IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 25, 2021, accepted June 7, 2021, date of publication June 11, 2021, date of current version June 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3088123

Modeling and Predictability Analysis on Channel Spectrum Status Over Heavy Wireless LAN Traffic Environment

YAFEI HOU[1], (Senior Member, IEEE), JULIAN WEBBER[2], (Senior Member, IEEE), KAZUTO YANO[3], (Member, IEEE), SHUN KAWASAKI[1], SATOSHI DENNO[1], (Member, IEEE), AND YOSHINORI SUZUKI[3]

[1]Natural Science and Technology, Institute of Academic and Research, Okayama University, Okayama 700-8530, Japan
[2]Graduate School of Engineering Science, Osaka University, Toyonaka 560-8531, Japan
[3]Wave Engineering Laboratory, Advanced Telecommunications Research Institute International, Kyoto 619-0288, Japan

Corresponding author: Yafei Hou (yfhou@okayama-u.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 20K04484.

ABSTRACT Using the real wireless spectrum occupancy status in 2.4 and 5 GHz bands collected at a railway station as representative of a heavy wireless LAN (WLAN) traffic environment, this paper studies the modeling of durations of busy/idle (B/I) status and its predictability based on predictability theory. We first measure and model the channel status in the heavy traffic environment over almost all of the WLAN channels at 2.4 GHz and 5 GHz bands in a busy (rush hour) period and non-busy period. Then, using two selected channels at 2.4 GHz and 5 GHz bands, we analyze the upper bound (UB) and lower bound (LB) of predictability of the busy/idle durations based on predictability theory. The analysis shows that the LB predictability of durations can be easily increased by changing their probability distribution. Based on this property, we introduce the data categorization (DC) method. By categorizing the busy/idle durations into different streams, the proposed data categorization can improve the prediction performance of some streams with large LB predictability, even if it employs a simple low-complexity auto-regressive (AR) predictor.

INDEX TERMS Spectrum usage model, heavy WLAN traffic environment, cognitive radio, predictability theory, auto-regressive predictor, data categorization.

I. INTRODUCTION

Wireless communication technology has become one of indispensable and integral parts for our society. With the increase of requirements for high capacity and large number access to support the coming of beyond 5G and coming 6G era [1], some efficient techniques have been proposed to further improve the spectrum efficiency which are being discussed in the standardizing process of the next generation wireless LAN [2]. One of the most intensively researched paradigms in wireless communications is cognitive radio (CR) system which configures dynamically to use the best wireless channels in its vicinity to avoid congestion and interference in a smart way [3], [4].

To achieve efficient spectrum usage, two relative research topics are important for CR systems [5]. Firstly, the channel usage status or pattern model of frequency bands needs to be modeled. An accurate and realistic spectrum occupancy pattern is essential and extremely useful in domain of dynamic spectrum access (DSA) or CR research. The correct pattern models can be utilized in design of the DSA system, system analysis, the implementation of simulation tools and the development of more efficient DSA techniques. The other important but challenging research topic is that the channel status of spectrum usage at allocated frequency bands needs to be correctly predicted using efficient prediction methods [6], [7].

For the model of spectrum usage, there are many investigations based on measurement campaigns in different real environments [8]–[11]. The busy and idle durations of an IEEE 802.11b WLAN operated at 2.4 GHz band has been described using a continuous-time semi-Markov chain (CTSMC) model where the data is obtained from measurements using a vector signal analyzer [12]. Some more realistic traffic sources

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Fadda.

such as File Transfer Protocol (FTP), Voice over Internet Protocol (VoIP) [13] and Hypertext Transfer Protocol (HTTP) [14] have been researched for the fitting of the spectrum occupancy model. They found that the idle sojourn time is fitted to a generalized Pareto distribution [13] or a hyper-Erlang distribution. However, they just considered an interference-controlled environment. In [14], the authors analyzed the distribution of busy and idle durations in a real indoor office environment with several wireless devices operated in the 2.4 GHz band. In [15], the wireless environments which include one-pair, two-pair, or three-pair 802.11 networks with a prefixed separation distance of 1 m over 2.4 GHz band have been investigated. In [16], the authors investigated the time-dimension models of spectrum usage for large systems from amateur systems (144–146 MHz) to open bands and ISM band (2.4–2.5 GHz). The authors of [17] researched the space-dimension models of spectrum usage and provided the closed-form relations between the average spectrum occupancy in terms of the duty circle and some simple operation parameters such as noise power, false-alarm probability. In [18], the authors have researched the deterministic-stochastic duty circle model for the empirical data captured in one university laboratory over the 2.4 GHz and 834–845 MHz bands. In [19], the authors characterized the radio frequency spectrum opportunities available in a common GSM (Global System for Mobile Communications) channel to support the operation of a CR network which employed a discrete Generalized Pareto (dGP) distribution. In [20], the 2.4 GHz spectrum in a hospital environment in which medical devices make wireless access has been characterized and modeled as a generalized extreme value (GEV) distribution.

These researches consider real environments and provide profound and significant results and deep understanding of spectrum access of CR networks. However, some researches only consider some specific traffic sources such as FTP, VoIP, HTTP, Video, etc., and cannot be directly used for some real scenarios where collected channel data includes multiple applications. In addition, most measurement campaigns, however, consider the wireless traffic involved with a small number of users or devices which generates a moderate or low wireless traffic. Until now, there has been seldom research considering heavy WLAN traffic environments. Compared with low WLAN traffic environments, for the heavy WLAN traffic environments, more spectrum resources are required and cognitive radio technique is extremely important and difficult to be realized [3]. Therefore, it is very valuable to research the modeling of wireless spectrum usage over heavy WLAN traffic environments.

On the other hand, compared with capturing channel statistical information, the prediction of channel status is more difficult and sometimes impossible. A tutorial paper [21] has summarized most of the existing prediction methods for optimization of wireless resource allocation. These research results have provided many efficient methods for usage of the cognitive radio system. However, these researches are, in the most part, considering the prediction of some key parameters such as channel occupancy ratio (COR) with a time resolution unit of either seconds, hours or days. If the system can correctly predict the start and end of channel busy or idle duration, the system can more efficiently utilize the available radio resources and improve spectrum efficiency. For example, a CR system can be designed to utilize idle periods scattered in multiple frequency bands by splitting one transmission packet into small sub-packets [22], [23] then transmitting on the multiple bands. However, for the channel status prediction over heavy WLAN traffic environments, it is more difficult than that of over low WLAN traffic environments because the channels in unlicensed bands are occupied by a huge number of different services such as audio, video or file transfer etc. Such unstable and disordered traffic data makes the prediction intangible and difficult.

Therefore, for the data captured from the heavy WLAN traffic environments, some fundamental questions for the prediction of busy or idle durations are that: to what degree is the time-series data predictable? Whether can we improve its predictability using some simple methods? For the predictability research, the references [24], [25] have investigated such issues for radio spectrum state over TV bands, ISM bands, cellular bands, GSM900 and GSM1800 downlink bands. However, until now, there have been no any researches on the spectrum status over heavy WLAN traffic environments and how to improve their lower bound value.

In this paper, we investigate the modeling of durations of busy/idle (B/I) spectrum status over a heavy WLAN traffic environment at a major railway station and its predictability based on predictability theory. The major contributions and novelties of this paper are as follows.

(1) This paper investigates the spectrum usage status and modeling analysis of wireless traffic data obtained at a railway station over almost all of WLAN channels at 2.4 GHz and 5 GHz bands. For each channel, we measured the spectrum usage status over 30 minutes and obtain enough data samples for the statistical analysis. The fitting modeling can provide realistic parameters for usage considering a scenario over heavy WLAN traffic environments.

(2) Our measurement utilized a commercially available wireless LAN frame acquisition and analysis tool. The tool is capable of capturing radio frames transmitted by IEEE 802.11 wireless LAN devices, and can measure received power, frame size, frame type and information on the IEEE 802.11 wireless LAN frame etc. The captured data information can provide more realistic parameters for modeling the spectrum status of heavy WLAN traffic environment.

(3) Using spectrum busy/idle duration data of two selected WLAN channels, we analyze the predictability of spectrum status over heavy WLAN traffic environment. The upper bound and lower bound of predictability are provided based on predictability theory. Although it is still unknown how to find a method to realize the

upper bound (UB) of prediction, the analysis shows that the lower bound (LB) of partial data can be easily increased by changing the probability distribution of busy/idle durations.

(4) Based on predictability analysis, we introduce the data categorization (DC) on the busy/idle durations. By setting the preceding status durations into different ranges, the upcoming status durations can be distributed into different categorization with different range. This categorization will change the value of the predictability of partial data. The proposed data categorization can improve the prediction performance of some streams which has better predictability than that of without DC even if the system employs a simple low-complexity auto-regressive (AR) predictor.

This paper is organized as follows. Section II describes the methodology utilized to capture and process the empirical data used for the validation of the mathematical models. The measurement environment and channel information are also introduced in this section. Section III examines the probability distribution models and the goodness of fit (GOF) metrics utilized to assess their suitability in fitting the obtained data. The mathematical models of channel occupancy status for both 2.4 GHz and 5 GHz bands will be also analyzed and explained in Section III. Then we analyze the predictability of the busy/idle durations of two selected channels in Section IV. The proposed data categorization (DC), which separates the busy/idle durations into several streams with different distribution property to improve the predictability of partial busy/idle data, is explained in Section V. The AR-based predictor with proposed data categorization method and its prediction performance are presented in Section VI. The paper ends with conclusions presented in Section VII.

II. MEASUREMENTS SETUP AND DATA CAPTURE

In this section, we first explain how the busy/idle durations are collected in a real heavy WLAN traffic environment.

A. MEASUREMENT SETUP

Measurements were carried out at a major railway station in Japan during peak hour periods, within both the 2.4 and 5 GHz bands. We selected 12 channels for measurements each at 2.4 GHz and 5 GHz bands for comparison as shown in Table 1. Measurements were conducted at the end of January 2017.

TABLE 1. Selected channels (CHs) for measurement.

2.4 GHz	CH 1–CH 9, CH 11–CH 13	
5 GHz	W52	CH 36, CH 40, CH 44, CH 48
	W53	CH 52, CH 56, CH 60, CH 64
	W56	CH 104, CH 108, CH 112, CH 116

The experiments were performed using commercially available wireless LAN frame acquisition and analysis software. The capture software is able to record radio frames transmitted by IEEE 802.11 wireless LAN devices in real-time, and can measure received power, frame-size, and modulation scheme. In order to simultaneously capture the data on all channels, we employed 5 PCs, each equipped with 3 wireless LAN systems with the interfaces, devices and modules. Information on the IEEE 802.11 wireless LAN frame, such as MAC information, can also be obtained. The measurement equipment was positioned in an aisle near the ticket-gate of a mainline railway station. When busy (around 18:30) there were several hundred people and streams of passengers constantly arriving and departing.

B. FRAME ESTIMATION PROCEDURE

The MAC addresses were first anonymized to an integer number in order to remove the identity of each STA. Frame duration is determined by the PHY preamble, MAC header and Data sections. Frame duration is the estimated over-the-air channel time which includes the PHY preamble, MAC header and Data sections. The Data section duration is estimated from the Length field (Bytes), data-rate, protocol type (IEEE 802.11a/b/n/ac), bandwidth, number of spatial-streams, and modulation and code scheme (MCS) data. The Length field is used to determine the number of OFDM symbols in the frame. In some corrupted frames the Data rate is not available and it is assumed that the Type is the same as the previous frame. The PHY header of 20 μs was appended for IEEE 802.11a/g frames and $32+(nss\times 4)$ μs for IEEE 802.11n frames where nss is the number of spatial streams. A MAC header of 30 Bytes and 34 Bytes was appended for IEEE 802.11a/g and IEEE 802.11n frames respectively.

When two or more uncoordinated STA attempt to transmit to different APs at the same time and same band with sufficient power, a frame collision will occur. If the interfering transmissions originate from enough distance the interference may be sufficiently low to not result in a frame check sequence (FCS) error. However, some transmissions will be corrupted, and these are reported by an FCS error when the cyclic redundancy check (CRC) fails.

