

# Comparative Analysis of Clustering Methods for Power Delay Profile in MMW Bands and In-Vehicle Scenarios

Radek Zavorka\*, Ales Prokes\*, Josef Vychodil\*, Tomas Mikulasek\*, Petr Horky\*, Christoph Mecklenbräuker†, Aniruddha Chandra‡, Jan Marcin Kelner§, Cezary Henryk Ziólkowski§

\*Department of Radio Electronics, Brno University of Technology, Brno, Czech Republic

†Institute of Telecommunications, TU Wien, Vienna, Austria

‡National Institute of Technology, Durgapur, India

§ Institute of Communications Systems, Faculty of Electronics, Military University of Technology, 00-908 Warsaw, Poland

e-mail: xzavor03@vutbr.cz

**Abstract**—The spatial statistics of radio wave propagation in specific environments and scenarios, as well as being able to recognize important signal components, are prerequisites for dependable connectivity. There are several reasons why in-vehicle communication is unique, including safety considerations and vehicle-to-vehicle/infrastructure communication.

The paper examines the characteristics of clustering power delay profiles to investigate in-vehicle communication. It has been demonstrated that the Saleh-Valenzuela channel model can also be adapted for in-vehicle communication, and that the signal is received in clusters with exponential decay. A measurement campaign was conducted, capturing the power delay profile inside the vehicle cabin, and the reweighted  $\ell_1$  minimization method was compared with the traditional k-means clustering techniques.

**Index Terms**—millimeter-waves; in-vehicle communication; clusters, k-means, reweighted  $\ell_1$  minimization

## I. INTRODUCTION

Channel modeling is a significant part of almost all communication systems. We need to understand how the signal propagates from point A to point B in a certain environment if reliable communication is required or new standards or frequency bands should be implemented [1].

A prominent stochastic model for indoor multipath propagation is the Saleh-Valenzuela (SV) model [2]. According to the SV model, the tap gain coefficients of the channel impulse response (CIR) between transmitter (TX) and receiver (RX) and also in individual clusters have exponential decrease. A ragged environment, propagation losses, and reflection potentially causes clusters of multi path components (MPC) from the received signal. The model was proposed for the indoor environment of mid-size rooms [3], but we show that the model is applicable to in-vehicle communication.

Given that modern car wire harnesses often weigh more than several tens of kilograms [4], millimeter-wave (MMW) communication can also be used as a wireless

bus to support non-critical safety functions. It is essential to systematically analyze the spatial signal distribution inside the passenger cabin in various conditions to ensure dependable connectivity.

In [5], researchers focused on examining and comparing delay spread and path loss at UWB 3–11 GHz and MMW 55–65 GHz. In the work [6], the in-vehicular wireless 60 GHz channel for the aircraft cabin is described. The measurement campaign outlined in the paper [7] was used to analyze the effects of driving-related vibrations and twisting on CIR and delay-Doppler spread (DDS) inside the car cabin.

Data clustering proves invaluable in the identification of outliers and determination of signal components, particularly within complex datasets. A wide array of methods, ranging from straightforward to highly intricate, have been harnessed for this purpose. Among these, the unsupervised k-means algorithm has gained notable popularity [8], [9]. Nevertheless, a critical aspect is specifying the optimal number of clusters that the algorithm should delineate.

An intriguing alternative lies in the Sparsity-based clustering technique, which adeptly segregates the Power Delay Profile (PDP) into clusters without necessitating a predetermined cluster count. Detailed discussions of this method can be found in the works [10], [11], grounded in the principles of reweighted  $\ell_1$  minimization.

### A. Contribution of the Paper

This paper is focused on analyzing clustering methods for measured millimeter waves in the delay domain in frequency band 55-65 GHz and intra-vehicle scenario. The benefits of paper can be divided into following areas:

- comparison of k-means method and Sparsity based clustering based on reweighted  $\ell_1$  minimization

- analyzing of PDP for in-vehicular communication from measurement campaign by the above mentioned algorithms
- validation of the SV model for and intra-vehicle communication at MMW band

The paper is organized as follows:

The standard SV channel model is presented in section II. The following sections go through clustering techniques, ranging from basic k-means in section V to more intricate Sparsity based clustering- Reweighted  $\ell_1$  minimization in section VI. Section VII of the paper regarding experimental measurement provides details on the measurement setup and the application of the aforementioned clustering algorithms to the observed data. Conclusion section sums up the paper.

