

Busy/Idle Duration Prediction for Video and Audio WLAN Traffics Using Autoregressive Predictor with Data Categorization

Yafei Hou*, Shun Kawasaki* and Satoshi Denno*

*Graduate School of Natural Science and Technology, Okayama University, Okayama City, Japan

Contact Author: yfhou@okayama-u.ac.jp

Abstract—Due to huge access from massive devices and peoples such as hospitals, railway stations and airports, wireless local area network (WLAN) is required to have high spectrum efficiency (SE). One of the most intensively researched techniques for wireless LAN systems is cognitive radio (CR) technique which is expected to solve such issue by modeling and predicting of channel status from the current statistics information of spectrum usage. In this paper, we investigate the prediction performance of busy/idle (B/I) duration of two major and widely used wireless services : video service; and audio service using an auto-regressive (AR) based predictor. We first investigate the modeling of their busy/idle duration and analyze their predictability based on predictability theory. Then, we categorize the durations of recent B/I statuses with their ranges to make the duration of the next status be distributed into different sets or streams with different ranges. From their predictability and prediction performance using the low-complexity AR-based predictor, we can confirm that data categorization can largely improve the prediction performance of partial time-series data.

Index Terms—Channel status prediction; WLAN traffic; Autoregressive predictor; Data categorization.

I. INTRODUCTION

The coming next-generation wireless communication network technology is supposed to hold diversified service requirements, applications and huge number of devices which is facilitating and reshaping our lives [1]. To meet this trend, wireless local area network (WLAN) is required to transfer ever-increasing amounts of data in home, office and in Wi-Fi hot-spots. Therefore, high spectrum efficiency (SE) is required for WLAN, and some efficient techniques have been prompted to further improve the SE [2]. One of the most intensively researched techniques for wireless communications is cognitive radio (CR) system which utilizes dynamically the best wireless channels resources in its vicinity to avoid interference or congestion in a smart way [3]. For this view, recent progress of using machine learning and data analysis on CR system has further cultivated the next generation CR system [4].

To achieve efficient spectrum utilization, the CR system needs to correctly capture current channel status and then to predict channel status before its wireless transmission. A tutorial paper [5] has summarized most of the existing prediction methods for the optimization of wireless resource allocation. These results have provided many efficient methods for the usage of CR 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 second, hour or day. If the system can correctly predict the start and end of channel busy or idle status, the system can more efficiently utilize the available radio resources and improve spectrum efficiency when channel is accessed by a large number of devices. For example, CR system can be designed to utilize idle periods scattered in multiple frequency bands by splitting one transmission packet into small sub-packets [6, 7] then transmitting on the multiple bands. It can largely improve the total spectrum efficiency especially for some heavy wireless traffic environments, where more spectrum resources are required and channel status prediction is extremely important but difficult to be realized [8].

Compared with capturing channel status, the prediction of channel status is more difficult and sometime impossible because the channel status is unstable in short period and the disordered traffic data makes the prediction intangible and difficult. The prediction performance of busy/idle durations using a non-linear regression function with probabilistic neural network (PNN) based predictor has been showed in [9]. The prediction performance is limited to a very short duration ahead.

In this paper, we focus on the predictability analysis [10, 11] on time-series data of two major and widely used wireless services: video service; and audio service using 802.11 WLAN Traffic, and show their prediction performance of AR based predictor with data categorization. We first investigate the distribution model of continuous busy/idle duration of video and audio service using 802.11 WLAN traffic. The similar work has been researched in [12] but only considered VoIP skype service and FTP service. In addition, we will consider two services on two different scenarios: the noninterference case, and interference case to find the impact from interference. In addition, by introducing the categorization of the durations of recent B/I statuses with their ranges, the duration of the next status will be distributed into different sets (streams) with different ranges. Based on this property, we evaluate the prediction performance of the low-complexity AR-based predictor with the categorization of the duration of recent busy/idle status. From their predictability and prediction performance using the low-complexity AR-based predictor, we

