

A Machine-Learning Approach for the Exemplar Extraction of mmWave Industrial Wireless Channels

Mohamed Kashef, Peter Vouras, Robert Jones, Richard Candell, and Kate A. Remley

Abstract– Industrial wireless channel modeling is essential for the development of industrial internet of things (IIoT) wireless systems. Moreover, millimeter-wave (mmWave) wireless bands have a high potential to be used for IIoT applications because of their high data-rates and the better applicability of having multiple antennas. As a result, we introduce an exemplar extraction approach to be applied on mmWave wireless channel measurements. A machine learning (ML) clustering scheme is used to divide the measured power-angle-delay-profiles into a number of groups with respect to the angle of arrival. Each of the groups is represented by a power-delay-profile (PDP) exemplar to provide a tractable way for testing and evaluation of mmWave IIoT wireless systems through compactly representing different groups based on their spatial characteristics. Hence, testing of wireless communications equipment can be performed over the exemplars to assess their spatial performance with a significantly reduced amount of data, allowing the development of lab-based device evaluation in a realistic, yet repeatable, test environment. Governing equations are provided in sufficient detail for users to implement the technique in their own labs.

Index Terms– Channel modeling, mmWave channels, clustering, exemplar channel, unsupervised learning, wireless systems

I. INTRODUCTION

IN future industrial internet of things (IIoT) systems, wireless-communication technologies are crucial in achieving the envisioned massive connectivity between various system components. Industrial physical environments are different than office and home environments which leads to different wireless channel characteristics [1], [2]. Hence, indoor industrial wireless channel models are being studied. In [3], four generic types are considered, namely, sparse and dense clutter environments with high and low base station positions. A survey of various 5G channel models for IIoT scenarios can be found in [4]. However, various industrial environments differ from each other. Hence, the assessment of IIoT wireless systems requires characterizing the channels of various environments [5]. This contribution discusses an approach for conducting three-dimensional reflective-channel measurements and extracting key propagation-channel features. These features can then be used to either develop simplified channel models or they can be replicated directly to test IIoT devices in controlled environments.

Compared to the sub-6 GHz wireless spectrum, the millimeter-wave (mmWave) bands offer larger available bandwidth and higher directivity for the same aperture size which enables improved received signal strength and spectral efficiency. As a result, the use of mmWave bands is considered for many new technologies. Recent measurement campaigns have studied the propagation channel at mmWave frequencies for both indoor and outdoor channels. For a sampling of the literature, see [6]–[16]. The short wavelength of mmWave allows using a large number of antennas in wireless devices and hence more spatial diversity can be achieved. Moreover, the mmWave signals are impacted by stronger atmospheric absorption and lower sensitivity by the smaller antennas and

hence its transmitted energy needs to be better managed. As a result, the spatial performance of various equipment in mmWave bands needs to be studied and wireless channels need to be modeled.

Many measurement campaigns have been performed for various industrial environments to understand the wireless communications behavior in different environments, especially in industrial environments where the density of metallic objects can be higher and various objects within the environment can be moving more frequently compared to home and office environments. Examples of these campaigns and their findings can be found in [17]–[26] and the references therein. The National Institute of Standards and Technology (NIST) conducted mmWave measurements in the highly reflective Central Utility Plant (CUP) at the Department of Commerce Boulder Laboratories in 2019. The system used by NIST consists of a vector-network-analyzer-based synthetic aperture system designed to capture spatial channel characteristics in the mmWave bands. The system is described in more detail below, and seminal work is presented in [16], [27]. Generally, the power delay profile (PDP) of a wireless channel captures the temporal characteristics of the channel due to multiple reflected signals, or multipath components (MPCs), arriving at the receiver [28]. In mmWave bands, directional PDPs capture the spatio-temporal variations of the channel.

In the literature, the problem of the classification and clustering of wireless channels deploying machine learning (ML) approaches has been investigated as in [29]–[33]. In these papers, both supervised and unsupervised learning were used for scenario identification where wireless channel scenario refers to a specific propagation environment such as the urban macrocell, urban satellite, indoor hotspot, etc, while each propagation environment can be further classified into line-of-sight (LoS) and non-line-of-sight (NLoS) propagation scenarios. The wireless channel characteristics of each scenario usually differ from the others dramatically. In this work, we are the first to study the spatial characteristics of mmWave channels through obtaining exemplars of various directional channel

M. Kashef, P. Vouras, and R. Candell are with Communications Technology Laboratory, NIST, Gaithersburg, Maryland, United States, email: {mohamed.kashef, peter.vouras, richard.candell}@nist.gov.

R. Jones and K. A. Remley are with Communications Technology Laboratory, NIST, Boulder, Colorado, United States, email: {robert.jones, kate.remley}@nist.gov.

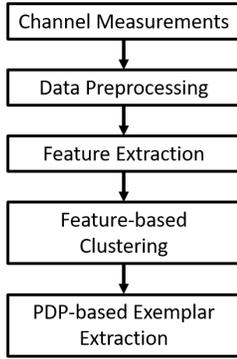


Fig. 1. An overview of the proposed approach which was initially introduced in [34]. An extended and improved version of it is described in this work.

groups that are obtained using a ML-based clustering. The contribution of the paper emphasizing the benefits of applying the proposed approach by presenting the results over real industrial channel measurements. We elaborate more on the related work in Section II.

An overview of the proposed approach is shown in Fig. 1. Generally, the channel’s power-angle-delay profile (PADP) is measured to characterize both the angle-of-arrival and time-of-arrival of received power. In this work, the measured PADPs are used as the input for the proposed approach. The result of the proposed approach allows the assessment of wireless systems over the extracted exemplars that represent the channels spatial characteristics. The main advantage of this approach is the ability of testing IIoT systems performance without exhaustively testing over all wireless-channel paths or evaluating the performance over a generic model that may not express an environment specifics. In this work, we chose a very challenging propagation scenario inside a utility plant that generates dense multipath reflections. Unsupervised ML was used to partition the measured PDPs into clusters that correspond to different directions and to extract canonical PDPs that embody salient features of each cluster.

In [34], a directional PDP exemplar extraction approach deploying unsupervised spectral clustering for PDP clustering is introduced. Three basic features was deployed, namely, Mean delay, root-mean-squared (RMS) delay spread, and total power. In this paper, an extended and improved version of the approach is introduced that allows for an increased number of PDP features to be used. The improved approach allows for extracting multiple exemplars in a single cluster based on the directional PDP temporal features. The generic channel extraction method allows the use of any clustering scheme.

Our paper is organized as follows: in Section II, we present a brief discussion of related work in the literature. In Section III, we describe the measured data, the data preparation stage, and the formal definition of the problem. In Section IV, we present the proposed approach including the feature definitions, the unsupervised learning clustering and the exemplar extraction phases. We then present the results in Section V, followed by conclusions and future direction in Section VI.

II. INDUSTRIAL ENVIRONMENT: RELATED WORK

In the context of fifth-generation (5G) networks, the mmWave frequency band (30 to 300 GHz) has emerged as a promising candidate for multi-Gb/s wireless connectivity due to availability of large bandwidth chunks [35], [36]. Since radio-interface diversity is important from a reliability perspective, using mmWave bands would certainly benefit IIoT applications as mmWave signal propagation is highly likely to be uncorrelated to sub-6 GHz signal propagation due to its differing propagation characteristics [4]. However, in [37], it was shown that using mmWave communications for Tactile Internet and similar high reliability applications is not straightforward. Specifically, the use of narrow-beamwidth directional antennas, which is necessary to combat high pathloss at mmWave frequencies [38], can result in link outages due to antenna misalignment. Thus, providing sustained high data-rate applications at mmWave frequencies is not a straightforward task. Also, testing and simulation need to be performed over the correct channel models for industrial environments.

Multiple works have discussed the deployments of mmWave in ultra-reliable low latency communications (URLLC) scenarios such as industrial applications. In [39], various challenges of achieving URLLC in mmWave frequency bands are highlighted. In [40], it is shown that cooperative networking through optimizing traffic allocations between microwave and mmWave cells can significantly improve the latency performance of mmWave-based heterogeneous networks. In [41], two strategies, namely, traffic dispersion and networking densification are proposed to reduce the end-to-end latency in mmWave wireless networks. In [42], the feasibility of using mmWave access for URLLC considering dynamic blockages is considered. It is shown that the optimal BS deployment is driven by reliability and latency constraints instead of coverage and rate requirements.

The reliability of the wireless service is mainly affected by the multipath fading in industrial environments [43]. Such fading effects are caused by the overall distribution of the various scatterers in the environment including the reflecting metal surfaces found in the specific environment under analysis in this work, which result in correlated temporal variations in the received signals in industrial wireless channels [44]. These correlated variations can be captured through studying both the envelope variations stochastic model and the time-varying channel impulse response (CIR) or alternatively, the PDP depending on the importance of the phase information for a certain study. Many works have argued that the fading distribution still follows a Rician distribution even with the moving scatterers in the environment [45]–[47]. However, this is only true with a large number of moving scatterers which is often not valid in field measurements [43]. On the other hand, obtaining an average CIR to model the correlated temporal variations cannot be performed over all the time-varying CIRs because of the different characteristics of these channels over time. However, these CIRs can be grouped, if possible, in order to obtain a CIR representation of each of these groups to model the correlated fading of the IIoT wireless channels.

