rate of the 5GHz band (about 280Mbps on average) is much higher than that of the 2.4GHz band (about 100Mbps on average). Another important reason is because of the interference of different bands, which will be described in Section III-B.

B. Behavior of Rogue APs

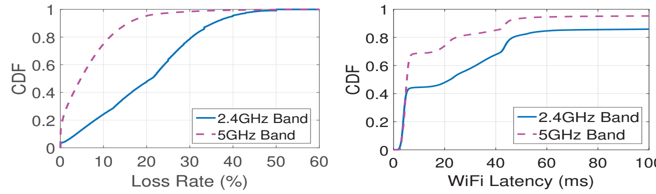
Before the deployment of TD_WLAN, many students are willing to access the Internet by self-built APs ($i.e.$, rogue APs). They are now legacy devices while many users still prefer to connect to them instead of APs in TD_WLAN. The random channel configurations for rogue APs bring great challenges to wireless management.

In Fig. 4(a), the solid line represents the CDF of the number of rogue APs detected by each AP in TD_WLAN and the dotted line represents the CDF of the number of APs that could detect the same rogue AP. Note that to accurately depict the impact of rogue devices on APs in TD_WLAN, we only focus



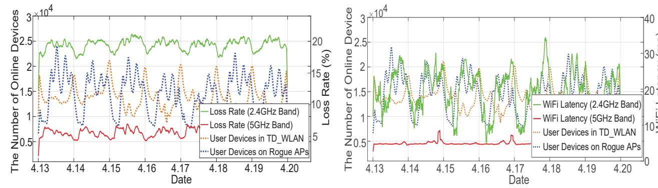
(a) (b)

Fig. 4. (a). The distributions for the number of rogue APs detected by each monitor AP (solid line) and the number of monitor APs that could detect the same rogue AP (dotted line). (b). The distribution for the RSSI of rogue APs.



(a) (b)

Fig. 5. (a). The distribution of the loss rate in MAC layer. (b). The distribution of the WiFi latency.



(a) (b)

Fig. 6. (a). The relationship between the number of online user devices and the loss rate. (b). The relationship between the number of online user devices and the WiFi latency.

on the rogue APs whose channels overlap with the channel of the corresponding AP. First, we observe that an AP can detect 12 rogue APs on average, which indicates that the channel contention could be very serious. Second, about 35% of rogue APs can be detected by more than 5 APs. Previous studies have pointed out that -85dBm is the threshold of RSSI for triggering the CSMA/CA mechanism [11], [27]. The value is applicable in TD_WLAN by confirming with network administrators. When the RSSI of rogue APs is lower than -85dBm, they will appear as hidden terminals and there is a high probability of packet collision. Fig. 4(b) shows the distribution for the RSSI of rogue APs. We can see that it is lower than -85dbm in about 80% of cases. The problem of hidden terminals in TD_WLAN is more typical.

As mentioned in Section II-C, when APs in TD_WLAN work in the monitor mode, they could obtain the working channels of all detected rogue APs based on the management frames such as *beacon*. We find that more than 70% of detected rogue APs only work in the 2.4GHz band, about 17% of them only work in the 5GHz band, and the rest work in both bands. We also observe that in the 2.4GHz band, only 40% of rogue APs switch the channels during the week, while in the 5GHz band, more than 83% of rogue APs tend to switch to channels with lower interference based on the DCA strategy. The above results indicate that the interference in the 5GHz band is much lower, which is consistent with Fig. 3(b).

C. Behavior of Non-WiFi Devices

Unlike WiFi devices, non-WiFi devices do not sense the medium before transmitting the energy [23], which could cause the RF interference and degrade the performance.

During the week, we detect 7 types of non-WiFi devices in total. They are respectively *Cordless Phone* (3.14%), *ZigBee Device* (1.14%), *Microwave Oven* (0.62%), *Bluetooth* (61.68%), *Game Controller* (12.90%), *Wireless Video* (7.95%) and *Baby Monitor* (12.57%). Bluetooth devices are dominant in TD_WLAN. We find that the frequency bands for many non-WiFi devices (*e.g.*, microwave oven, baby monitor, *etc.*) may overlap with any channel in the 2.4GHz band, which could easily lead to interference. In addition, we observe that game controllers and wireless videos may also work in the 5GHz band. However, the proportion of non-WiFi devices in the 5GHz band is very small (less than 4%). In TD_WLAN, a monitor AP could detect 2 non-WiFi devices on average, which reflects the prevalence of non-WiFi devices.

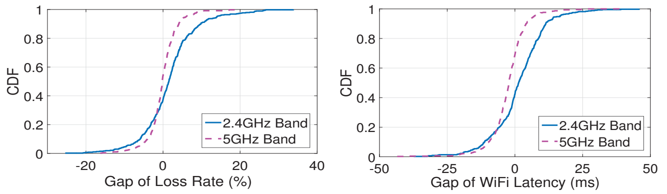
IV. Wireless Performance Analysis

In this section, we systematically study the effect of the device association behavior and the DCA strategy on the wireless performance and quantify the performance degradation caused by different types of interference in TD_WLAN, which provides guidance for wireless optimization in Section V.

