Channel-Adaptive Transmission Protocols for Smart Grid IoT Communication

Sharda Tripathi and Swades De, *Senior Member, IEEE*

***Abstract*—This article presents a new paradigm for channel dynamics adaptive transmission of intermittent data in smart grid IoT communication networks, wherein novel channel prediction frameworks using stochastic modeling as well as data-driven learning of channel variability are proposed. A probing-based transmission is also proposed as a benchmark. These prediction frameworks are complemented with an adaptive channel coding scheme to increase the transmission reliability of time-critical grid monitoring data over a wireless channel. Through analyzing the prediction and packet loss performance at varying SNR and fading conditions, it is noted that the stochastic modeling framework is efficient when the fading correlation in the channel is high while the learning-based approach is more adaptive to channel dynamics as the correlation reduces. The proposed frameworks are easily implementable on low-cost end nodes, owing to the optimal selection of parameters for low runtime complexity. When compared to probing-based data transmission for a given fading in the channel, the packet loss probability of the learning-based transmission closely matches while with stochastic model loss probability is found to be 12.3% higher. However, their respective signaling overheads are 38% and 98% lower with respect to the probing-based approach, which is a significant gain at the cost of marginally additional computation complexity.**

***Index Terms*—Adaptive coding, Gaussian process regression, IoT data communication protocols, resource efficiency, smart grid communication, wireless channel prediction.**

I. Introduction

INFORMATION and communication technology has a pivotal role to play in efficiency and reliability enhancement of the IoT networks. A pertinent emerging application is of wide-area situational awareness in the smart grid. It is supported by a pervasive monitoring system comprising of advanced sensing equipment such as phasor measurement units (PMUs), which facilitate real-time collection and exchange of synchrophasor data over the communication network for gridwise protection and control [1]. In the existing literature, several wired (power line and fiber optic) and wireless (3G cellular, IP-based, ZigBee, Wi-Fi, *Z*-wave, WiMAX, 3GPP LTE, LoRa, and NB-IoT) communication networking technologies are suggested for smart grid data communication [2], [3]. However, considering the development of dynamic network architecture and intelligent algorithms that cater to the data flow requirements of advanced smart grid features, particularly, self-healing, consumer friendliness, optimal resource usage, and resilience to cyber attacks, it is likely that future smart grid communication will evolve as a hybrid of contemporary protocols and technologies.

A. Related Works and Motivation

In recent years, power utility vendors have been densely deploying PMUs as IoT devices which generate data streams of multiple parameters for complete observability of the grid. It has been noted that even in the current deployment regime, the annual data volume generated by PMUs is in peta-Bytes [4].

Conventionally, the transmission of PMU data for real-time wide-area monitoring and control is periodic, with the streaming rate of each node on the order of a few tens of kb/s. However, as an emerging IoT application, data communication protocols in wide-area monitoring are continuously evolving to meet the grid-level QoS requirements in a more dynamic environment. Also, as suggested in [5], higher rate synchrophasors are necessary for precise power grid system monitoring and control. For handling massive data generated by a multitude of such IoT devices, intelligent data pruning without sacrificing on the information content, at the device level, network level, as well as in the cloud, for optimal utilization of resources (e.g., transmission energy, communication bandwidth, and cloud storage space) is of contemporary research interest [6]–[8]. A study in [9] has proposed a paradigm for 5G intelligent Internet of Things to process big data intelligently and optimize the usage of communication channels. Specifically, for communication of PMU data in a smart grid IoT network, a few recent works have studied intelligent data pruning at the PMU level [10], [11] as well as at the PDC level [12]–[14]. Similar observations have been made in [15] and [16] on the data generated from smart electric meters, which indicate that even low rate telemetry data generated from a large number of such IoT devices would lead to a massive requirement of communication bandwidth for transmission. Although smart metering is a less time-critical application, concern on the aggregated

Manuscript received November 3, 2019; revised March 1, 2020, March 31, 2020, and April 27, 2020; accepted April 29, 2020. This work was supported in part by the Department of Science and Technology, International Bilateral Cooperation Division under Grant INT/UK/P-153/2017, and in part by the Science and Engineering Research Board, DST, under Grant CRG/2019/002293. *(Corresponding author: Swades De.)*

Sharda Tripathi is with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Turin, Italy (e-mail: sharda2309@gmail.com).

Swades De is with the Department of Electrical Engineering and the Bharti School of Telecommunication, Indian Institute of Technology Delhi, New Delhi 110016, India (e-mail: swadesd@ee.iitd.ac.in).

Digital Object Identifier 10.1109/JIOT.2020.2992124

data transmission bandwidth requirement calls for a similar measure of data pruning at the source.

It is notable from the works in [10]–[14] that the device-level intelligent processing before transmission benefits in not only reduced data processing at the PDC level but also optimizing the device-level and channel resources. Intelligent processing at the source also relaxes the necessity of streaming-based reporting the PMU data to the PDC. To this end, Das and Sidhu [10] proposed the sparse data reporting rate by a compressive sampling of the available PMU data. Likewise, in [11], a learning-based online pruning is employed on the sampled data at the PMU before transmission. In this approach, data transmission from a *smart PMU* is initiated on "event-driven" basis, only when there is a need to retrain the support vector regression model at the PDC for interpolating the missing (not transmitted) samples and this is largely governed by the underlying power-grid dynamics. As a consequence of intelligent processing before transmission, although the data generation process at the PMU is a regular (periodic) event, a sparse/sporadic PMU data transmission behavior is observed.

Intuitively, in such a scenario with *smart IoT devices* with sparse/sporadic data transmission requirements, conventional link layer strategies would be inefficient because it is difficult to characterize the data arrival process. Besides, the response of the wireless channel to the dynamic data flow may not be a known distribution. Thus, for a sporadic IoT data communication, novel techniques are required to be devised that are application aware and yet independent of the channel dynamics. The proposed channel-aware data transmission schemes in this article are targeted to address reliable data transmission while meeting the time-criticality constraints in such cases.

To enhance the reliability of transmission as well as the efficiency of radio resource allocation, knowledge of the wireless channel state (CS) at the packet transmission instant is of paramount interest in the sporadic IoT data communication. Though early works do not clearly address sporadic data communication, approaches for prediction of CSs for continuous data availability were widely investigated. These primarily include autoregressive (AR) model-based linear prediction algorithms [17] and channel prediction in short-term fading [18]. The AR model was also used in [19] and [20] for channel prediction in vehicular *ad hoc* networks and millimeter-wave MIMO OFDM systems. Though AR model-based prediction has better performance, it assumes channel variations to be wide-sense stationery, which may not be true in reality.

To track the variability of channel coefficients, adaptive Kalman filter [21] and spatiotemporal AR models [22] were suggested for enhanced prediction quality. In [23], reduced-rank channel prediction was proposed for limited feedback time-variant channels. Although prediction accuracy in these works is better in comparison to linear predictors, they require the knowledge of channel statistics as additional overhead. To address this, a polynomial fitting-based predictor using channel measurements was considered in [24] for future channel predictions, however, it has limited prediction range especially in MIMO channels. Consequently, the sum-of-sinusoids method was proposed in [25] and [26] to reliably predict channel over a long prediction range. For the fast-varying nonstationary channel, the first-order Taylor expansion-based model [27] was shown to predict reliably with marginally increased complexity in comparison to traditional channel prediction approaches.

Another more efficient category of channel prediction methods is based on nonlinear transformations. These include the use of discrete cosine transforms [28], compressed sensing [26], neural networks [29], [30], and deep learning [31]–[33]. Compared to neural networks, learning-based prediction methods exhibit faster learning speed and higher convergence precision. The approach in [34] proposed to train the echo state network for short-term prediction of the Ricean-fading scenarios, and it was shown to perform better in comparison to AR and signal processing approaches. It may be noted that transform-based approaches require complex matrix operations which are high in computational complexity. The deep learning networks proposed so far also have complex structures with multiple layers and require extensive training data in terms of volume, granularity, and feature set. Besides, unlike continuous data availability, channel prediction for sporadic smart grid data communication is challenging owing to uncertainty about the underlying temporal process during blind intervals due to the sparsity of channel gain observations.

As the 5G-IoT communication systems are evolving, smart PMUs (smart IoT devices)[1] will support intelligent processing in addition to sensing and transmission of data in a dynamic environment. In view of limited applicability of the existing methodologies to smart IoT networks, in this article, channel-adaptive transmission based on simple yet efficient channel prediction frameworks using stochastic modeling, data-driven learning, and probing of wireless channel are proposed for reliable transmission of sporadic but time-critical PMU data.

B. Main Contributions

In this article, novel protocols are proposed for channel-adaptive transmission of sporadic but time-critical smart IoT data using stochastic modeling, learning, and probing-based estimate of CS. The main contributions of this article are as follows.

1) A novel stochastic modeling framework based on the characterization of rate of change of wireless fading channel is proposed to estimate the CS for transmission of sporadically available smart PMU data.
2) A novel data-driven framework based on the Gaussian process regression is also proposed to dynamically learn the CS in desired transmission slots using channel gains from previous packet transmission slots.
3) An adaptive transmission scheme for time-critical PMU data is introduced based on the proposed channel prediction frameworks and using adaptive channel coding.