## II. SALEH-VALENZUELA MODEL

Channel models for indoor communication consist of a large number of multipath components caused by reflections, scattering, and diffraction. Given below is the general, complex, low-pass channel impulse response:

$$h(\tau) = \sum_{n=1}^N \beta_n \exp(j\phi_n) \delta(\tau - \tau_n), \quad (1)$$

where  $\beta_n$  and  $\phi_n$  are the amplitude gain and phase of the  $n$ th path, respectively.  $N$  is the total number of propagation paths.  $\tau_n$  is the propagation delay (arrival time) of the  $n$ th path.  $\delta(\cdot)$  is the Dirac delta function. [2], [10], [12]

CIR according to SV model [2] describing indoor channel is expressed by

$$h(\tau) = \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \beta_{kl} \exp(j\theta_{kl}) \delta(t - T_l - \tau_{kl}), \quad (2)$$

where  $\beta_{kl}$  and  $\theta_{kl}$  are the amplitude gain and phase of the  $k$ th path within the  $l$ th cluster, respectively.  $T_l$  is delay of  $l$ th cluster arrival.  $\tau_{kl}$  is the excess arrival delay of the  $k$ th path within the  $l$ th cluster. [2]

The most popular approach based on paper [2] assumes a monotonically decreasing received power with the function of  $T_t$  and  $\tau_{kl}$  and average power of  $\beta_{kl}$  is described by equation

$$\overline{\beta_{kl}^2} = \overline{\beta_{0,0}^2} \exp\left(-\frac{T_l}{\Gamma}\right) \exp\left(-\frac{\tau_{kl}}{\gamma}\right), \quad (3)$$

where  $\overline{\beta_{0,0}^2}$  is the average power gain of the first path of the first cluster, and  $\Gamma$  and  $\gamma$  are power-delay time constants for the clusters and the paths, respectively [2]. The Figs. 1 and 2 represent the described model.

## III. CHANNEL DESCRIPTION

In-vehicle radio channel measurements are typically assumed to exhibit time-invariant characteristics. So, in order to properly define the radio channel, the CTF is as follows:

$$H(k) = s_{21}(k), \quad (4)$$

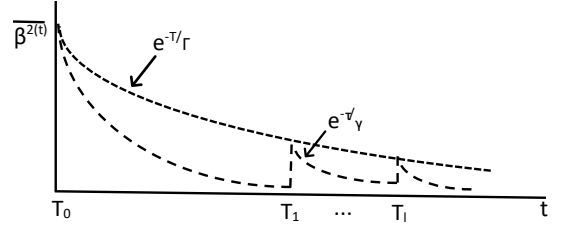


Fig. 1. SV channel model - Exponentially decaying ray and cluster average powers [2]

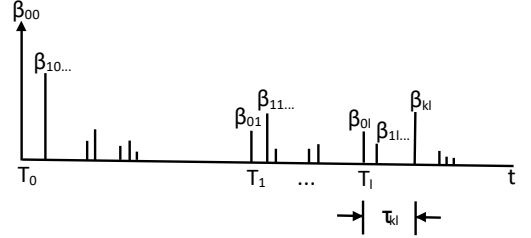


Fig. 2. SV channel model - A realization of the impulse response [2]

where  $k$  is the measurement index that can be linked to a specific frequency. We transform the CTF into a CIR using an inverse Fourier transform in accordance with:

$$h(n) = \sum_{k=0}^{N-1} H(k) e^{jkn2\pi/N}, \quad (5)$$

where, in the delay domain,  $n$  denotes a discrete time.

The PDP is then determined as the average of a several realizations in accordance with:

$$P(n) = E\{|h(n)|^2\}. \quad (6)$$

## IV. METHODS OF RADIO SIGNAL CLUSTERING

In order to cluster radio signals, a variety of methods are available. From fundamental machine learning techniques like k-means [13] and enhanced k-PowerMeans [14] up to sparsity-based techniques derived from reweighted  $\ell_1$  minimization [10], [11].