Many measurement campaigns have been performed in

various industrial environments to capture the characteristics of industrial wireless channels at RF frequencies between 700 MHz and 5.8 GHz. Examples of these campaigns and their findings can be found in [17]–[26] and the references therein. The National Institute of Standards and Technology (NIST) conducted RF propagation measurements at three selected sites of different classes of industrial environments. The CIRs for various measurement points were collected and used to obtain various metrics such as the path loss, delay spread, and K factor for various industrial wireless settings [17]. Also, the propagation properties in an automobile welding factory were analyzed using measurement data and the path-loss exponent values for line of sight (LOS) and non-LOS (NLOS) channels were calculated [19]. The K-factor and delay spread were also calculated to characterize the multipath impacts. Narrow-band propagation measurements performed in five factories led to the conclusion that the path loss is log-normally distributed with Rician fading in the measured environments [20]. Similar parameters were calculated in four indoor environments including a large industrial hall at 1.9 GHz. The cross-correlation between these parameters was analyzed, where positive cross-correlation between the shadowing and delay spread was observed [22]. The propagation in industrial settings for 900 MHz was considered in [23], where it was found that environments with heavy clutter have the highest path losses. Moreover, the indoor radio propagation was studied in a representative factory automation cell where industrial robots are controlled [24]. During the measurements, the robots were in motion and executed a typical pick-and-place process. From the recorded data, the multipath components were detected to reconstruct the power delay profile, which was used for delay analysis. The obtained delay values provide input for latency-optimized design of the transmitted waveform. In [25], two industrial environments were considered which are highly absorbent and highly reflective, with respect to radio wave propagation. The results show that different degradation sources exist in various industrial scenarios and hence, wireless solutions with different fundamental properties to combat degradation sources must be chosen for each of these environments to ensure high reliability. Similar criteria were considered in [21] in addition to measuring the CIR at four representative locations for the cases of LOS and NLOS with heavy or light surrounding clutter. The power delay profile of industrial wireless measurements in a factory hall has been modeled to follow a generalization of Saleh-Valenzuela model [26]. Moreover, for non-line-of-sight scenarios at larger distances, several hundred multipath components are collected to capture 50% of the available energy.

In conclusion, the average CIR or the PDP are only evaluated in a few works in the literature where they are generally calculated at microwave frequencies and for a stationary setting of the industrial environment. As a result, our work, recognizably, performs channel representation to study directional impact on mmWave channels which is crucial to the use of mmWave bands in advanced communications systems and get the best benefits from spatial diversity in these systems. Also, our work deploys advances in ML to perform channel modeling such that a large number of features can be used in

characterizing the mmWave directional links. In the next section, we describe the mmWave synthetic aperture measurement system used to collect data in a reflective industrial setting.

III. MEASUREMENTS AND DATA PREPARATION

In order to illustrate the method, we collected data in a highly reflective, 3D spatial channel in an industrial environment. In this section, we describe the synthetic-aperture measurement system, the industrial environment, and the data collection and preparation schemes.

A. Synthetic Aperture Measurement System

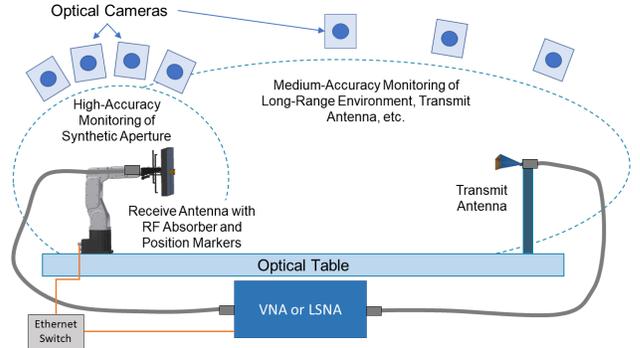


Fig. 2. Block diagram of the synthetic aperture system used in this work.

Synthetic-aperture systems, along with the associated post-processing algorithms are capable of making high-resolution measurements of multipath in static wireless communication channels [8], [48], [49]. An estimate of the wireless channel impulse response derived from a synthetic-aperture system provides information on the source of signal echoes caused by reflections, the extent of random multipath caused by diffuse scattering and diffraction, and the amount of shadowing effects or signal blocking created by objects in the physical scene.

In a synthetic aperture system used for sounding communication channels, the delay and spatial characteristics of the channel are extracted in post-processing from a sequence of digitized measurements made using a single receiver connected to a receive (probe) antenna. The probe is moved using a precise positioner to different locations in space and at each location the receiver acquires samples of the electromagnetic fields propagating across the observation plane of the aperture.

The NIST synthetic aperture system implementation includes a fixed transmit antenna, a mechanical positioner that moves the receive antenna, and a vector network analyzer (VNA) to transmit and receive RF waveforms. The system, deployed in a laboratory environment, is illustrated in the block diagram of Fig. 2. The positioner consists of a robotic arm that allows for 3D synthetic-aperture scans with different geometries and polarizations. The probe antenna is mounted onto the tip of the robotic arm. The position accuracy of the robotic arm is below 100 μm , as determined by an optical camera system that estimates the actual positions of the robot. The positioning repeatability of the probe antenna as it moves along the specified spatial sampling lattice is nominally 5 μm [50]. See [48] for more detail on the hardware, and [16] for verification of the angular resolution of the system in a laboratory environment.

B. Environment and Data Collection



Fig. 3. Data collection in the NIST Central Utility Plant (CUP). The receive antenna is shown at lower right. It is located in front of a piece of RF absorber and mounted to a robot arm that scans to form the synthetic aperture array. The same measured data was used in assessing the initial version of the proposed approach in [34].

Electromagnetic environments in industrial and factory settings are characterized by strong specular and persistent diffuse multipath from a variety of sources. Establishing a reliable wireless link in industrial locations is difficult and requires accurately characterizing the propagation environment via channel sounding. To recreate the challenging scattering environment of a factory floor, data was collected in the CUP at the Department of Commerce laboratories in Boulder, CO. The CUP has many large metal structures including steam pipes, boiler tanks, and equipment control racks as visible in Fig. 3. To create a synthetic aperture array for sounding the wireless channel, a VNA was used to transmit and receive the sounding signal which consists of a sequence of narrowband sinusoidal tones between 26.5 and 40 GHz in 10 MHz increments. The VNA was set on a small rack placed behind the receive antenna in an unobtrusive location. A WR-28 receive antenna horn operating in the band from 26.5 to 40 GHz was mounted on a robotic arm which moved to precise signal sampling locations in space so as to create the synthetic aperture. The VNA was configured for a dynamic range of approximately 90 dB.

The WR-28 transmit antenna horn was pointed directly towards a metal tank and a control panel as depicted in Fig. 4. The WR-28 horn antenna is linearly polarized with 17 dBi of gain and has a $23^\circ/24^\circ$ one-sided 3 dB beamwidth in the E/H planes. The transmit horn was mounted on a small floor stand and oriented with an elevation angle of approximately 15° . The receive horn used to generate the synthetic aperture was also pointed towards the control panel in a bistatic configuration and with an orientation that reduced the line-of-sight coupling between transmit and receive antennas as much as possible. The spatial sample points of the synthetic aperture [51] were located on a 35-by-35 planar grid with 3.75mm spacing that corresponds to $\lambda/2$ at 40 GHz. The minimum two-sided beamwidth of the beamformed output of the synthetic aperture array is 2.9° when measured in the boresight direction at 40 GHz. Note that because the synthetic array is square, the mainbeam also has a square shape and the azimuth and elevation beamwidths are equal. In the post-processing, directional PDPs were computed that provide

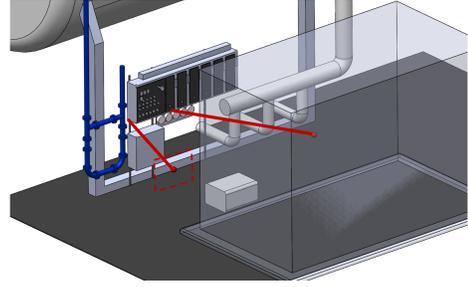


Fig. 4. CAD drawing of the CUP. Red lines indicate boresight directions for the transmit and receive antennas. The transparent rectangular prism corresponds to the boiler that is opposite the control panel. The dotted red rectangle denotes the planar synthetic aperture. The white rectangular prism on the floor represents the VNA.

received signal power as a function of delay for specified beam pointing directions in space. The PDPs can be used to compute channel propagation statistics, such as RMS delay spread, as a function of angle.

Channel sounding measurements were performed inside the CUP at the NIST Boulder campus. This location contains heavy industrial machinery, including metal piping, steam tanks, and ancillary equipment, that creates a very dense multipath scattering environment. The result is that signal propagation conditions closely resemble the wireless environment in an automated industrial setting, such as an automobile factory, where there are robots, tools, and metal parts clustered together. The measured data shows long fading time constants for diffuse multipath energy with strong discrete multipath components interspersed throughout.

C. Data Preparation and Resulting Data

During the experiment, the VNA measured S_{21} parameters in 10 MHz increments between 26.5 and 40 GHz at every spatial location in the planar array sampling lattice. Using true-time delay beamforming, the measured S_{21} parameters were coherently processed to steer the array main beam and to generate directional PDPs. The details of data processing and the corresponding uncertainties are described in [48], [51]. A frequency-invariant taper was computed and applied across the aperture in post-processing to reduce the sidelobe levels of the array response in the boresight direction. For other desired beam-steering directions, an additional linear phase taper was also applied across the aperture. After coherently combining the product of measured S_{21} values and the complex taper weights applied across all the spatial samples of the synthetic aperture, an inverse Fourier transform was used to generate each directional PDP by transforming the frequency domain data to the temporal domain. Note that beam-steering directions were chosen systematically based on an algorithm in [52] such that all the beams overlap at the 3-dB beamwidth. This rigorous approach accounts for the fact that the width of a scanned beam increases in proportion to the product of the cosines of the azimuth and elevation angles.

IV. CLUSTERING AND EXEMPLAR EXTRACTION

We denote the PADD instants by $h(\theta, \phi, \tau)$, where θ and ϕ denote the azimuth and elevation of the angle of arrival, respectively, and τ is the delay. The collected data are sampled versions of the PADD, where $\theta \in \Theta$ and $\phi \in \Phi$ such that Θ and Φ are the sets that contain the discrete values of θ and ϕ on the measurement grid. The set, \mathcal{H} , of the PDPs $h(\theta, \phi, \tau)$ for all fixed combinations of θ and ϕ is the input for the approach. The output is N disjoint groups $\mathcal{H}_i, i \in \{1, N\}$ using unsupervised ML clustering based on a set of features. Each group is represented by an exemplar PDP denoted by $\hat{h}_i(\tau)$, which represents the corresponding group \mathcal{H}_i .