[1] Hereafter, "smart PMU" and "PMU" are used interchangeably throughout this article.

4) A probing-based transmission is also proposed which is considered as the benchmark for comparing the stochastic model-based and learning-based approaches. The results demonstrate that the stochastic modeling framework is efficient when the fading correlation in the channel is high, while the learning-based approach is more adaptive to channel dynamics as the correlation reduces. Furthermore, for a given channel fading condition, the packet loss probability of the learning-based transmission closely matches with the benchmark scheme, while with the stochastic model-based prediction, the loss probability is found to be 12.3% higher. However, the respective signaling overheads are 38% and 98% lower with respect to the benchmark.

Unlike the state of the art, the proposed channel-adaptive sporadic data transmission schemes are independent of channel stationarity and do not require the knowledge of fading distribution. Temporal channel variability is exploited in the proposed schemes to predict CS for increasing the throughput with optimal resource usage. Furthermore, training length and the number of CSs can be optimally chosen for reduced runtime complexity, thereby enabling implementability in low-cost IoT end nodes such as PMUs. To the best of our knowledge, such a comprehensive framework keen on reliable data transmission in the emerging smart IoT communication context has not been studied so far.

C. Paper Organization

The remainder of this article is organized as follows. In Section II, the system model and protocol description are presented. Stochastic, learning, and probing-based framework for channel prediction are proposed in Section III, followed by an elaboration of the proposed channel-aware data transmission scheme in Section IV. The numerical results based on large-scale simulations are discussed in Section V. Finally, this article is concluded in Section VI.

II. System Model and Protocol Description

A. System Model

A wide area smart grid communication network deals with time-critical health parameters that are monitored and hierarchically transmitted from the PMUs to the remotely located phasor data concentrator (called super PDC) via local PDC for fast decision/actuation capability. The access network architecture of a wide-area situational awareness system shown in Fig. 1 resembles a typical IoT network wherein the PMUs are the end nodes. Motivated by the studies in [10] and [11], the PMUs are considered to have some intelligence whereby the redundant sampled data are pruned before transmitting over the communication channel. Thus, instead of periodic/streaming-based transmission (as in conventional PMUs), the smart PMUs communicate sporadically to the local PDC in an event-driven mode, thereby reducing the communication bandwidth as well as node-level energy requirements. PMU to the data network connectivity is considered to be over a wireless link. Appropriate multiaccess protocols, such

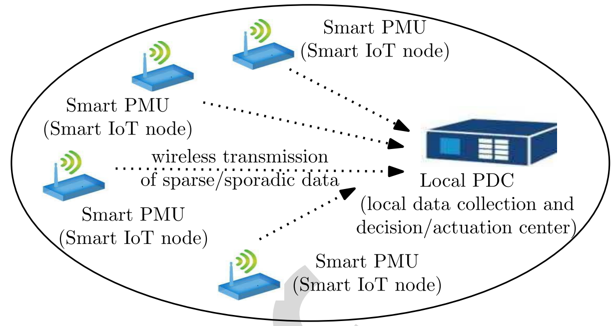


Fig. 1. Wireless IoT network with smart PMUs (smart IoT nodes) for smart grid monitoring and control.

as beaconing or polling-based approach, are considered for many such PMUs to communicate to the local PDC.

In this article, a point-to-point communication scenario is considered where sporadically available data from a PMU is transmitted over the fading wireless channel. Specifically, efficient channel prediction strategies that complement the channel encoding and physical transmission are sought for sporadic data transmission over dynamic communication. The temporal channel variations are characterized by the product $f_D T_s$ [35], where f_D is the Doppler frequency corresponding to the relative velocity of the receiver and T_s is the symbol duration. For a slow-fading scenario, the process is very correlated ($f_D T_s < 0.1$), while fast fading corresponds to consecutive channel samples being almost independent ($f_D T_s > 0.2$). It is assumed that the communication process is slotted. The slot size is of one symbol duration, over which the channel is considered to remain invariant. However, a PMU packet transmission may comprise of several symbols during which the channel gain may vary. Also the transmitter-to-receiver propagation delay is assumed to be negligibly small.

B. Protocol Description

The CS information (CSI) is primarily required at the instant when a PMU packet is available for transmission. With intermittently available PMU data, the interval between one batch of packets to the next is random. Therefore, the channel knowledge from the last transmitted batch is not applicable for the next batch of transmission. To this end, three approaches are proposed. The first two approaches estimate the CS, respectively, using stochastic and learning-based models, whereas in the third approach CSI is collected by probing the channel immediately before the transmission. Due to the time-critical nature of the PMU data, a CSI-aware forward error correction (FEC) mechanism using Reed–Soloman (RS) codes is introduced for appropriately protecting the PMU data from prospective errors. Note that retransmission of unsuccessful packets is considered impractical, as smart grid monitoring data have strict latency constraints. The proposed adaptive coding assigns redundant symbols to the packet in accordance with the current CS such that any packet transmitted during a given estimated channel condition is provided with sufficient redundancy for its successful delivery at the receiver.

Furthermore, based on the received signal quality, the receiver sends ACK/NAK along with useful channel information.

III. Proposed Channel State Prediction Analysis

Considering a slotted time interval scenario, where due to sporadic PMU data for transmission, the interval n between a batch of packets to the next is a discrete random variable. Owing to time criticality of PMU data, at the beginning of a new batch of transmission, i.e., at the end of the nth slot, the CS needs to be accurately estimated so that appropriate FEC overhead can be incorporated for the successful delivery of the packets. In this section, three proposed frameworks for the estimation of CS at the transmitter using stochastic, learning, and probing-based models are analyzed for channel-adaptive data transmission and their respective computation complexities are discussed.

A. Stochastic Modeling Framework

Let the received complex signal over the wireless communication channel at time t be $R(t) = R_I(t) + jR_Q(t)$, where $R_I(t)$ and $R_Q(t)$, respectively, represent the in-phase and quadrature components, and $j = \sqrt{-1}$. Then, the received signal envelope $Z(t) \triangleq |R(t)| = \sqrt{{R_I}^2 + {R_Q}^2}$. Following the observation in [36], the rate of change of signal envelope with respect to time $\dot{Z}(t) \triangleq (d/dt)Z(t) = \lim_{\triangle t \to 0}[(Z(t+\triangle t) - Z(t))/\triangle t]$ is a zero-mean Gaussian random variable irrespective of the underlying distribution of fading channel, i.e., $\dot{Z}(t) \sim \mathcal{N}(0, \dot{\sigma})$. This property of $\dot{Z}(t)$ is used here to estimate the probability distribution of n-slot ahead CS, $Z(t + nT_s)$, given that the current CS $Z(t)$ is known. Here, T_s is the slot duration.

Using the Taylor series expansion, signal envelope $Z(t+T_s)$ in the next slot is given by

$$Z(t + T_s) = Z(t) + \dot{Z}(t) \cdot T_s + \ddot{Z}(t) \cdot \frac{{T_s}^2}{2!} + \cdots. \tag{1}$$

Since $T_s \ll 1$, applying the first-order approximation to (1), we have

$$Z(t + T_s) \approx Z(t) + \dot{Z}(t) \cdot T_s. \tag{2}$$

Denoting $Z(t)$ as the signal envelope in slot 0, i.e., $Z(t) \equiv Z(0)$, we have $Z(t + nT_s) \equiv Z(n)$. Accordingly, $\dot{Z}(t) \cdot T_s$ being the temporal variation of $Z(t)$ in the next time slot, it is denoted as $\delta Z(1)$. Following [36], $\delta Z(1) \sim \mathcal{N}(0, \dot{\sigma_1})$, where $\dot{\sigma_1} = T_s \cdot \dot{\sigma}$. However, $Z(1) = Z(0) + \delta Z(1)$ being the signal envelope in slot 1, $Z(1) \geq 0$. Therefore, $\delta Z(1) \in [-Z(0), \infty)$. In other words, the distribution of $\delta Z(1)$ is truncated Gaussian [18], which is obtained as

$$f_{\delta Z(1)}(\alpha) = \begin{cases} \frac{1}{1-\Phi_1\left(-\frac{Z(0)}{\dot{\sigma_1}}\right)} \frac{1}{\sqrt{2\pi}\dot{\sigma_1}} e^{\left(\frac{-\alpha^2}{2\dot{\sigma_1}^2}\right)}, & \text{if } -Z(0) \leq \alpha \\ 0, & \text{elsewhere} \end{cases} \tag{3}$$

where $\Phi_1(\beta) = \int_{-\infty}^{\beta}(1/\sqrt{2\pi})e^{-(t^2/2)}dt$ is the cumulative distribution function (CDF) of standard univariate normal distribution.