Synthetic signals were generated in MATLAB for validating the clustering approaches. Figure 3 shows an exemplary realization of a synthetic signal. There is a transmit signal and received signal that have been impacted by noise, and a signal that has been reconstructed via reweighted  $\ell_1$  minimization (explained below).

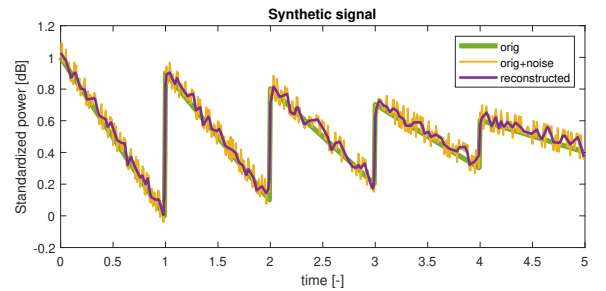


Fig. 3. Synthetic signal for verification of clustering algorithms

## V. K-MEANS CLUSTERING ALGORITHM

Most likely one of the most well-known unsupervised clustering algorithms, applied in many fields [8], [13]. The number of expected clusters must be specified for proper working. The basic idea behind the algorithm is to find a predetermined number of centroids in the space among points. Measured data from the graph is separated into distinct clusters. Centroids are placed in order to reduce the distance between measured points and clusters [15].

Simulations of the given algorithm are shown in Fig. 4. It is obvious that k-means clustering produces inaccurate results because includes the tail of the previous cluster in the new one.

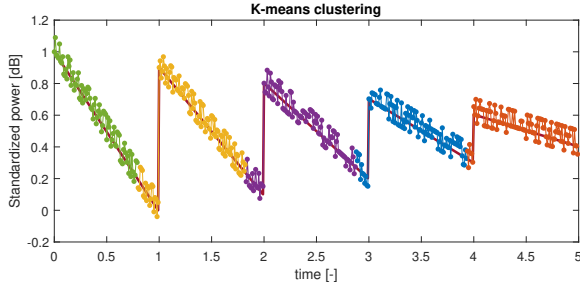


Fig. 4. Clustering based on k-means algorithm on synthetic signal

## VI. SPARSITY BASED CLUSTERING- REWEIGHTED L1 MINIMIZATION

The SV model is based on the assumption that multipath propagation occurs in clusters with an exponential decrease in received power. Therefore, the proposed approach partitions the measured PDP into several segments, or distinct clusters. The concept of the algorithm proceeds in two steps [10]:

- reconstruction of transmit signal  $\hat{P}$  from measured PDP.
- identification of the clusters from the shape of  $\hat{P}$ .

The following non-convex optimization problem is formulated from the aforementioned concept [16]:

$$\min_{\hat{P}} \left\| P - \hat{P} \right\|_2^2 + \lambda \left\| \Omega_2 \cdot \Omega_1 \cdot \hat{P} \right\|_0, \quad (7)$$

where  $P = [P(0), \dots, P(N-1)]^T$  denotes the received signal and similarly for the reconstructed signal  $\hat{P}$ , and  $\lambda$  is a positive regularization parameter. The feasible set  $\mathcal{X}$  for the minimization problem (7) is

$$\mathcal{X} = \{ \hat{P} \in \mathbb{R}_0^N \mid \left\| \Omega_2 \cdot \Omega_1 \cdot \hat{P} \right\|_1 \leq L_{\max} \}. \quad (8)$$

The finite-difference operator  $\Omega_1$  is used for estimating the slope of  $\hat{P}$ . The turning point, where the slope changes significantly, is determined using the finite-difference operator  $\Omega_2$ , represented by the filter [1, -1], defined in [10] as:

$$\Omega_2 = \begin{bmatrix} 1 & -1 & 0 & \dots & \dots & 0 \\ 0 & 1 & -1 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & 1 & -1 & 0 \\ 0 & 0 & \dots & \dots & 1 & -1 \end{bmatrix}_{(N-2) \times (N-1)} \quad (9)$$

The reweighted  $\ell_1$  norm minimization strategy [17] is used by [10] to numerically approximate a sparse solution and a more accurate clustering result.