A. Feature Extraction

In this subsection, we describe the various types of features that can represent the characteristics of mmWave directional PDPs. In this work, we introduce four different types of considered features, namely, PDP-based, discrete MPC-based, frequency-correlation-based, and diffuse multipath-based features. Moreover, we give examples of the features used within each type and their extraction methodology.

The PADD instant for a certain pair of θ and ϕ is defined as

$$h(\theta, \phi, \tau) = \sum_{l=0}^{L-1} \alpha_l \delta(\tau - \tau_l), \quad \theta \in \Theta, \phi \in \Phi, \quad (1)$$

where α_l is the power gain for the l -th path, τ_l is the path arrival time, L is the number of the arrival paths, and $\delta(\tau)$ is the Dirac function.

In the rest of this subsection, we describe the different types of features and their importance. We also give examples of the features within each group that will be used in the results. However, these groups include more features than the defined ones and can be used either combined or separately depending on the clustering needs for each specific case.

1) *PDP-based Features*: This set of features characterizes the complete characteristics for a directional PDP including all the MPCs after preprocessing. This may include the total power, the mean delay, and the RMS delay spread. All of the features are evaluated for a single PDP at a certain pair of θ and ϕ . We will drop the argument to simplify the expressions. The total power of the PDP, G , is evaluated as

$$G = \sum_{l=0}^{L-1} \alpha_l. \quad (2)$$

The mean delay is the first moment of the power delay profile and is evaluated as follows

$$\tau_{\text{mean}} = \frac{1}{G} \sum_{l=0}^{L-1} \alpha_l \tau_l. \quad (3)$$

The RMS delay spread is the second moment of the power delay profile and is evaluated as follows

$$\tau_{\text{RMS}} = \sqrt{\frac{1}{G} \left(\sum_{l=0}^{L-1} \alpha_l \tau_l^2 - \tau_{\text{mean}}^2 \right)}. \quad (4)$$

The estimated noise level for a PDP is the average power level of the first received L_n samples that does not include any of the transmitted signal. The value of L_n corresponds to the number of samples within the time-of-flight interval before the transmitted signal has arrived at the receive antenna. The estimated noise level for a PDP is evaluated as follows

$$N_{\text{est}} = \frac{1}{L_n} \sum_{l=0}^{L_n-1} \alpha_l. \quad (5)$$

2) *Discrete-MPC-based Features*: The discrete MPCs are defined as the peaks of a directional PDP which mainly consist of the high-power components corresponding to the main reflectors in an environment. The discrete MPCs, in many cases, carry most of the power of the PDP and are the ones that are detected by the receiver, depending on its sensitivity threshold. Hence, the discrete-MPC-based features often represent the behavior of a directional channel with respect to its strongest multipath.

In order to obtain the discrete MPCs, an adaptive threshold peak detection scheme is deployed, where the output set of delay indices are included in the set \mathcal{D} . The peak detection technique used in this work follows the same idea as a constant false alarm rate (CFAR) detector [53], [54]. In this technique, a PDP sample is compared to a scaled version of the maximum of the averages of both the leading and lagging time windows on the PDP samples. A PDP sample is considered a discrete MPC whenever it exceeds the corresponding threshold. In our work, we used a window size of 20 delay indices for both the lagging and leading time windows and we used a scaling factor of 4.15 dB to be multiplied by the maximum value of the window averages.

The number of samples in the leading and lagging time windows determines how quickly the adaptive threshold will react to nonstationary environments with abrupt transitions in the ambient energy level. The background energy level includes thermal noise and diffuse multipath. Long leading windows imply that the detection threshold will start increasing far in advance of any actual changes to the background energy. Likewise, long lagging windows imply that the detection threshold does not quickly ‘forget’ previous increases in the ambient energy and drops slowly to the correct value even after a large bump in background energy level has dissipated. In our scenario, we determined that 20 samples (which corresponds to 1.5 nsec or equivalently 1.5 feet) provided a well-tuned trade-off between responsiveness to the nonstationary environment and the accuracy of the background energy level estimated using the sample average of the leading and lagging time windows.

The 4.15 dB scale factor determines the height of the detection threshold above the background energy level which includes thermal noise and diffuse multipath scattering. The precise threshold value can be set rigorously by considering the Neyman-Pearson (NP) criteria for maximizing the probability of detection given a fixed probability of false alarm. While it is possible to apply the NP criteria to channel sounding data and compute optimal detection thresholds based on nominal probability density functions, such as the Rician, in our case it was necessary to tune parameters more closely

to the local reality rather than to the theoretical derivation. Therefore, we searched exhaustively through many different values of the scale factor and found that 4.15 dB set the detection threshold such that an inclusive set of strong specular multipath components was detected without excessive false alarms due to diffuse multipath or thermal noise.

The reliable detection of discrete MPCs is important since the number, delay, and power of discrete MPCs is an input parameter in the calculation of channel statistics such as RMS delay spread. These statistics are sensitive to the power level and clustering of discrete MPCs which makes it imperative that diffuse multipath energy is consistently excluded. Therefore, the heuristic experiments used to determine the correct scaling factor consisted of minimizing how much diffuse multipath energy was detected and misclassified as due to a discrete MPC. The value of 4.15 dB for the scale factor reliably detected the largest discrete MPCs while neglecting almost all the diffuse multipath samples.

Hence, we can define the total power of the discrete MPCs as follows

$$G_d = \sum_{k \in \mathcal{D}} \alpha_k. \quad (6)$$

The mean delay of the discrete MPCs is the first moment of the delays and is evaluated as follows

$$\tau_{\text{mean,d}} = \frac{1}{G_d} \sum_{k \in \mathcal{D}} \alpha_k \tau_k. \quad (7)$$

The RMS delay spread of the discrete MPCs is the second moment of the delays and is evaluated as follows

$$\tau_{\text{RMS,d}} = \sqrt{\frac{1}{G_d} \left(\sum_{k \in \mathcal{D}} \alpha_k \tau_k^2 - \tau_{\text{mean,d}}^2 \right)}. \quad (8)$$

Another important group of features is the discrete MPCs signal-to-noise ratio (SNR) features. Various discrete-MPC-based SNR values may include the maximum, the minimum, the average, and the dynamic range of SNR of discrete MPCs. In order to evaluate this group of features, we first evaluate the set of SNR_d , that includes the SNR values for each individual MPC within the set \mathcal{D} as follows

$$SNR_d(k) = \frac{\alpha_k}{N_{\text{est}}}, \quad k \in \mathcal{D}. \quad (9)$$

In the results presented here, we consider two SNR features: the average SNR and the SNR dynamic range of the discrete MPCs. The average SNR is evaluated as follows

$$SNR_{\text{mean,d}} = \frac{1}{|\mathcal{D}|} \left(\sum_{k \in \mathcal{D}} SNR_d(k) \right), \quad (10)$$

where $|\cdot|$ is the total number of elements in the set.

Also, the SNR dynamic range is defined as follows

$$SNR_{\text{range,d}} = \frac{\max(SNR_d(k))}{\min(SNR_d(k))}, \quad k \in \mathcal{D}. \quad (11)$$

3) *Frequency-Correlation-based Features*: The frequency correlation describes the similarity between frequency components in the frequency response. The coherence bandwidth feature represents the bandwidth over which the channel can be considered "frequency flat." Beyond this bandwidth, time-varying signal distortion may occur and hence a simple increase in SNR cannot decrease the error probability.

This group of features is defined by the single-sided frequency offset in the frequency domain autocorrelation function that corresponds to a drop of x dB from the peak. We start by evaluating the magnitude of the complex autocorrelation function which is then normalized with respect to the peak value to compare the resulting features for all the directional PDPs. In this work, we will consider a 3 dB drop in the autocorrelation function as defining the coherence bandwidth. We will further calculate the autocorrelation using two different frequency ranges, namely, 26.5 to 40 GHz and 28 to 29 GHz. The narrower frequency range better represents the characteristics of a typical communications channel.

4) *Diffuse-multipath-based features*: This set of features characterizes the amount and type of the scattered power in a channel. Diffuse multipath may affect device performance by raising the effective noise floor, depending on receiver sensitivity and type of signals that are transmitted. These features may include the total amount of diffuse power, its SNR, and the fading slope. In order to define the set of delay indices of the diffuse components, we eliminate the discrete MPCs and the adjacent L_D samples in each of the leading and lagging side of the discrete MPCs. As a result, the set of diffuse delay indices, denoted by \mathcal{F} , is defined as $\{l : l \leq d - L_D \ \& \ l \geq d + L_D, \ \forall d \in \mathcal{D}\}$. In this work, we select $L_D = 18$ heuristically to remove the impact of discrete components on the adjacent diffuse components.

Our analysis of measured data drew a clear distinction between channel statistics computed from discrete multipath samples due to specular scattering versus the statistics of diffuse multipath samples that are akin to spatial noise. The method employed for computing diffuse multipath statistics was to excise the discrete multipath samples from the data before computing features such as diffuse multipath power, etc. However, even discrete MPCs created by specular returns have a finite duration and are not delta functions. Even though the duration of discrete MPCs will vary, we found that by setting the excise parameter to $L_D = 18$, nearly all the energy due to discrete MPCs was reliably removed from the adjoining diffuse multipath or thermal noise samples. Therefore, no samples due to discrete MPCs would be available to corrupt the channel statistics computed for diffuse multipath when $L_D = 18$.