Let the channel fading state be characterized by L amplitude levels with boundary values at $[Z_i, Z_{i+1})\ \forall i = 0, \ldots, L-1$. With the current time slot, i.e., the last transmission slot of current data batch denoted as slot 0, the probability that the received signal strength in current time slot, $Z(0)$ is in level i is used to estimate the current CS as $\psi_i(0) = \Pr\{Z_i \leq Z(0) \leq Z_{i+1}\} = \Pr\{CS(0) = i\}$. Given that the received signal envelope in the current slot is $Z(0)$, the current CS is known, i.e., $\psi_i(0) = 1$. The probability that the channel in the next time slot belongs to any state i, $\psi_i(1) = \Pr\{Z(1) \in [Z_i, Z_{i+1})\}$ is $\Pr\{Z_i \leq Z(0) + \delta Z(1) \leq Z_{i+1}\}$ which is evaluated as

$$\psi_i(1) = p\{Z_i \leq Z(0) + \delta Z(1) \leq Z_{i+1}\} \quad \forall i = 0, \ldots, L-1. \tag{4}$$

$$= \int_{Z_i - Z(0)}^{Z_{i+1} - Z(0)} f_{\delta Z(1)}(\alpha) d\alpha$$

$$= \frac{\Phi_1\left(\frac{Z_{i+1}-Z(0)}{\dot{\sigma_1}}\right) - \Phi_1\left(\frac{Z_i - Z(0)}{\dot{\sigma_1}}\right)}{1 - \Phi_1\left(-\frac{Z(0)}{\dot{\sigma_1}}\right)}. \tag{5}$$

Thus, for continuous packet transmission within a batch, (5) gives the probability distribution of CS in the next time slot. With the distribution of CS i in the next time slot known $\forall i = 0, \ldots, L-1$, the channel is estimated to be in a state having the highest probability in that slot. Mathematically

$$\mathrm{CS}(1) = i \text{ such that,}$$
$$\psi_i(1) = \max\{\psi_j(1) \quad \forall j = 0, \ldots, L-1\}. \tag{6}$$

It may be recalled that due to sporadic PMU data, interbatch arrival duration n is blind, where the received signal envelope, and hence actual CS, is unknown due to the absence of any transmission. Thus, it is required to estimate n slot ahead CS $\mathrm{CS}(n)$ from the knowledge of CS at the transmitter during slot 0, i.e., $\mathrm{CS}(0)$. From $\mathrm{CS}(0)$, the first slot, i.e., $\mathrm{CS}(1)$ marks the starting of blind interval of duration n slots when no packets are transmitted. Using (5), $\psi_i(1)$ gives the probability distribution of CSs in the first blind slot. For successive blind slots, the probability that the CS is in one of the L levels is estimated from Bayes' rule by iteratively conditioning the CS distribution in the next slot on the probabilistic CSs in current slot. Mathematically, for any blind slot $\kappa \in (2, \ldots, n)$, the probability distribution of CSs is expressed as

$$\psi_i(\kappa) = \sum_{j=0}^{L-1} p\{\mathrm{CS}(\kappa) = i | \mathrm{CS}(\kappa - 1) = j\}$$
$$\times\ p\{\mathrm{CS}(\kappa - 1) = j\} \tag{7}$$

$$= \sum_{j=0}^{L-1} \frac{\Phi_1\left(\frac{Z_{i+1}-\bar{Z}_j}{\dot{\sigma_\kappa}}\right) - \Phi_1\left(\frac{Z_i - \bar{Z}_j}{\dot{\sigma_\kappa}}\right)}{1 - \Phi_1\left(-\frac{\bar{Z}_j}{\dot{\sigma_\kappa}}\right)} \cdot \psi_j(\kappa - 1)$$
$$\forall i = 0, \ldots, L-1 \tag{8}$$

where $\dot{\sigma_\kappa} = {T_s}^\kappa \cdot \dot{\sigma}$ denotes the variance of the probability distribution function of $\dot{Z}(t)$ in the κth slot and $\bar{Z}_j = (Z_j + Z_{j+1})/2$ is the mean value of signal envelope in the jth level. Consequently, using (6), the probability distribution of CSs at the end of n-slot blind interval when a new batch of packets is available for transmission is given by

$$\text{CS}(n) = i \text{ such that,}$$
$$\psi_i(n) = \max\{\psi_j(n) \quad \forall j = 0, \ldots, L-1\}. \tag{9}$$

The proposed stochastic framework for CS estimation is simple and oblivious to the distribution of the underlying fading model. However, being a first-order model, its efficacy in predicting rapidly varying CSs, especially during the blind intervals, is limited. To this end, a data-driven framework using the Gaussian process regression is proposed in the next section to predict CSs for the sporadic PMU data transmission process.

B. Learning-Based Framework

In contemporary research, data-driven techniques are widely investigated to support the diverse requirements of next-generation wireless networks. Here, since the availability of packets for transmission at the PMU is intermittent in nature and comprises of several blind intervals over a period of time, the intuition for proposing a learning-based framework is to *learn* the instantaneous channel gain at the packet transmission instant using previous channel gains when the packet transmission has occurred and accordingly choose optimal redundancy. The proposed model for CS prediction is based on the Gaussian process regression. As a special case of the Bayesian probabilistic inference, it can model complex time sequences in the presence of incomplete information through kernel modifications [37]. Hence, suitable for long-term forecasting in the sporadic communication scenarios as considered in this article.

Denoting the last transmission slot of current data batch as slot 0, it is required to predict channel gain $x(n)$ for the estimation of CS $\text{CS}(n)$ at the end of n-slot blind interval. Let $\{X_A = x(0), x(-1), \ldots, x(-(a-1))\}$ be the time sequence of channel gains corresponding to slots in which packets are previously transmitted. It may be noted that due to sporadic PMU data, $x(i)s$ need not be regularly sampled. Since there are missing values in X_A corresponding to slots in which no packet is transmitted, we drop the slot index for ease of notation and redenote $\{X_A = x_{n-1}, x_{n-2}, \ldots, x_{n-a}\}$ such that the latest observed channel gain values required for predicting channel gain at the end of n-slot blind interval $x(n)$ are denoted as $x(0) \equiv x_{n-1}, x(-1) \equiv x_{n-2}$, and so on. Likewise, also denoting $x(n)$ as x_n for further analysis in the proposed learning-based framework.

The predicted instantaneous channel gain $\hat{x_n}$ is assumed to be a nonlinear function of its feature vector x_{F_n}, comprising of optimal number of lagged channel gain samples d. Consequently, for regression analysis, the training set is structured as $\{(x_{F_{n-1}}, x_{n-1}), \ldots, (x_{F_{n-a}}, x_{n-a})\} \subset \mathbb{R}^d \times \mathbb{R}$. The input space is d-dimensional such that $x_{F_i} = \{x_{i-1}, x_{i-2}, \ldots, x_{i-d}\}$. Considering the regression model

$$x_n = f\left(x_{F_n}\right) + \epsilon_n \tag{10}$$

where f is a function that maps the input x_{F_n} to the label x_n, and $\epsilon_n \sim \mathcal{N}(0, \sigma^2)$. From the theory of the Gaussian process regression [37], function f is a random variable characterized by the Gaussian process with 0 mean and covariance kernel function $\mathcal{K}(x_{F_n}, x'_{F_n})$, i.e.,

$$f\left(x_{F_n}\right) \sim \mathcal{GP}\left(0, \mathcal{K}\left(x_{F_n}, x'_{F_n}\right)\right). \tag{11}$$

To deduce f, prior over function f is updated into a posterior through the likelihood function. Denoting all input vectors as feature matrix $X_F = \{x_{F_{n-1}}, x_{F_{n-2}}, \ldots, x_{F_{n-a}}\}^T$ and outputs as label vector $X_\alpha = \{x_{n-1}, x_{n-2} \cdots x_{n-a}\}^T$. Following (11), prior over f is expressed as

$$p(f|X_F) \sim \mathcal{N}\left(f|0, \mathcal{K}\left(x_{F_n}, x'_{F_n}\right)\right). \tag{12}$$

Assuming likelihood $p(X_\alpha|f)$ to be also a Gaussian function such that the mean of likelihood is centered around arbitrary f

$$p(X_\alpha|f) \sim \mathcal{N}\left(X_\alpha|f, \sigma^2 I\right). \tag{13}$$

From Bayes' inference, posterior over function f, $p(f|X_F, X_\alpha) \propto p(X_\alpha|f)p(f|X_F)$. Since both prior and likelihood are Gaussian, posterior over f is also a Gaussian distribution. Using (12) and (13), we have

$$p(f|X_F, X_\alpha) \sim \mathcal{N}\left(f|\tilde{\mu}, \tilde{\sigma}^2\right) \tag{14}$$
$$\tilde{\mu} = \mathcal{K}\left(x_{F_n}, x'_{F_n}\right)\left[\mathcal{K}\left(x_{F_n}, x'_{F_n}\right) + \sigma^2 I\right]^{-1} X_\alpha$$
$$\tilde{\sigma}^2 = \mathcal{K}\left(x_{F_n}, x'_{F_n}\right)\left[\mathcal{K}\left(x_{F_n}, x'_{F_n}\right) + \sigma^2 I\right]^{-1} \sigma^2 I.$$

For predicting through the Gaussian process regression, it is required to evaluate the predictive posterior which essentially predicts over all possible fs weighted by posterior in (14) as

$$p\left(\hat{x}_n|x_{F_n}, X_F, X_\alpha\right) = \int p\left(\hat{x}_n|x_{F_n}, f, X_F\right) \cdot p(f|X_F, X_\alpha) df \tag{15}$$

where $\hat{x}_n$ is the predicted value corresponding to the label x_n. The predictive posterior is again a Gaussian given by

$$p\left(\hat{x}_n|x_{F_n}, X_F, X_\alpha\right) \sim \mathcal{N}\left(\hat{x}_n|\hat{\mu}, \hat{\sigma}^2\right) \tag{16}$$
$$\hat{\mu} = \mathcal{K}\left(\hat{x}_n, X_F\right)\left[\mathcal{K}(X_F, X_F) + \sigma^2 I\right]^{-1} X_\alpha$$
$$\hat{\sigma}^2 = \mathcal{K}\left(\hat{x}_n, \hat{x}_n\right) - \mathcal{K}\left(\hat{x}_n, X_F\right) \times \left[\mathcal{K}(X_F, X_F) + \sigma^2 I\right]^{-1} \mathcal{K}\left(X_F, \hat{x}_n\right).$$

Predicted value $\hat{x}_n$ is the mean of this predictive distribution. Thus, for fading channel characterized by L levels with channel gain boundaries demarcated as $[X_j, X_{j+1}) \ \forall j = 0 \cdots L-1$, CS at the end of the nth time slot is given by

$$\text{CS}(n) = j, \text{ if } \left\{X_j \le \hat{x}_n < X_{j+1}\right\}. \tag{17}$$

In this article, training and predictions of channel samples using the Gaussian process regression is performed using statistics and machine learning toolbox in MATLAB 2018b.