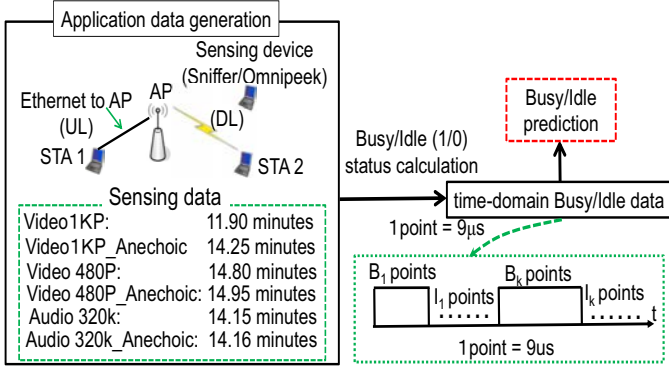


Fig. 1. Data capture for busy/idle prediction.

can confirm that data categorization can largely improve the prediction performance of some streams.

II. DATA COLLECTION IN REAL ENVIRONMENT AND THE MODEL OF BUSY / IDLE DURATION DISTRIBUTION

A. The experiments and data capture

We conducted several wireless traffic experiments to obtain the data of spectrum usage as shown in Fig. 1. The experiments of video services with two kinds resolution as 1KP and 480P, and audio service with average transmit speed as 320 kbps using WLAN system over 5 GHz band have been operated in an indoor environment (interference case). As a comparison, we also got the data of spectrum usage in an anechoic chamber (noninterference case) for showing the impact of interference. The data process is shown in Fig. 1. Frame header data was recorded using commercial sniffing software on Channel 36 in W52 at the 5 GHz band. The frame arrival time, data-rate, and length were extracted and the number of data bits per symbol, bandwidth and standard (IEEE 802.11b/g/n) were obtained from the data-rate information based on the IEEE 802.11-2016 standards [13] using orthogonal frequency division multiplexing (OFDM). The frame duration was estimated from the required number of OFDM symbols after adding the MAC header with PHY preamble. The busy/idle (B/I) sequence was then generated using a granularity of $9\mu s$ per slot following the current WLAN standards.

B. The model of Busy / idle duration distribution

The corresponding empirical CDF and probability density function (PDF) are calculated for fitting the probability distribution models. The major simple distribution models used in this paper are listed in Table V. The distribution models considered in this paper are exponential (EX), generalized Pareto (GP), Log-normal (LN), log-logistic (LLG), generalized extreme value (GEV), gamma (GM) and Inverse Gaussian (Inv-Gaussian) distributions.

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. [14], the authors have considered methods of moment (MOM) inference scheme for estimating

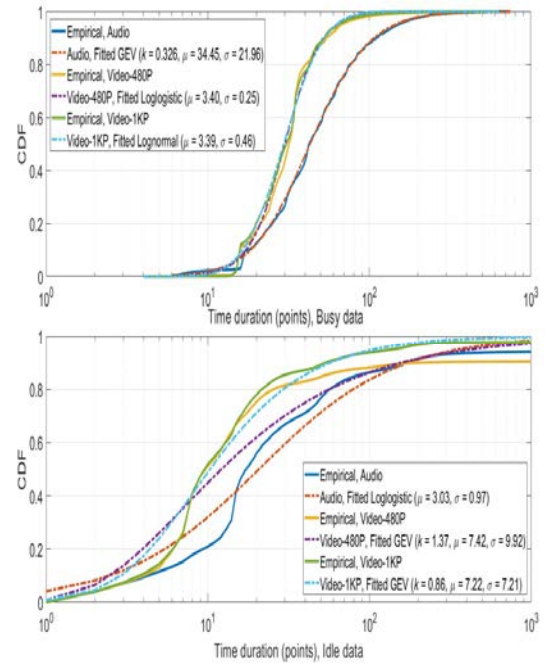


Fig. 2. The model fitting for busy/idle duration (noninterference case).

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 select one distribution model with minimum Kolmogorov-Smirnov (KS) distance between the empirical data and model, which is similar to that in [14].