A. Impact of Device Association Behavior

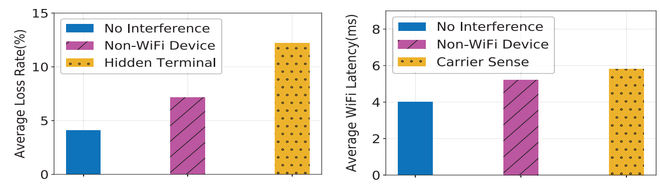
We plot Fig. 5(a) and Fig. 5(b) to show the distributions of the loss rate in MAC layer and the WiFi latency defined in Section II-D. We can see that compared to the 2.4GHz band, the average loss rate and the average WiFi latency in the 5GHz band are much lower, which benefits from the fewer rogue and non-WiFi devices and the higher transmission rate in the 5GHz band. However, as mentioned in Section III-A, more devices tend to be associated with the 2.4GHz band for the dual-band SSID (*i.e.*, *T-University*), which motivates us to improve performance by optimizing the band selection.

Fig. 6 shows the relationship between the number of online devices and wireless performance. Here online devices refer to user devices associated with APs in TD_WLAN or rogue APs. Because the trend of online devices in the 2.4GHz band is consistent with that in the 5GHz band, we only show the total number of online devices for brevity. In Fig. 6(a), an interesting phenomenon is the trend of the loss rate is negatively correlated with the trend of user devices associated with APs in TD_WLAN while is positively correlated with the trend of user devices associated with rogue APs. The increase in the number of user devices on rogue APs may cause a large probability of packet collision, which inevitably increases the loss rate. However, the deployed locations and parameter settings (*e.g.*, transmission power) of APs in TD_WLAN are carefully chosen based on the engineering measurement during the deployment (Note that the dynamic Transmit Power Control (TPC) in TD_WLAN is not enabled. The transmission



(a) (b)

Fig. 7. (a). The CDF for the gap of the loss rate. (b). The CDF for the gap of the WiFi latency.



(a) (b)

Fig. 8. (a). The effect of the hidden terminal interference and the RF interference on the average loss rate. (b). The effect of the carrier sense interference and the RF interference on the average WiFi latency.

power is fixed based on the engineering measurement), which could ensure the interference (including the hidden terminal interference) across them is as low as possible. Therefore, the loss rate will reduce when user devices migrate from rogue APs to APs in TD_WLAN. In Fig. 6(b), we observe that in the 2.4GHz band, the WiFi latency may become poor when there are a large number of user devices on either rogue APs or APs in TD_WLAN, while in the 5GHz band, the WiFi latency is very low and has no obvious fluctuations. We can conclude that *the association behavior of user devices has a great impact on the performance*.

B. Effectiveness of the DCA Strategy

In TD_WLAN, the DCA strategy is enabled for all ACs by default, which could dynamically configure channels for APs based on the carrier sense interference in different channels. When the interference utilization exceeds a preset threshold, ACs will guide APs to switch to one of the orthogonal channels with the least interference. In this section, we evaluate its effectiveness in large-scale dual-band wireless networks.

To this end, we turn off the DCA strategy and further collect one week of data from April 20th, 2019 to April 26th, 2019. For a given AP, we define the **performance gap** as the difference between the average value of a performance metric in the week with the DCA strategy and that in the week without the DCA strategy. The performance gap less than 0 means that the DCA strategy improves corresponding performance metric. Fig. 7(a) and Fig. 7(b) respectively show the distributions of the gap of the loss rate and the gap of the WiFi latency in different bands. In the 2.4GHz band, after the DCA strategy is enabled, although the loss rate for 10% of APs reduces by at least 5%, the loss rate for another 20% of APs increases by more than 5%, and the average loss rate increases by about 2%. Besides, the average WiFi latency also increases by about 1ms. In the 5GHz band, the DCA strategy has no significant effect on the loss rate but reduces the average WiFi latency by 1ms. Note that the average WiFi latency in the 5GHz band without the DCA strategy is about 6ms, which means that the DCA strategy reduces the WiFi latency by 16.7% on average. Therefore, we can conclude that *the DCA strategy has potential to optimize the performance in the 5GHz band while may lead to the performance degradation in the 2.4GHz band in TD_WLAN*. The reasons are as follows: (1) Most DCA strategies are based on greedy methods [28]–[30]. Each AP will be assigned a locally optimal channel. Due to limited spectrum resources (especially in dense deployment scenarios), some APs may benefit from the strategy while the performance for other APs may degrade seriously. However, there are more available channels in the 5GHz band, which makes the DCA strategy have a positive impact to some extent. (2) The DCA strategy only considers the carrier sense interference instead of the hidden terminal interference. Even though APs can switch to channels with less carrier sense interference, the loss rate will increase if there are more hidden terminals working on the newly switched channels, which is typical in the 2.4GHz band because of the prevalence of rogue and non-WiFi devices.

C. Impact of Different Types of Interference

In this section, we further quantify the impact of different types of interference on the performance in TD_WLAN. We focus on three types of interference, namely hidden terminal interference caused by rogue APs with RSSI below -85dBm, carrier sense interference caused by neighbor 802.11 networks and RF interference caused by non-WiFi devices.

The characteristics of the three types of interference are as follows: **(1)** Hidden terminal interference mainly leads to packet loss. Therefore, we use the loss rate as the evaluation metric. We define that hidden terminal interference may occur when at least one rogue AP with RSSI below -85dBm is detected. To eliminate the impact of poor channel quality on packet loss, we only consider the cases with SNR larger than 30dB, because previous studies have indicated that this value could ensure good access quality [11], [31]. **(2)** Carrier sense interference generally does not increase the loss rate due to the CSMA/CA mechanism, while it will increase the transmission time in WiFi hops. We use the WiFi latency to characterize its effect and define that carrier sense interference may occur when the interference utilization is larger than 0. **(3)** Non-WiFi devices do not sense the medium when transmitting the energy. RF interference could lead to packet loss if wireless data is transmitted in corresponding channels. Furthermore, it will occupy spectrum resources and increase the channel contention. As a result, both the loss rate and the WiFi latency are used to evaluate the effect. We determine whether RF interference occurs based on the active time of non-WiFi devices, which can be inferred by the detected time of them.