Then, the total diffuse power is defined as follows

$$G_{\text{diff}} = \sum_{l \in \mathcal{F}} \alpha_l. \quad (12)$$

In this work, we consider the average diffuse power as a feature of the directional PDPs as follows

$$G_{\text{mean,diff}} = \frac{1}{|\mathcal{F}|} G_{\text{diff}}. \quad (13)$$

The average diffuse SNR is evaluated as follows

$$SNR_{\text{mean,diff}} = \frac{G_{\text{mean,diff}}}{N_{\text{est}}}. \quad (14)$$

Another important feature, which combines the impact of both the discrete MPCs and the diffuse power of the directional PDPs, is the K-factor. In the present work, it is defined by the ratio of the power in the discrete MPCs to the power in the diffuse multipath. In this work, it is calculated for each directional PDP as follows

$$K_{\text{mean}} = \frac{\frac{1}{|\mathcal{D}|} G_d}{G_{\text{mean,diff}}}. \quad (15)$$

Note that all power-related features can be utilized in the machine learning algorithms in their absolute values or their corresponding logarithmic values in dB. In this work, we consider all of them in dB through applying the function $10\log(*)$. The logarithmic scale is more representative of the way these features would affect a communications system.

B. Data Clustering

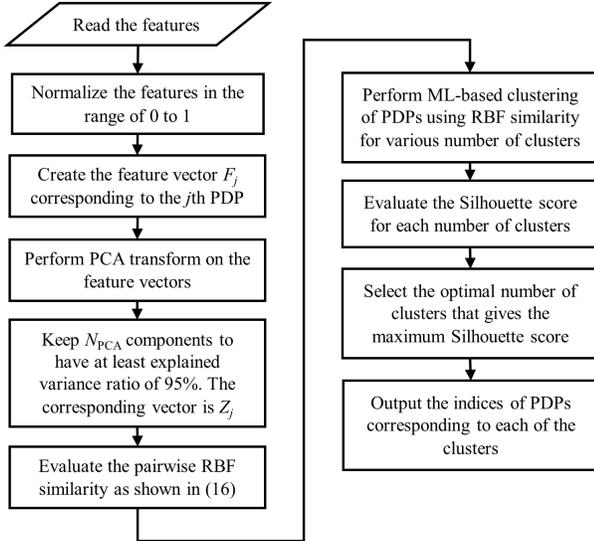


Fig. 5. Flowchart of the data clustering process.

The first step in clustering is to normalize the features to the range of 0 to 1. The proposed approach can be used with a large number of features where some or all of them can be correlated to a certain degree. Hence, we perform a principal component analysis (PCA) transform on the data projecting the feature vectors on an orthonormal space of uncorrelated principal components [55]. We select the number of the principal components to keep a certain level of the explained variance ratio which is defined to be the percentage of variance that is attributed by each of the selected principal components.

Then, we perform ML unsupervised clustering over the transformed vectors in each beam direction. We have used the Scikit-learn implementation for various clustering algorithms

[56]. The input vectors to the clustering algorithm is the transformed vectors to the selected principal components. Each of these vectors corresponds to a PDP in a specific beam direction defined by a pair of azimuth and elevation angles, by Z_j , where j is the index of the PDP.

A clustering algorithm requires a similarity metric. In this work, we use the radial basis function (RBF) similarity which is the negative exponential of Euclidean distance. The pairwise RBF similarity is evaluated as follows

$$SR_{jk} = \exp(-\gamma \|Z_j - Z_k\|^2), \quad (16)$$

where $\|Z_j - Z_k\|$ is the Euclidean distance between two vectors and γ is the negative exponential weighting factor. The RBF similarity is evaluated for all pairs of the feature vectors and the resulting values are used as input to the clustering algorithm.

Finally, the number of the clusters, if required, is obtained through a recursive search for the maximum Silhouette score [57]. The Silhouette score measures how closely-related an object is to its own cluster against the other clusters. We repeat clustering over various values of the number of clusters, N , and keep the clusters that achieve the highest Silhouette score. We present the flow of data clustering process and a summary of steps in Fig. 5.

C. Multiple Exemplar Extraction

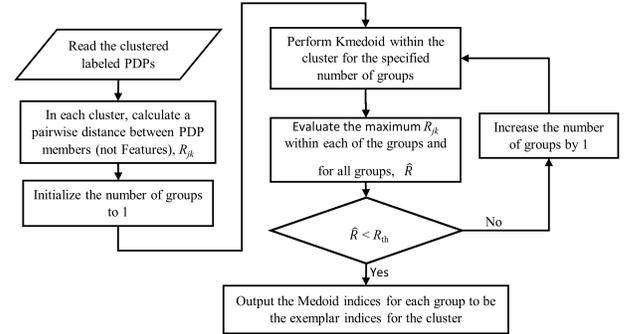


Fig. 6. Flowchart of the exemplar extraction process.

The last phase of the approach is to extract a number of exemplar PDPs from each of the clusters. These exemplars are members of their corresponding cluster of PDPs. In this phase, we consider various pairwise distance metrics between PDPs within the same cluster. This distance metric can be a simple Euclidean, Manhattan distance, or a time correlation between a pair of PDPs. Generically, we denote the distance between two PDPs of indexes j and k as R_{jk} .

The goal of the multiple exemplar extraction process is to choose the number of groups within each cluster such that $R_{jk} < R_{\text{th}}$, where R_{th} is a tuning threshold to define the size of a group of PDPs to be represented by a single exemplar. As a result, selecting R_{th} to be very large leads to have a single exemplar for each cluster, while having it too small could lead to have all clustered channel instants presented as exemplars. In this stage, we performed a second level of clustering after

the initial feature-based clustering. This second level depends on the selected distance metric between the PDP values at all the delay bins such that an exemplar represents a group of similar PDPs in both features and temporal shape.

In this step, we deploy a Kmedoids clustering algorithm [58] to obtain the index of the corresponding exemplar. The exemplar within each of the groups is defined as the PDP with shortest mean R_{jk} to other group members. We loop on the number of groups within each cluster to ensure that the minimum number of groups within each cluster maintains $R_{jk} < R_{th}$. We present the flow of exemplar extraction process and a summary of steps in Fig. 6.

Finally, we present a summary of all the steps of the proposed approach with a brief description in Table I. We describe the four major processing steps and the deployed algorithm in each of them.

TABLE I
SUMMARY OF THE PROPOSED APPROACH

Data Processing	Brief Description
Feature Extraction	We apply the feature definitions in Sec. IV-A to the directional PDP data such that we obtain a high dimensional feature vector corresponding to each of the PDPs.
Feature Processing	We process the feature vectors through normalizing their the feature values and perform PCA transform to reduce the size of the vectors.
PDP Clustering	We perform unsupervised ML-based clustering on the transformed feature vectors to produce groups of directional PDPs with similar characteristics.
Exemplar Extraction	In order to extract exemplars for each of the PDP groups, we perform ML-based Kmedoids clustering for the PDP waveforms of each PDP group to get the medoid PDPs as the exemplars.

V. RESULTS

In this section, we present the results of various stages of the exemplar extraction algorithm to illustrate the process and validate the ability of the approach towards characterizing the measured data. Moreover, we offer a comparison between various ML clustering schemes and their impact on the performance of the exemplar extraction approach. The selected features that are used in the results are listed in Table II. The index column values refer to the corresponding features in the following results.

In the following subsections, we present the results of various stages of the proposed approach which are the PCA stage, clustering and exemplar stage, the obtained exemplar PDPs, and the resulting cluster statistics. Later, we compare various clustering schemes by comparing the results of the whole approach when deploying different clustering techniques. We ran the proposed approach on a computer with Intel(R) Xeon(R) E-2186G CPU @ 3.80GHz and 64.0 GB of RAM using WinPython, and it took almost 30 seconds for this specific measured dataset to obtain the clusters and exemplar PDPs from the evaluated feature vectors.

TABLE II
LIST OF FEATURES USED IN THE PRESENT WORK

Index	Description
1	The total output power (dB)
2	The mean delay (nsec)
3	The RMS delay (nsec)
4	The estimated noise level (dB)
5	The total power (dB) of the discrete MPCs
6	The average SNR (dB)
7	The SNR dynamic range (dB)
8	The average K factor (dB)
9	The mean delay (nsec) of the discrete MPCs
10	The RMS delay (nsec) of the discretized MPCs
11	The frequency offset (GHz) for input signal 26.5 to 40 GHz
12	The frequency offset (GHz) for input signal 28 to 29 GHz
13	Number of threshold crossings above the noise threshold.
14	The estimated power of diffuse multipath (dB)
15	The estimated ratio of diffuse power to noise power (dB)

A. PCA Results

The “explained variance ratio” is the percentage of the variance that is attributed to each of the principal components. The greater the explained variance ratio is the more important this component is with respect to the clustering algorithm. In this work, we used six principal components to keep a total of 95% of the explained variance ratio.

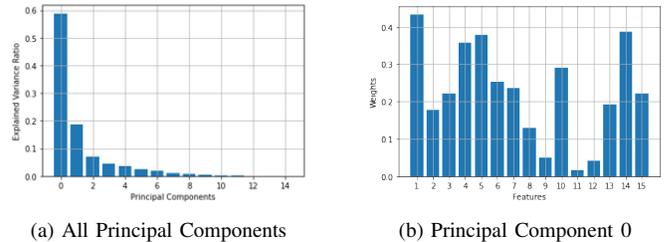


Fig. 7. Explained variance ratios for all of the obtained principal components and the explained variance ratios for all of the features listed in Table II for Principal Component 0.

In Fig. 7(a), the explained variance ratio is shown against the principal components. Moreover, the first principal component indexed by 0 contains almost 59% of the explained variance ratios. In Fig. 7(b), we show the weights of all of the features in the principal component 0 only while not showing the rest of the principal components for space limitation and because the principal component 0 contains most of the explained variance ratio in our data. Hence by analyzing the weights of the features in this principal component, we can estimate the importance of each feature on the resulting clusters. As an example, the feature 11, namely, the frequency offset (GHz) for input signal from 26.5 to 40 GHz, has a small weight and hence a low impact on the resulting clusters, whereas feature 14, the diffuse multipath power, has a strong impact.