Figure 2 and 3 show the model fitting results of busy/idle duration for the noninterference case and interference case, respectively. For the noninterference case in Fig. 2, busy duration data can be well-fitted using some simple models than that of the idle duration data, especially for the data of audio service. Figure 2 also shows that the data duration from video service has more concentrated pattern than that from audio service. The reason is that the audio service includes packets with different length. On the other hand, the idle durations of both services appear more randomness which makes the model fitting be low accuracy. As shown in Fig. 3, when the interference occurs, the fitting error is increased for the busy durations of both service. Idle data from interference can be fitted with higher accuracy than that of the noninterference case. The results imply that all services from noninterference case has the fixed traffic pattern which make CDF of the idle durations have more randomness than that of the interference case. This make the model fitting more difficulty.

III. PREDICTABILITY ANALYSIS FOR TIME-SERIES DATA

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

TABLE I
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	$f_{GP}(x; \mu, \sigma, \xi) = \begin{cases} \left(\frac{\xi(x-\mu)}{\sigma} + 1\right)^{-\frac{\xi+1}{\xi}} & \xi \neq 0, \\ e^{-\frac{(x-\mu)}{\sigma}} & \xi = 0. \end{cases}$	$F_{GP}(x; \mu, \sigma, \xi) = \begin{cases} 1 - \left(\frac{\sigma + \xi(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}} & \xi \neq 0, \\ 1 - e^{-\frac{(x-\mu)}{\sigma}} & \xi = 0. \end{cases}$
Log-normal (LN)	$f_{LN}(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$	$F_{LN}(x; \mu, \sigma) = \frac{1}{2} \operatorname{erfc}\left(-\frac{\ln x - \mu}{\sigma\sqrt{2}}\right)$
LogLogistic (LLG)	$f_{LLG}(x; \mu, \sigma) = \frac{\left(\frac{\mu}{x}\right)^{\sigma} \left(\frac{x}{\mu}\right)^{(\sigma-1)}}{\left(1 + \left(\frac{x}{\mu}\right)^{\sigma}\right)^2}$	$F_{LLG}(x; \mu, \sigma) = \frac{\left(\frac{x}{\mu}\right)^{\sigma}}{1 + \left(\frac{x}{\mu}\right)^{\sigma}}$
Generalized Extreme Value (GEV)	$s = \frac{x-\mu}{\sigma}$ $f_{GEV}(x; \mu, \sigma, \xi) = \begin{cases} (1 + \xi s)^{-\frac{1}{\xi}} \exp(-(1 + \xi s)^{-\frac{1}{\xi}}) & \xi \neq 0, \\ \exp(-s) \exp(-\exp(-s)) & \xi = 0. \end{cases}$	$F_{GEV}(x; \mu, \sigma, \xi) = \begin{cases} \exp(-(1 + \xi s)^{-\frac{1}{\xi}}) & \xi \neq 0, \\ \exp(-\exp(-s)) & \xi = 0. \end{cases}$
Gamma (GM)	$f_{GM}(x; \alpha, \beta) = \frac{1}{\beta^{\alpha} \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$	$F_{GM}(x; \alpha, \beta) = \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}$
Inv-Gaussian (IG)	$f_{IG}(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right]; \Phi(x): \text{Normal distr.}$	$F_{IG}(x; \mu, \lambda) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + e^{\left(\frac{2\lambda}{\mu}\right)} \Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right)$

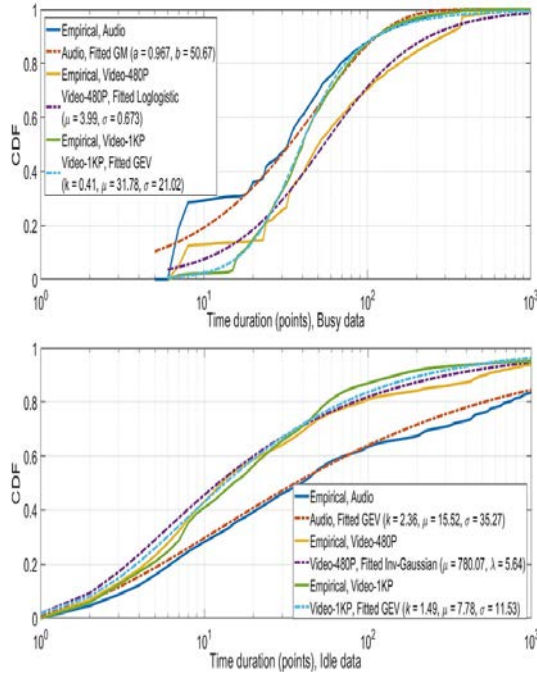


Fig. 3. The model fitting for busy/idle duration (interference case).

quantify the degree of predictability for the real-world time-series data [11, 12]. Here we give a simple explanation for this predictability analysis which can decide the fundamental limits as the performance bounds of predictability.