The reweighted  $\ell_1$  minimization [17] is applied to the optimization problem (7) for iteratively increasing the sparsity of the solution by the weighted norm. The algorithm consists of the following phases [10].

- 1) Set the iteration count  $m$  to zero and the initial weights  $w_n^{(0)} = 1$ ,  $n = 1, \dots, N$ . In the reweighted  $\ell_1$  minimization, the weight parameter is later employed to ensure a sparse solution.
- 2) The weighted  $\ell_1$  minimization problem is as follows:

$$\begin{aligned} \hat{P}^{(m)} &= \arg \min_{\hat{P} \in \mathbb{R}_0^N} \left\| P - \hat{P} \right\|_2 \\ \text{s.t. } &\left\| W^m \cdot \Omega_2 \cdot \Omega_1 \cdot \hat{P} \right\|_1 \leq L_{\max} \end{aligned} \quad (10)$$

where  $L_{\max}$  is the upper bound on the number of clusters in a PDP.  $W^{(m)}$  is a matrix with weights at diagonal.

- 3) The weights are updated according to

$$w_n^{(m+1)} = \frac{1}{\hat{P}^{(m)}(n) + \epsilon}, \quad (n = 1, \dots, N). \quad (11)$$

In order to maintain stability and make sure that a zero-valued component in  $\hat{P}^{(m)}$  does not strictly preclude a nonzero estimate at the following phase, the constant  $\epsilon$  is utilized. Epsilon can be selected as the smallest value permitted in the computing environment or a value significantly smaller than the anticipated nonzero magnitudes of  $P$ . In our case for testing, the  $\epsilon$  was equal to  $10^{-9}$ .

- 4) When the method has run for the specified number of iterations  $M$  or when the weights have converged, the algorithm terminates. If not, increment  $m$  and move on to step 2.

The recovered  $\hat{P}$  signal (red color) for the simulation case is depicted in Fig. 3.

It is feasible to locate the beginnings of clusters by using the recovered signal  $\hat{P}$ , which is piecewise linear. The following vector

$$\Phi = \left[ \Omega_2 \cdot \Omega_1 \cdot \hat{P} \right]_{(N-2) \times 1}, \quad (12)$$

denotes the pivot point of  $\hat{P}$  where the slope dramatically changes and can be used to pinpoint clusters with accuracy. In Fig. 6 there is a synthetic signal from Fig. 3 divided into clusters automatically by presented sparsity

based clustering algorithm which uses reweighted  $\ell_1$  minimization according gained decision boundary shown in Fig. 5

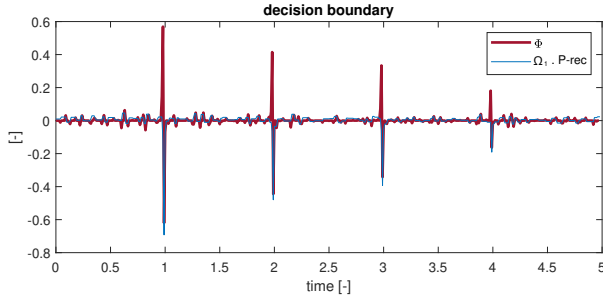


Fig. 5. Decision boundary for synthetic signal according to  $\Phi$  by reweighted  $\ell_1$  minimization

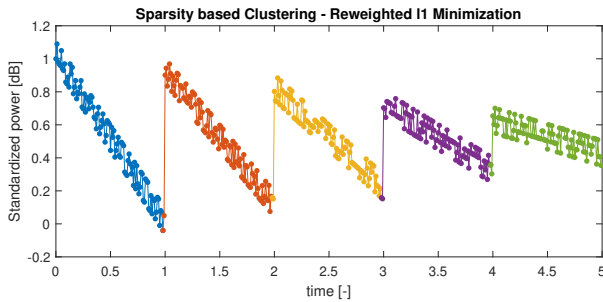


Fig. 6. Synthetic signal divided into clusters by based clustering algorithm with the use of reweighted  $\ell_1$  minimization

## VII. EXPERIMENTAL MEASUREMENTS

The measurements were taken as part of a measuring campaign [18] to analyze in-vehicle communication and signal propagation inside car cabin.