B. Clustering and Exemplar Extraction

In this section, we show the clustering and exemplar extraction results as a function of beam direction and various

data features. Spectral clustering, as an example of graph-based clustering schemes, is used in the results obtained in this section. However, a comparison of various ML clustering types is provided in Sec. V-E.

In Fig. 8, we draw the clustering results with the exemplars marked on the figures. In this case, we found that the optimal number of clusters is five, where the Silhouette score had a maximum value of 0.34. In Fig. 9, the features of the resulting clusters are plotted for four scatter-plot sub-figures: a) mean delay (ns) against RMS delay spread (ns), b) The average K Factor (dB) against the total power of discretely (dB), c) the diffuse power to noise ratio (dB) against the total diffuse power (dB), and d) the frequency offset when input signal at 28 to 29 GHz (GHz) against the average K factor. All cluster members are shown with the same color and the corresponding exemplars are marked by a specific shape. We are showing a sample of the used features not all of the features because of the space limitation.

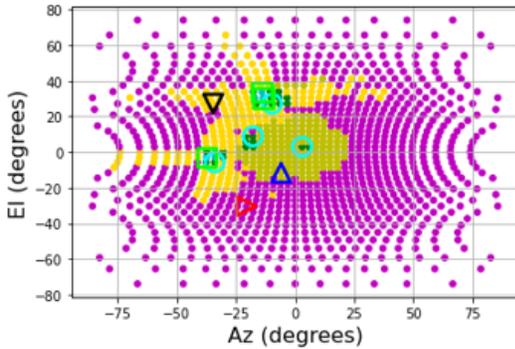


Fig. 8. Clustered data with marked exemplars against the beam directions. Each cluster members have the same color with the corresponding exemplars marked by a specific shape.

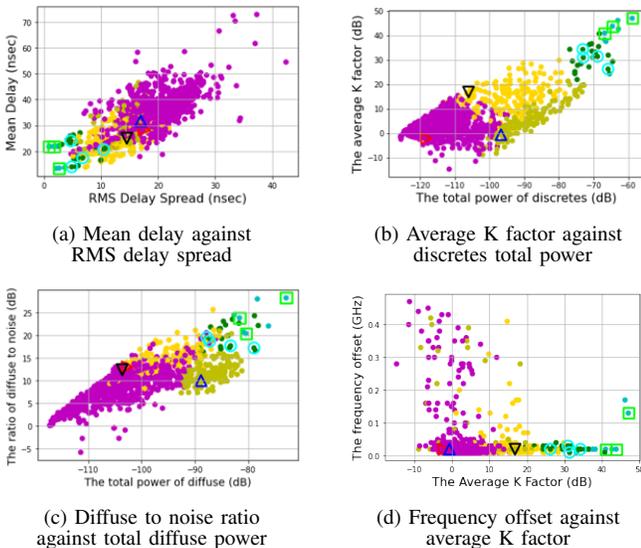


Fig. 9. Clustered data with marked exemplars against a selected set of the features that are used for clustering.

These two figures show how the exemplars represent various features. We start with the small cluster shown by dark green

data points that is characterized by these characteristics: low delay, low RMS delay spread, high power of discrete MPCs, high K factor, low frequency offset for 3 dB power drop, and high diffuse power. The cluster with cyan data points has similar attributes, although its PDPs have higher total discrete power, higher average K factor values, lower mean delay and shorter RMS delay spreads compared to the PDPs of the cluster with the dark green points. These two clusters physically represent the directions where the metallic reflectors exist in the environment and hence higher numbers of discrete MPCs exist. Moreover, although, this dark green cluster is coherent in the feature space, its corresponding data points occur in different beam directions. This is captured through the four different exemplars marked with the circles in Figs. 8 and 9.

Another cluster is the biggest cluster with magenta data points, which is characterized by high delay, high RMS delay spread, low power of discrete MPCs, low K factor, and low diffuse power. This cluster represents the directions where no reflector exists and, hence, the received power is mainly from the diffuse components of the reflected signals in other directions. The other two clusters, namely, the clusters with golden and greenish-yellow data points, are characterized by PDPs with mid-range values for various features that are not in the direction of the main reflectors but closer to these directions than the PDPs of the diffuse-multipath cluster with magenta data points. The main difference between these two clusters is that the cluster with golden data points has a higher average K factor and less diffuse power compared to the cluster with greenish-yellow data points.

C. Exemplar PDPs

In this section, we examine the resulting exemplars for one of the clusters. We compare them to the benchmark of the average PDP of the corresponding cluster data points. As an example, we show in Fig. 10 the exemplars of the dark green cluster. In this example, we used the Euclidean distance between the power of the discrete MPCs of the PDPs as the pairwise distance for the exemplar extraction stage and we set the value of R_{th} to 0.0003.

In these exemplars, we notice that the average clustered PDP captures the overall delay characteristics showing multiple peaks at different delay bins. However, realistically, these peaks are received from different directions. Hence, the exemplars extracted by our proposed approach reveal the directional channel performance. In the cluster with green data points, four exemplars are obtained which are marked with the cyan circles in Figs. 8 and 9. In Fig. 10, the four exemplars are shown where the highest discrete MPC in each exemplar PDP exists at a different delay bin and reflected from a different metallic surface in the environment. Hence, we find that although the clustered data points have similar feature attributes, they have different temporal profiles, and they should be represented with different exemplar PDPs.

The extracted exemplars capture all of the groups of the PDPs within a cluster that have similar features but different temporal profiles, and hence a wireless device or system can

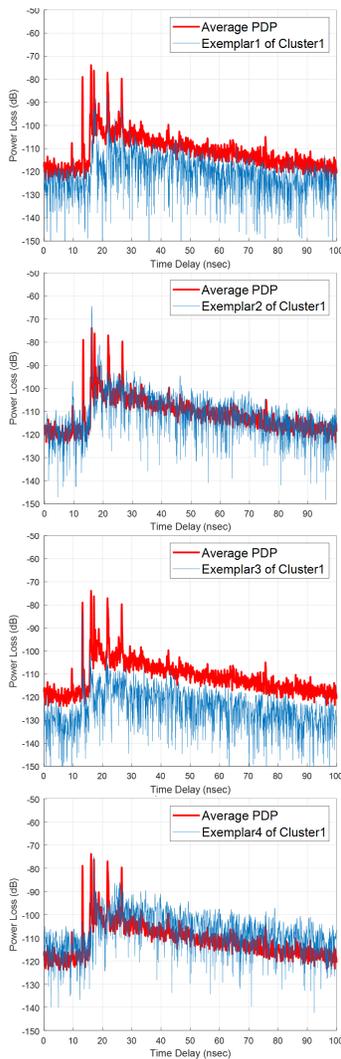


Fig. 10. Exemplar PDPs corresponding to the beam directions marked with \bigcirc in Fig. 8.

be tested efficiently utilizing these exemplar PDPs without the need to test over all measured PDPs in an environment. This testing can be either performed in a laboratory environment where only these exemplars are replicated or in the original environment while testing is performed in the direction corresponding to the exemplars, but without the need of testing in all directions.

D. Resulting Clusters Statistics

We next present probability distribution of the histograms of the resulting clusters for a selected set of features. The goal of this result is to validate the clustering process output by illustrating in Fig. 11 that clusters are distinguishable from each other through their features. However, as shown in the PCA analysis, the impact of some features can be lower compared to others. As an example, the greenish yellow and the gold clusters can be distinguished through the total power. However, the average K factor and the RMS delay spread have similar statistical behaviour of these clusters.

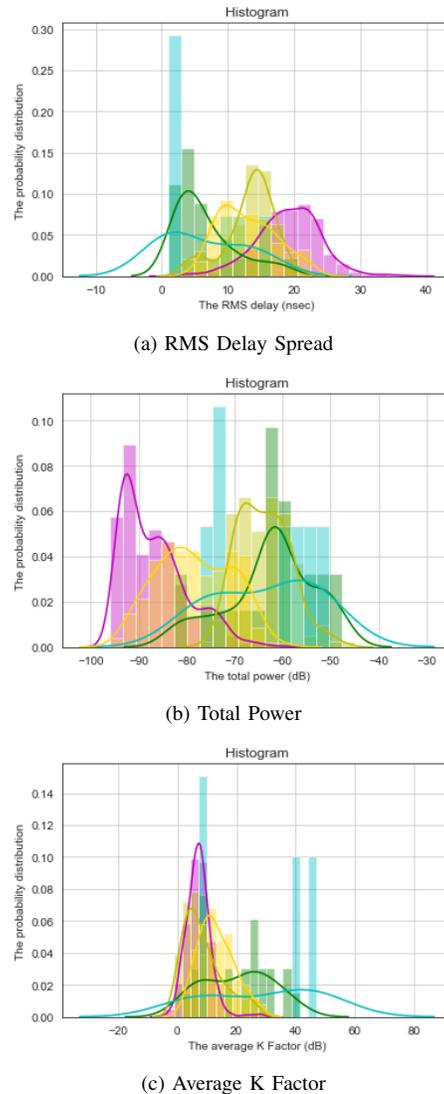


Fig. 11. The histograms of a selected set of features for the resulting clusters. The bar colors are corresponding to the colors of the clusters shown in Fig. 8.

E. Comparison of Clustering Schemes

In the previous results, we have deployed a graph-based clustering scheme, namely, spectral clustering. In this section, we present examples of clustering results for various types of clustering categories and comment on deploying them for the proposed exemplar extraction approach. We use the same set of features and the Silhouette score for the obtained number of clusters, if needed.