Let us suppose there is a random variate X with M kind of values. Therefore, its entropy $S(X)$ can be represented as

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

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

be represented as

$$S(\mathbf{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} S(x_1, x_2, \dots, x_n). \quad (2)$$

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

$$S(\mathbf{X}') = \lim_{n \rightarrow \infty} S(x_n | x_{n-1}, x_{n-2}, \dots, x_1). \quad (3)$$

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

$$\begin{aligned} S(\mathbf{X}) &= \lim_{n \rightarrow \infty} \frac{1}{n} S(x_1, x_2, \dots, x_n) \\ &= \lim_{n \rightarrow \infty} S(x_n | x_{n-1}, x_{n-2}, \dots, x_1) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n S(x_i | h_{i-1}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n S(i). \end{aligned} \quad (4)$$

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

The n -th length average entropy $S(\mathbf{X})$ is difficult to be obtained because it is decided by the value of $[x_1, x_2, \dots, x_n]$ 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 [14]. On the other hand, when the correlation property between the time-series data is not considered, the n -th length average entropy can be simply calculated as $S^{Unc}(\mathbf{X})$ only using the PDF of time-series data as

$$S^{Unc}(\mathbf{X}) = - \sum_{i=1}^{\infty} x(i) \log(f(x_i)). \quad (5)$$

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

$$P_r = \text{Prob}\{\hat{x}_n = x_n | h_{n-1}\}. \quad (6)$$

Therefore, the predictability of one time-series data is the maximum value of all prediction methods as

$$\Pi(h_{n-1}) = \sup\{\hat{x}_n = x_n | h_{n-1}\}. \quad (7)$$

TABLE II
THE ENTROPY AND PREDICTABILITY PROBABILITY (NONINTERFERENCE CASE).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (Video 1KP)	4.44	5.61	0.58	0.44	393
Idle (Video 1KP)	4.34	5.34	0.69	0.61	2301
Busy (Video 480P)	4.36	5.46	0.57	0.44	295
Idle (Video 480P)	4.94	6.19	0.67	0.57	4766
Busy (Audio 320K)	5.20	6.72	0.51	0.33	404
Idle (Audio 320K)	5.37	6.77	0.61	0.49	2361

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

$$\Pi = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Pi(i), \quad (8)$$

with $\Pi(i) \triangleq f(h_{n-1})\Pi(h_{n-1})$.

Finally, we can build the relationship between the predictability probability and the entropy of time-series data using Fano inequality as

$$\begin{aligned} S(x_n|h_{n-1}) &\leq -[p \log_2 p + (1-p) \log_2(1-p)] \\ &\quad + (1-p) \log_2(M-1) \\ &\triangleq S_F(p) = S_F(\Pi(h_{n-1})), \end{aligned} \quad (9)$$

where p is $\Pi(h_{n-1})$, and M is the number of different value of x . Using this relationship, the low bound and upper bound of predictability probability for time-series data can be represented as

$$\begin{aligned} S^{Real} &\leq S_F(\Pi^{Real}) \\ &= -[\Pi^{Real} \log_2 \Pi^{Real} + (1 - \Pi^{Real}) \log_2(1 - \Pi^{Real})] \\ &\quad + (1 - \Pi^{Real}) \log_2(M-1), \end{aligned} \quad (10)$$

$$\begin{aligned} S^{Unc} &\leq S_F(\Pi^{Unc}) \\ &= -[\Pi^{Unc} \log_2 \Pi^{Unc} + (1 - \Pi^{Unc}) \log_2(1 - \Pi^{Unc})] \\ &\quad + (1 - \Pi^{Unc}) \log_2(M-1). \end{aligned} \quad (11)$$

Using Eqs.(10) and (11), it is easy to find the low bound Π^{Unc} and up bound Π^{Real} of predictability probability when their entropy S^{Unc} and S^{Real} can be calculated. 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 algorithm [1] to find the relation between the busy and idle durations and then calculate the value of S^{Real} . 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. (5) to calculate the value.