C. Probing-Based Framework

It may be noted that data transmission based on CS prediction requires a computational overhead at the transmitting node in terms of execution of stochastic as well as learning-based prediction models. To this end, a probing-based data transmission approach is proposed here wherein, if the channel is being used for data transmission after a long time interval, a probing packet is first transmitted to estimate the CS. In the case of successive packet transmissions, feedback from the receiver is collected at the transmitter to update CS. Thus, the number of probing packets required is equal to the number of blind intervals encountered during sporadic PMU packet transmissions. Also, maximum redundancy is assigned to the probing query and response in order to ensure their successful reception. This probing-based data transmission scheme appears to be a more intuitive approach for the sporadic data communication scenario since the channel knowledge is based on immediate probing feedback, it is more accurate compared to the estimated knowledge in stochastic modeling and learning-based modeling. Consequently, the packet loss probability in probing-based data transmission is considered as a benchmark for comparing the performance of the previous two-channel prediction-based data transmission schemes.

D. Complexity of the Proposed Channel Prediction Algorithms

The evaluation of probabilistic CS distribution in the stochastic modeling framework as proposed in Section III-A primarily includes the computation of standard univariate normal CDF, $\Phi_1(\beta)$ in (5), and identifying the state having highest probability in (6). $\Phi_1(\beta)$ can be derived from the error function as $\Phi_1(\beta) = (1/2)(1 - \text{erf}(-\beta/\sqrt{2}))$. For the purpose of numerical computation, let the error function be represented as $y = \text{erf}(\beta)(1 + \delta)$. It is found that in existing software tools, such as MATLAB and Mathematica, δ is assumed to be in the order of 10^{-7} and the evaluation of y is based on rational approximation as suggested in [38]. Thus, the computation of $\Phi_1(\beta)$ is of constant complexity. Furthermore, for L CSs, identifying the CS having highest probability requires $\mathcal{O}(L)$ computations. Consequently, the net complexity is $3L\mathcal{O}(1) + \mathcal{O}(L) \sim \mathcal{O}(L)$. For this article, a fixed number of CSs are considered, thus computation complexity of the proposed stochastic framework is essentially constant.

To analyze the complexity of CS prediction using the learning-based approach as proposed in Section III-B, training of the regression model, prediction of channel gains, and identifying CS from predicted channel gains are the essential steps. For each of these, computation complexities are, respectively, found to be $\mathcal{O}(a^3)$ [37], $b\mathcal{O}(a^3)$, and $\mathcal{O}(a)$, where a denotes the training length and b is the number of step-ahead predictions. Thus, the net computation complexity of the proposed learning-based framework is on the cubic order of training length.

The computation complexity of the probing-based data transmission is negligible as no intelligent signal processing is required at the transmitter to know the current CS.

IV. Proposed Channel-Adaptive Transmission

It may be noted that unlike wireline Ethernet protocols, in case of wireless transmission, channel uncertainties, such as fading and interference need to be carefully addressed to meet the required QoS. Broad guidelines for PMU data communication methods using IP over Ethernet in a client–server format are defined in the IEEE standard C37.118. However, to the best of our knowledge, no standard protocols specifically defined for handling the vagaries of communicating PMU data over wireless channel exist in the literature. To this end, in this article, we have complemented the proposed channel prediction techniques based on stochastic modeling, data-driven learning, and probing with a channel-aware data transmission scheme, wherein the knowledge of predicted CS is exploited to adaptively choose the channel coding parameters for efficient and reliable transmission of time-critical PMU data over the wireless channel. In this section, the adaptive scheme for sporadic but time-critical PMU data transmission based on the proposed CS prediction frameworks, as discussed in Section III, is presented along with the performance indices.

A. Channel-Adaptive Transmission Scheme

A flowchart representation of the channel-adaptive transmission scheme using the proposed stochastic, learning, and probing-based CS prediction frameworks are shown, respectively, in Fig. 2(a)–(c). An adaptive block coding is chosen in the transmission approach because of the time-critical nature of PMU data, wherein retransmission of lost packets is not feasible. RS code is a linear nonbinary block code, suited for correction of burst errors over wireless channels [39]. It is denoted as RS(c, k) with both c and k represented by m bit symbols such that for every k information symbols, $c - k$ parity symbols are appended to create c symbol codeword. For a given (c, k) block, the RS decoder can correct up to $(c - k)/2$ symbol errors of m bits each. The transmission scheme primarily includes the prediction of current CS using the proposed frameworks and the selection of appropriate block length c_i for the packet transmission. Here, the subscript i corresponds to the fading level i $\forall i = 0, \ldots, L - 1$. At the receiver, a packet is successfully received if the number of erroneous symbols e is within the error correction capability of the code; else the packet is dropped. Thus, the proposed adaptive coding responds to current CS by appropriately choosing c_i, for attaining a high packet success rate with far less bandwidth requirement, unlike fixed-rate code where error correction is always intended for the worst case scenario.

As discussed in Section III-D, since the complexity of the proposed learning-based channel prediction algorithm is $\mathcal{O}(a^3)$, the length of the training set is limited to optimum training length (OTL) such that the prediction is statistically reliable and computationally practical. The selection of OTL and other parameters of the learning model is further discussed in Section V. To build the training set for subsequent prediction, selective slotwise channel gains from the current packet transmission duration are communicated to the transmitter once decoding is completed at the receiver. In contrast, CSI of the latest slot only is required in the