Fig. 8(a) shows the effect of the hidden terminal interference and the RF interference on the average loss rate. When hidden terminals and active non-WiFi devices are not detected, the average loss rate is very low (about 4%). Hidden terminal interference and RF interference could increase the average loss rate by 8% and 3% respectively. Fig. 8(b) shows the effect of the carrier sense interference and the RF interference on the average WiFi latency. They could increase the average WiFi latency by 45% and 30% respectively.

V. Wireless Performance Optimization

Based on the measurement, we propose a novel strategy and evaluate its effectiveness by actual deployment in this section.

A. Objectives and Ideas

The design of the performance optimization strategy has two objectives: *(1) Improve the overall loss rate and WiFi latency by optimizing the spectrum resources. (2) Achieved by easy-to-operate parameter configurations in commodity ACs.*

We have shown that although the 5GHz band outperforms the 2.4GHz band, there are more devices associated with the 2.4GHz band for the dual-band SSID. Therefore, a natural idea for the strategy is to migrate devices that tend to be associated with the 2.4GHz band to the 5GHz band through band navigation. However, blind navigation is unreasonable because the performance could be even more unsatisfactory in the 5GHz band ($e.g.$, RSSI is very low). In many cases, it is also necessary to navigate devices in the 5GHz band to the 2.4GHz band. We need to infer the reasons that affect performance and the scenarios of poor performance in different bands (Section V-B) and design the strategy to perform *intelligent band navigation* in the association phase (Section V-C). In addition to the band selection, to further improve the access quality, the strategy could *optimize the channel configurations* in different bands (Section V-D).

B. Constructing DAG for Causal Inference

To achieve the intelligent navigation, we need to find the main factors affecting performance in different bands. In this paper, we consider the following 7 factors: Channel Utilization ($\boldsymbol{CU}$), Interference Utilization ($\boldsymbol{IU}$), $\boldsymbol{RSSI}$, $\boldsymbol{SNR}$, the number of associated devices ($\boldsymbol{Sta_{num}}$), the number of hidden terminals ($\boldsymbol{HT_{num}}$), and the number of carrier sense devices ($\boldsymbol{CS_{num}}$), because they characterize the current network status when devices are associated and ACs could obtain them in time to achieve the navigation. The 7 factors are not independent, therefore, we need to consider their mutual influence when locating the reasons that affect performance.

Bayesian Network (BN) [32], [33] is an effective tool for causal inference, the form of which is a DAG composed of all factors, where nodes represent different factors and edges represent causality (*parent nodes are potential reasons that affect child nodes*). It could achieve probability inference by modeling the joint probability density distribution across all factors. Based on the idea of BN, we expect to infer the reasons that affect the loss rate and WiFi latency and the scenarios of poor performance. However, constructing the DAG for probability inference is generally very challenging, which is an NP-hard problem. To solve the problem, we propose an exact and efficient DAG construction algorithm.

(1) DAG Construction Algorithm

Assume that the weighted adjacency matrix of the DAG is denoted as $\boldsymbol{A} \in \mathbb{R}^{d\times d}$, where d is the number of nodes. In this problem, the number of nodes is 9, including 7 selected factors and 2 performance metrics ($i.e.$, loss rate and WiFi latency). We could use a mapping function $\mathcal{MAP}$ to convert $\boldsymbol{A}$ into a binary adjacency matrix ($i.e.$, $\mathcal{MAP}(\boldsymbol{A}) \in \{0,1\}^{d\times d}$), which is defined as follows: if $\boldsymbol{A_{ij}} \neq 0$, there exists a directed edge between node i and node j; otherwise, there is no directed edge. Let $\boldsymbol{X} \in \mathbb{R}^{n\times d}$ denote n data samples and assume that $\boldsymbol{X}$ is n observations of the *random vector* $X = (X_1, ..., X_d)$. Given $\boldsymbol{X}$, our goal is to learn an optimal DAG $\mathcal{MAP}(\boldsymbol{A}) \in S\_DAG$ to model the relationship across d nodes, where S_DAG represents the set of the binary adjacency matrix for all DAGs.

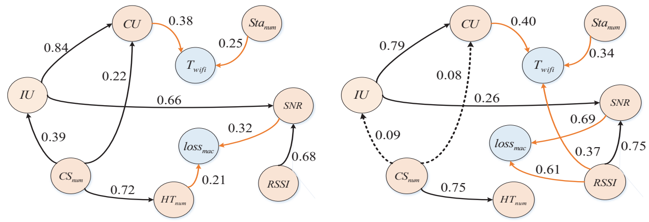
In a DAG, each node depends on its parent nodes. We could use the generalized linear model to capture the relationship and learn the graph structure, $i.e.$, $\mathbb{E}(X_i|\{X_j\}_{j\in Par(i)}) = f(a_i^T X)$, where $Par(i)$ represents the set of parent nodes for node i and a_i represents the i^{th} column of $\boldsymbol{A}$. In this problem, we treat $\boldsymbol{A}$ as the parameter to be learned. For simplicity, we adopt *linear regression* ($i.e.$, f) to model the relationship between child nodes and parent nodes. Therefore, the corresponding objective function is $l(\boldsymbol{A}) = \frac{1}{2n}||\boldsymbol{X} - \boldsymbol{XA}||^2$. The problem to be solved can be formulated as follows:

$$\begin{aligned} &\min_{\boldsymbol{A}\in\mathbb{R}^{d\times d}} F(\boldsymbol{A}) = l(\boldsymbol{A}) + \lambda||\boldsymbol{A}||_1 \\ &s.t. \quad \mathcal{MAP}(\boldsymbol{A}) \in S\_DAG \end{aligned} \tag{2}$$

where $||\boldsymbol{A}||_1$ is the *l1-regularization* term and λ is the balance hyperparameter. Although $F(\boldsymbol{A})$ is continuous, the constraint is discrete, which makes the problem difficult to solve. Our idea is to convert the discrete constraint $\mathcal{MAP}(\boldsymbol{A}) \in S\_DAG$ into a continuous and smooth equality constraint $g(\boldsymbol{A}) = 0$. $g(\boldsymbol{A}) = 0$ needs to ensure that $\boldsymbol{A}$ is acyclic and the derivative of $g(\boldsymbol{A})$ $w.r.t.$ $\boldsymbol{A}$ is easy to compute. Here we could let $g(\boldsymbol{A}) = tr(e^{\boldsymbol{A}\odot\boldsymbol{A}}) - d = 0$ (proof is omitted due to the space constraint). The derivative of $g(\boldsymbol{A})$ $w.r.t.$ $\boldsymbol{A}$ is $\nabla g(\boldsymbol{A}) = (e^{\boldsymbol{A}\odot\boldsymbol{A}})^T \odot 2\boldsymbol{A}$, which is easy to compute. The continuous optimization problem under the smooth equality constraint can be solved by the augmented Lagrangian method [34], which could construct the DAG for causal inference.

(2) Constructing DAGs in Different Bands

Based on the above algorithm, we respectively construct DAGs in different bands for causal inference. For effective model training, we additionally collect data in TD_WLAN for one month before April 13th, 2019. The construction results are shown in Fig. 9. The value on the edge represents the weight, which measures the strength of the corresponding relationship. Note that the construction results are not sensitive to the time length of training data (We respectively take one week, two weeks, and three weeks of data to train the model and only the weights of the constructed DAGs have negligible differences). We can see that the potential causality across 7 factors is basically the same in different bands. Compared to



(a) 2.4GHz (b) 5GHz

Fig. 9. DAGs in different bands (black arrows represent the causality across factors, orange arrows represent the causality between factors and performance metrics and dotted arrows represent the relationship is weak).

the 2.4GHz band, the relationship from CS_{num} to CU and IU in the 5GHz band is relatively weak (dotted arrows), $i.e.$, the corresponding weight of the learned DAG is low (lower than 0.1). It is because the number of rogue devices is small and its impact on the channels is not obvious.

The main difference of DAGs in the 2.4GHz band and 5GHz band is the impact of factors on performance metrics. We can see that in the 2.4GHz band, the main factors affecting the loss rate are HT_{num} and SNR, and the main factors affecting the WiFi latency are Sta_{num} and CU. While in the 5GHz band, the main factors affecting the loss rate are SNR and $RSSI$, and the main factors affecting the WiFi latency are Sta_{num}, CU and $RSSI$. The learning results reflect the difference in characteristics of different bands. Based on the DAGs, we could infer the scenarios of poor performance in different bands and design the strategy to achieve the intelligent band navigation, which will be discussed in Section V-C.

C. Intelligent Band Navigation

For intelligent band navigation, the optimization strategy should achieve the following two goals: **(1)** If devices only support one band, the navigation should not be performed to ensure smooth network connection. **(2)** When a device initiates the association request, if the loss rate or WiFi latency in the current band is not ideal with a high probability, while the other band could potentially improve the performance, the strategy should guide it to associate with the other band.

The first goal can be achieved by building the **device information list**, which records the band information of devices that ever appear in the network based on historical association data. For devices that only support one band, ACs do not perform band navigation, and directly establish connections. The device information list updates in real time during the operation of TD_WLAN. For new devices not in the device information list, if they only support one band, the navigation may fail. At this time, ACs will allow them to access in the supported band and further update the list.

To achieve the second goal, we need to find the scenarios of poor performance in different bands, $i.e.$, *the band navigation conditions*, which can be inferred by the DAGs constructed in Section V-B. Previous works [10], [18] have indicated that the user experience will be seriously affected when the WiFi latency exceeds 30ms, therefore, we select 30ms as the threshold for evaluating the quality of the WiFi latency. For the loss

TABLE II

BAND NAVIGATION CONDITIONS IN DIFFERENT BANDS

Band	Class	Conditions
2.4GHz	$condition_{loss}$	$(40 \leq HT_{num} < 80$ and $12dB < SNR < 25dB)$, $(SNR \leq 12dB)$, $(HT_{num} \geq 80)$
	$condition_{latency}$	$(CU \geq 80\%)$, $(Sta_{num} > 8)$ $(40\% \leq CU < 80\%$ and $Sta_{num} > 2)$, $(30\% \leq CU < 40\%$ and $Sta_{num} > 5)$
5GHz	$condition_{loss}$	$(SNR < 15dB)$, $(RSSI < -80dBm)$
	$condition_{latency}$	$(CU \geq 80\%)$, $(50\% \leq CU < 80\%$ and $Sta_{num} > 5)$, $(Sta_{num} > 15)$, $(RSSI < -80dBm)$