Table 2 and Table 3 show the values of S^{Real} , S^{Unc} , Π^{Real} , Π^{Unc} and M of busy and idle durations for the noninterference

TABLE III
THE ENTROPY AND PREDICTABILITY PROBABILITY (INTERFERENCE CASE).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (Video 1KP)	4.69	6.58	0.59	0.38	561
Idle (Video 1KP)	4.79	6.07	0.66	0.56	4284
Busy (Video 480P)	3.88	6.97	0.68	0.35	635
Idle (Video 480P)	5.65	7.46	0.60	0.45	3582
Busy (Audio 320K)	3.35	6.10	0.72	0.42	469
Idle (Audio 320K)	6.52	8.81	0.56	0.38	6670

case and interference case, respectively. The values show that the predictability probability for noninterference case has larger value than that of interference case except the busy data of audio service. In addition, the predictability for methods only using PDF information can reach about 30%–50%. When using the relation information among the data, such value can be improved to 50%–70%. On the other hand, it also shows that the centralized type of PDF has larger S^{Unc} than that of flat type from both figures and Eq. (5).

IV. AR BASED PREDICTOR USING DURATION CATEGORIZATION

A. Auto-regressive model

Autoregressive (AR) model is a representation of a type of random process [15]. The autoregressive model specifies that the output variable depends linearly on its own previous values and on a stochastic term; thus the model is in the form of a stochastic difference equation. We use X_i to represent the busy or idle points ($9 \mu s/\text{point}$) at i -th continuous busy or idle duration. The AR predictor with order p can be represented as

$$X_{i+1} = a_1 X_i + a_2 X_{i-1} + \dots + a_p X_{i-p+1}. \quad (12)$$

Here a_1 to a_p are calculated using training data so that they give the solution as the least squares for linear regression. In this paper, we assume that the AR predictor uses the previous 100 busy/idle duration values to calculate a_1 to a_p once for predicting the next 200 busy/idle duration values.

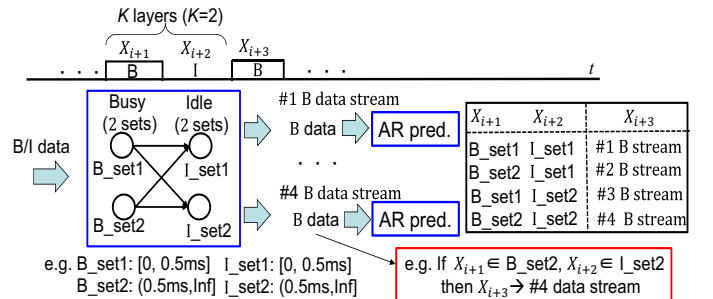


Fig. 4. AR based predictor with B/I duration categorization.

The time-series of traffic data usually has some correlation properties that can be used for improving the prediction performance. In addition, for each service such as audio, video or FTP etc., the data packet size has some fixed pattern or distribution. Therefore, for some time-series data, the busy and

TABLE IV
THE ENTROPY AND PREDICTABILITY PROBABILITY (VIDEO 480P,
INTERFERENCE CASE).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (All data)	3.88	6.97	0.68	0.35	635
Busy (Set (1, 1))	5.37	6.85	0.52	0.35	540
Busy (Set (1, 2))	1.85	4.00	0.86	0.66	507
Busy (Set (2, 1))	6.38	7.86	0.41	0.23	584
Busy (Set (2, 2))	6.27	7.92	0.43	0.23	605
Idle (All data)	5.65	7.46	0.60	0.45	3582
Idle (Set (1, 1))	4.97	6.20	0.64	0.54	2700
Idle (Set (1, 2))	5.45	6.61	0.60	0.50	2304
Idle (Set (2, 1))	6.80	9.44	0.50	0.26	3076
Idle (Set (2, 2))	5.30	6.50	0.58	0.46	1173

idle duration exhibits some correlation with high probability. Therefore, using the relationship of busy/idle duration or its status transition probability, the prediction performance of AR based predictor would be improved. On the other hand, from Eq. (5), if the PDF of time-series data is a centralized type, the entropy S^{Unc} will be a small value which makes the predictability probability Π^{Unc} be larger.