The transmit and receive open-ended waveguide antennas are used for measurements inside a mid-sized vehicle Skoda Octavia III. The RX antenna is positioned on the front windshield adjacent to the rear-view mirror. The TX antenna is positioned in the middle of the back seat. A complete description of the measurement campaign may be found in [18].

### A. 55-65 GHz measurement apparatus

In the 55-65 GHz frequency range, the transmission coefficient between two antennas is measured using the R&S ZVA67 four-port VNA. A broadband power amplifier (QuinStar QPW-50662330, measured gain of 35 dB in the band of interest) was used on the transmitting side of the measurement setup to increase the system dynamic range. Antennas for transmitting and receiving signals were WR15 open waveguide, which has a radiation pattern identical to [18] (similarly as in [19]). This measurement setup has a system dynamic range of about 50 dB. The IF bandwidth of 100 Hz and the VNA output power of 5 dBm were selected.

A 10 MHz frequency step is used during the measurement, which is carried out in the frequency domain. The

inverse Fourier transform with Blackman windowing is used to convert the channel frequency response into the time domain. Phase-stable coaxial cables were utilized to prevent the TX antenna motions from reducing the observed phase accuracy. The power amplifier was calibrated together with the all four ports. In order to equalize the VNA transmission in the forward direction, the TX and RX antennas flanges were coupled together. The open-ended waveguide utilized for the experiment at 55-65 GHz has a radiation pattern that is  $120^\circ$  wide in the E-plane and  $60^\circ$  wide in the H-plane (for 3 dB decay).

### B. Application clustering methods on measured PDP based on in-vehicle communication

A significant number of PDPs in a variety of scenarios were gathered during the measurement campaign outlined above. It is straightforward to show that the channel inside the car cabin and the measurement inside the room are quite comparable, allowing an SV model to be used. Therefore, in-vehicle PDP can be clustered using the suggested approaches.

The illustrative PDPs from measurement configuration are shown in Fig. 7, Scenario 1- the transceiver in the location behind the driver, the trunk door was open and car cabin without passengers and Fig. 8 Scenario 2- with open windows instead trunk door. The reconstructed signal  $\hat{P}$  is shown with red color, and it is used to detect crucial turning points by using the sparsity-based clustering-reweighted  $\ell_1$  minimization approach.

Decision boundary  $\Phi$  according to (12) is shown in Fig. 7 and 8 below, respectively. There is no denying that some pikes are noticeably higher than others. Once the decision level is set to  $-0.35$ , value optimized for this measurement scenario, clusters can be found and the PDP divided into clusters of MPC.

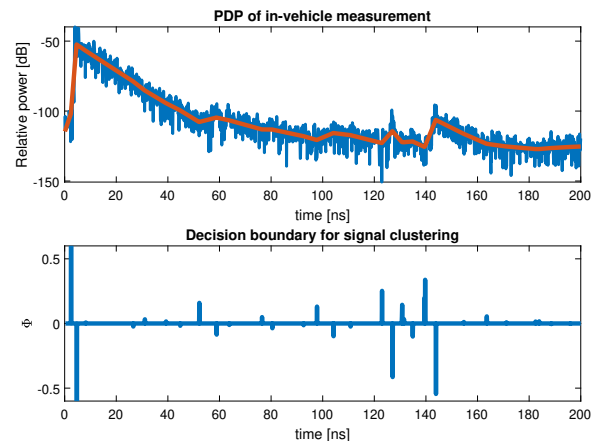


Fig. 7. PDP of measured in-vehicle communication (Scenario 1) (blue) and reconstructed  $\hat{P}$  signal (red) using reweighted  $\ell_1$  minimization and Decision boundary for clustering of measured signal

PDPs clustered using k-means and sparsity-based clustering, respectively, are displayed in Figs. 9 and