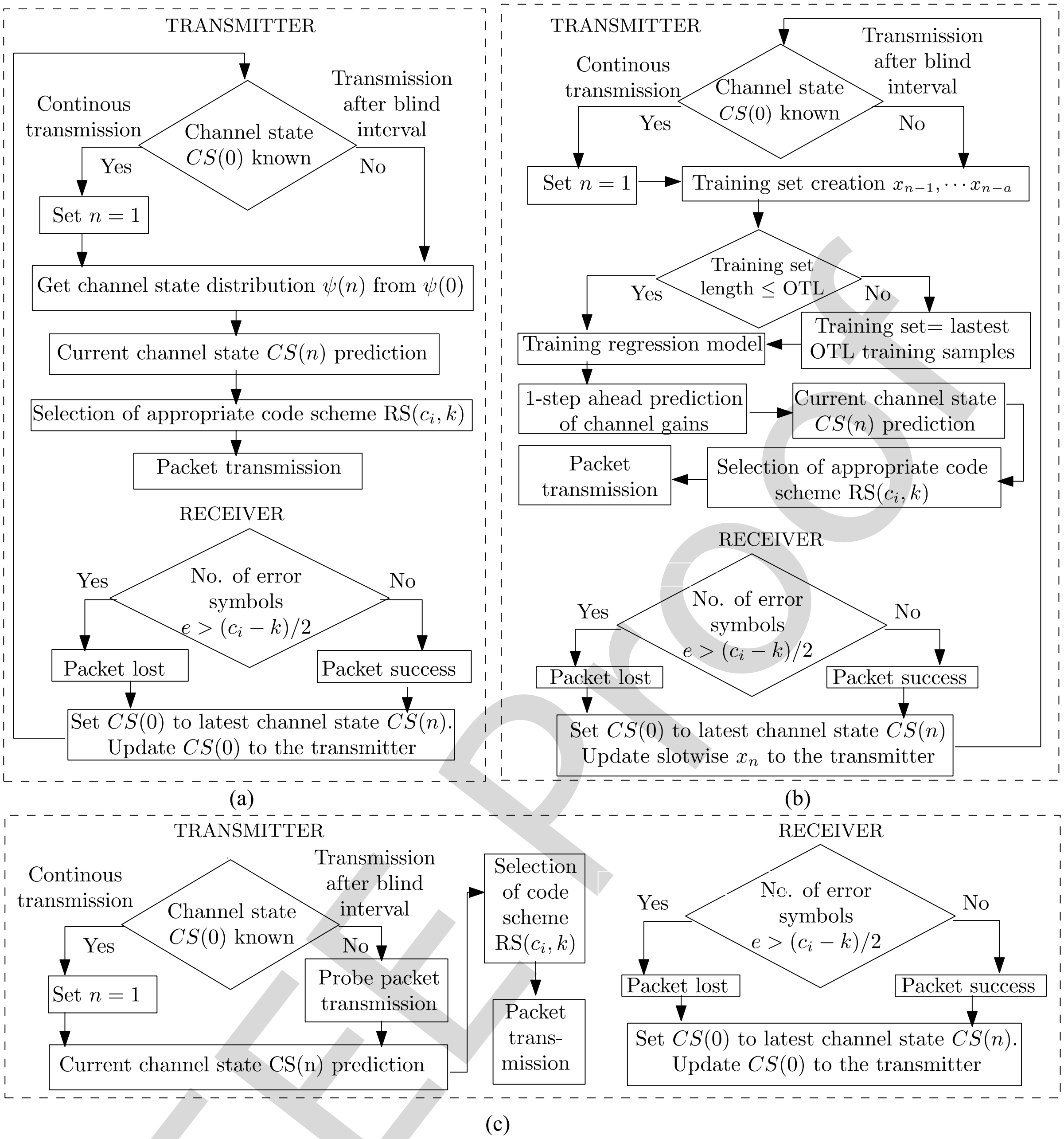


Fig. 2. Channel-aware transmission schemes for time-critical PMU data based on channel estimation using: (a) stochastic modeling; (b) learning; and (c) probing-based approaches.

case of the stochastic modeling framework, whereas in the probing-based approach, the CSI of the last slot is collected immediately before the data packet transmission. Thus, the devised transmission schemes aim at maximizing the PMU packet success probability by exploiting the historical channel information, thereby increasing the reliability of grid operation.

B. Performance Indices

The proposed stochastic modeling, learning, and probing-based channel prediction frameworks as developed in Section III are studied by numerical simulations. Also, to verify the analytical performance, the channel-adaptive transmission schemes for each of the proposed approaches are studied over simulated fading channels in MATLAB. The performance is quantified using the following indices.

1) *False Prediction Probability p_f:* It is defined as the ratio of predicted CSs not matching with actual CSs over the total number of packets transmitted N_p, over a sufficiently large time interval Δ, i.e., $\lim_{\Delta\to\infty} p_f =$ (number of mismatched predictions)$/N_p$.
2) *Symbol Error Probability p_{se}:* It is the ratio of the number of symbols with received SNR below threshold SNR over the total number of symbols transmitted during the interval Δ. Let the number of erroneous symbols and the total number of transmitted symbols be N_{se} and N_s, respectively. Then, $p_{se} = \lim_{\Delta\to\infty} N_{se}/N_s$.

TABLE I
VARIATION OF COMMUNICATION SYSTEM PERFORMANCE WITH THE STRUCTURE OF THE RS CODE AT SNR = 10 dB AND $f_D = 50$ Hz

Number of bits per symbol, m	Number of information symbols, k	Maximum block size, c_{max}	Symbol error probability, p_{se}	Packet loss probability, p_l	Bandwidth consumed, BW_c (bps)
8	40	255	0.3939	0.05	2004.5
10	32	1023	0.3927	0.0073	1.009×10^4
12	27	4095	0.3941	3.67×10^{-4}	4.852×10^4
14	23	16383	0.3943	0	2.264×10^5
16	20	65535	0.3942	0	1.035×10^6

3) *Packet Loss Probability p_l:* It may be recalled that if the number of erroneous symbols in a packet encoded with RS(c_i, k) code exceeds $(c_i - k)/2$, it is considered to be lost. Here, c_i denotes the block size chosen for a packet transmitted during the fading level i, $0 \leq i \leq L-1$. If the number of packets lost over time interval Δ is N_{pl}, then $p_l = \lim_{\Delta\to\infty} N_{pl}/N_p$.
4) *Bandwidth Consumption BW_c:* It is the amount of data transmitted over the wireless link during time interval Δ. If the jth PMU packet is encoded as a block of length $c_i(j)$, then $BW_c = \lim_{\Delta\to\infty} \sum_{j=1}^{N_p} c_i(j)/\Delta$.
5) *Signaling Overhead O_s:* It is the average number of additional bits transmitted per symbol over the wireless link for enhancing the performance of a chosen communication protocol. This includes the probing overhead and the feedback counts at the transmitter. Denote the number of probing packets and count of feedback collected over time interval Δ as N_{prob} and N_{fb}, respectively. Then, $O_s = \lim_{\Delta\to\infty} 2mc_{\max}N_{\text{prob}}/\sum_{j=1}^{N_p} c_i(j)$, where $c_{\max} = 2^m - 1$ is the maximum block length that can be assigned using RS(c, k) code with both c and k being m bit symbols.

V. RESULTS AND DISCUSSION

In this section, first the structure of the channel-adaptive RS coding scheme is presented. Subsequently, the prediction quality of the proposed stochastic modeling and learning-based frameworks is discussed. Next, the prediction and packet loss performance of the proposed stochastic modeling and learning-based channel prediction approaches are compared with the probing-based transmission approach for different SNR values and fading coefficients, followed by a discussion on their overhead requirements.

An example of the Rayleigh fading wireless channel is considered for numerical performance studies. Typical system parameters considered are: symbol duration $T_s = 1$ ms, carrier frequency $f_c = 900$ MHz, threshold SNR = 7 dB, and PMU packet size = 40 B. The fading wireless channel is characterized by three states. Accordingly, three coding schemes: RS(c_0, k), RS(c_1, k), and RS(c_2, k) are used. CS boundaries in this article are set at 10 and 25 dB. Hence, the CS in a slot is either 0, or 1, or 2, respectively, when the received symbol SNR in that interval is <10 dB, between 10 and 25 dB, and >25 dB. It is observed that the prediction performance of the proposed stochastic modeling framework and learning-based framework exhibits similar behavior irrespective of the choice of CS boundaries.

A. Choice of Adaptive RS Coding Parameters

For RS(c_i, k) code, the error correction capability is governed by block size c_i. Recall that for a given (c_i, k) block, the RS decoder can correct up to $(c_i - k)/2$ symbol errors of m bits each. During the worst CS, the block size is chosen to be $c_{\max} = 2^m - 1$ in order to provide maximum error protection. Likewise for the best CS, minimum block size $c_{\min}$ is specified. From large-scale simulations of the proposed channel-adaptive transmission schemes over a wireless fading channel, it is identified that during the best CS, the desired value of $c_{\min}$ is at least 50 symbols for required quality of service at the PDC. For simplicity, a packet transmitted in an intermediate CS is assigned a block size of $(c_{\max} + c_{\min})/2$. Thus, faithful recovery of erroneous packets during different CSs is governed by the selection of parameter m. In Table I, variation of communication system performance with different RS code parameters is presented for SNR = 10 dB and $f_D = 50$ Hz. It may be observed that with an increasing value of m, the symbol error probability remains almost the same due to fixed SNR, however, it adds more redundant symbols to the transmitted packets. Consequently, the packet loss probability eventually drops and bandwidth consumption increases. For the required quality of service, the packet loss probability is set at about 10^{-4}. Accordingly, $m = 12$ is chosen for further performance studies.

B. Channel State Prediction Using the Stochastic Framework

Using the stochastic modeling framework proposed in Section III-A, CS estimation during the simulation of the sporadic communication scenario between PMU and PDC is studied. In Fig. 3, predicted CSs are compared against actual CSs in the corresponding slots for SNR = 10 dB and $f_D = 50$ Hz. It may be noted that owing to the sporadicity of data transmission instants, the samples are not equispaced. The gap between some of the consecutive samples represents the blind intervals during which no PMU packets were available for transmission. It may be noted that the CS prediction for the current slot using stochastic modeling is based only on the state in the previous slot. Consequently, it is observed from Fig. 3 that the stochastic predictions follow the change in actual CS with a lag of one sample. In the case of sustained

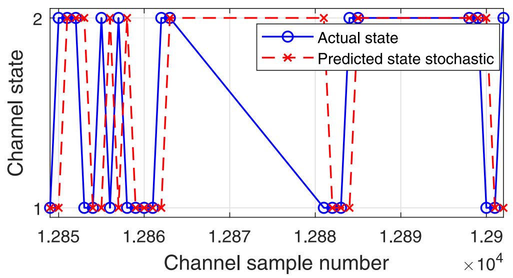
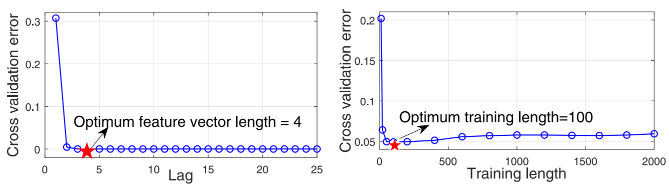


Fig. 3. Predicted CS using the stochastic modeling approach with respect to actual CS, at SNR = 10 dB and $f_D = 50$ Hz.



(a) (b)

Fig. 4. Optimum parameter selection for the learning-based model. (a) Feature vector length. (b) Training set length.

CS over consecutive slots, stochastic predictions exactly match with the actual CS.