In Fig. 12, we show the clustering results using a Kmeans algorithm as an example for centroid-based clustering [59]. The use of centroid-based clustering results in more uniform-sized clusters based on the features used. As a result, the low-power cluster from the spectral clustering result, the magenta cluster in Fig. 8, has split into two clusters using the Kmeans algorithm. As well, the two high-power clusters, the cyan and dark green clusters in Fig. 8, are combined into one in the Kmeans algorithm. Physically, splitting the noise-like magenta cluster in Fig. 8 into two, artificially distinguishes them and

TABLE III

COMPARISON OF THE PERFORMANCE OF VARIOUS ML CLUSTERING ALGORITHMS AND THE CASE OF NO CLUSTERING WHERE THE AVERAGING IS USED TO OBTAIN A SINGLE EXEMPLAR. WE SHOW THE OPTIMAL VALUES IN **BOLD FONT**.

	Spectral Clustering	DBSCAN	BIRCH	Kmeans	Averaging
Silhouette Score	0.34	0.32	0.38	0.26	–
Mean of RMS delay spread std. (nsec)	2.78	7.23	3.5	3.29	5.58
Mean of mean delay std. (nsec)	5.05	10.22	5.57	5.83	8.19
Mean of total power std. (dB)	4.51	11.47	5.17	4.26	10.60
Mean of average K factor std. (dB)	4.75	10.15	5.86	5.15	7.54
Mean of total diffuse power std. (dB)	4.08	9.61	4.40	3.72	8.93
Mean of mean distance to exemplar	0.39	0.68	0.48	0.42	0.83
Mean of median distance to exemplar	0.39	0.67	0.46	0.40	0.83

would not represent the most efficient device test scenario. Likewise, combining the PDPs with the strong discrete MPCs may diminish the number of important test cases. Hence, using Kmeans does not add new information compared to the spectral clustering with respect to the exemplar extraction.

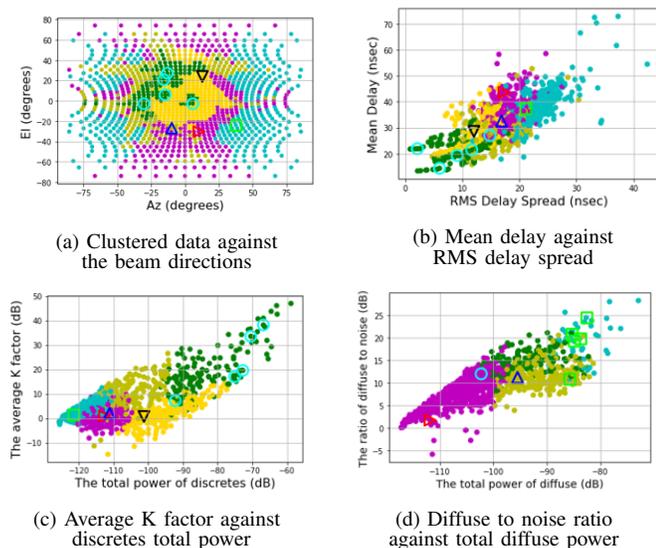


Fig. 12. Clustered data with marked exemplars using Kmeans algorithm

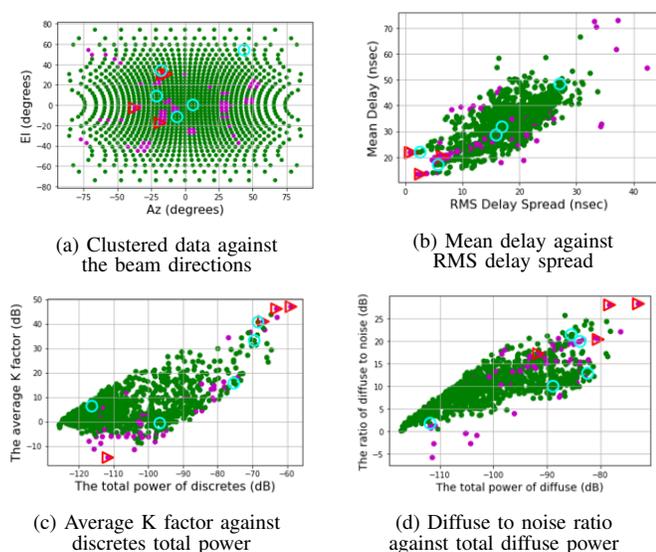


Fig. 13. Clustered data with marked exemplars using DBSCAN algorithm

In Fig. 13, we show the clustering results using a DBSCAN algorithm as an example for density-based clustering [59]. The use of density-based clustering results in clusters that are separated by low density regions in the feature space of the data points. Clearly, in this case, the data points are not well separated by low density regions, and hence, only two clusters are obtained, specifically, one for the connected data points and one for the outliers. We were able to extract few of the exemplars from the connected data points cluster. However, this type of clustering is not generally suitable for the problem of wireless channel modeling because of the nature of data being inseparable.

In Fig. 14, we show the clustering results using a BIRCH algorithm as an example for hierarchical clustering [59]. Generally, the use of an optimized hierarchical clustering algorithms should lead to good clustering results based on the distance metric we used. However, in this case, it has a very close performance to the spectral clustering case while not being able to distinguish the two high power clusters. As a result, a few exemplars are missed in this case using BIRCH clustering algorithm.

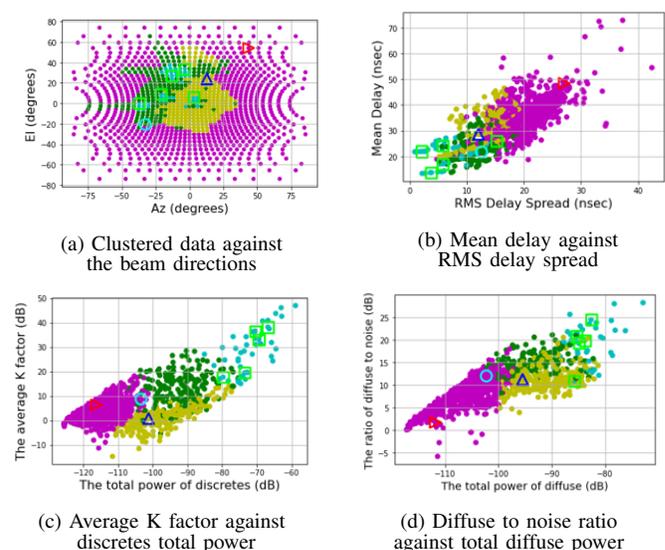


Fig. 14. Clustered data with marked exemplars using Birch algorithm

Finally, we introduce a quantitative comparison of the performance of various ML clustering algorithms in Table III. We compare the four examples of the four ML clustering categories to each other and to the the averaging PDP rep-

resentation technique where a single exemplar is obtained by evaluating the mean of all measured PDPs in all directions. In the first row in the table, the Silhouette score is an indication of the clustering quality with respect to the used features. We notice that Spectral Clustering, DBSCAN, and Birch algorithms have higher scores. However, DBSCAN obtained only two clusters to achieve this score because of the data being inseparable and hence does not capture the various characteristics of the environment.

In the following five rows, we compare the mean over all clusters of the standard deviation (std.) of the various features for the elements within a cluster. In order to calculate these values for a feature, we evaluate the standard deviation for each of the clusters using only its members. Then, we get the numerical mean over all the clusters depending on the number of clusters used for each clustering algorithm. In case of the averaging approach, the corresponding value is the standard deviation of all the measured elements. By comparing these values, we notice that the use of Spectral Clustering, BIRCH, and Kmeans has lowered the mean of the standard deviation of all features significantly compared to the original data standard deviation. This indicates that the clusters capture the channel features with less variation than simple averaging would.

Finally in the last two rows, we compare the mean over all exemplars of the mean and median of feature-based distance between each exemplar to its represented elements in the measurements. In this last comparison, we notice that Spectral Clustering has a better performance compared to other clustering schemes followed by the performance BIRCH and Kmeans algorithms. As a result, the selection of a machine learning clustering algorithm in the first phase of the exemplar extraction algorithm depends on one of the above criteria or on the overall behaviour of all of them. In this specific set of measurements, the overall performance led us to select Spectral Clustering as the main clustering scheme for the exemplar extraction approach.

VI. CONCLUSION

In this paper, we have demonstrated a method to characterize the spatial properties of wireless channels in industrial environments. The proposed approach serves as a way to compactly represent various feature groups to facilitate wireless system testing in such environments. Specifically, we introduced an approach for directional PDP exemplar extraction from measured data for a static, highly reflective channel. The approach deploys unsupervised ML clustering of PDPs and uses various types of channel features for exemplar extraction. We have shown that the proposed approach achieved an average feature-based distance between the exemplars and the corresponding PDPs that is 47% of the mean distance between the average PDP to all the measured directional PDPs. This demonstrates that the extracted exemplars better represent the measured channels as compared to the common procedure of averaging of all of the directional PDPs. Extracting exemplars that represent the key features allows the test and assessment of wireless equipment over the exemplars without the need to test over all of the different instances of wireless channel paths

or to evaluate the performance over a generic model that does not capture the specifics of a certain environment.

Furthermore, the studied use case of an industrial environment with various metallic reflecting surfaces was found to have a wide range of spatial wireless channel characteristics such that a change of the mmWave directional receive antenna may lead to a totally different received signal. Hence, to operate in such an environment, a wireless node should be tested under various types of channel characteristics. Future work in this area will follow two main directions: first, more work needs to be performed to standardize spatial wireless channel models to be able to study future mmWave system performance. Second, methods need to be developed to test and assess mmWave wireless equipment over various spatial channel characteristics. This work helps in both of these directions by providing a tool to characterize wireless channels based on various temporal and spatial features and allowing test over the obtained exemplars without the need to test over all of the different instances of wireless channel directions.