In normal operation with most network interface controllers (NICs), a packet received with an FCS error will not be passed on to the operating system. However, the capture software and wireless adapter in monitoring mode allowed this data to be captured and processed. It should also be noted that as the wireless adapter is in a different physical location to the AP, it is possible that the particular packet was not actually corrupted at the AP.

As the wireless adapters can only capture one frame at a time, when two transmissions occur, only one is able to be captured. Therefore, in that respect the number of transmissions is under-reported. However, the STA that was unable to complete its transmission will retry a transmission later, and this should eventually be captured by the wireless adapter and therefore included in the statistics.

C. NON-IEEE 802.11 FRAMES AND VERIFICATION

Radio signals other than the IEEE 802.11 wireless LAN frame such as generated by a microwave oven, a weather radar, a Bluetooth device or such like cannot be captured with the software. For this reason, radio data of the same frequency band was captured using a radio measuring equipment (I/Q data recorder) that has a maximum reception bandwidth up to 200 MHz. An RF circuit corresponding to 2.4 GHz, or 5 GHz band is connected to the measuring equipment so as to receive and record all the radio signals in the band. In addition, we visualized the radio resource usage status by converting the recorded radio signal data into a spectrogram. The configuration of the measuring device is depicted in Fig. 1 and a photograph of the measurement system on location is shown in Fig. 2.

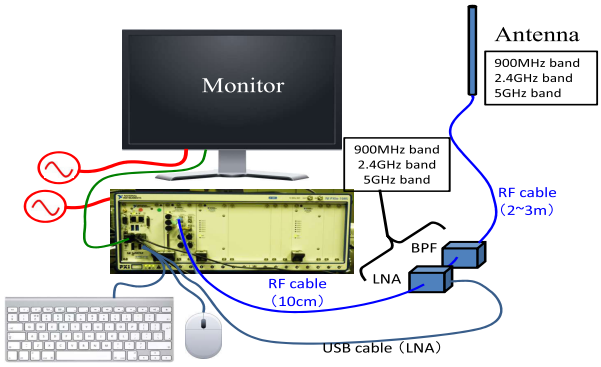


FIGURE 1. The configuration of the RF power measuring device.

FIGURE 2. Measurement system set-up on location.

The captured signal was also used to validate the results obtained by the capture software. The channel occupancy ratio was computed in 1 minute intervals and compared with the results from the frame capture. It was confirmed that there was a high correlation between the occupancy results obtained using the two methods.

D. SPECTROGRAMS OF THE MEASURED DATA

The spectrograms of the measured data over the 2.4 GHz and 5 GHz bands are shown in Fig. 3 and Fig. 4, respectively.

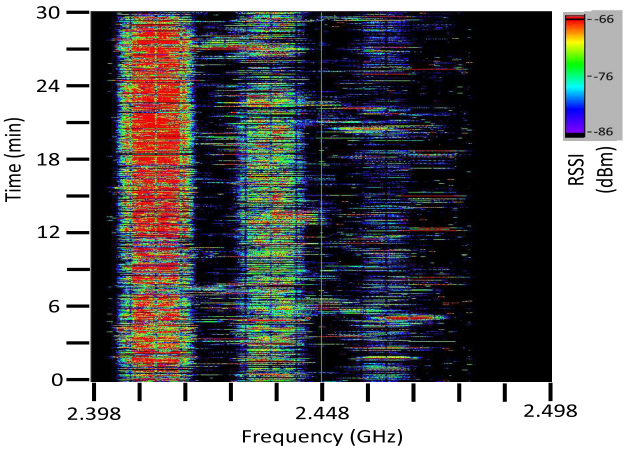


FIGURE 3. Spectrogram of the measured data at a railway station over 2.4 GHz band.

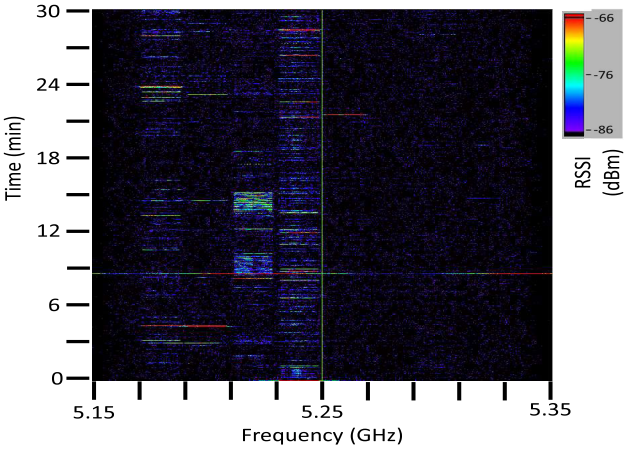


FIGURE 4. Spectrogram of the measured data at a railway station over 5 GHz band.

As shown in both figures, there are more wireless traffic over the 2.4 GHz bands than over the 5 GHz bands. On the other hand, even for the channels over the busy 2.4 GHz band, some channels have more traffic than others. The results reflect that the existing channel allocation methods are not fully efficient and a new method need to be found to mitigate the usage imbalance.

III. MATHEMATICAL MODELS OF CHANNEL OCCUPANCY STATUS FOR 2.4 GHz AND 5 GHz BANDS

After capturing the empirical channel occupancy data, we converted all data into binary patterns and then calculated the length of the continuous busy and idle durations with a small resolution of $9\mu s$ which accords with IEEE WLAN Standard [26]. It should be noted that the durations of busy and idle are continuous which cannot be an integral multiplication of $9\mu s$ because of wireless channel multipath fading and reflection or other factors. The empirical cumulative distribution function (CDF) and probability density function (PDF) of busy or idle durations are calculated for the

TABLE 2. Probability distribution models for fitting.

Distribution model	PDF $f(x;\text{parameters})$	CDF $F(x;\text{parameters})$
Exponential (EX)	$f_{EX}(x;\lambda)=\begin{cases}\lambda e^{-\lambda x} & x\geq 0,\\ 0 & x<0.\end{cases}$	$F_{EX}(x;\lambda)=\begin{cases}1-e^{-\lambda x} & x\geq 0,\\ 0 & x<0.\end{cases}$
Generalized Pareto (GP)	$f_{GP}(x;\mu,\sigma,\xi)=\begin{cases}(\frac{\xi(x-\mu)}{\sigma}+1)^{-\frac{\xi+1}{\xi}} & \xi\neq 0,\\ e^{-\frac{(x-\mu)}{\sigma}} & \xi=0.\end{cases}$	$F_{GP}(x;\mu,\sigma,\xi)=\begin{cases}1-\left(\frac{\sigma+\xi(x-\mu)}{\sigma}\right)^{\frac{-1}{\xi}} & \xi\neq 0,\\ 1-e^{-\frac{(x-\mu)}{\sigma}} & \xi=0.\end{cases}$
Log-normal (LN)	$f_{LN}(x;\mu,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right)$	$F_{LN}(x;\mu,\sigma)=\frac{1}{2}\operatorname{erfc}\left(-\frac{\ln x-\mu}{\sigma\sqrt{2}}\right)$
Logistic (LG)	$f_{LG}(x;\mu,\sigma)=\frac{e^{-\frac{x-\mu}{\sigma}}}{\frac{x-\mu}{\sigma}\left(1+e^{-\frac{x-\mu}{\sigma}}\right)^2}$	$F_{LG}(x;\mu,\sigma)=\frac{1}{2}+\frac{1}{2}\tanh\left(\frac{x-\mu}{2\sigma}\right).$
Generalized Extreme Value (GEV)	$s=\frac{x-\mu}{\sigma}$ $f_{GEV}(x;\mu,\sigma,\xi)=\begin{cases}(1+\xi s)^{\frac{-1-\xi}{\xi}}\exp(-(1+\xi s)^{\frac{-1}{\xi}} & \xi\neq 0,\\ \exp(-s)\exp(-\exp(-s)) & \xi=0.\end{cases}$	$s=\frac{x-\mu}{\sigma}$ $F_{GEV}(x;\mu,\sigma,\xi)=\begin{cases}\exp(-(1+\xi s)^{\frac{-1}{\xi}} & \xi\neq 0,\\ \exp(-\exp(-s)) & \xi=0.\end{cases}$
gamma (GM)	$f_{GM}(x;\alpha,\beta)=\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}$	$F_{GM}(x;\alpha,\beta)=\frac{\gamma(\alpha,\beta x)}{\Gamma(\alpha)}$
Weibull (WB)	$f_{WB}(x;\lambda,k)=\begin{cases}\frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}e^{-(x/\lambda)^k} & x\geq 0,\\ 0 & x<0.\end{cases}$	$F_{WB}(x;k,\lambda)=1-e^{-(x/\lambda)^k}$

fitting of the mathematical distribution models. In this paper, we will show the fitting results using some major simple distribution models listed in Table 2.

A. PROBABILITY MODELS AND GOODNESS OF FIT (GOF) METRICS

The major simple distribution models considered in this paper are exponential (EX), generalized Pareto (GP), Log-normal (LN), Logistic (LG), generalized extreme value (GEV), gamma (GM) and Weibull (WB) distribution. EX distribution is largely utilized to obtain the validity of the widely used Markov chain model for traffic analysis and system design. However, many researchers reported that EX model is not accurate in many real environments because the wireless traffic appears to have some self-similarity with heavy-tailed trend [27]. Therefore, GP distribution is largely employed for traffic analysis. Similar to GP distribution, LN, GM and WB distributions can be used for fitting wireless traffic with heavy-tailed trend.

Pareto distribution, GP, GM and WB distributions have been used for modeling wireless traffic. In addition, extreme value theory (EVT) has been used as a rational framework for the problem of burst prediction which needs only a subset of the data to work on. In the burst prediction based on EVT, GEV distribution is used for modeling different types of traffic and throughput [28], [29]. Therefore, we also select GEV distribution as one of candidates for model fitting. The significance of parameters in each distribution can be found in [30].

We utilize a technique based on maximum likelihood estimation (MLE), which is widely adopted as an efficient inference technique to calculate the distribution parameters from empirical data. In Ref. [16], the authors have considered methods of moment (MOM) inference scheme for estimating the distribution parameters and compared with that of MLE-based method. The results showed that the MLE-based method generally outperforms that of MOM scheme. Therefore, in this paper, we choose MLE-based method for calculating the distribution parameters.

To show the suitability of the fitting, we use three GOF metrics as Kolmogorov-Smirnov (KS) distance, Kullback-Leibler (KL) divergence [31] and Bhattacharyya distance [32]. These GOF metrics can provide some numerical values to show the matched level for a certain probability model fitting for the whole range of busy or idle durations. The superscript of *sym* is to be B or I. That corresponds to busy or idle duration data, respectively.

The KS distance D_{KS}^{sym} between the CDF model of busy or idle duration $F_{fit}^{sym}(L)$ and empirical CDF of busy or idle duration $F_{emp}^{sym}(L)$ with duration length L ($L>0$) [point] can be denoted as

$$D_{KS}^{sym}(F_{fit}^{sym}(L),F_{emp}^{sym}(L))=\max_L\{|F_{fit}^{sym}(L)-F_{emp}^{sym}(L)|\}. \quad (1)$$

KS distance is typically used in the context of a non-parametric test which performs a hypothesis test to check the test distribution is the same as the reference. KS distance has a symmetric property that the distance from one distribution to another is equal to the distance from the latter to the former. In addition, if D_A, D_B and D_C are distributions, KS distance has a property of triangle inequality as

$$D_{KS}(D_B,D_C)\leq D_{KS}(D_B,D_A)+D_{KS}(D_A,D_C). \quad (2)$$

Those properties make the KS distance powerful and convenient for the fitting of distribution model.