Based on this idea, we proposed a new AR predictor with B/I duration categorization as shown in Fig. 4 for an example. Here the busy duration data X_{i+3} will be categorized according to the set categorization of previous busy duration X_{i+1} and idle duration X_{i+2} . Since busy duration X_{i+1} and idle duration X_{i+2} have 2 sets, therefore the next busy duration data X_{i+3} will be separated into four busy data streams. For example, if $X_{i+1} \in B_set2$ and $X_{i+2} \in I_set2$, then the idle data X_{i+3} will be put into the 4-th stream. After that, the AR based prediction is operated in each stream.

The idea can be extended for K layers. The captured busy and idle duration data is firstly processed by a B/I duration categorization with K layers and each layer has S_i ($i = 1, \dots, K$) sets. Therefore, the busy or idle duration can be allocated into S ($S = \prod_{i=1}^K S_i$) different streams. Then, for each busy or idle stream, the AR based algorithm is utilized for the next duration prediction.

To evaluate the prediction performance of the proposed method, we use B/I data of Video-480P and Audio-320K at interference case. The B/I duration categorization uses previous busy X_{i+1} and idle X_{i+2} for categorizing the busy duration streams with $K=2$. The set range of each layer is shown in Fig. 4 with $[0, 0.5ms]$ and $(0.5ms, inf)$ for both idle duration and busy duration. To show the effectiveness of data categorization, we also show the values of S^{Real} , S^{Unc} , Π^{Real} , Π^{Unc} and M of busy or idle durations in each stream.

Table IV gives the values of S^{Real} , S^{Unc} , Π^{Real} , Π^{Unc} and M of the busy and idle durations of each stream after data categorization for the video data. It shows that data categorization makes each stream have different predictability probability. For the busy data, the second stream has larger Π^{Unc} than that of no data categorization which shows this time-series busy data is easier to be predicted than others. The similar results can also be found for idle data with data