C. Channel State Prediction Using the Learning-Based Framework

As discussed in Section III-D, runtime of the learning-based model for CS estimation using the Gaussian process regression is influenced by the length of the training set used in the prediction model. Besides, the length of the input feature vector comprising of lagged channel samples is another user-defined parameter in the model implementation. In this article, v-fold cross-validation error of the Gaussian process regression model is used to decide the optimum value of feature vector and training length. Fig. 4(a) and (b), respectively, captures the variation of mean cross-validation error versus lag value and training length, for 20 Rayleigh channel generation instances at SNR = 10 dB and $f_D = 50$ Hz. It may be observed from the plots that with increasing lag and training length, mean cross-validation does not improve beyond a certain value. This saturation point is chosen as optimum for learning-based model implementation. Specifically, for a given channel condition, optimum feature vector length and OTL are found to be 4 and 100 samples, respectively.

Using the Gaussian process regression model with OTL, 1-step ahead channel gain predictions using optimum feature vector length as input are made for every slot during the PMU packet transmission duration. Predicted CSs with respect to the actual CSs in the corresponding slots during large-scale simulations of the learning-based approach are shown in Fig. 5. It is observed that unlike the stochastic modeling approach, CS changes are better traced in the learning framework. Thus, false predictions with the learning-based approach are comparatively rare.

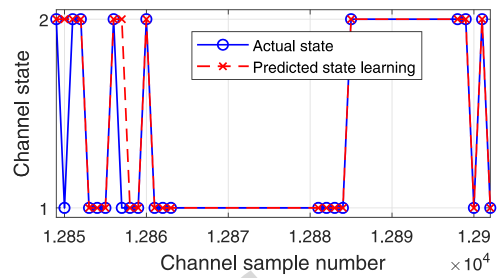


Fig. 5. Predicted CS using the learning-based prediction model with respect to actual CS, at SNR = 10 dB and $f_D = 50$ Hz.

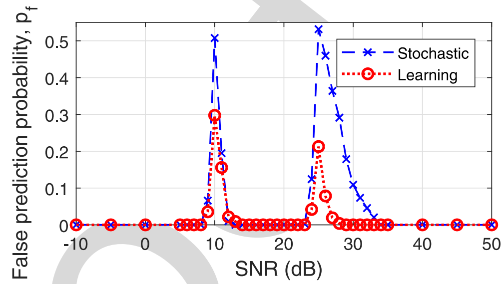


Fig. 6. Variation of false prediction probability of stochastic modeling and learning-based framework with SNR at $f_D = 50$ Hz.

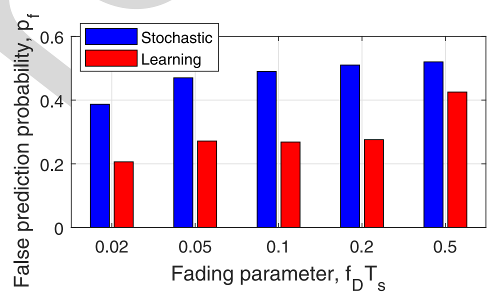


Fig. 7. Variation of false prediction probability of stochastic modeling and learning-based framework with fading at SNR = 10 dB.

D. Comparison of False Prediction Probability

Variation of false prediction probability p_f with increasing values of average SNR in the fading channel is presented in Fig. 6 for the proposed stochastic modeling and learning-based framework at $f_D = 50$ Hz. It is observed that the prediction accuracy in each case is sensitive to the SNRs located in the vicinity of CS boundaries. This behavior is observed because in these regions, the actual value of parameters that identify the CS (i.e., received signal envelope for stochastic modeling and channel gain in case of the learning-based approach) has a small separation margin from the boundary values. Thus, even a small prediction error may lead to false identification of CS. It is found that mean p_f of the learning-based model over different average SNR at $f_D = 50$ Hz is 70% lower with respect to the stochastic modeling framework. Additionally, for CS boundaries at 10 and 25 dB, the prediction accuracy of the learning-based model is higher, respectively by, 42% and 58%.

In Fig. 7, p_f of the stochastic modeling and learning-based frameworks with varying channel fading parameter $f_D T_s$ are presented at an average SNR = 10 dB. Note that $f_D T_s < 0.2$

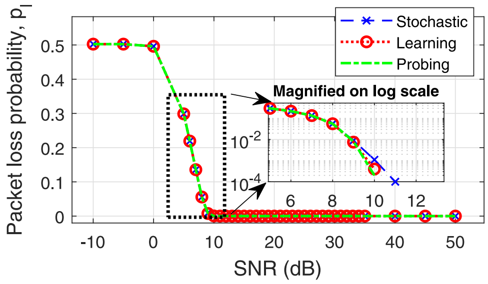


Fig. 8. Comparison of the packet loss probability of learning and stochastic modeling-based frameworks with respect to probing-based transmission at different SNR and $f_D = 50$ Hz.

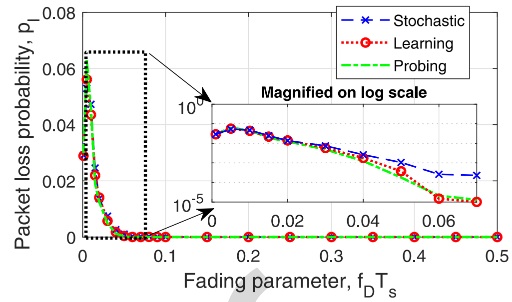


Fig. 9. Packet loss probability comparison of learning-based and stochastic modeling-based frameworks with respect to probing-based transmission at different fading parameter and SNR = 10 dB.

signifies a slow fading channel. Consequently, the successive channel samples are highly correlated, leading to a higher accuracy in CS predictions. On the contrary, for $f_D T_S > 0.2$ consecutive channel samples are almost temporally independent, thereby deteriorating the prediction accuracy. Hence, an increasing trend of p_f is observed in Fig. 7. Nevertheless, the learning-based model outperforms the stochastic modeling framework in terms of false prediction probability p_f. It is found that at 10 dB SNR, mean p_f over different values of fading parameter for learning-based model is 39% lower compared to the stochastic modeling framework.

Remark 1: In dynamic channel conditions, the learning-based model is able to follow channel dynamics more closely compared to the stochastic modeling-based framework.

E. Comparison of Packet Loss Probability

In this section, the channel-adaptive data transmission scheme based on the proposed stochastic modeling, learning, and probing-based CS estimation frameworks is simulated for varying average SNR and fading conditions, and their relative performances are compared with respect to packet loss probability p_l, respectively, in Figs. 8 and 9.

It is observed from Fig. 8 that at very low SNRs (< 0 dB), the channel is mostly unusable and p_l is high irrespective of the CS estimation approach. However, as the channel condition improves, p_l eventually drops close to 0. A detailed view on the log scale reveals that as compared to the stochastic modeling framework, p_l obtained using the learning-based model is close to the benchmark probing-based transmission. During the transition region, mean p_l of stochastic modeling, learning, and probing-based approaches is observed to be, respectively, 0.089, 0.078, and 0.079. The stochastic modeling framework has higher packet loss probability p_l due to high false prediction probability p_f. Numerically, with respect to the probing-based approach, at $f_D = 50$ Hz, mean packet loss probability over varying SNR for learning and stochastic model-based schemes is higher by 1.2% and 12.3%, respectively.

Fig. 9 shows loss performance p_l of the proposed stochastic modeling, learning, and probing-based transmission schemes at different values of channel fading parameter $f_D T_s$ at SNR = 10 dB. With increasing $f_D T_s$, p_l rapidly decays to 0 despite high false prediction probability p_f in the fast-fading scenarios. This behavior is primarily due to the efficacy of RS codes in handling fast fading. Without RS coding, the probing-based approach will benefit in the fast fading environment, where the prediction capability of the stochastic and learning-based frameworks gradually reduce due to decreasing correlation in channel samples. However, with the proposed channel-adaptive transmission scheme using RS coding, probing-based data transmission helps only over a small fading window.

For the slowly varying channel, the size of burst error is larger and may exceed the error correction capability of the code even after using maximum redundancy. A detailed view of p_l variation reveals that the performance of the channel-adaptive transmission scheme using stochastic modeling, learning, and probing-based CS estimation is alike for $f_D T_s < 0.02$. Thus, if the channel is highly correlated, the stochastic modeling framework, which is relatively simpler in terms of computation complexity and inexpensive due to the minimum feedback requirement is equally efficient. Consequently, learning and probing-based approaches may not be required at all in this region. However, with increasing $f_D T_S$, the prediction accuracy of the stochastic model deteriorates, while learning and probing-probing-based approaches adapt to channel dynamics.

Remark 2: With increasing average SNR and fading conditions in the channel, the performance of the learning-based approach closely matches with the probing-based approach and is better in comparison to the stochastic framework. However, stochastic modeling-based channel prediction benefits the system in case of a slowly varying channel.

F. Overhead Analysis

Signaling and computational overheads of the proposed adaptive transmission schemes are studied here.