Kullback-Leibler (KL) divergence can be given as

$$D_{KL}^{sym}=\sum_{k=1}^{P}f_{emp}^{sym}(L_k)\ln\left(\frac{f_{emp}^{sym}(L_k)}{f_{fit}^{sym}(L_k)}\right)+\sum_{k=1}^{P}f_{fit}^{sym}(L_k)\ln\left(\frac{f_{fit}^{sym}(L_k)}{f_{emp}^{sym}(L_k)}\right). \quad (3)$$

$f_{emp}^{sym}(L_k)$ and $f_{fit}^{sym}(L_k)$ are empirical PDF and evaluated PDF model of busy or idle duration data. P is the number

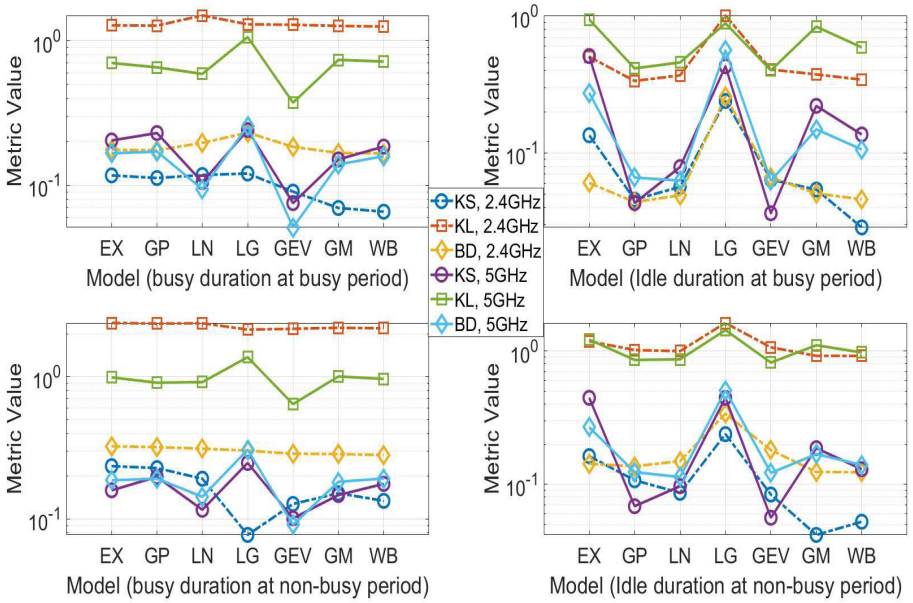


FIGURE 5. GOFs results for the fitting of channel occupancy status using data over all channels of 2.4 GHz and 5 GHz bands.

of different data. The KL divergence is typically used in information-theoretic settings, hypothesis testing, or even Bayesian settings, to measure the information cross entropy change between distributions before and after applying some inference. KL divergence provides the information entropy loss level when the wrong fitting model is selected. This metric is often utilized for machine learning and wireless system capacity optimization.

The Bhattacharyya distance D_B^{sym} as the GOF metric for PDF model $f_{sit}^{sym}(L_k)$ and empirical PDF $f_{emp}^{sym}(L_k)$ can be represented as

$$D_B^{sym} = -\ln\left(\sum_{k=1}^{P}\sqrt{f_{fit}^{sym}(L_k)f_{emp}^{sym}(L_k)}\right). \quad (4)$$

Different with the KS distance, Bhattacharyya distance is usually used to measure the similarity of two probability distributions. Different with KS distance, it grows depending on the difference between the standard deviations. Using Eq. (3) and Eq. (4), it is easy to deduce the relationship between the KL divergence and Bhattacharyya distance as

$$D_{KL}^{sym} \geq 2D_B^{sym}. \quad (5)$$

Due to this property, compared with KL divergence, Bhattacharyya distances of several fitting distribution functions have small difference among them.

From Eqs. (1), (3) and (4), we can find that a small value of these GOF metrics means that the probability distributions of mathematical models are well-fitted to that of the empirical data. When the values are zero, both mathematical models and distribution of empirical data are totally identical.

B. MODELS FITTING USING GOF METRICS FOR ALL CHANNELS OVER 2.4 GHz AND 5 GHz BANDS

In this section, we discuss suitable probability models for fitting the busy and idle durations. The major process was handled as follows. First, each CDF/PDF model in Table 2 is fitted to the empirical CDF/PDF of busy and idle durations captured at each channel based on MLE. Then the resulting GOF metrics, i.e. KS distance, KL divergence and Bhattacharyya distance, are evaluated between each empirical CDF or PDF model of busy and idle durations and all listed CDF or PDF models. Finally, the best fitting model with the smallest GOFs are selected as the probability models of the busy or idle durations.

Since a user can select any channels over 2.4 GHz band or 5 GHz band, to get the correctness of our fitting models for channel busy or idle durations, we utilize all used data of channels over 2.4 GHz band and over 5 GHz band to find general models. The results of three GOF metrics are shown in Fig. 5. For KS value, the largest values are smaller than 0.08 which show good fitting results for data. On the other hand, the data obtained at busy period has similar well-fitted results. That is, both busy and idle durations appear a Weibull distribution over 2.4 GHz and generalized extreme value distribution over 5 GHz band. On the other hand, over 5 GHz band, all results show that generalized extreme value distribution is a generally well-fitted distribution model for all busy and idle duration data. This is due to the fact that compared with that of 2.4 GHz, the wireless traffic over 5 GHz is low regardless of the busy or non-busy period.

Table 3 concludes the fitting distribution models and the parameters of the distribution models using data from all

TABLE 3. Parameters values of fitting results [point].

Bands	Target	Railway station			
		Busy period		Non-busy period	
		Model	Parameter values	Model	Parameter values
2.4 GHz	Busy	WB (k, λ)	201.61, 1.16	LG (μ, σ)	212.80, 70.86
	Idle	WB (k, λ)	436.72, 0.74	GM (α, β)	0.53, 2772.40
5 GHz	Busy	GEV (μ, σ, ξ)	0.97, 7.05, 11.62	GEV (μ, σ, ξ)	1.24, 7.72, 11.11
	Idle	GEV (μ, σ, ξ)	1.66, 48.31, 27.96	GEV (μ, σ, ξ)	1.67, 113.48, 65.18

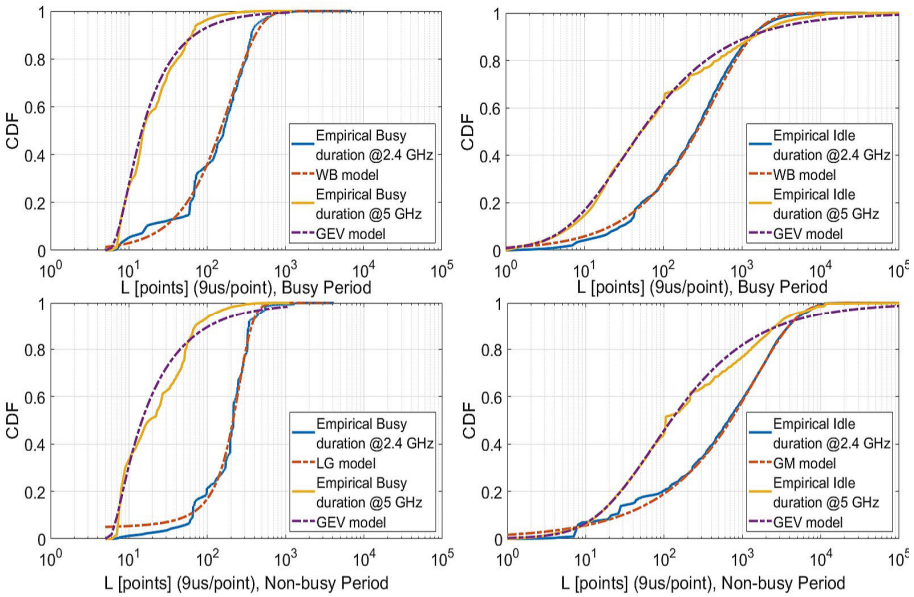


FIGURE 6. Fitting results of channel status using data over all channels of 2.4 GHz and 5 GHz bands.

channels at 2.4 GHz band and 5 GHz band. Based on the results of Table 3, we use Fig. 6 to show the fitting results of channel status over 2.4 GHz and 5 GHz bands. As shown in Fig. 6, the selected distribution models are well-fitted to that of empirical busy and idle durations, especially for the data of idle durations which supports our proposed fitting models in Table 3.

IV. PREDICTABILITY ANALYSIS FOR TIME-SERIES BUSY AND IDLE DATA

After we obtain the statistics of busy or idle durations, a fundamental question for the prediction of time-series data is that: to what degree is the time-series data predictable? Regarding to this question, a methodology of using statistical entropy measures and Fano inequality have been proposed to quantify the degree of predictability for the real-world time-series data [24], [25]. In the following sections, we selected the busy and idle durations of Channel 1 over 2.4 GHz band and Channel 36 over 5 GHz band for analyzing their predictability and showing how to improve their predictability.

First, we give a simple explanation for the concept of predictability which can decide the fundamental limitations such as the performance bounds of the prediction method. Let us suppose there is a random variate X with M kind of values. Therefore, its entropy $S(X)$ can be represented as

$$S(X) = -\sum_{i=1}^{M} f(x_i) \log(f(x_i)), \tag{6}$$

where $f(x)$ is the probability density function (PDF) of X. In addition, for an n-length time-series data $\mathbf{X}$ as $[x_1, x_2, \ldots, x_n]$, its average entropy can be represented as

$$S(\mathbf{X}) = \lim_{n\to\infty} \frac{1}{n} S(x_1, x_2, \ldots, x_n). \tag{7}$$

Let us also define a conditional entropy as $S(\mathbf{X}')$ with equation as

$$S(\mathbf{X}') = \lim_{n\to\infty} S(x_n | x_{n-1}, x_{n-2}, \ldots, x_1). \tag{8}$$

When $n \to \infty$, the average entropy $S(\mathbf{X})$ is usually equal to conditional entropy $S(\mathbf{X}')$ which is represented as

$$\begin{aligned} S(\mathbf{X}) &= \lim_{n\to\infty} \frac{1}{n} S(x_1, x_2, \ldots, x_n) \\ &= \lim_{n\to\infty} S(x_n | x_{n-1}, x_{n-2}, \ldots, x_1) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} S(x_i | \mathbf{h}_{i-1}) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} S(i). \end{aligned} \tag{9}$$

Here $\mathbf{h}_{i-1}$ and $S(i)$ are given as $\mathbf{h}_{i-1}$=$[x_{i-1}, x_{i-2}, \ldots, x_1]$ and $S(i) = S(x_i|\mathbf{h}_{i-1})$, respectively.

The n-length average entropy $S(\mathbf{X})$ is difficult to be obtained because it is determined by the value of $[x_1, x_2, \ldots, x_n]$ and their joint PDF. Usually, such entropy can be approximately calculated as $S^{Real}(\mathbf{X})$ using an iterative method called as Lempel-Ziv algorithm [33]. On the other hand, when the correlation property between the time-series data is not considered, the n-length average entropy can be simply calculated as $S^{Unc}(\mathbf{X})$ only using the PDF of time-series data as

$$S^{Unc}(\mathbf{X}) = -\sum_{i=1}^{\infty} f(x_i)\log(f(x_i)). \tag{10}$$

The $S^{Real}(\mathbf{X})$ and $S^{Unc}(\mathbf{X})$ are two different metrics of data entropy. The value of $S^{Real}(\mathbf{X})$ considers the interrelationships on the timeline of time-series data which is shown as the definition of $\mathbf{h}_{i-1}$. In other words, the entropy difference as $(S^{Unc}(\mathbf{X}) - S^{Real}(\mathbf{X}))$ can be regarded as the additional information certainty obtained from analyzing the correlation properties of time-series data. If data has no correlation among the data, the value of $S^{Real}(\mathbf{X})$ will be the same to the value of $S^{Unc}(\mathbf{X})$.

For the predictability of time-series data, we define P_r as the accuracy probability of one prediction method which is represented as

$$P_r = \text{Prob}\{\hat{x}_n = x_n|\mathbf{h}_{n-1}\}. \tag{11}$$

We suppose that there exists all prediction methods which can achieve different prediction accuracy P_r. Therefore, the predictability of one time-series data is the maximum value among all prediction methods as

$$\Pi(h_{n-1}) = \sup\{\text{Prob}[\hat{x}_n = x_n|\mathbf{h}_{n-1}]\}. \tag{12}$$

For n-length time-series data, the average predictability probability Π is represented as

$$\Pi = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \Pi(i), \tag{13}$$

with $\Pi(i) \triangleq f(\mathbf{h}_{i-1})\Pi(\mathbf{h}_{i-1})$.