rate in MAC layer, we choose 30% as the evaluation threshold. This value is set by network administrators in TD_WLAN based on experience. On this basis, the problem of obtaining the band navigation conditions can be formalized as finding conditions $condition_{loss}$ and $condition_{latency}$ respectively in the 2.4GHz band and 5GHz band to meet:

$$\begin{aligned} &p(loss_{mac} \geq 30\% \mid condition_{loss}) \geq p_{threshold} \\ &p(T_{wifi} \geq 30ms \mid condition_{latency}) \geq p_{threshold} \end{aligned} \tag{3}$$

where $p_{threshold}$ is the probability threshold. It means that we need to infer the conditions ($i.e.$, $condition_{loss}$ and $condition_{latency}$) under which the probability of the WiFi latency larger than 30ms or the loss rate larger than 30% exceeds $p_{threshold}$. In this paper, $p_{threshold}$ is set to 0.5, which can be set flexibly by network administrators in different network environments based on their experience. For example, if users are not sensitive to performance requirements, the threshold can be set higher. The problem can be solved effectively by *bayesian inference* based on DAGs. The inference results in both bands are summarized in the third column of Table II. Correspondingly, the strategy is designed as follows:

Under the premise of satisfying the first goal, the strategy will evaluate the potential performance of the current band in which a device initiates the association request. If both $condition_{loss}$ and $condition_{latency}$ in the current band are not satisfied, it means that the potential performance is ideal and there is no need to navigate the device. However, if at least one of $condition_{loss}$ and $condition_{latency}$ is satisfied, the strategy needs to evaluate the potential performance in the other band. If both $condition_{loss}$ and $condition_{latency}$ in the other band are not satisfied, the strategy will attempt to navigate the device to the new band; otherwise, the strategy will not navigate the device.

We develop the strategy based on the band navigation implementation widely supported by mainstream vendors ($e.g.$, Cisco, Huawei, $etc.$) [35], [36]. Specifically, when a device initiates the association request in one band, if the strategy determines that it should be navigated to the other band, the corresponding AC will reject the request in the current band. After that, if the device initiates the association request in the other band, the AC will directly establish connection with it. However, if the device still insists on initiating the association request in the current band after three consecutive rejections, the navigation will fail and the AC will allow it to access in the current band. The usual practice of vendors is to reject three to five times [35], [36]. To prevent the degradation of the user experience, we choose to reject at most three times.

D. Heuristic Channel Optimization

To further improve the access quality in different bands, based on the measurement results, the proposed strategy also achieves the heuristic channel optimization as follows:

In the 5GHz band, compared to the 2.4GHz band, there are more orthogonal channels but lower interference caused by rogue devices and non-WiFi devices, which makes the channel resources not be fully utilized in the practical network operation. Besides, because of the intelligent band navigation, the number of associated devices in the 5GHz band would increase (it will be shown in Section V-E), which potentially increases the channel contention. To make full use of the spectrum resources and reduce the transmission time of data frames, in the 5GHz band, we enable *channel bonding*, which is peculiar to 802.11n and 802.11ac. It could extend the channel width by bonding the adjacent channels into a channel. To ensure the sufficient number of channels in the dense deployment of TD_WLAN, we just bond two adjacent 20MHz channels into a single 40MHz channel. In fact, bonding two adjacent 20MHz channels is suitable because the bandwidth of the campus network edge router of T university is limited and the network administrators also limit the effective throughput of each AP during the operation of TD_WLAN. In addition, based on the conclusions in Section IV-B, we enable the DCA strategy in the 5GHz band, because it is helpful to mitigate the carrier sense interference to some extent.

In the 2.4GHz band, compared to the channel contention, the impact of the hidden terminal interference is more serious. We attempt to mitigate its impact. Considering that the DCA strategy may have negative effects in the 2.4GHz band, we turn off it after configuring initial channels for all APs and then fine-tune the channels. The fine-tuning targets are APs whose average loss rate in MAC layer is greater than 30% (the value is consistent with that in Section V-C) during the week from April 20th, 2019 to April 26th, 2019. Because the DCA strategy only considers the carrier sense interference, the main basis for fine-tuning is the number of hidden terminals ($i.e.$, rogue APs with whose RSSI is lower than -85dBm) detected by APs under different channels during the week. We adopt the heuristic way to switch these target APs to the channels with the least number of hidden terminals.

E. Optimization Strategy Evaluation

To evaluate the effectiveness of the proposed strategy, in the week from May 11th, 2019 to May 17th, 2019, we deploy it in ACs and compare the wireless performance with that in the week from April 13th, 2019 to April 19th, 2019. Note that ACs in TD_WLAN provide the parameter configuration files for network administrators to specify the channel bonding strategy, channels of all managed APs and band navigation conditions, etc. The optimization strategy would be further enabled on the basis of these configuration files.