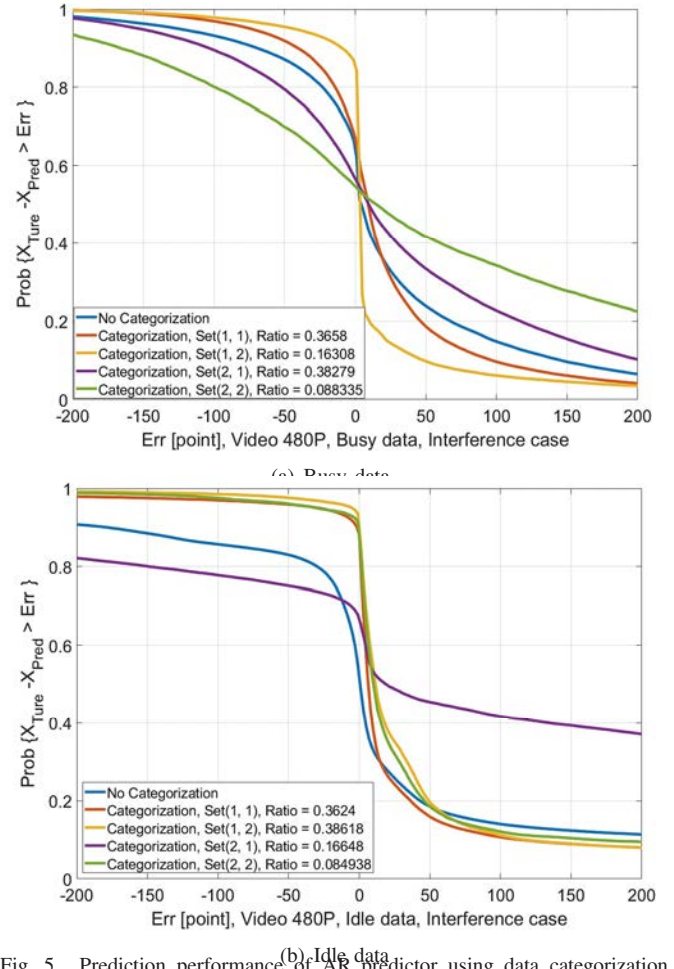


Fig. 5. Prediction performance of AR predictor using data categorization (Video 480P at interference case).

categorization. Fig. 5 shows the complementary cumulative distribution function (CCDF) distribution of prediction error of the busy and idle data in each stream. From Fig. 5, the second stream which occupied about 16.3% busy data shows the better prediction accuracy than that of all busy data without data categorization.

Table V and Fig. 6 shows the similar results for the busy and idle durations of each stream after data categorization for the audio data. The second stream of busy data is easier to be predicted than other streams and all busy data without data categorization. Therefore, as shown in both figures and tables, the proposed data categorization method can make partial the time-series data easier to be predicted even if system employs one low-complexity AR predictor. In addition, it also shows that the value Π^{Unc} which only related to the PDF of data seems to be a good indicator for the predictability of time-series data.

V. CONCLUSION

This paper has investigated the predictability of busy/idle duration of video and audio services using an auto-regressive (AR) based predictor. The modeling of their busy/idle duration

TABLE V
THE ENTROPY AND PREDICTABILITY PROBABILITY (AUDIO 320K,
INTERFERENCE CASE).

Data type	S^{Real}	S^{Unc}	Π^{Real}	Π^{Unc}	M
Busy (All data)	3.35	6.10	0.72	0.42	469
Busy (Set (1, 1))	5.10	6.44	0.52	0.35	359
Busy (Set (1, 2))	1.99	3.76	0.84	0.67	360
Busy (Set (2, 1))	5.91	7.31	0.43	0.24	392
Busy (Set (2, 2))	5.61	7.12	0.47	0.28	411
Idle (All data)	6.52	8.81	0.56	0.38	6670
Idle (Set (1, 1))	5.81	7.46	0.59	0.45	3264
Idle (Set (1, 2))	5.65	6.96	0.56	0.43	1501
Idle (Set (2, 1))	7.64	10.31	0.45	0.20	4086
Idle (Set (2, 2))	5.62	6.95	0.53	0.39	893

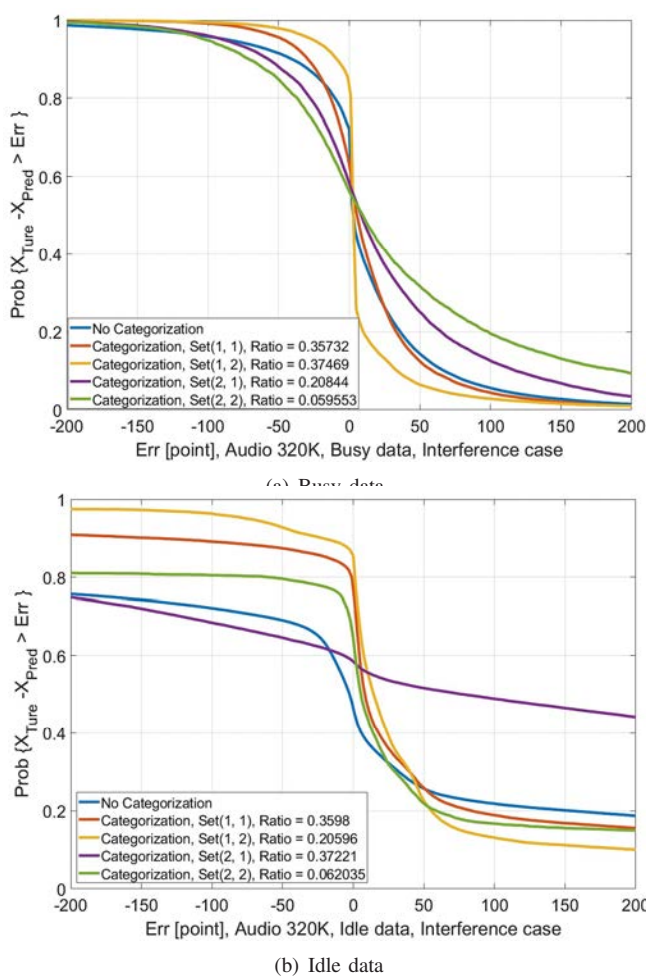


Fig. 6. Prediction performance of AR predictor using data categorization (Audio 320K at interference case).