Finally, we can build the relationship between the predictability probability and the entropy of time-series data using Fano inequality. From the information theory, the relationship between the conditional entropy and the prediction error probability $P_e = \text{Prob}\{x \neq \hat{x}\}$ can be represented using Fano inequality as

$$S(x_n|\mathbf{h}_{n-1}) \leq S(P_e) + P_e\log_2(M-1), \tag{14}$$

and $S(P_e)$ can be represented as

$$S(P_e) = -P_e\log_2(P_e) - (1-P_e)\log_2(1-P_e). \tag{15}$$

Let p as the value of $\Pi(\mathbf{h}_{n-1})$ (Eq. (12)) to represent the prediction accuracy of the best one among all prediction methods. For the best prediction method, we can let P_e as $(1-p)$. Therefore, using Eq. (14), the following equation can be obtained.

$$\begin{aligned} S(x_n|\mathbf{h}_{n-1}) &\leq -[p\log_2 p + (1-p)\log_2(1-p)] \\ &\quad + (1-p)\log_2(M-1) \\ &\triangleq S_F(p) = S_F(\Pi(\mathbf{h}_{n-1})), \end{aligned} \tag{16}$$

where M is the number of different value of x. $S_F(p)$ function is concave and monotonically decreases with p. Using this relationship, the lower bound and upper bound of predictability probability for time-series data can be represented as

$$\begin{aligned} S^{Real} &\leq S_F(\Pi^{Real}) \\ &= -[\Pi^{Real}\log_2\Pi^{Real} + (1-\Pi^{Real})\log_2(1-\Pi^{Real})] \\ &\quad + (1-\Pi^{Real})\log_2(M-1), \end{aligned} \tag{17}$$

$$\begin{aligned} S^{Unc} &\leq S_F(\Pi^{Unc}) \\ &= -[\Pi^{Unc}\log_2\Pi^{Unc} + (1-\Pi^{Unc})\log_2(1-\Pi^{Unc})] \\ &\quad + (1-\Pi^{Unc})\log_2(M-1). \end{aligned} \tag{18}$$

Using Eqs.(17) and (18), it is easy to find the lower bound Π^{Unc} and upper bound Π^{Real} of predictability probability when their entropy S^{Unc} and S^{Real} can be calculated. For the detail process, readers can find further information in [24] and [25]. However, it should be noted that Π^{Unc} and Π^{Real} just provide the level or degree of difficulty of predictability probability but not the actual prediction method. For some time-series data, it is perhaps impossible to find an efficient method to achieve the value as Π^{Real}.

In this paper, we use a low-complexity Lempel-Ziv (LZ) algorithm [33] to find the relation between the busy and idle durations and then calculate the value of S^{Real}. LZ algorithm is based on the Lempel-Ziv compression algorithm. For a time series of length n, the entropy can be estimated as

$$S = \left(\frac{1}{n}\sum_i L_i\right)^{-1} ln(n) \tag{19}$$

where L_i is the longness of the shortest substring starting at position i which does not previously appear from position 1 to $i-1$. The estimated entropy converges to the real entropy of the time series when n approaches to infinity. It should be noted, that due to limited memory and huge size of data, the Lempel-Ziv algorithm just utilizes partial data for the S^{Real} calculation. For S^{Unc}, we just use the PDF of the busy and idle durations and Eq. (10) to calculate the value.

Table 4 shows the values of S^{Real}, S^{Unc}, Π^{Real}, Π^{Unc} and M of busy and idle durations for Channel 1 and Channel 36, respectively. Π^{Real} and Π^{Unc} show that the predictability

TABLE 4. The entropy and predictability probability.

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (Channel 1)	5.639	7.404	0.530	0.343	922
Idle (Channel 1)	6.587	8.346	0.470	0.292	1525
Busy (Channel 36)	3.949	5.413	0.656	0.503	574
Idle (Channel 36)	6.784	8.918	0.510	0.330	3901

probability for Channel 36 over 5 GHz has larger value than that of Channel 1 over 2.4 GHz band. In addition, for both idle and busy durations at Channel 1, the LB of predictability for methods only using PDF information can reach about 29%–35%. When using the relation information among the data, the predictability can be improved to 47%–53% as UB values. On the other hand, for busy and idle duration data captured over 5 GHz band, both LB and UB values of predictability are larger than that of over 2.4 GHz band. From Fig. 4 and Fig. 5, it is easy to know that the spectrum usage over 5 GHz band is rarer than that of 2.4 GHz band which makes the distribution properties of both idle and busy durations at 5 GHz band more concentrated than that of 2.4 GHz band. Therefore, the predictability is higher than that of over 2.4 GHz band.

V. PREDICTABILITY IMPROVEMENT WITH DATA CATEGORIZATION

From the previous section, we can find that the correlation property highly decides the predictability of time-series data. If there is no correlation among the succeeding data, such as time-series data of additive white Gaussian noise, it is not possible to find a larger predictability than that of Π^{Unc}. In other words, the predictability difference as (Π^{Real}-Π^{Unc}) is from the correlation property among the succeeding data. Therefore, how to exploit the correlation property becomes a challenging and difficult target which involves in the data volume, dimensions and length of time-series data.

It is difficult to find the correlation property of data in any dimensions. However, it is possible to derive the relationship among data with limited length. In this section, we first show such relationship of the busy/idle duration data of the previous section. Then a method named as data categorization is explained to improve the predictability of the partial busy and idle duration data.

A. ADAPTIVE KERNEL DENSITY ESTIMATION

To find the relationship among succeeding status durations, here we use adaptive kernel density estimation (AKDE) technique to estimate the joint probability distribution of succeeding status durations. AKDE is an adaptive method to approximate the joint distribution by adjusting its sampling resolution in order to reduce the approximation error. Here, we just give a brief introduction of AKDE. The reader can find its tutorial introduction in reference [34]. In addition, more research results on the relation of joint distribution of succeeding busy/idle durations has been reported in our paper as [35].

For density estimation, usually a histogram graph can provide its density result but with many discontinuous points. Therefore, some smooth symmetric functions are employed to interpolate the adjacent density points to make the curve more smooth. These smooth symmetric functions are termed as kernel function. There are many types of kernel function such as fisher kernel, polynomial kernel etc. These functions provide different approximate capability by adjusting the smallest interpolation resolution which is termed as bandwidth **H**.

Let us consider the analysis of a L-variate data $(x_1, \ldots, x_L)$ from an unknown density function, $f(\mathbf{x})$, where $\mathbf{x} \in \Re^L$. Usually, a normal kernel with Gaussian distribution is utilized for density estimate. For AKDE, it adaptively adjusts the smooth bandwidth h to achieve the smallest distance between the estimated density $\hat{f}(\mathbf{x})$ and the true density $f(\mathbf{x})$. Usually, the mean integrated square error (MISE) is used for this likelihood criterion as

$$\mathrm{MISE}(\hat{f}) = \mathbf{E}\left\{\int \left[(f(\mathbf{x}, \mathbf{H}) - \hat{f}(\mathbf{x}))^2\right] \mathrm{d}\mathbf{x}\right\} \tag{20}$$

where matrix **H** is a smooth bandwidth matrix for adaptive adjustment, and $\mathbf{E}\{.\}$ is the mean function over all **x**.

For multivariate density estimation with a normal kernel, $K \sim N(0, \Sigma)$, the AKDE with n samples $\mathbf{x}_i \in \Re^L$ $(i = 1, \ldots, n)$ is given by

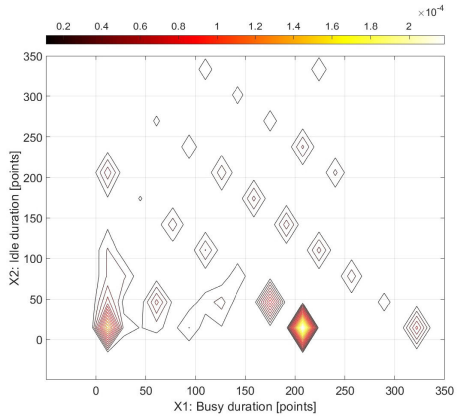
$$\hat{f}(\mathbf{x}) = \frac{1}{n(2\pi)^{L/2}|\Sigma|^{1/2}} \sum_{i=1}^{n} e^{[-\frac{1}{2}(\mathbf{x}-\mathbf{x}_i)'\Sigma(\mathbf{x}-\mathbf{x}_i)]}. \tag{21}$$

It should be noted that AKDE can find an approximate density function for the data but neglects some small values in a joint distribution to minimize the scale value of MISE. In addition, due to adjusting bandwidth **H** and interpolation, some parts of the density function may be out of the original data range.

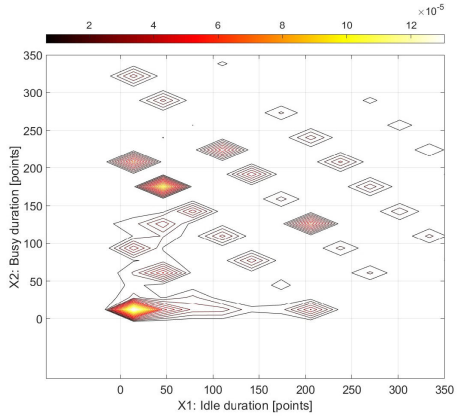
Figures 7 and 8 show the 2-D probability distribution obtained by AKDE for estimating the joint density of the busy and idle duration data captured on a railway station over Channel 1 at the 2.4 GHz band and over Channel 36 at the 5 GHz band. To find the relationship between the busy and idle durations, we generate two distributions. The first one (x_i^{Type} = Busy, x_{i+1}^{Type} = Idle) means that the first data is busy duration and the second data is idle duration following the previous busy one. The other one (x_i^{Type} = Idle, x_{i+1}^{Type} = Busy) is the similar process but the first data is idle duration and the second data is the following busy duration. It should be noted that two types have different meanings. The first type reflects how long the channel will be idle after transmitting a signal with a specific busy duration, and the second one shows the signal duration after a specific idle duration. Therefore, it is expected that there is some difference between the two distributions.

From both figures, we can find that over Channel 1 at the 2.4 GHz band, the joint density has more separated areas than that of over Channel 36 at the 5 GHz band. The reason is mainly due to that more wireless traffic and types of services utilize the channel over the 2.4 GHz band than that of the 5 GHz band. Different type of service diversifies the busy-idle or idle-busy durations and creates more patterns.

Another interesting result is that data categorization as $\{x_{i+1}|x_i\}$ has different range. For example with Fig. 7(a), when busy duration is below 50 points, the idle duration is mainly below 240 points. On the other hand, when busy duration is between 150 points to 250 points, the idle duration

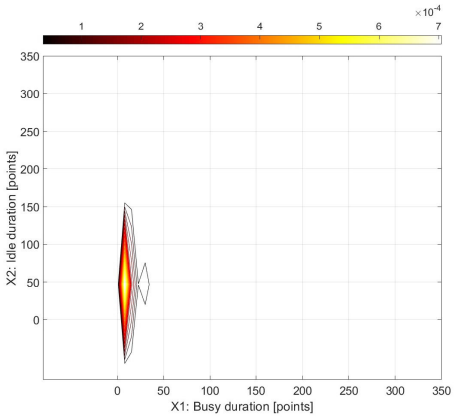


(a) 2-D probability distribution of (X1: x_i^{Type} = Busy, X2: x_{i+1}^{Type} = Idle) obtained by AKDE.

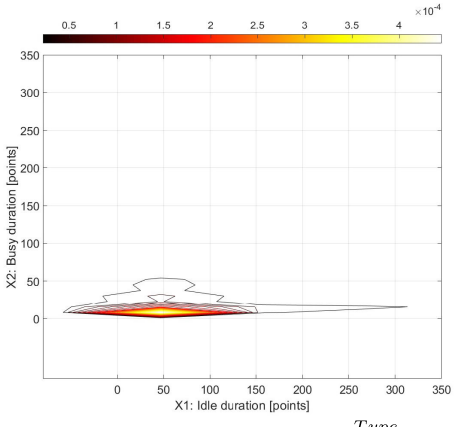


(b) 2-D probability distribution of (X1: x_i^{Type} = Idle, X2: x_{i+1}^{Type} = Busy) obtained by AKDE.