Generally, the characteristics of the environment have consequences for the design of the wireless control links used in automated factory settings. The communication protocols must support ultra-reliable data transfers to mitigate any safety risks. Since 5G/6G networks in the mid to high frequency bands (28-100 GHz) will most likely continue to use multi-carrier modulation schemes such as orthogonal frequency division multiplexing (OFDM), the duration of guard intervals between transmitted symbols may have to be increased to account for the long multipath fading time constants. The analysis presented in this paper supports a rigorous and systematic approach to validate or adjust the relevant communication protocols for dense multi-user scenarios specified in new and emerging wireless standards, such as IEEE 802.11ax (WiFi 6).

ACKNOWLEDGMENT

We would like to thank Sudantha Perera for his contributions in the data measurements and the insights of deploying the results in testing environments.

DISCLAIMER

Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

REFERENCES

- [1] K. S. Low, N. Win, and M. J. Er, "Wireless Sensor Networks for Industrial Environments," 2005.
- [2] M. Cheffena, "Propagation channel characteristics of industrial wireless sensor networks [wireless corner]," *IEEE Antennas and Propagation Magazine*, vol. 58, pp. 66–73, Feb. 2016.
- [3] 3GPP, "Addition of indoor industrial channel model to Doc. 38.901 - Study on channel model for frequencies from 0.5 to 100 GHz," Tech. Rep. R1-1909807, 3rd Generation Partnership Project (3GPP), 09 2019.
- [4] T. Jiang, J. Zhang, P. Tang, L. Tian, Y. Zheng, J. Dou, H. Asplund, L. Raschkowski, R. D'Errico, and T. Jämsä, "3GPP standardized 5G channel model for IIoT scenarios: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8799–8815, 2021.

- [5] R. Candell, M. Kashef, Y. Liu, K. B. Lee, and S. Foufou, "Industrial wireless systems guidelines: Practical considerations and deployment life cycle," *IEEE Industrial Electronics Magazine*, vol. 12, pp. 6–17, Dec. 2018.
- [6] S. Ranvier, M. Kyro, K. Haneda, T. Mustonen, C. Icheln, and P. Vainikainen, "VNA-based wideband 60 GHz MIMO channel sounder with 3-D arrays," in *2009 IEEE Radio and Wireless Symposium*.
- [7] G. Santella and E. Restuccia, "Analysis of frequency domain wide-band measurements of the indoor radio channel at 1, 5.5, 10 and 18 GHz," in *Proceedings of GLOBECOM'96. 1996 IEEE Global Telecommunications Conference*, vol. 2, pp. 1162–1166 vol.2, 1996.
- [8] A. W. Mbugua, W. Fan, Y. Ji, and G. F. Pedersen, "Millimeter wave multi-user performance evaluation based on measured channels with virtual antenna array channel sounder," *IEEE Access*, vol. 6, pp. 12318–12326, 2018.
- [9] M. Kyrö, S. Ranvier, V.-M. Kolmonen, K. Haneda, and P. Vainikainen, "Long range wideband channel measurements at 81–86 GHz frequency range," in *Proceedings of the Fourth European Conference on Antennas and Propagation*, pp. 1–5, 2010.
- [10] S. Salous, V. Degli Esposti, F. Fuschini, R. S. Thomae, R. Mueller, D. Dupleich, K. Haneda, J.-M. Molina Garcia-Pardo, J. Pascual Garcia, D. P. Gaillot, S. Hur, and M. Nekovee, "Millimeter-wave propagation: Characterization and modeling toward fifth-generation systems. [wireless corner]," *IEEE Antennas and Propagation Magazine*, vol. 58, no. 6, pp. 115–127, 2016.
- [11] O. H. Koymen, A. Partyka, S. Subramanian, and J. Li, "Indoor mm-wave channel measurements: Comparative study of 2.9 GHz and 29 GHz," in *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2015.
- [12] G. R. MacCartney, T. S. Rappaport, S. Sun, and S. Deng, "Indoor office wideband millimeter-wave propagation measurements and channel models at 28 and 73 GHz for ultra-dense 5G wireless networks," *IEEE Access*, vol. 3, pp. 2388–2424, 2015.
- [13] A. Bamba, F. Mani, and R. D'Errico, "Millimeter-wave indoor channel characteristics in V and E bands," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 10, pp. 5409–5424, 2018.
- [14] D. Dupleich, R. Müller, S. Skoblikov, C. Schneider, J. Luo, G. Del Galdo, and R. Thomä, "Multi-band indoor propagation characterization by measurements from 6 to 60 GHz," in *2019 13th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, 2019.
- [15] P. Zhang, J. Li, H. Wang, H. Wang, and W. Hong, "Indoor small-scale spatiotemporal propagation characteristics at multiple millimeter-wave bands," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 12, pp. 2250–2254, 2018.
- [16] D. Guven, B. F. Jamroz, J. Chuang, C. Gentile, R. D. Horansky, K. A. Remley, D. F. Williams, J. T. Quimby, A. J. Weiss, and R. Leonhardt, "Methodology for measuring the frequency dependence of multipath channels across the millimeter-wave spectrum," *IEEE Open Journal of Antennas and Propagation*, vol. 3, pp. 461–474, 2022.
- [17] R. Candell, K. A. Remley, J. Quimby, D. Novotny, A. Curtin, P. Papazian, G. Koepke, J. Diener, and M. Kashef, "Industrial wireless systems: Radio propagation measurements." NIST Technical Note 1951, National Institute of Standards and Technology, Gaithersburg, MD, 2017.
- [18] M. Cheffena, "Propagation channel characteristics of industrial wireless sensor networks [wireless corner]," *IEEE Antennas and Propagation Magazine*, vol. 58, pp. 66–73, Feb. 2016.
- [19] K. Zhang, L. Liu, C. Tao, K. Zhang, Z. Yuan, and J. Zhang, "Wireless Channel Measurement and Modeling in Industrial Environments," *Advances in Science, Technology and Engineering Systems Journal*, vol. 3, no. 4, pp. 254–259, 2018.
- [20] T. S. Rappaport and C. D. McGillem, "UHF fading in factories," *IEEE Journal on Selected Areas in Communications*, vol. 7, pp. 40–48, Jan. 1989.
- [21] A. Miaoudakis, A. Lekkas, G. Kalivas, and S. Koubias, "Radio channel characterization in industrial environments and spread spectrum modem performance," in *2005 IEEE Conference on Emerging Technologies and Factory Automation*, vol. 1, pp. 87–93, Sep. 2005.
- [22] C. Oestges, D. Vanhoenacker-Janvier, and B. Clerckx, "Channel characterization of indoor wireless personal area networks," *IEEE Transactions on Antennas and Propagation*, vol. 54, pp. 3143–3150, Nov. 2006.
- [23] E. Tanghe, W. Joseph, L. Martens, H. Capoen, K. V. Herwegen, and W. Vantomme, "Large-scale fading in industrial environments at wireless communication frequencies," in *2007 IEEE Antennas and Propagation Society International Symposium*, pp. 3001–3004, June 2007.
- [24] B. Hofheld, D. Wieruch, L. Raschkowski, T. Wirth, C. Pallasch, W. Herfs, and C. Brecher, "Radio channel characterization at 5.85 GHz for wireless M2M communication of industrial robots," in *2016 IEEE Wireless Communications and Networking Conference*, pp. 1–7, April 2016.
- [25] J. Ferrer-Coll, P. Angskog, J. Chilo, and P. Stenumgaard, "Characterisation of highly absorbent and highly reflective radio wave propagation environments in industrial applications," *IET Communications*, vol. 6, pp. 2404–2412, Oct. 2012.
- [26] J. Karedal, S. Wyne, P. Almers, F. Tufvesson, and A. F. Molisch, "Statistical analysis of the UWB channel in an industrial environment," in *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004*, vol. 1, pp. 81–85, Sep. 2004.
- [27] A. J. Weiss, D. F. Williams, J. Quimby, R. Leonhardt, T. Choi, Z. Cheng, K. A. Remley, A. Molisch, B. Jamroz, J. Rezac, P. Vouras, and C. Zhang, "Large-signal network analysis for over-the-air test of up-converting and down-converting phased arrays," in *2019 IEEE MTT-S International Microwave Symposium (IMS)*, pp. 622–625, 2019.
- [28] B. Soret, M. C. Aguayo-torres, and J. T. Entrambasaguas, "Capacity with explicit delay guarantees for generic sources over correlated Rayleigh channel," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1901–1911, June 2010.
- [29] J. Zhang, L. Liu, Y. Fan, L. Zhuang, T. Zhou, and Z. Piao, "Wireless channel propagation scenarios identification: A perspective of machine learning," *IEEE Access*, vol. 8, pp. 47797–47806, 2020.
- [30] M. I. AlHajri, N. T. Ali, and R. M. Shubair, "Classification of indoor environments for IoT applications: A machine learning approach," *IEEE Antennas and Wireless Propagation Letters*, vol. 17, no. 12, pp. 2164–2168, 2018.
- [31] C. Huang, A. F. Molisch, R. Wang, P. Tang, R. He, and Z. Zhong, "Angular information-based NLOS/LOS identification for vehicle to vehicle MIMO system," in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2019.
- [32] H. Li, Y. Li, S. Zhou, and J. Wang, "Wireless channel feature extraction via GMM and CNN in the tomographic channel model," *Journal of Communications and Information Networks*, vol. 2, pp. 41–51, Mar. 2017.
- [33] Y. Yu, F. Liu, and S. Mao, "Fingerprint extraction and classification of wireless channels based on deep convolutional neural networks," *Neural Processing Letters*, vol. 48, pp. 1767–1775, Feb. 2018.
- [34] M. Kashef, P. Vouras, R. Jones, R. Candell, and K. A. Remley, "Temporal exemplar channels in high-multipath environments," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4545–4549, 2021.
- [35] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, 2014.
- [36] K. Chandra, R. V. Prasad, B. Quang, and I. G. M. M. Niemegeers, "CogCell: cognitive interplay between 60 GHz picocells and 2.4/5 GHz hotspots in the 5G era," *IEEE Communications Magazine*, vol. 53, no. 7, pp. 118–125, 2015.
- [37] K. C. Joshi, S. Niknam, R. V. Prasad, and B. Natarajan, "Analyzing the tradeoffs in using millimeter wave directional links for high data-rate tactile internet applications," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1924–1932, 2020.
- [38] W. Roh, J. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, 2014.
- [39] R. Ford, M. Zhang, M. Mezzavilla, S. Dutta, S. Rangan, and M. Zorzi, "Achieving ultra-low latency in 5G millimeter wave cellular networks," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 196–203, 2017.
- [40] G. Yang, M. Xiao, M. Alam, and Y. Huang, "Low-latency heterogeneous networks with millimeter-wave communications," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 124–129, 2018.
- [41] G. Yang, M. Xiao, and H. V. Poor, "Low-latency millimeter-wave communications: Traffic dispersion or network densification?," *IEEE Transactions on Communications*, vol. 66, no. 8, pp. 3526–3539, 2018.
- [42] I. K. Jain, R. Kumar, and S. S. Panwar, "The impact of mobile blockers on millimeter wave cellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 4, pp. 854–868, 2019.
- [43] Q. Zhang, Q. Zhang, W. Zhang, F. Shen, T. H. Loh, and F. Qin, "Understanding the temporal fading in wireless industrial networks: Measurements and analyses," in *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Oct. 2018.
- [44] B. Soret, M. C. Aguayo-torres, and J. T. Entrambasaguas, "Capacity with explicit delay guarantees for generic sources over correlated