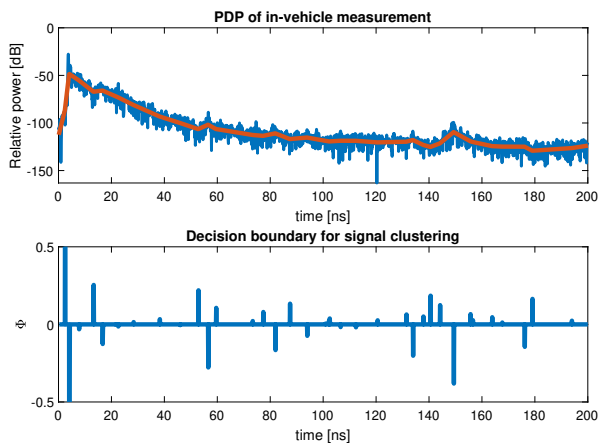


Fig. 8. PDP of measured in-vehicle communication (Scenario 2) (blue) and reconstructed  $\hat{P}$  signal (red) using reweighted  $\ell_1$  minimization and Decision boundary for clustering of measured signal

10. The common k-means approach is obviously inappropriate for time domain data since it divides PDP into a horizontal plane rather than separating clusters into a vertical plane. Compared to that, Sparsity based clustering segregated the signal with excellent precision in accordance with the decision boundary in Figs. 7 and 8 using Reweighted  $\ell_1$  minimization.

In scenario 1, points that indicate the beginning of new clusters have a markedly higher intensity than other points. This is due to the fact that the decrease in PDP is practically exponential (a straight line on a logarithmic scale) and exhibits minimal deviation. Additionally, the reconstructed signal for scenario 2 appears to be very smooth, but as can be seen in the graph of the decision boundary for scenario 2, the signal  $\Phi$  contains more intense spikes, making it crucial to set the horizontal level for point selection very carefully.

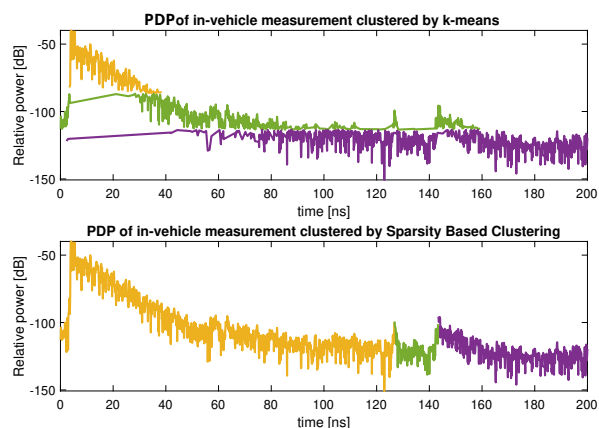


Fig. 9. PDP of in-vehicle measurement (Scenario 1) clustered by k-means algorithm and Sparsity Based Clustering- Reweighted  $\ell_1$  minimization

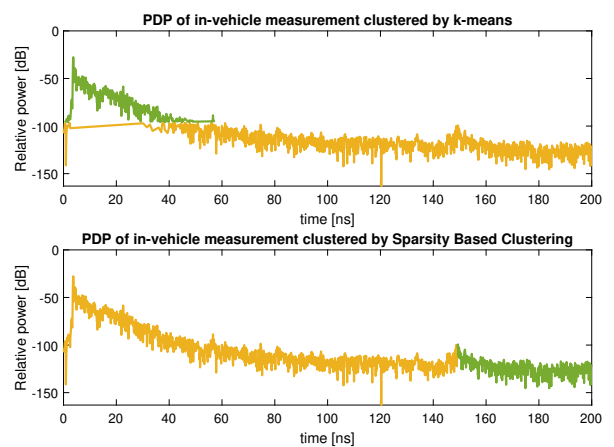


Fig. 10. PDP of in-vehicle measurement (Scenario 2) clustered by k-means algorithm and Sparsity Based Clustering- Reweighted  $\ell_1$  minimization

## VIII. CONCLUSIONS

We investigated clustering for partitioning MPCs into distinct segments. The clustering of the PDP is highly advantageous for estimating easily interpretable channel parameters. PDP data was collected during an extensive measurement campaign focused on in-vehicle communication at 60 GHz. The delay domain representation of the PDP was subjected to both k-means and sparsity-based clustering techniques. The k-means technique is less suited for analyzing time domain data, primarily due to its horizontal rather than vertical delay data clustering. However, the sparsity-based clustering approach proved highly effective for accurately segmenting real-world PDP data obtained from in-vehicle measurements. It was confirmed that the Saleh-Valenzuela model is adequate for in-vehicle communication at 60 GHz.