and their predictability performance based on predictability theory have been studied. Then, the data categorization method has been proposed to improve the predictability of the partial busy/idle durations. The proposed method categorizes the durations of recent B/I statuses with their ranges to make the duration of the next status be distributed into different sets or streams with different ranges. From their predictability and prediction performance using the low-complexity AR-based

predictor, we can confirm that data categorization can largely improve the prediction performance of some streams.

ACKNOWLEDGMENT

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

REFERENCES

- [1] A. Kamilaris and A. Pitsillides, "Mobile phone computing and the Internet of Things: a survey," *IEEE Internet Things Journal*, vol. 3, no. 6, pp. 885–898, Dec. 2016.
- [2] A. Ijaz et al., "Enabling massive IoT in 5G and beyond systems: PHY radio frame design considerations," *IEEE Access*, vol. 4, pp. 3322–3339, 2016.
- [3] R. Struzak, "Cognitive radio, spectrum, and evolutionary heuristics," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 166–171, June 2018.
- [4] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, April 2017.
- [5] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 4, pp. 648–664, Dec. 2018.
- [6] N. Egashira, K. Yano, S. Tsukamoto, J. Webber, M. Sutoh, Y. Amezuwa, and T. Kumagai, "Integrated synchronization scheme for WLAN systems employing multiband simultaneous transmission," *Proc. IEEE Wireless Commun. and Netw. Conf.*, San Francisco, CA, pp. 1–5, March 2017.
- [7] 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," *Proc. 1st Int. Conf. on Artif. Intell. in Info. and Commun.*, Okinawa, pp. 22–27, Feb. 2019.
- [8] R. Struzak, "Cognitive radio, spectrum, and evolutionary heuristics," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 166–171, June 2018.
- [9] J. Webber, A. Mehdodniya, Y. Hou, K. Yano and T. Kumagai, "Study on idle slot availability prediction for WLAN using a probabilistic neural network," *Proc. 23rd Asia-Pacific Conf. on Commun.*, Perth, WA, pp. 1–6, Dec. 2017.
- [10] G. Ding et al., "On the limits of predictability in real-world radio spectrum state dynamics: from entropy theory to 5G spectrum sharing," *IEEE Communications Magazine*, vol. 53, no. 7, pp. 178–183, July 2015.
- [11] 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.
- [12] A. Gupta, S. Agarwal and S. De, "A new spectrum occupancy model for 802.11 WLAN traffic," *IEEE Communications Letters*, vol. 20, no. 12, pp. 2550–2553, Dec. 2016.
- [13] IEEE, "802.11-2016 - IEEE Standard for Information Technology - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE, 2016.
- [14] L. B. Miguel and F. Casadevall, "Time-domain model of spectrum usage for the analysis, design, and simulation of cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 5, pp. 2091–2104, June 2013.
- [15] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications: with R Examples*. Springer, 2010.

Yafei Hou received his Ph.D. degrees from Fudan University, China and Kochi University of Technology (KUT), Japan in 2007. He is Assistant Professor at the Graduate School of Natural Science and Technology, Okayama University, Japan. He is also a guest research scientist at Wave Engineering Laboratories, ATR Institute International, Japan from October 2016.

Shun Kawasaki received his B.S. degree from Okayama University, Japan in 2019. Now he is working for his MS degree.

Satoshi Denno received the M.E. and Ph.D degrees from Kyoto University, Kyoto, Japan in 1988 and 2000, respectively. He is a full professor at graduate school of natural science and technology, Okayama University, Japan.