For packet transmission, required channel overhead comprises of its block size and the corresponding signaling. It may be noted that signaling overhead varies with fading conditions in the channel, while for a given fading, the block size is chosen based on average SNR. Thus, a variation of signaling overhead with fading parameter and bandwidth consumption with SNR for adaptive data transmission schemes using proposed stochastic modeling, learning, and probing-based channel prediction frameworks are shown in Figs. 10 and 11, respectively. From Fig. 10, it can be observed that signaling overhead required for the stochastic model is minimum owing to the requirement of only previous slot CSI for

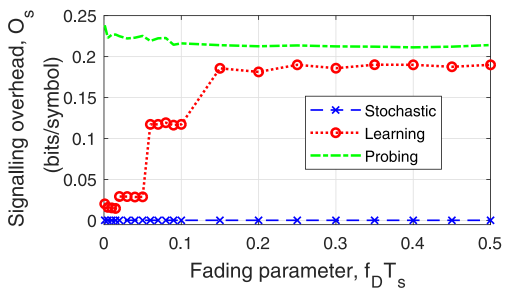


Fig. 10. Comparison of signaling overhead of the learning-based framework and stochastic modeling with respect to probing-based transmission, with varying fading parameter, SNR = 10 dB.

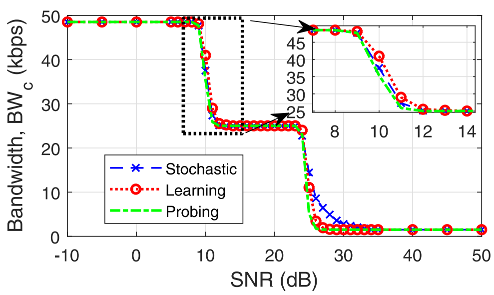


Fig. 11. Comparison of bandwidth consumption of learning-based and stochastic modeling-based frameworks with respect to probing-based transmission at different SNR and $f_D = 50$ Hz.

prediction of current CS. On the contrary, the training length of the learning-based prediction model increases with the fading parameter, thus adding to the signaling overhead. In the case of probing-based data transmission, requirement of signaling overhead is maximum as the probing query and its response are assigned the maximum number of redundant symbols for their successful reception. It is evaluated that for a fixed number of blind intervals at SNR = 10 dB, the mean signaling overhead requirement of the probing-based data transmission scheme exceeds by 65.8% and 11.2% with respect to data transmission using the proposed learning model for, respectively, slow and fast varying channel. In comparison to data transmission using stochastic modeling, the signaling overhead of probing-based data transmission is almost 98% higher.

Furthermore, from the bandwidth consumption plot in Fig. 11, it can be noted that BW_c in the worst CS (SNR < 10 dB) is highest owing to the largest size of transmission block in the channel-adaptive coding scheme, followed by intermediate state (10 dB $\leq$ SNR $\leq$ 25 dB), and least BW_c in the best state (SNR > 25 dB). Moreover, it can be observed that due to better prediction accuracy in the vicinity of channel boundaries (see Fig. 6), BW_c in the learning-based approach is optimized to suit the channel conditions. For instance, from the magnified subplot in Fig. 11, BW_c of the learning-based model is higher compared to the stochastic framework when more symbols are expected to be in error in order to maintain low packet loss and *vice versa*. Numerically, with respect to the probing-based approach, at $f_D = 50$ Hz, mean BW_c over varying SNR for learning and stochastic model-based schemes is higher by 2.3% and 4%, respectively, which is only marginal.

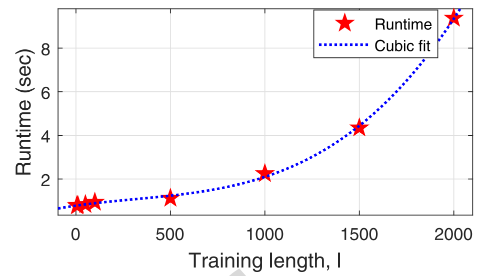


Fig. 12. Variation of runtime with training length in learning-based CS prediction framework.

It may be recalled from Section III-D that the computational complexity (i.e., runtime) of the stochastic modeling and learning-based framework is, respectively, constant and $\mathcal{O}(a^3)$, where a is the training length. Mean runtime of the proposed stochastic modeling is found to be 0.0038 s/packet. In Fig. 12, variation of runtime with training length for transmission of 100 packets using learning-based CS estimation is presented. The cubic nature of runtime variation with training length as studied in Section III-D is validated in this plot using curve fitting. The parameters of curve fitting as obtained are: runtime, $\tau(a) = \lambda_1 a^3 + \lambda_2 a^2 + \lambda_3 a + \lambda_4$, where $\lambda_1 = 0.6655$, $\lambda_2 = 0.8557$, $\lambda_3 = 0.9921$, and $\lambda_4 = 1.397$; goodness of fit, $R^2 = 0.9992$, root mean-square error (RMSE) = 0.011. It may be recalled that computation complexity of data transmission using probing-based CS estimation is negligible.

Remark 3: The computation complexity of the learning-based prediction model is higher. However, for varying channel conditions, it incurs far less signaling overhead and has comparable packet loss performance compared to the benchmark probing-based transmission. Also, runtime complexity as well as signaling overhead of the stochastic modeling framework are significantly low, though it incurs somewhat higher packet losses, especially in more dynamic channels.

G. Delay Investigation

For the real-time implementation of the proposed channel-aware transmission protocols, the computation capability of commercially available hardware PMUs can be augmented using a secondary processor such as Raspberry Pi (RPi) on which the stochastic modeling, data-driven framework, and probing-based channel prediction models are configured. From the networking literature, it is known that the delay incurred for the successful reception of a packet comprises of processing, transmission, and propagation delay. It may be recalled from Section II that in this article, a point-to-point communication scenario is considered where the PMU packets are transmitted to the nearest PDC over a single-hop wireless communication network such that the propagation delay is negligibly small. Besides, with the use of 4G technologies, such as LTE having a typical uplink rate of 50–100 Mb/s, the transmission time of PMU data packet is on the order of microseconds, which is insignificant. Also, due to the time-critical nature of PMU data, retransmissions are not considered. Consequently, the primary component of delay involved in the transmission of time-critical PMU data is the execution

TABLE II
PROCESSING DELAY OF THE ADAPTIVE TRANSMISSION STRATEGIES

Channel prediction model	Processing delay
Probing	1.2 ms per packet
Stochastic	3.4 ms per packet
Learning ($f_D T_s \leq 0.02$)	9.41 ms per packet
Learning ($0.02 < f_D T_s \leq 0.05$)	12.83 ms per packet
Learning ($0.05 < f_D T_s \leq 0.1$)	18.88 ms per packet
Learning ($0.1 < f_D T_s$)	25.38 ms per packet

time of the proposed channel-aware transmission strategies on the augmented secondary processor. It may be noted that since the Linux-based operating system is supported on most secondary processors, the proposed channel-aware transmission strategies are executed in a Python-based environment to have an estimate of the processing delay. In Table II, code simulation times observed during the execution of channel-aware transmission framework using probing, stochastic modeling, and data-driven learning-based prediction model, respectively, on Python 3.7.4 running on Intel i7 processor @ 2.4 GHz and 8-GB RAM are presented.

To meet the QoS criterion of smart grid monitoring and control, it is required that the incurred delay is within the acceptable latency limits. This is typically in the range of 20 ms–10 s and varies with the kind of application feeding on the data [40]. From Table II, it is observed that for the probing-based approach and stochastic modeling-based prediction, the respective processing delays are well within the minimum acceptable latency threshold. In case of the learning model, the latency bound is easily met for slowly varying channel ($f_D T_s \leq 0.1$), while for fast variations ($f_D T_s > 0.1$), processing delay is on the same order as the minimum latency threshold. It is notable from [11] that the delay in learning-based pruning is around 12 ms. Since smart grid networks with fixed deployed PMUs and PDCs are expected to experience very little mobility in the environment (equivalently low value of $f_D T_s$), delay in learning-based channel adaptation is typically less than 10 ms. Hence, the total processing time in data pruning and learning-based channel adaptation is expected to be closely around the minimum delay limit.

Remark 4: Execution of the proposed channel-adaptive transmission protocols in a Python-based environment indicates that the proposed stochastic modeling-based as well as probing-based approaches require negligible additional processing delay at the smart PMU node. It is also found that the proposed learning-based approach can be effectively implemented at a minor cost of adding secondary processing and storage capabilities, and the total data handling delay at the smart PMU is closely comparable to the required latency constraint for delivery of time-critical PMU data.

VI. CONCLUSION

To summarize, in this article, novel strategies have been proposed for channel-aware transmission of sporadic but time-critical PMU data in smart grid IoT networks. It has been demonstrated that by exploiting temporal correlation in the wireless channel, the proposed techniques, especially learning-based prediction can effectively follow channel variability leading to accurate CS prediction in the required transmission slots. In comparison to the benchmark probing-based data transmission scheme, at $f_D = 50$ Hz, mean packet loss probability over varying SNR for the stochastic modeling and learning-based transmission exceed by 12.3% and 1.2%, respectively, though their corresponding signaling overhead requirements are 98% and 38% lower.

With this article, we anticipate that augmenting the smart IoT devices, such as smart PMUs, with node-level intelligence in terms of channel awareness and adaptive data transmission capability will significantly contribute to efficient handling of big data footprints in future IoT communications.

REFERENCES

[1] J. De La Ree, V. Centeno, J. S. Thorp, and A. G. Phadke, "Synchronized phasor measurement applications in power systems," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 20–27, Jun. 2010.