## ACKNOWLEDGMENT

This work was supported by the Czech Science Foundation, Lead Agency Project No. 22-04304L, Multi-band prediction of millimeter-wave propagation effects for dynamic and fixed scenarios in rugged time varying environments. The infrastructure of the SIX Center was used.

## REFERENCES

- [1] A. F. Molisch, *Channel Models*, 2011, pp. 125–143.
- [2] A. Saleh and R. Valenzuela, “A statistical model for indoor multipath propagation,” *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 2, pp. 128–137, 1987.
- [3] A. Meijerink and A. F. Molisch, “On the physical interpretation of the saleh-valenzuela model and the definition of its power delay profiles,” *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 9, pp. 4780–4793, 2014.
- [4] G. Leen and D. Heffernan, “Expanding automotive electronic systems,” *Computer*, vol. 35, no. 1, pp. 88–93, 2002.
- [5] J. Blumenstein, T. Mikulasek, A. Prokes, T. Zemen, and C. Mecklenbrauker, “Intra-vehicular path loss comparison of UWB channel for 30–11 GHz and 55–65 GHz,” in *2015 IEEE International Conference on Ubiquitous Wireless Broadband (ICUBW)*, 2015, pp. 1–4.

- [6] M. Peter, W. Keusgen, A. Kortke, and M. Schirmacher, "Measurement and analysis of the 60 GHz in-vehicular broadband radio channel," in *2007 IEEE 66th Vehicular Technology Conference*, 2007, pp. 834–838.
- [7] A. Prokes, J. Vychodil, M. Pospisil, J. Blumenstein, T. Mikulasek, and A. Chandra, "Time-domain nonstationary intra-car channel measurement in 60 GHz band," in *2016 International Conference on Advanced Technologies for Communications (ATC)*, 2016, pp. 1–6.
- [8] K. P. Sinaga and M.-S. Yang, "Unsupervised k-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020.
- [9] Z. Wang, Y. Zhou, and G. Li, "Anomaly detection by using streaming k-means and batch k-means," in *2020 5th IEEE International Conference on Big Data Analytics (ICBDA)*, 2020, pp. 11–17.
- [10] R. He, W. Chen, B. Ai, A. F. Molisch, W. Wang, Z. Zhong, J. Yu, and S. Sangodoyin, "On the clustering of radio channel impulse responses using sparsity-based methods," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 6, pp. 2465–2474, 2016.
- [11] —, "A sparsity-based clustering framework for radio channel impulse responses," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, 2016, pp. 1–5.
- [12] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York: Mc Graw-Hill Education, 2007.
- [13] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.
- [14] R. He, B. Ai, A. F. Molisch, G. L. Stuber, Q. Li, Z. Zhong, and J. Yu, "Clustering enabled wireless channel modeling using big data algorithms," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 177–183, 2018.
- [15] S. Na, L. Xumin, and G. Yong, "Research on k-means clustering algorithm: An improved k-means clustering algorithm," in *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, 2010, pp. 63–67.
- [16] R. Giryes, M. Elad, and A. M. Bruckstein, "Sparsity based methods for overparameterized variational problems," 2014. [Online]. Available: <https://arxiv.org/abs/1405.4969>
- [17] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted  $\ell_1$  minimization," 2007. [Online]. Available: <https://arxiv.org/abs/0711.1612>
- [18] J. Blumenstein, A. Prokes, A. Chandra, T. Mikulasek, R. Marsalek, T. Zemen, and C. Mecklenbräuker, "In-vehicle channel measurement, characterization, and spatial consistency comparison of 30–11 GHz and 55–65 GHz frequency bands," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 3526–3537, 2017.
- [19] H. Xu, V. Kukshya, and T. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 620–630, 2002.