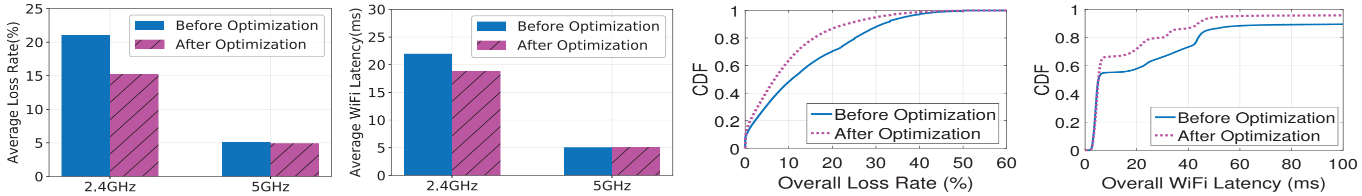
The experimental results show that under the intelligent band navigation, the number of devices migrating from the 2.4GHz band to the 5GHz band is more than that migrating from the 5GHz band to the 2.4GHz band. Before the optimization, for the dual-band SSID, the proportion of devices associated with the 2.4GHz band is about 60%. However, after the optimization, the proportion of devices associated with the 2.4GHz band reduces to about 44%, of which about 35% do not trigger the band navigation ($i.e.$, only support the 2.4GHz band or continue to initiate the association request in the 2.4GHz band after three consecutive rejections), and the rest (about 9%) are navigated from the 5GHz band to the 2.4GHz band. In the measurement study, the band navigation in two bands is triggered in about 33.6% of cases.

Fig. 10(a) and Fig. 10(b) respectively show the change of the average loss rate and the average WiFi latency in different bands. We can see that although the proportion of devices associated with the 5GHz band increases, the performance in the 5GHz band remains at an ideal level, which implies that the performance for devices migrating from the 2.4GHz band to the 5GHz band improves a lot without affecting devices that previously tend to be associated with the 5GHz band. Meanwhile, in the 2.4GHz band, the average loss rate reduces by 6% and the average WiFi latency reduces by 3ms. As mentioned in Section V-A, **our final objective** is to improve the overall wireless performance (do not distinguish bands) in TD_WLAN, which is more direct to reflect the change of the access quality of devices. Therefore, we plot Fig. 10(c) and Fig. 10(d) to compare the overall loss rate and WiFi latency for the two weeks. We observe that the average loss rate is about 10% before the optimization. While after the optimization, the average loss rate is about 6%, which means that the overall packet loss reduces by 40% ($(10\% - 6\%)/10\%$) on average. Besides, after the optimization, the WiFi latency for more than 60% of devices is below 5ms. At the 75th percentile, the WiFi latency reduces from about 40ms to 20ms. The proportion of high WiFi latency ($e.g.$, larger than 30ms) reduces greatly. We also plot Fig. 11(a) and Fig. 11(b) to compare the trends of the overall loss rate and WiFi latency with the change of time in the two weeks. We can see that after the optimization, the loss rate and WiFi latency reduce obviously, which further demonstrates the effectiveness of the optimization strategy.

VI. Discussion

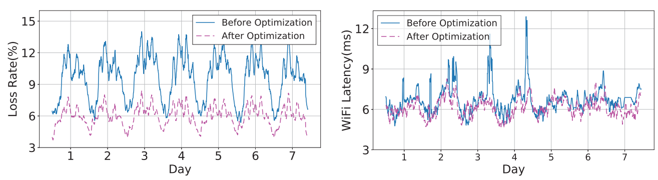
In this section, we would like to discuss the novelty of the intelligent band navigation and the generalization of the measurement study and optimization strategy.

Novelty: In this paper, our contribution is not to propose the concept of band steering, but to provide a novel solution to intelligently guide when to perform band navigation. To the best of our knowledge, few academic works try to improve the wireless network performance by optimizing the band selection. Although band steering has been widely supported by many mainstream vendors as a mature product feature, most vendors adopt relatively simple heuristics to obtain navigation conditions [36]–[38], which are generally not good



(a) (b) (c) (d)

Fig. 10. Performance comparison before and after the optimization. (a). The average loss rate in 2.4GHz and 5GHz bands. (b). The average WiFi latency in 2.4GHz and 5GHz bands. (c). The distribution of the overall loss rate in both bands. (d). The distribution of the overall WiFi latency in both bands.



(a) (b)

Fig. 11. (a). Trends of the loss rate before and after the optimization. (b). Trends of the WiFi latency before and after the optimization.

choices because they do not consider the impact of different factors on the performance in two bands. For example, Cisco proactively navigates devices from the 2.4GHz band to the 5GHz band under the premise of load balance [37]. Aruba takes account of device load and RSSI at the same time when navigating devices to the 5GHz band [38]. In contrast, navigation conditions in our method are obtained intelligently by a data-driven way, which could reflect the effect of different factors on wireless performance more effectively.

Generalization: We mainly discuss the generalization of the data collection method, measurement results and the proposed optimization strategy. (1) Because commercial ACs generally support SNMP and the proposed data collection method in Section II-A can be implemented easily as extensions of the SNMP service, it is not vendor specific and can be applied in other wireless networks as a general data collection method. (2) As one of the largest dual-band wireless networks, TD_WLAN is representative for characterizing the behavior and performance of wireless networks. Differences in the association behavior in two bands, the interference of rogue and non-WiFi devices and the effect of DCA strategy revealed by measurement results in TD_WLAN also generally exist in other types of dual-band wireless networks, especially in large-scale deployment scenarios. (3) The proposed optimization strategy essentially obtains band navigation conditions and channel configurations through a data-driven approach, which provides a new optimization idea and can be generalized to other types of wireless networks. However, differences in network environments could make the final strategy different (*e.g.*, DAGs in Fig. 9, band navigation conditions in Table II, *etc.*), which needs case-by-case analysis.

VII. Related Work

Many studies characterize the wireless behavior and performance in different environments by deploying customized gateways or dedicated measurement hardware [6]–[10], [17], [18], [39]. For example, Grover *et al.* and Patro *et al.* present the studies on the usage and experience of home wireless networks by deploying customized APs [7], [8]. However, these studies are limited by the scale of measurements.