FIGURE 7. Probability distribution of 2-D succeeding B/I duration obtained by AKDE (railway station @ Channel 1, 2.4 GHz band).



(a) 2-D probability distribution of (X1: x_i^{Type} = Busy, X2: x_{i+1}^{Type} = Idle) obtained by AKDE.



(b) 2-D probability distribution of (X1: x_i^{Type} = Idle, X2: x_{i+1}^{Type} = Busy) obtained by AKDE.

FIGURE 8. Probability distribution of 2-D succeeding B/I duration obtained by AKDE (railway station @ Channel 36, 5 GHz band).

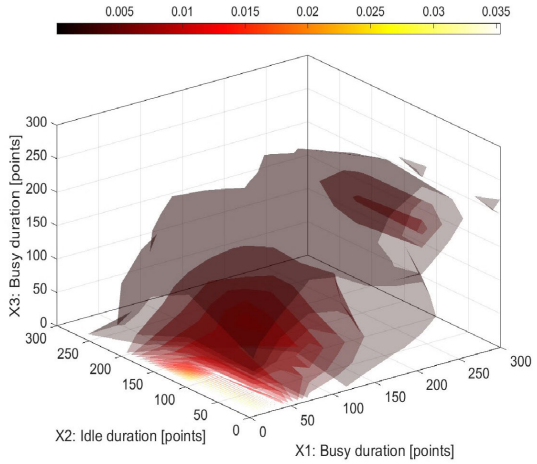
is mainly below 450 points which has a larger range on data duration.

To further show the relationship of recent status durations and the next status duration, we show the probability distribution of 3-D succeeding status duration obtained by AKDE in Fig. 9 for Channel 1 in the 2.4 GHz band and Fig. 10 in Channel 36 in the 5 GHz band. We also generate two distributions for each channel. The first one as (x_i^{Type} = Busy, x_{i+1}^{Type} = Idle, x_{i+2}^{Type} = Busy) is for two continuous channel statuses (x_i, x_{i+1}) and the upcoming status (x_{i+2}) is busy. The other one is with (x_i^{Type} = Idle, x_{i+1}^{Type} = Busy, x_{i+2}^{Type} = Idle) for two continuous channel durations (x_i, x_{i+1}) with the next one (x_{i+2}) as idle. Both figures show the data categorization as $\{x_{i+2}|x_{i+1}, x_i\}$ will have different range especially for the data from Channel 1 in the 2.4 GHz band. For example with Fig. 9(a), when busy duration (x_i) is below 50 points and idle duration (x_{i+1}) is below 100 points, the busy duration (x_{i+2}) is mainly below 100 points. When increasing the range of x_i and x_{i+1}, the value of busy duration (x_{i+2}) is distributed in a larger range. In addition, the value of idle duration (x_{i+2}) in Fig. 9(b) mainly distributes a larger range in each data category than that of busy duration in Fig. 9(a). It should be noted that each categorized distribution in the 5 GHz band usually has a smaller range compared with that of data in the 2.4 GHz band.

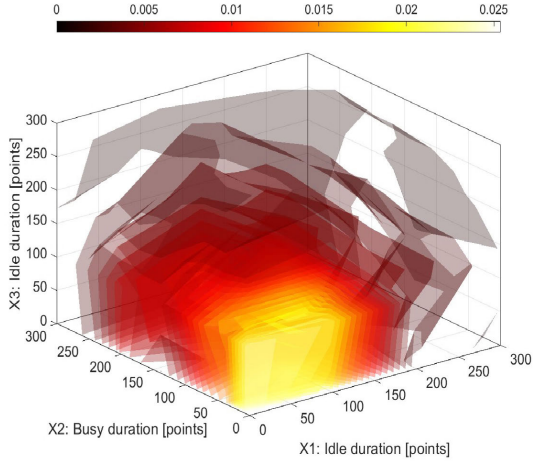
From the relationships of 2-D and 3-D succeeding status durations, by setting the previous status durations into different ranges, the upcoming status durations can be distributed into different categorization with different range. This data categorization can be used for improving the prediction accuracy for some categorizations where durations are distributed with a smaller range.

B. DATA CATEGORIZATION

Although the predictability just provides a measurement scale to show whether data is easy to be predicted or not, it cannot provide how to realize its upper bound or even lower bound with a specific prediction method. For some data, it is not possible to find one best prediction method to achieve its upper bound. Usually, the lower bound of predictability only depends on the PDF of data, the higher LB means that data is seldom changed with large range. In this case, the data is easy

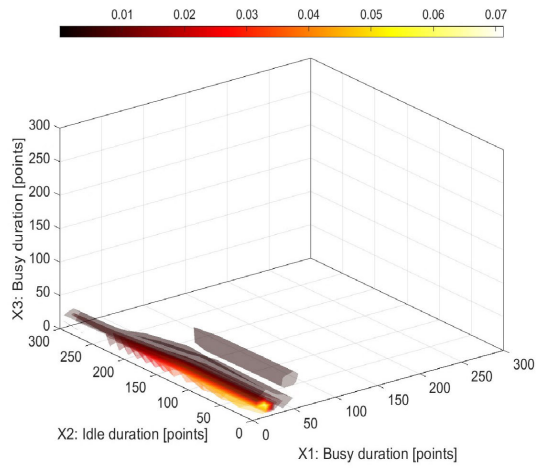


(a) 3-D probability distribution of (X1: x_i^{Type} = Busy, X2: x_{i+1}^{Type} = Idle, X3: x_{i+2}^{Type} = Busy) obtained by AKDE.

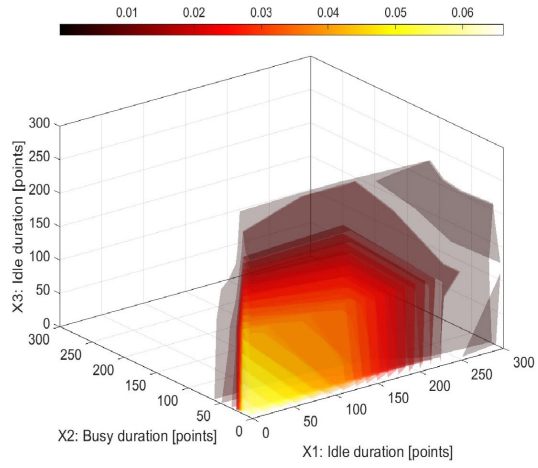


(b) 3-D probability distribution of (X1: x_i^{Type} = Idle, X2: x_{i+1}^{Type} = Busy, X3: x_{i+2}^{Type} = Idle) obtained by AKDE.

FIGURE 9. Probability distribution of 3-D succeeding status duration obtained by AKDE (railway station @ Channel 1, 2.4 GHz band).



(a) 3-D probability distribution of (X1: x_i^{Type} = Busy, X2: x_{i+1}^{Type} = Idle, X3: x_{i+2}^{Type} = Busy) obtained by AKDE.



(b) 3-D probability distribution of (X1: x_i^{Type} = Idle, X2: x_{i+1}^{Type} = Busy, X3: x_{i+2}^{Type} = Idle) obtained by AKDE.

FIGURE 10. Probability distribution of 3-D succeeding status duration obtained by AKDE (railway station @ Channel 36, 5 GHz band).

for prediction. Therefore, if the lower bound of predictability of time-series data is enlarged, the prediction accuracy of some prediction methods might be improved. However, for some prediction methods which explore the time relationship among the time-series data, it is hard to further improve their prediction accuracy.

From the previous sections, by setting the preceding status durations into different ranges, the upcoming status durations can be distributed into different categorization with different range. This categorization will change the value of the predictability of partial data. Based on this idea, we propose data categorization to separate time-series data into different streams. The major idea is shown in Fig. 11 using busy/idle duration data as one example. The captured time-series busy or idle duration data x_i can be divided into four streams according to the sets (S_1, S_2) that the previous values x_{i-2} and x_{i-1} belong to. As the example shown in Fig. 11, both busy and idle durations have two sets as (B_set1, B_set2) and (I_set1, I_set2) with different non-overlap ranges,

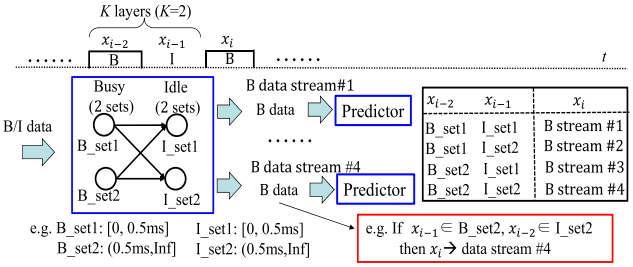


FIGURE 11. Data categorization for busy and idle durations.

respectively. The busy duration X_i will be separated to the busy stream # 4 if previous busy duration X_{i-2} and idle duration X_{i-1} belong to B_set2 and I_set2, respectively.

The idea can be extended to K layers. The captured busy and idle durations are firstly processed by a B/I duration categorization with K layers and each layer has S_i $(i = 1, \ldots, K)$ sets. Therefore, there are S_{all} $(S_{all} = \prod_{i=1}^{i=K} S_i)$ different streams. By setting the ranges of sets and number of layers

and sets, the data categorization can separate all time-series data into different streams with different statistical properties and predictability.

Figure 12 shows how the proposed method works for the prediction of partial time-series data (stream #1). When the condition that $(x_{i-1}, x_i) \in (B_{set1}, I_{set1})$ is met, the prediction method is activated for the predicted busy duration x_{i+1}. The previous n-length busy data ($[x_{k-n+1}^{B\#1}, x_{k-n+2}^{B\#1}, \ldots, x_k^{B\#1}]$) which belongs to stream #1 has been saved for prediction in advance. The prediction process includes two steps:

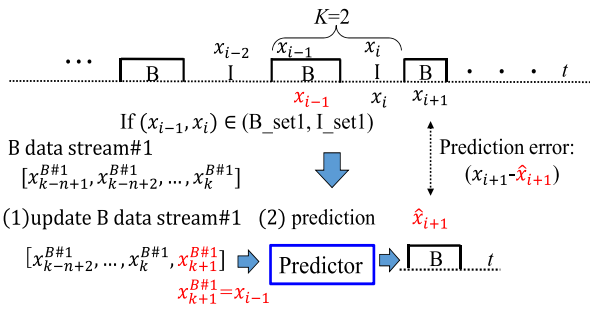


FIGURE 12. Example using data categorization for partial data prediction (stream # 1).

(1) update the n-length busy data as $[x_{k-n+2}^{B\#1}, x_{k-n+3}^{B\#1}, \ldots, x_k^{B\#1}, x_{k+1}^{B\#1}]$ with $x_{k+1}^{B\#1} = x_{i-1}$.
(2) predict the next busy duration $\hat{x}_{i+1}$ using one specific prediction algorithm.

It should be noted that data categorization is not limited to the continuous succeeding data such as using n-length data $\mathbf{X}_n$ ($[x_{i-n+1}, x_{i-n+2}, \ldots, x_i]$) for estimating the data x_{i+1}. It can be upgraded to utilize $\mathbf{X}_n$ for predicting the l-th coming data x_{i+l} if the correlation between $\mathbf{X}_n$ and x_{i+l} exists and the correlation property is clarified.

C. PREDICTABILITY IMPROVEMENT USING DATA CATEGORIZATION

To show the effectiveness of data categorization, we set the layer number K as 2. Both busy and idle durations are categorized with two sets which is similar to that of Fig. 11. We also consider two configurations of set ranges which are shown in Table 5, and they are named as CASE I and CASE II, respectively. CASE II has the larger range than that of CASE I to show the impact on predictability from the different statistical property of each stream. After data categorization, each busy/idle stream is calculated with the values of S^{Real}, S^{Unc}, Π^{Real}, Π^{Unc} and M using LZ algorithm and their PDFs, respectively. In addition, to show how the DC changes the values of S^{Unc} and Π^{Unc}, the PDF of each busy or idle stream after DC process is also provided.