[2] R. Ma, H.-H. Chen, Y.-R. Huang, and W. Meng, "Smart grid communication: Its challenges and opportunities," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 36–46, Mar. 2013.

[3] Y. Li, X. Cheng, Y. Cao, D. Wang, and L. Yang, "Smart choice for the smart grid: Narrowband Internet of Things (NB-IoT)," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1505–1515, Jun. 2018.

[4] R. Arghandeh *et al.*, "Data mining techniques and tools for synchrophasor data," North Amer. Electricity Rel. Corporat., Princeton, NJ, USA, Rep. PNNL-28218, Jan. 2019.

[5] J. O. Fernandez, "The Virginia Tech calibration system," M.S. thesis, Virginia Polytechnic Inst., State Univ., Blacksburg, VA, USA, 2011.

[6] S. Tripathi and S. De, "Data-driven optimizations in IoT: A new frontier of challenges and opportunities," *CSI Trans. ICT*, vol. 7, no. 1, pp. 35–43, Mar. 2018.

[7] M. Ghorbanian, S. H. Dolatabadi, and P. Siano, "Big data issues in smart grids: A survey," *IEEE Syst. J.*, vol. 13, no. 4, pp. 4158–4168, Dec. 2019.

[8] V. Gupta, S. Tripathi, and S. De, "Green sensing and communication: A step towards sustainable IoT systems," *J. Indian Inst. Sci.*, to be published.

[9] D. Wang, D. Chen, B. Song, N. Guizani, X. Yu, and X. Du, "From IoT to 5G I-IoT: The next generation IoT-based intelligent algorithms and 5G technologies," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 114–120, Oct. 2018.

[10] S. Das and T. S. Sidhu, "Application of compressive sampling in synchrophasor data communication in WAMS," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 450–460, Feb. 2014.

[11] S. Tripathi and S. De, "Dynamic prediction of powerline frequency for wide area monitoring and control," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 2837–2846, Jul. 2018.

[12] P. H. Gadde, M. Biswal, S. Brahma, and H. Cao, "Efficient compression of PMU data in WAMS," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2406–2413, Sep. 2016.

[13] J. Khan, S. Bhuiyan, G. Murphy, and J. Williams, "Data denoising and compression for smart grid communication," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 2, no. 2, pp. 200–214, Jun. 2016.

[14] V. Loia, S. Tomasiello, and A. Vaccaro, "Fuzzy transform based compression of electric signal waveforms for smart grids," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 1, pp. 121–132, Jan. 2017.

[15] S. Tripathi and S. De, "An efficient data characterization and reduction scheme for smart metering infrastructure," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4300–4308, Oct. 2018.

[16] M. R. Chowdhury, S. Tripathi, and S. De, "Adaptive multivariate data compression in smart metering Internet of Things," *IEEE Trans. Ind. Informat.*, early access, Mar. 17, 2020, doi: 10.1109/TII.2020.2981382.

[17] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proc. IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec. 2007.

[18] P. Mukherjee, D. Mishra, and S. De, "Exploiting temporal correlation in wireless channel for energy-efficient communication," *IEEE Trans. Green Commun. and Netw.*, vol. 1, no. 4, pp. 381–394, Dec. 2017.

[19] F. Zeng, R. Zhang, X. Cheng, and L. Yang, “Channel prediction based scheduling for data dissemination in VANETs,” *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1409–1412, Jun. 2017.
[20] C. Lv, J. Lin, and Z. Yang, “Channel prediction for millimeter wave MIMO-OFDM communications in rapidly time-varying frequency-selective fading channels,” *IEEE Access*, vol. 7, pp. 15183–15195, 2019.
[21] A. Heidari, A. K. Khandani, and D. Mcavoy, “Adaptive modeling and long-range prediction of mobile fading channels,” *IET Commun.*, vol. 4, no. 1, pp. 39–50, Jan. 2010.
[22] L. Liu, H. Feng, T. Yang, and B. Hu, “MIMO-OFDM wireless channel prediction by exploiting spatial–temporal correlation,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 310–319, Jan. 2014.
[23] Z. Xu, M. Hofer, and T. Zemen, “A time-variant channel prediction and feedback framework for interference alignment,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5961–5973, Jul. 2017.
[24] R. O. Adeogun, P. D. Teal, and P. A. Dmochowski, “Extrapolation of MIMO mobile-to-mobile wireless channels using parametric-model-based prediction,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 10, pp. 4487–4498, Oct. 2015.
[25] H. P. Bui, Y. Ogawa, T. Nishimura, and T. Ohgane, “Performance evaluation of a multi-user MIMO system with prediction of time-varying indoor channels,” *IEEE Trans. Antennas Propag.*, vol. 61, no. 1, pp. 371–379, Jan. 2013.
[26] S. Uehashi, Y. Ogawa, T. Nishimura, and T. Ohgane, “Prediction of time-varying multi-user MIMO channels based on DOA estimation using compressed sensing,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 565–577, Jan. 2019.
[27] W. Peng, M. Zou, and T. Jiang, “Channel prediction in time-varying massive MIMO environments,” *IEEE Access*, vol. 5, pp. 23938–23946, 2017.
[28] J. F. Schmidt, J. E. Cousseau, R. Wichman, and S. Werner, “Low-complexity channel prediction using approximated recursive DCT,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 10, pp. 2520–2530, Oct. 2011.
[29] T. Ding and A. Hirose, “Fading channel prediction based on combination of complex-valued neural networks and chirp Z-transform,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 9, pp. 1686–1695, Sep. 2014.
[30] S. Navabi, C. Wang, O. Y. Bursalioglu, and H. C. Papadopoulos, “Predicting wireless channel features using neural networks,” in *Proc. IEEE ICC*, May 2018, pp. 1–6.
[31] J. Joo, M. C. Park, D. S. Han, and V. Pejovic, “Deep learning-based channel prediction in realistic vehicular communications,” *IEEE Access*, vol. 7, pp. 27846–27858, 2019.
[32] C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, “Channel state information prediction for 5G wireless communications: A deep learning approach,” *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 1, pp. 227–236, Jan.–Mar. 2020.
[33] Y. Sui, W. Yu, and Q. Luo, “Jointly optimized extreme learning machine for short-term prediction of fading channel,” *IEEE Access*, vol. 6, pp. 49029–49039, 2018.
[34] Y. Zhao, H. Gao, N. C. Beaulieu, Z. Chen, and H. Ji, “Echo state network for fast channel prediction in Ricean fading scenarios,” *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 672–675, Mar. 2017.
[35] M. Zorzi, R. R. Rao, and L. B. Milstein, “ARQ error control for fading mobile radio channels,” *IEEE Trans. Veh. Technol.*, vol. 46, no. 2, pp. 445–455, May 1997.
[36] S. L. Cotton, “Second-order statistics of $\kappa - -\mu$ shadowed fading channels,” *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8715–8720, Oct. 2016.
[37] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, Jan. 2006.
[38] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. New York, NY, USA: Dover, 1965.
[39] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[40] *IEEE Standard for Synchrophasor Data Transfer for Power Systems*, IEEE Standard C37.118.2-2011, pp. 1–53, Dec. 2011.

Sharda Tripathi received the B.Tech. degree in electronics and communication engineering from Rajiv Gandhi Technical University, Bhopal, India, in 2007, the M.Tech. degree in digital communication engineering from the Department of Electronics and Telecommunication Engineering, Maulana Azad National Institute of Technology, Bhopal, in 2011, and the Ph.D. degree in electrical engineering from the Indian Institute of Technology Delhi, New Delhi, India, in 2019.

She is currently a Postdoctoral Research Fellow with the Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy. Her current research interests include cross-layer data-driven resource optimization strategies for IoT communications and applications, stochastic analysis and performance modeling of dynamic communication over wireless channels, queueing theoretic modeling of communication networks, and application of machine learning for predictive modeling of nonstationary processes.

Swades De (Senior Member, IEEE) received the B.Tech. degree in radio physics and electronics from the University of Calcutta, Kolkata, India, in 1993, the M.Tech. degree in optoelectronics and optical communication from the Indian Institute of Technology (IIT) Delhi, New Delhi, India, in 1998, and the Ph.D. degree in electrical engineering from the State University of New York at Buffalo, Buffalo, NY, USA, in 2004.

From 2004 to 2007, he was a Tenure-Track Assistant Professor of electrical and computer engineering with the New Jersey Institute of Technology, Newark, NJ, USA. He is currently a Professor with the Department of Electrical Engineering, IIT Delhi. He worked as an ERCIM Postdoctoral Researcher with ISTI-CNR, Pisa, Italy, in 2004, and has nearly five years of industry experience in India on telecom hardware and software development from 1993 to 1997 and in 1999. His research interests are broadly in communication networks, with emphasis on performance modeling and analysis. Current directions include energy harvesting sensor networks, broadband wireless access and routing, cognitive/white-space access networks, smart grid networks, and IoT communications.

Prof. De currently serves as an Area Editor for IEEE COMMUNICATIONS LETTERS and *Computer Communications* (Elsevier), and an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE NETWORKING LETTERS, and *IETE Technical Review Journal*.