There also exist a few studies that conduct large-scale measurements on wireless networks [4], [11]–[16], [40]. For example, Biswas *et al.* describe the wireless behavior over a large cohort of wireless networks [12]. Pefkianakis *et al.* explore the performance in home wireless networks by collecting data from a large ISP [13]. Although these studies provide valuable insights on the wireless behavior and performance, most of them mainly focus on the 2.4GHz band [11], [13]–[15]. While in current wireless networks, most devices support dual-band. Therefore, we need to fully understand the behavior and performance in both bands for wireless optimization. Although some research has begun to explore the characteristics of 802.11ac networks [4], [12], they still lack a comprehensive comparison between 2.4GHz and 5GHz bands. Besides, little research explores to improve the performance in large-scale dual-band wireless networks. Our work fills the gaps.

Many existing wireless performance optimization methods need to install applications in user devices or update software in APs [4], [41], [42], which suffer from poor scalability and are difficult to manage. In this paper, we propose an effective optimization strategy that can be achieved by practical parameter configurations in commodity ACs.

VIII. Conclusion

In this paper, we conduct a comprehensive measurement on the wireless behavior and performance in a large-scale dual-band wireless network, and further propose a novel strategy to improve the network performance by intelligent band navigation and heuristic channel optimization. The actual deployment shows that the proposed strategy effectively reduces the loss rate and the WiFi latency. We believe that our work is a meaningful step towards understanding and improving the performance in large-scale dual-band wireless networks.

IX. ACKNOWLEDGEMENT

We appreciate the feedback by the anonymous ICNP reviewers and thank Aaron Striegel for the caring and shepherding of the paper. This work is supported by the National Key Research and Development Program of China (No. 2020YFE0200500). JingAn Xue is the corresponding author.

References

[1] X. Chen, R. Jin, K. Suh, B. Wang, and W. Wei, "Network performance of smart mobile handhelds in a university campus WiFi network," in *Proceedings of the 2012 Internet Measurement Conference*. ACM, 2012, pp. 315–328.

[2] "The most popular statistics about mobile devices," [EB/OL], 2019, https://www.statista.com/stats/mobile%20devices.

[3] "Number of wireless local area network (WLAN) connected devices worldwide from 2016 to 2021 (in billions)," [EB/OL], 2019, https://www.statista.com/statistics/802706/world-wlan-connected-device.

[4] A. Bhartia, B. Chen, F. Wang, D. Pallas, R. Musaloiu-E, T. T.-T. Lai, and H. Ma, "Measurement-based, practical techniques to improve 802.11ac performance," in *Proceedings of the 2017 Internet Measurement Conference*. ACM, 2017, pp. 205–219.

[5] "Facts and stats about Wi-Fi," [EB/OL], 2019, https://worldwifiday.com/about-us/facts/.

[6] S. Sundaresan, N. Feamster, and R. Teixeira, "Measuring the performance of user traffic in home wireless networks," in *International Conference on Passive and Active Network Measurement*. Springer, 2015, pp. 305–317.

[7] A. Patro, S. Govindan, and S. Banerjee, "Observing home wireless experience through WiFi APs," in *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*. ACM, 2013, pp. 339–350.

[8] S. Grover, M. S. Park, S. Sundaresan, S. Burnett, H. Kim, B. Ravi, and N. Feamster, "Peeking behind the NAT: an empirical study of home networks," in *Proceedings of the 2013 Conference on Internet Measurement Conference*. ACM, 2013, pp. 377–390.

[9] M. Rodrig, C. Reis, R. Mahajan, D. Wetherall, and J. Zahorjan, "Measurement-based characterization of 802.11 in a hotspot setting," in *Proceedings of the 2005 ACM SIGCOMM Workshop on Experimental Approaches to Wireless Network Design and Analysis*. ACM, 2005, pp. 5–10.

[10] C. Pei, Y. Zhao, G. Chen, R. Tang, Y. Meng, M. Ma, K. Ling, and D. Pei, "WiFi can be the weakest link of round trip network latency in the wild," in *IEEE INFOCOM 2016-IEEE Conference on Computer Communications*. IEEE, 2016, pp. 1–9.

[11] K. Sui, Y. Zhao, D. Pei, and L. Zimu, "How bad are the rogues' impact on enterprise 802.11 network performance?" in *IEEE INFOCOM 2015-IEEE Conference on Computer Communications*. IEEE, 2015, pp. 361–369.

[12] S. Biswas, J. Bicket, E. Wong, R. Musaloiu-e, A. Bhartia, and D. Aguayo, "Large-scale measurements of wireless network behavior," in *ACM SIGCOMM Computer Communication Review*, vol. 45, no. 4. ACM, 2015, pp. 153–165.

[13] I. Pefkianakis, H. Lundgren, A. Soule, J. Chandrashekar, P. Le Guyadec, C. Diot, M. May, K. Van Doorselaer, and K. Van Oost, "Characterizing home wireless performance: The gateway view," in *IEEE INFOCOM 2015-IEEE Conference on Computer Communications*. IEEE, 2015, pp. 2713–2731.

[14] D. Kotz and K. Essien, "Analysis of a campus-wide wireless network," *Wireless Networks*, vol. 11, no. 1-2, pp. 115–133, 2005.

[15] D. Da Hora, K. Van Doorselaer, K. Van Oost, and R. Teixeira, "Predicting the effect of home Wi-Fi quality on QoE," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 944–952.

[16] X. Wei, N. C. Valler, H. V. Madhyastha, I. Neamtiu, and M. Faloutsos, "A behavior-aware profiling of handheld devices," in *IEEE INFOCOM 2015-IEEE Conference on Computer Communications*. IEEE, 2015, pp. 846–854.