Figures 13 and 14 show the PDF of busy and idle duration captured from Channel 1 of 2.4 GHz band with DC for CASE I and CASE II, respectively. For comparison, the PDF results of all busy or idle duration data without DC are also provided in each figure. From all figures, it can be found that DC diversifies the PDF of each stream.

TABLE 5. Set range settings [point].

CASE		Channel 1@2.4 GHz	Channel 36@5 GHz
CASE I	$B\_set1; B\_set2$	[0, 40);[40, ∞]	[0, 20);[20, ∞]
	$I\_set1; I\_set2$	[0, 150);[150, ∞]	[0, 20);[20, ∞]
CASE II	$B\_set1; B\_set2$	[0, 100);[100, ∞]	[0, 40);[40, ∞]
	$I\_set1; I\_set2$	[0, 100);[100, ∞]	[0, 40);[40, ∞]

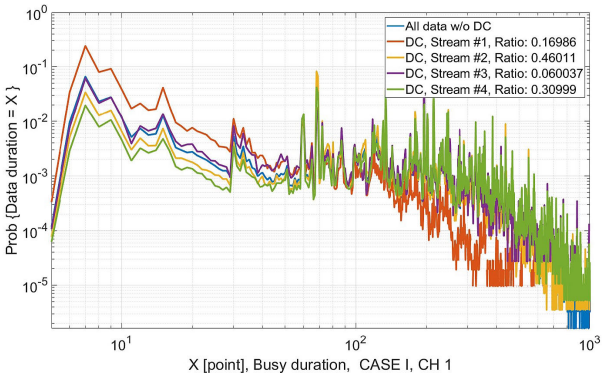
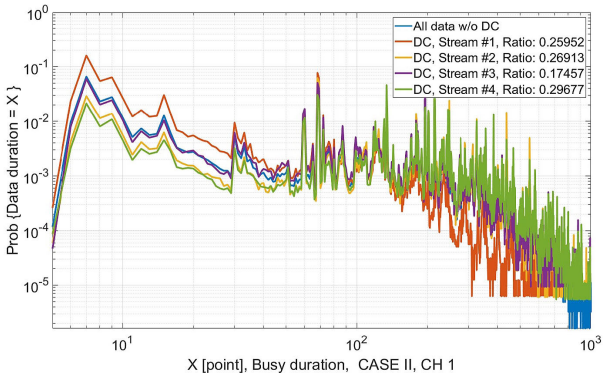


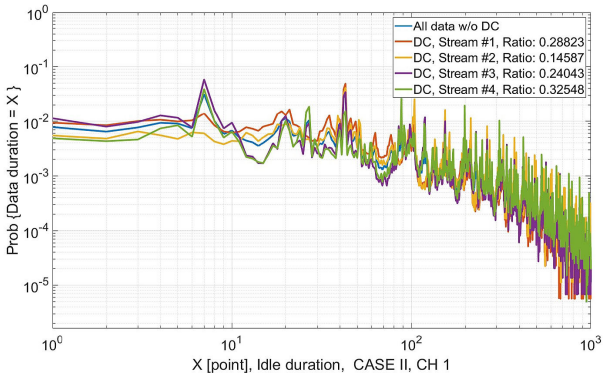
FIGURE 13. PDF of busy / idle duration with and without DC (railway station @ Channel 1, 2.4 GHz band, CASE I).

From the Eq. (6), the concentrated PDF which distributed with large values over a concentrated range usually has the smaller entropy than that of non-concentrated one. Using Fig. 13(a) as an example, compared with that of busy duration without DC, the stream #1, which occupies about 17% of all busy durations, shows a concentrated PDF property. Therefore, it should have a smaller S^{Unc} and then a larger lower bound Π^{Unc} than that of busy durations without DC. With the same reason, it also can be found that DC deteriorates the lower bound Π^{Unc} of other three streams. On the other hand, when the range of each set is changed, the PDF of each stream is also adjusted as shown in both Figs. 13 and 14 which makes different Π^{Unc} for each stream.

Table 6 and Table 7 show the entropy and predictability probability of busy/idle duration data obtained from Channel 1 after data categorization with CASE I and CASE II, respectively. As a comparison, the entropy and predictability of busy and idle data without DC are listed in both tables. From both tables, which are in accord with the previous figures and analysis, it can be found that the



(a) Busy duration



(b) Idle duration

FIGURE 14. PDF of busy / idle duration with and without DC (railway station @ Channel 1, 2.4 GHz band, CASE II).

TABLE 6. The entropy and predictability probability with data categorization (Channel 1, CASE I).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (all data)	5.639	7.404	0.530	0.353	922
Busy (Stream #1)	4.241	5.423	0.640	0.510	532
Busy (Stream #2)	5.679	7.260	0.520	0.345	868
Busy (Stream #3)	5.639	7.404	0.530	0.343	902
Busy (Stream #4)	6.309	7.820	0.470	0.310	892
Idle (all data)	6.587	8.346	0.470	0.292	1525
Idle (Stream #1)	6.508	8.057	0.470	0.310	1347
Idle (Stream #2)	6.137	7.624	0.500	0.350	1258
Idle (Stream #3)	6.486	8.173	0.475	0.305	1461
Idle (Stream #4)	6.800	8.466	0.455	0.290	1510

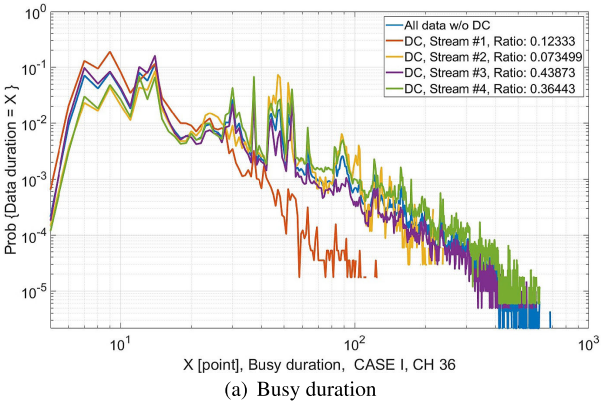
TABLE 7. The entropy and predictability probability with data categorization (Channel 1, CASE II).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (all data)	5.639	7.404	0.530	0.343	922
Busy (Stream #1)	4.448	5.988	0.630	0.470	649
Busy (Stream #2)	6.153	7.585	0.480	0.320	943
Busy (Stream #3)	5.837	7.333	0.510	0.340	877
Busy (Stream #4)	6.354	7.831	0.460	0.310	907
Idle (all data)	6.587	8.346	0.470	0.292	1525
Idle (Stream #1)	6.192	7.840	0.500	0.330	1282
Idle (Stream #2)	6.422	8.009	0.480	0.320	1415
Idle (Stream #3)	6.691	8.312	0.460	0.300	1516
Idle (Stream #4)	6.927	8.570	0.445	0.280	1520

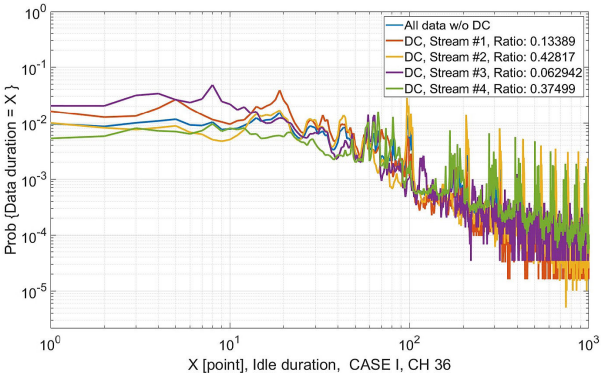
proposed method can increase lower bound Π^{Unc} of the first stream, especially for busy data.

However, although DC can increase the lower bound Π^{Unc} of some streams by adjusting their PDF properties, it is hard to improve their UB values Π^{Real}. The major reason is that the value of Π^{Real} is mainly decided by the correlation properties among the time-series data, and DC is difficult to deepen these correlation properties in each stream by setting different layer number or set ranges. On the contrary, inappropriate ranges of sets sometimes make the predictability of time-series data decrease because DC weakens the correlation properties among data in such case.

Figures 15 and 16 show the PDF of busy and idle duration captured from Channel 36 of 5 GHz band with DC for CASE I and CASE II, respectively. These figures show the similar results to that of Figs. 13 and 14. Table 8 and Table 9 show the entropy and predictability probability of busy/idle duration data obtained from Channel 36 after DC with CASE I and CASE II, respectively. Compared with the busy data of Channel 1 over 2.4 GHz band, busy durations are short which are also shown in Figs. 13 and 14. The probability distribution of busy duration will be more concentrated than that of over 2.4 GHz band. Therefore, both S^{Real} and S^{Unc} of busy data is smaller than that of data obtained over 2.4 GHz band. In addition, using data categorization, the similar results to that of Table 6 and Table 7 are obtained.



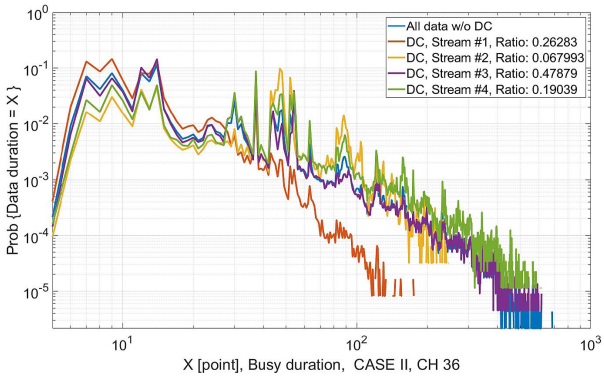
(a) Busy duration



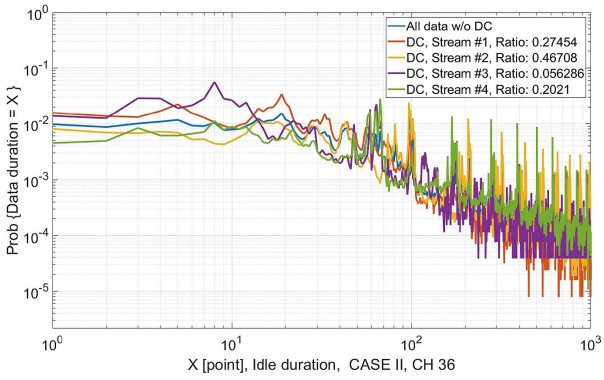
(b) Idle duration

FIGURE 15. PDF of busy / idle duration with and without DC (railway station @ Channel 36, 5 GHz band, CASE I).

From the results of all tables, we can find that the proposed DC can change the Π^{Real} and Π^{Unc}, especially for Π^{Unc},



(a) Busy duration



(b) Idle duration

FIGURE 16. PDF of busy / idle duration with and without DC (railway station @ Channel 36, 5 GHz band, CASE II).

TABLE 8. The entropy and predictability probability with data categorization (Channel 36, CASE I).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (all data)	3.949	5.413	0.656	0.503	574
Busy (Stream #1)	3.223	3.976	0.670	0.570	119
Busy (Stream #2)	3.771	4.932	0.680	0.550	458
Busy (Stream #3)	4.483	5.691	0.550	0.390	223
Busy (Stream #4)	4.997	6.162	0.560	0.430	537
Idle (all data)	6.784	8.918	0.510	0.330	3901
Idle (Stream #1)	6.092	7.721	0.550	0.400	2568
Idle (Stream #2)	6.171	7.949	0.540	0.370	2279
Idle (Stream #3)	6.418	8.209	0.530	0.370	2860
Idle (Stream #4)	7.498	9.741	0.470	0.270	3794

of selected partial data from all time-series data. Therefore, the proposed method improves the prediction accuracy of these partial data using some simple prediction methods.

VI. BUSY/IDLE DURATION PREDICTION USING AUTO-REGRESSIVE PREDICTOR WITH DATA CATEGORIZATION

In this section, we will use a simple auto-regressive (AR) predictor to compare the prediction accuracy of time-series data with and without the proposed data categorization.