Rayleigh channel,” *IEEE Transactions on Wireless Communications*, vol. 9, pp. 1901–1911, June 2010.

- [45] X. Zhao, Q. Han, X. Liang, B. Li, J. Dou, and W. Hong, “Doppler spectra for F2F radio channels with moving scatterers,” *IEEE Transactions on Antennas and Propagation*, vol. 64, pp. 4107–4112, Sep. 2016.
- [46] J. B. Andersen, J. O. Nielsen, G. F. Pedersen, G. Bauch, and G. Dietl, “Doppler spectrum from moving scatterers in a random environment,” *IEEE Transactions on Wireless Communications*, vol. 8, pp. 3270–3277, June 2009.
- [47] T. Aulin, “A modified model for the fading signal at a mobile radio channel,” *IEEE Transactions on Vehicular Technology*, vol. 28, pp. 182–203, Aug. 1979.
- [48] A. J. Weiss, J. Quimby, R. Leonhardt, B. Jamroz, D. Williams, K. Remley, P. Vouras, and A. Elsherbeni, “Configuration and control of a millimeter-wave synthetic aperture measurement system with uncertainties,” in *2020 95th ARFTG Microwave Measurement Conference (ARFTG)*, pp. 1–4, 2020.
- [49] J. Hejlselbaek, W. Fan, and G. F. Pedersen, “Ultrawideband vna based channel sounding system for centimetre and millimetre wave bands,” in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, 2016.
- [50] “MECA500 (R3) User Manual: Robot Firmware 7.0.6, Document Revision F.” Mecademic Inc, 2018.
- [51] P. Vouras, B. Jamroz, J. Quimby, A. Weiss, D. Williams, R. Leonhardt, D. Guven, R. Jones, J. Kast, and K. Remley, “Wideband synthetic aperture millimeter-wave spatial channel measurements with uncertainties,” *in review*, 2020.
- [52] G. V. T. W. M. Waters, D. P. Patel, “Optimum number of faces of a volume-scanning active array radar,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, July 1998.
- [53] V. G. Hansen and J. H. Sawyers, “Detectability loss due to “greatest of” selection in a cell-averaging cfar,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-16, no. 1, pp. 115–118, 1980.
- [54] Y.-r. Zhang, M.-g. Gao, and Y.-j. Li, “Performance analysis of typical mean-level cfar detectors in the interfering target background,” in *2014 9th IEEE Conference on Industrial Electronics and Applications*, pp. 1045–1048, 2014.
- [55] S. Meng, Y. Fu, T. Liu, and Y. Li, “Principal component analysis for clustering temporomandibular joint data,” in *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, vol. 1, pp. 422–425, 2015.
- [56] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [57] “Online, sklearn.metrics.silhouette_score, URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html.” (visited: 2020-09-23).
- [58] “Online, kmedoids Class Reference, URL: https://pyclustering.github.io/docs/0.9.0/html/d0/dd3/classpyclustering_1_1cluster_1_1kmedoids_1_1kmedoids.html.” (visited: 2021-06-03).
- [59] “Online, Scikit-learn User Guide: Clustering, URL: <https://scikit-learn.org/stable/modules/clustering.html>.” (visited: 2021-06-03).



Mohamed Kashef received the B.Sc. and M.S. degrees (with Hons.) in electronics and electrical communications engineering from the Cairo University, Cairo, Egypt, in June 2006 and August 2009, respectively, and the Ph.D. degree in electrical engineering from the University of Maryland at College Park, College Park, MD, USA, in 2013. He is currently a Research Scientist at the National Institute of Standards and Technology (NIST). His current focus is on industrial wireless systems including wireless networks deployment, channel modelling,

applying artificial intelligence for data analysis and test methods for wireless networks in industrial scenarios. His research interests also include wireless communication systems and networks, visible light communication networks, and optimization of stochastic systems.



Peter Vouras founded and currently serves as Chair of the IEEE Signal Processing Society Synthetic Aperture Technical Working Group (TWG). Officially launched in April 2020, the TWG pursues collaborative activities relevant to all aspects of synthetic aperture applications such as radar, channel sounding, sonar, radiometry, medical imaging, and others. Mr. Vouras also serves as Chair of the newly created IEEE Signal Processing Society Synthetic Aperture Standards Committee (SASC). The SASC works to develop high-quality technical standards that define best practices for image formation using synthetic apertures in radar, channel sounding, and sonar. Mr. Vouras is a co-lead editor for the upcoming Special Issue on Recent Advances in Wideband Signal Processing for Classical and Quantum Synthetic Apertures to be published in the IEEE Journal on Selected Topics in Signal Processing. Mr. Vouras is also co-organizer of the Asilomar 2022 special session on Wideband Synthetic Aperture Processing and Applications.

Mr. Vouras obtained a M.S.E degree in Electrical and Computer Engineering from the Johns Hopkins University, Homewood Campus, Baltimore, MD in 2001, a B.S. degree in Electrical and Computer Engineering from George Mason University, and B.A. degrees in Economics and Foreign Affairs from the University of Virginia. He is currently with the U.S. Department of Defense and was previously with the Wireless Networks Division of the Communications Technology Laboratory at the National Institute of Standards and Technology (NIST) in Gaithersburg, MD (2016 to 2022). From 1996 to 2016, Mr. Vouras was with the Radar Division at the Naval Research Laboratory (NRL) in Washington, D.C.



Robert D. Jones received dual B.S. degrees in electrical and mechanical engineering in 2019, and the M.S. degree in electrical engineering in 2021 from the Colorado School of Mines, Golden, CO, USA, where he is currently pursuing his Ph.D. degree. Since 2017, he has been a Student Researcher with the National Institute of Standards and Technology (NIST) at Boulder, CO, USA. His current research interests include modeling of III-V transistors, on-wafer metrology, and active device design at mmwave and terahertz frequency bands.



Richard Candell has over twenty years of experience in wireless systems engineering with extensive experience in the design and evaluation of wireless communications systems. Dr. Candell spent twelve years developing, testing, and deploying secure wireless technologies for commercial and defense applications. He served as the lead systems engineer in developing spread spectrum interference cancellation and performance evaluation strategies for satellite ground stations and mobile phased array beam steering transceivers. He holds patents in successive

interference cancellation and transmission burst detection applied to spread-spectrum satellite communications signals. He holds a Ph.D. in Computer Science from the University of Burgundy, Dijon, France. He also holds a BS and MS degree in Electrical Engineering from The University of Memphis. He joined the National Institute of Standards and Technology (NIST) in the US in 2014 where he leads the Industrial Wireless Systems research laboratory. He is a member of the IEEE Industrial Electronics Society and the Robotics and Automation Society. His current research interests include the performance of mobile robotic, manufacturing, and safety applications when deployed with wireless networks as the primary mode of communications. Dr. Candell was the primary author of the Guide to Industrial Wireless Systems Deployments (NIST AMS 300-4) and he serves as the Chair of the IEEE P1451.5p Wireless Performance Assessment and Measurement Working Group and the NIST Industrial Wireless System technical interest group.



Kate A. Remley (S'92-M'99-SM'06-F'13) was born in Ann Arbor, MI. She received the Ph.D. degree in Electrical and Computer Engineering from Oregon State University, Corvallis, in 1999.

From 1983 to 1992, she was a Broadcast Engineer in Eugene, OR, serving as Chief Engineer of an AM/FM broadcast station from 1989-1991. In 1999, she joined the RF Technology Division of the National Institute of Standards and Technology (NIST), Boulder, CO, as an Electronics Engineer. She is currently the Leader of the Metrology for Wireless

Systems Project at NIST, where her research activities include development of calibrated measurements for microwave and millimeter-wave wireless systems and standardized over-the-air test methods for the wireless industry.

Dr. Remley is a Fellow of the IEEE and was the recipient of the Department of Commerce Bronze and Silver Medals, an ARFTG Best Paper Award, the NIST Schlichter Award, and is a member of the Oregon State University Academy of Distinguished Engineers. She was the Chair of the MTT-11 Technical Committee on Microwave Measurements (2008-2010), the Editor-in-Chief of IEEE Microwave Magazine (2009-2011), and Chair of the MTT Fellow Evaluating Committee (2017-2018). She was a Distinguished Lecturer for the IEEE Electromagnetic Compatibility Society (2016-2017) and is the co-Technical Program Chair for the IEEE International Microwave Symposium IMS2022.