[17] K. Papagiannaki, M. Yarvis, and W. S. Conner, "Experimental characterization of home wireless networks and design implications," in *IEEE INFOCOM 2006-IEEE Conference on Computer Communications*. Citeseer, 2006, pp. 1–13.

[18] S. Sundaresan, N. Feamster, R. Teixeira, Y. Grunenberger, D. Papagiannaki, and D. Levin, "Characterizing home network performance problems," 2013.

[19] Y. Yiakoumis, M. Bansal, A. Covington, J. van Reijendam, S. Katti, and N. McKeown, "BeHop: A testbed for dense WiFi networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 18, no. 3, pp. 71–80, 2015.

[20] D. Perkins and E. McGinnis, *Understanding SNMP MIBs*. Prentice-Hall, Inc., 1997.

[21] M. Zhou, M. Ma, Y. Zhang, K. SuiA, D. Pei, and T. Moscibroda, "EDUM: Classroom education measurements via large-scale WiFi networks," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 316–327.

[22] H. Wang, J. Wang, W. Dang, J. Xue, and F. Li, "Squeezing the gap: An empirical study on DHCP performance in a large-scale wireless network," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 1628–1636.

[23] S. Rayanchu, A. Patro, and S. Banerjee, "Airshark: Detecting non-WiFi RF devices using commodity WiFi hardware," in *Proceedings of the 2011 Internet Measurement Conference*. ACM, 2011, pp. 137–154.

[24] I. Papapanagiotou, E. M. Nahum, and V. Pappas, "Configuring DHCP leases in the smartphone era," in *Proceedings of the 2012 ACM Conference on Internet Measurement Conference*. ACM, 2012, pp. 365–370.

[25] N. Ding, A. Pathak, D. Koutsonikolas, C. Shepard, Y. C. Hu, and L. Zhong, "Realizing the full potential of PSM using proxying," in *IEEE INFOCOM 2012-IEEE Conference on Computer Communications*. IEEE, 2012, pp. 2821–2825.

[26] "802.11 power save - association ID in TIM field triggers STA wake up?" [EB/OL], 2017, https://networkengineering.stackexchange.com/questions/45513/802-11-power-save-doze-mode-associationid-in-tim-field-triggers-sta-wake-up.

[27] N. Ahmed, "Interference management in dense 802.11 networks," 2009.

[28] V. Shrivastava, S. K. Rayanchu, S. Banerjee, and K. Papagiannaki, "PIE in the sky: Online passive interference estimation for enterprise WLANs." in *NSDI*, vol. 11, 2011, pp. 25–25.

[29] P. Wertz, M. Sauter, F. Landstorfer, G. Wolfle, and R. Hoppe, "Automatic optimization algorithms for the planning of wireless local area networks," in *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004*, vol. 4. IEEE, 2004, pp. 3010–3014.

[30] S. Chieochan, E. Hossain, and J. Diamond, "Channel assignment schemes for infrastructure-based 802.11 WLANs: A survey," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 1, pp. 124–136, 2010.

[31] "How to define minimum SNR values for signal coverage," [EB/OL], 2019, http://www.wireless-nets.com/resources/tutorials/define_SNR_values.html.

[32] L. E. Sucar, "Probabilistic graphical models," *Advances in Computer Vision and Pattern Recognition. London: Springer London. doi*, vol. 10, pp. 978–1, 2015.

[33] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. MIT press, 2009.

[34] A. Nemirovski, *Optimization II: Standard Numerical Methods for Nonlinear Continuous Optimization*, 1999.

[35] "Enable the band navigation for H3C [EB/OL]," 2017, https://www.tracert.com/h3c-wireless-spectrum-navigation.html.

[36] "Configuring band steering [EB/OL]," 2021, https://support.huawei.com/enterprise/en/doc/EDOC1000154662/2ef16c0.

[37] "Load balancing and band select on the Cisco wireless LAN controller [EB/OL]," 2021, https://community.cisco.com/t5/wireless-mobility-documents/load-balancing-and-band-select-on-the-cisco-wireless-lan/ta-p/3128513#toc-hId–1214102629.

[38] "Band steering - Aruba networks ," [EB/OL], 2021, https://www.arubanetworks.com/techdocs/ArubaOS_63_Web_Help/Content/ArubaFrameStyles/ARM/Band_Steering.htm.

[39] Y.-C. Cheng, J. Bellardo, P. Benkö, A. C. Snoeren, G. M. Voelker, and S. Savage, "Jigsaw: Solving the puzzle of enterprise 802.11 analysis," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 4, pp. 39–50, 2006.

[40] B. Alipour, L. Tonetto, A. Y. Ding, R. Ketabi, J. Ott, and A. Helmy, "Flutes vs. Cellos: Analyzing mobility-traffic correlations in large WLAN traces," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 1637–1645.

[41] K. Sui, M. Zhou, D. Liu, M. Ma, D. Pei, Y. Zhao, Z. Li, and T. Moscibroda, "Characterizing and improving WiFi latency in large-scale operational networks," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2016, pp. 347–360.

[42] C. Pei, Y. Zhao, Y. Liu, K. Tan, J. Zhang, Y. Meng, and D. Pei, "Latency-based WiFi congestion control in the air for dense WiFi networks," in *2017 IEEE/ACM 25th International Symposium on Quality of Service (IWQoS)*. IEEE, 2017, pp. 1–10.