A. AUTO-REGRESSIVE (AR) PREDICTOR

Auto-regressive (AR) model is a representation of a type of random process [36]. The AR model specifies that the current value is only related to its own previous values and a stochastic term; thus the model is in the form of a stochastic difference equation.

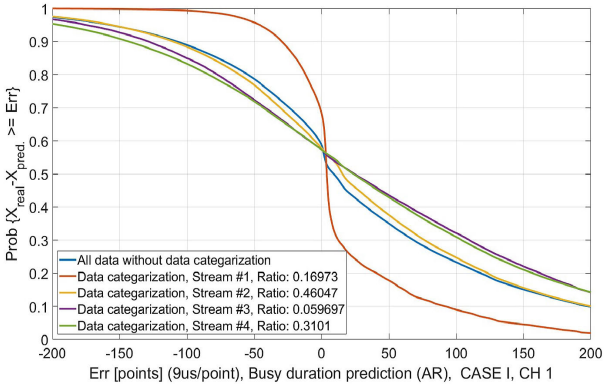
The AR-based predictor with order p can be represented as

$$\hat{X}_{i+1} = a_1 X_i + a_2 X_{i-1} + \ldots a_p X_{i-p+1}. \quad (22)$$

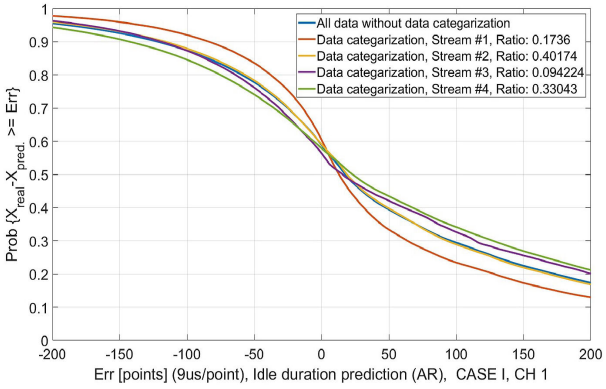
Here we use X_i to represent the duration of busy or idle status. The parameters $[a_1, \ldots, a_p]$ are calculated using training data so that they give the solution as the least squares for linear regression. The reader can find more details of AR model in [36].

TABLE 9. The entropy and predictability probability with data categorization (Channel 36, CASE II).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (all data)	3.949	5.413	0.656	0.503	574
Busy (Stream #1)	3.266	4.215	0.670	0.550	148
Busy (Stream #2)	4.148	5.301	0.645	0.515	487
Busy (Stream #3)	4.608	5.807	0.550	0.390	250
Busy (Stream #4)	5.160	6.332	0.540	0.410	552
Idle (all data)	6.784	8.918	0.510	0.330	3901
Idle (Stream #1)	6.329	7.949	0.540	0.400	3233
Idle (Stream #2)	6.131	7.979	0.540	0.370	2173
Idle (Stream #3)	6.937	8.969	0.500	0.325	3798
Idle (Stream #4)	7.234	9.375	0.480	0.300	3874

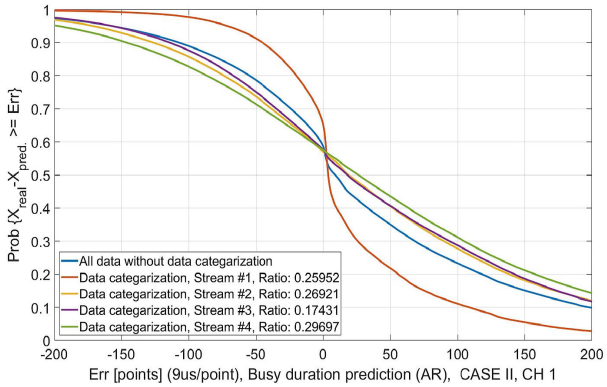


(a) Busy duration

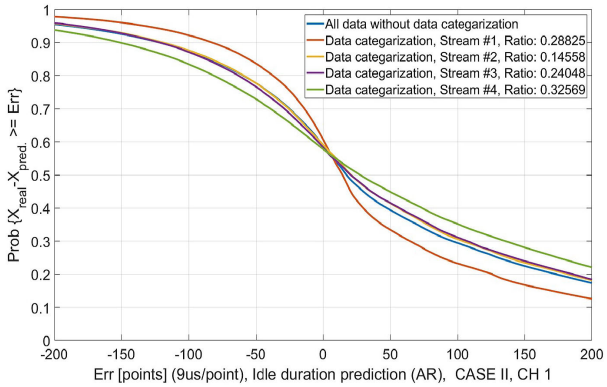


(b) Idle duration

FIGURE 17. Busy/idle duration prediction using the proposed method (railway station @ Channel 1, 2.4 GHz band, CASE I).

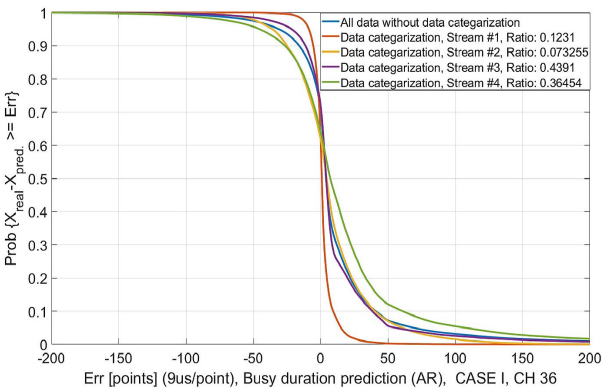


(a) Busy duration

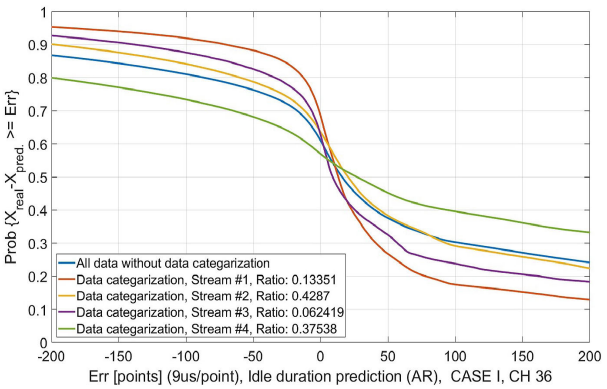


(b) Idle duration

FIGURE 18. Busy / idle duration prediction using the proposed method (railway station @ Channel 1, 2.4 GHz band, CASE II).



(a) Busy duration



(b) Idle duration

FIGURE 19. Busy / idle duration prediction using the proposed method (railway station @ Channel 36, 5 GHz band, CASE I).

The AR predictor for the busy/idle duration prediction with data categorization is almost identical structure as shown in Fig. 12. Here AR predictor is operated in each stream, independently. During the training process, k-th data stream utilizes its previous M durations to calculate the parameters $[a_1^k, \ldots, a_p^k]$. Then the following N durations are predicted using the similar equation as Eq. (22) for each data stream. For comparison, we also provide the prediction performance of busy or idle durations without data categorization, which means only one data stream with all busy or idle durations are used for training and prediction process.

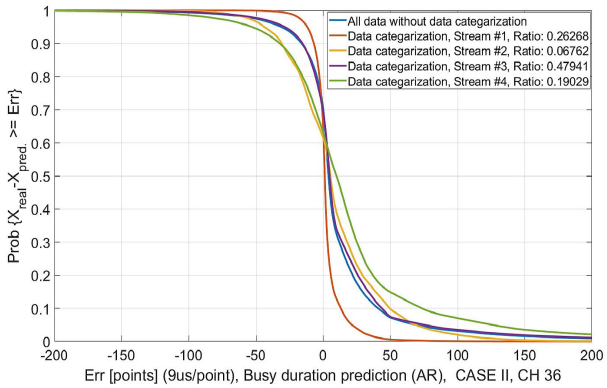
B. PREDICTION PERFORMANCE OF AR USING DATA CATEGORIZATION

We assume that the AR predictor uses the durations of the previous 200 busy/idle durations to calculate a_1 to a_p once for predicting the durations of the upcoming 200 busy/idle statues ($M = 200, N = 200$). In preliminary simulations, we have set the AR order p with different value and found that AR predictor with $p = 2$ can get the best prediction performance. In this paper, we just provide the results of AR predictor with $p = 2$. For data categorization, we utilize the same configurations of set ranges, which are shown in Table 5, as CASE I and CASE II. For the results of each stream, we also provide the ratio of the data in this stream to all data to show the percentage of data with different prediction performance.

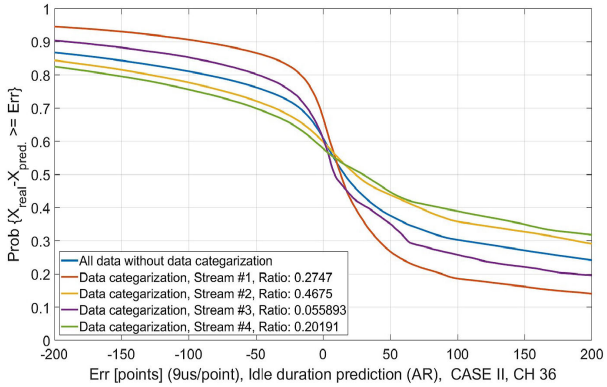
The evaluated prediction performances are shown in Fig. 17 for busy /idle duration prediction for CASE I and in Fig. 18 for CASE II on Channel 1 over the 2.4 GHz band. The x-axis shows the prediction error (Err) and y-axis shows the complementary cumulative distribution function (CCDF) of prediction error which shows that the prediction error is not smaller than Err.

Both figures show that the proposed method can differentiate the prediction performance of each stream, especially for busy duration prediction. The proposed data categorization shows better prediction performance for the stream #1 with 17% and 26% total data for CASE I and CASE II than that of without categorization process, respectively. For one example, the proposed DC of stream #1 can improve about 37% and 32% prediction accuracy when Err is ranged among $[-15, 15]$ points. On the other hand, for idle duration prediction, two range settings cannot improve the prediction performance. However, from Table 6 and Table 7, we can find the prediction results match highly with the predictability results, especially with Π^{Unc}. The proposed data categorization can improve the prediction accuracy of partial data even if system employs a simple AR predictor.

The prediction performances are shown in Fig. 19 for busy/idle duration prediction for CASE I and in Fig. 20 for CASE II on Channel 36 over the 5 GHz band. For data captured from Channel 36 in 5 GHz band, generally,



(a) Busy duration



(b) Idle duration

FIGURE 20. Idle duration prediction using the proposed method (railway station @ Channel 36, 5 GHz band, CASE II).

busy duration prediction performs reasonably well and our proposed method can further improve its accuracy for the stream #1. The major reason is that the busy duration on Channel 36 distributes within a small range and data categorization will further reduce its range. For idle duration prediction, the stream #1 of both CASE I and CASE II evidently shows better prediction performance than that of without DC. From the occurrence ratio of each stream, all figures show that our proposed method can improve the prediction performance for over about 20% of the total idle data and over about 12% of the total busy data. From Table 8 and Table 9, we can also find the prediction results match highly with the predictability results. The proposed data categorization can improve the prediction accuracy of a part of data.

It should be noted that it is still a challenge to find the suitable parameters K, S_i at the i-th layer and the range of each set to optimize the predictability and prediction performance because the real-environment traffic data includes many service combinations and mixed traffic pattern. In addition, how to find the optimal prediction method to achieve the upper bound of predictability is still open issue. These issues are our ongoing research topics.

VII. CONCLUSION

This paper first investigated the spectrum occupancy status and its modeling analysis for a heavy WLAN traffic environment at a major railway station. Our measurements can provide more realistic parameters for considering scenarios over heavy WLAN traffic environments. The fitting models of busy and idle duration can be used as a complement for building useful models such as interference model, simple multi-band CR spectrum occupancy model for CR system research. We then analyzed the predictability analysis on busy/idle durations of two selected channels at the 2.4 GHz and 5 GHz bands. From the predictability analysis, we proposed a method called data categorization to separate all busy/idle durations into several streams with different distribution property. This separation facilitates some streams to be more easily predicted. The results show that, even using a simple AR based predictor, data categorization can improve the prediction accuracy of some streams with a high predictability.

ACKNOWLEDGMENT

The authors would like to thank Ministry of Internal Affairs and Communications of Japan for providing the data for research purpose in this paper.

REFERENCES

[1] A. Ijaz, L. Zhang, M. Grau, A. Mohamed, S. Vural, A. U. Quddus, M. A. Imran, C. H. Foh, and R. Tafazolli, "Enabling massive IoT in 5G and beyond systems: PHY radio frame design considerations," *IEEE Access*, vol. 4, pp. 3322–3339, 2016.

[2] E. Khorov, I. Levitsky, and I. F. Akyildiz, "Current status and directions of IEEE 802.11be, the future Wi-Fi 7," *IEEE Access*, vol. 8, pp. 88664–88688, 2020.

[3] R. Struzak, "Cognitive radio, spectrum, and evolutionary heuristics," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 166–171, Jun. 2018.

[4] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

[5] M. Song, C. Xin, Y. Zhao, and X. Cheng, "Dynamic spectrum access: From cognitive radio to network radio," *IEEE Wireless Commun.*, vol. 19, no. 1, pp. 23–29, Feb. 2012.

[6] H. Sun, A. Nallanathan, C.-X. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 74–81, Apr. 2013.

[7] C. Sun, G. P. Villardi, Z. Lan, Y. D. Alemseged, H. N. Tran, and H. Harada, "Optimizing the coexistence performance of secondary-user networks under primary-user constraints for dynamic spectrum access," *IEEE Trans. Veh. Technol.*, vol. 61, no. 8, pp. 3665–3676, Oct. 2012.

[8] M. H. Islam, C. L. Koh, S. W. Oh, X. Qing, Y. Y. Lai, C. Wang, Y.-C. Liang, B. E. Toh, F. Chin, G. L. Tan, and W. Toh, "Spectrum survey in singapore: Occupancy measurements and analyses," in *Proc. 3rd Int. Conf. Cognit. Radio Oriented Wireless Netw. Commun. (CrownCom)*, Singapore, May 2008, pp. 1–7.

[9] L. Stabellini, "Quantifying and modeling spectrum opportunities in a real wireless environment," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Sydney, NSW, Australia, Apr. 2010, pp. 1–6.

[10] S. Contreras, G. Villardi, R. Funada, and H. Harada, "An investigation into the spectrum occupancy in japan in the context of TV white space systems," in *Proc. 6th Int. ICST Conf. Cognit. Radio Oriented Wireless Netw. Commun.*, Osaka, Japan, 2011, pp. 341–345.

[11] Y. Chen and H.-S. Oh, "A survey of measurement-based spectrum occupancy modeling for cognitive radios," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 848–859, 1st Quart., 2016.

[12] S. Geirhofer, L. Tong, and B. Sadler, "A measurement-based model for dynamic spectrum access in WLAN channels," in *Proc. MILCOM*, Oct. 2006, pp. 1–7.

[13] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in WLAN channels: Empirical model and its stochastic analysis," in *Proc. 1st Int. Workshop Technol. Policy Accessing Spectr. (TAPAS)*, 2006, pp. 1–10.

[14] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 66–72, May 2007.
[15] S. A. Rajab, W. Balid, and H. H. Refai, "Toward enhanced wireless coexistence in ISM band via temporal characterization and modelling of 802.11b/g/n networks," *Wireless Commun. Mobile Comput.*, vol. 16, no. 18, pp. 3212–3229, Nov. 2016.
[16] L. B. Miguel and F. Casadevall, "Time-domain model of spectrum usage for the analysis, design, and simulation of cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2091–2104, Jun. 2013.
[17] L. B. Miguel and F. Casadevall, "Space-domain model of spectrum usage for cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 306–320, Jan. 2017.
[18] K. Umebayashi, M. Kobayashi, and M. Lopez-Benitez, "Efficient time domain deterministic-stochastic model of spectrum usage," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1518–1527, Mar. 2018.
[19] M. O. Al Kalaa, W. Balid, H. H. Refai, N. J. LaSorte, S. J. Seidman, H. I. Bassen, J. L. Silberberg, and D. Witters, "Characterizing the 2.4 GHz spectrum in a hospital environment: Modeling and applicability to coexistence testing of medical devices," *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 1, pp. 58–66, Feb. 2017.
[20] M. Luis, R. Oliveira, R. Dinis, and L. Bernardo, "RF-spectrum opportunities for cognitive radio networks operating over GSM channels," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 731–739, Dec. 2017.
[21] N. Bui, M. Cesana, S. A. Hosseini, Q. Liao, I. Malanchini, and J. Widmer, "A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1790–1821, 3rd Quart., 2017.
[22] N. Egashira, K. Yano, S. Tsukamoto, J. Webber, M. Sutoh, Y. Amezawa, and T. Kumagai, "Integrated synchronization scheme for WLAN systems employing multiband simultaneous transmission," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, San Francisco, CA, USA, Mar. 2017, pp. 1–5.
[23] K. Yano, N. Egashira, J. Webber, M. Usui, and Y. Suzuki, "Achievable throughput of multiband wireless LAN using simultaneous transmission over multiple primary channels assisted by idle length prediction based on PNN," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Naha, Okinawa, Feb. 2019, pp. 22–27.
[24] G. Ding, J. Wang, Q. Wu, Y.-D. Yao, R. Li, H. Zhang, and Y. Zou, "On the limits of predictability in real-world radio spectrum state dynamics: From entropy theory to 5G spectrum sharing," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 178–183, Jul. 2015.
[25] J. Sun, L. Shen, G. Ding, R. Li, and Q. Wu, "Predictability analysis of spectrum state evolution: Performance bounds and real-world data analytics," *IEEE Access*, vol. 5, pp. 22760–22774, 2017.
[26] *IEEE Standard for Information Technology—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Standard IEEE 802.11-2016, 2016.
[27] M. Wellens, J. Riihijärvi, and P. Mähönen, "Empirical time and frequency domain models of spectrum use," *Phys. Commun.*, vol. 2, nos. 1–2, pp. 10–32, Mar. 2009.
[28] M. Uchida, "Traffic data analysis based on extreme value theory and its applications," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2007, pp. 1–7.
[29] C. Liu, Y. Shu, J. Liu, and O. W. W. Yang, "Application of extreme value theory to the analysis of wireless network traffic," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2007, pp. 486–491.
[30] P. Sahoo, *Probability and Mathematical Statistics*. Louisville, KY, USA: Univ. of Louisville, 2013.
[31] W. H. Press, S. A. Teukolsky, W. T. Vertterling, and B. P. Flannery, *Numerical Recipes, The Art of Scientific Computing*, 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[32] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, no. 1, pp. 99–109, 1943.
[33] I. Kontoyiannis, P. H. Algoet, Y. M. Suhov, and A. J. Wyner, "Nonparametric entropy estimation for stationary processes and random fields, with applications to English text," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1319–1327, May 1998.
[34] D. Scott and S. Sain, "Multi-dimensional density estimation," *Handbook of Statistics*, vol. 24, pp. 229–261, Apr. 2005, doi: 10.1016/S0169-7161(04)24009-3.
[35] Y. Hou, K. Yano, S. Denno, and Y. Suzuki, "A study on joint probability distribution property of busy/idle duration and its impact on the performance of auto-regressive based predictor over real environmental channel," *IEICE Technique Rep.*, vol. 118, no. 475, pp. 89–96, Mar. 2019.
[36] R. Shumway and D. Stoffer, *Time Series Analysis and Its Applications: With R Examples*. New York, NY, USA: Springer-Verlag, 2010.

YAFEI HOU (Senior Member, IEEE) received the dual Ph.D. degree from Fudan University, China, and the Kochi University of Technology (KUT), Japan, in 2007. He was a Postdoctoral Research Fellow with Ryukoku University, Japan, from August 2007 to September 2010. He was a Research Scientist with the Wave Engineering Laboratories, ATR Institute International, Japan, from October 2010 to March 2014. He became an Assistant Professor at the Graduate School of Information Science, Nara Institute of Science and Technique, Japan, from April 2014 to March 2017. He became an Assistant Professor at the Graduate School of Natural Science and Technology, Okayama University, Japan, since April 2017. He was a Guest Research Scientist with the Wave Engineering Laboratories, ATR Institute International, from October 2016 to March 2021. His research interests include communication systems, wireless networks, and signal processing. He is also a member of IEICE. He received the Institute of Electronics, Information and Communication Engineers (IEICE) Communications Society best paper awards in 2016 and 2020, and the Best Tutorial Paper Award, in 2017.

JULIAN WEBBER (Senior Member, IEEE) received the M.Eng. degree from the University of Bristol, U.K., in 1996, and the Ph.D. degree from the University of Bristol, in 2004. He worked with Texas Instruments Europe, from September 1996 to October 1998. He was a Research Fellow with the University of Bristol, from November 2001 to August 2007, and with Hokkaido University, Japan, from September 2007 to March 2012. He was a Research Scientist with the Wave Engineering Laboratories, ATR Institute International, Japan, from April 2012 to March 2018. He became an Assistant Professor with Osaka University, since April 2018, and a Guest Research Scientist with ATR. His research interests include signal processing, and wireless communication systems design and implementation. He is also a member of IEICE.

KAZUTO YANO (Member, IEEE) received the B.E. degree in electrical and electronic engineering and the M.S. and Ph.D. degrees in communications and computer engineering from Kyoto University, in 2000, 2002, and 2005, respectively. He was a Research Fellow with the Japan Society for the Promotion of Science (JSPS), from 2004 to 2006. In 2006, he joined the Advanced Telecommunications Research Institute International (ATR). He is currently the Head of the Department of Wireless Communication Systems, Wave Engineering Laboratories, ATR. His research interests include space-time signal processing for interference suppression, MIMO transmission, and PHY/MAC cross-layer design of wireless communication systems for ISM bands. He is also a Senior Member of IEICE. He received the Institute of Electronics, Information and Communication Engineers (IEICE) Communications Society Best Tutorial Paper Award, in 2017, and the ICAIIC 2019 Excellent Paper Award, in 2019.

SHUN KAWASAKI received the B.S. and M.S. degrees from Okayama University, Japan, in 2019 and 2021, respectively. His research interests include signal processing, and wireless communication systems.

SATOSHI DENNO (Member, IEEE) received the M.E. and Ph.D. degrees from Kyoto University, Kyoto, Japan, in 1988 and 2000, respectively. He joined the NTT Radio Communications Systems Laboratories, Yokosuka, Japan, in 1988. In 1997, he was seconded to the ATR Adaptive Communications Research Laboratories, Kyoto. From 2000 to 2002, he worked with NTT DoCoMo, Yokosuka. In 2002, he moved to DoCoMo Communications Laboratories Europe GmbH, Germany. From 2004 to 2011, he worked as an Associate Professor with Kyoto University. Since 2011, he has been a Full Professor with the Graduate School of Natural Science and Technology, Okayama University. From the beginning of his research career, he has been involved in the research and development of digital mobile radio communications. In particular, he has considerable interests in channel equalization, array signal processing, space time codes, spatial multiplexing, and multimode reception. He received the Excellent Paper Award from IEICE, in 1995.

YOSHINORI SUZUKI received the B.E., M.E., and Ph.D. degrees from Tohoku University, Sendai, in 1993, 1995, and 2005, respectively. He joined the NTT Wireless Systems Laboratories, in 1995. Since then, he involved in researching microwave signal processing techniques for satellite onboard applications and onboard multiple beam antenna feed techniques. He worked as a part-time Lecturer with Niigata University, from 2012 to 2014. From 2013 to 2014, he was in charge of sales engineering of satellite communication services in NTT Software Corporation (currently NTT TechnoCross Corporation). Since then, he was a Research Engineer with the NTT Access Network Service Systems Laboratories working on future mobile satellite communication systems. Since June 2018, he has been involved in the research of innovative radio communication systems with the ATR Wave Engineering Laboratories, Kyoto, Japan.