Outdoor mmWave Channel Propagation Models using Clustering Algorithms

Bogdan Antonescu
Institute for the Wireless IoT
Northeastern University
Boston, USA
antonescu.b@husky.neu.edu

Miead Tehrani Moayyed
Institute for the Wireless IoT
Northeastern University
Boston, USA
tehranimoayyed.m@husky.neu.edu

Stefano Basagni
Institute for the Wireless IoT
Northeastern University
Boston, USA
basagni@ece.neu.edu

***Abstract*—This paper concerns the task of generating simpler yet accurate mmWave channel models based on clustering all multipath components arriving at the receiver. Our work focuses on 28 GHz communications in urban outdoor scenarios simulated with a ray-tracer tool. We investigate the effectiveness of k-means and k-power-means clustering algorithms in predicting the optimal number of clusters by using cluster validity indices (CVIs) and score fusion techniques. Our results show how the joint use of these techniques generate accurate approximation of the mmWave large-scale and small-scale channel models, greatly simplifying the complexity of analyzing large amount of rays at any receiver location.**

***Index Terms*—mmWave, channel propagation models, clustering algorithms, cluster validity indices**

I. Introduction

One of the many disruptive technologies of future 5G wireless standards is represented by *millimeter-wave* (mmWave) communications in the 30–300 GHz band. Despite the fact that mmWave provides the high-speed data transfers and ubiquitous connectivity with very low latency responses, transmissions at these frequencies suffer from high propagation loss, sensitivity to blockage, atmospheric attenuation and diffraction loss. Solving these problems starts with a clear understanding of the *radio channel propagation models* in the mmWave band, which can be obtained by extensive experimental measurements (via steerable antennas and channel sounders) or via software ray-tracing simulations.

In this paper we are concerned with the task of generating simpler yet accurate mmWave radio channel models by using the ray tracing method. As a first task, we investigate the role of *clustering algorithms* in grouping the large amount of rays arriving at the receiver site. We observe that a clever use of clustering greatly simplifies ray-based channel analysis without loosing necessary accuracy. We use two variants of the well known k-means clustering algorithm in which the Euclidean distance metric is replaced with the multipath component distance (MCD). The result is a multi-dimensional space created by the channel parameters—Time-of-Arrival (ToA), azimuth and elevation of the Angle-of-Arrival (AoA) and Angle-of-Departure (AoD)—that define the multipath components (MPCs) received at each location. The clustering algorithms take this space as input, and group all MPCs in various clusters. To quantify the goodness of their solution, we use *cluster validity indices* (CVIs) and *score fusion* techniques. In the end, we combine all CVIs in an *ensemble*, to obtain a better predictor of clustering quality comparing with the one provided by each CVI taken separately.

Our simulations concern a mmWave urban outdoor scenario with multiple receiver locations. Scenario and simulations are generated and performed with a professional software ray-tracer tool, namely, Wireless InSite by Remcom. MATLAB code is used to implement the clustering algorithms applied to the estimated channel parameters as well as for the validation of their results. Finally, we process the inter- and intra-cluster parameters that describe the clusters and the rays in each cluster, and we investigate the effect of the two clustering solutions on the generation of channel propagation models. We find that the statistics extracted from the clustered solution represent an accurate approximation of the values estimated by the ray tracer without clustering. Thus, the large-scale and small-scale channel models that we generate could be used by a wireless network architect to get a first order of magnitude estimate of both path loss and root mean square (RMS) delay spread, to forecast network coverage and maximum data rate in the channel without using a ray-tracer or measuring channel parameters with expensive dedicated hardware.

The rest of the paper is organized as follows. Section II reviews briefly clustering concepts and validation techniques used in our research, and provides links for their full description. Section III introduces the outdoor scenario, and presents the results of the two variants of the k-means algorithm, the validation of their results and the effect of the clustering solutions to channel modeling. Section IV concludes the paper.

II. Clustering and Validation Techniques

Transmissions in radio channels are characterized by large scale effects (e.g., path loss, shadowing loss) and small-scale effects (e.g., fading loss). One way to model small-scale fading is through *multipath components* with randomly distributed amplitudes, phases and angles-of-arrival (AoA) that combine at the receiver causing the received signal to distort or fade. The other factor that influences the small-scale propagation channel model is the *Doppler spread* due to the mobility and speed of transmitter (Tx) and receiver (Rx) (such as in cars, on people and on other moving objects). To evaluate

978-1-7281-4905-9/20/$31.00 ©2020 IEEE

channel time-dispersive properties, the *RMS delay spread* of the received rays is considered a good indicator as it provides a measure of the severity of *Inter Symbol Interference* (ISI). This translates to more complicated schemes for symbol recovery at the receiver and in limitations of the maximum achievable channel data rate. As a rule of thumb, an RMS delay spread ten times smaller than the transmitted symbol time period guarantees no requirement for an ISI equalizer at the receiver.

In this paper, we are concerned with the *multipath* problem and with the analysis of sorting and grouping the rays received from multiple paths into clusters. We use well known *center-based* clustering algorithms that partition input data around few centroids or central points [1]. Among the most common algorithms in this class, k-means [2] and one of its variants, k-power-means, are used in many studies [3], [4], [5], [6]. For the cluster definition for mmWave MPCs and for the description of the clustering process using k-means and k-power-means with the *multipath component distance* (MCD) metric [7], the reader is referred to one of our previous papers [8].

Clustering as an *unsupervised* pattern classification method partitions the elements in a data set into clusters by identifying similar values of the parameters that characterize them. In our case, these elements are the arriving MPCs at each receiver point. They are characterized by various radio channel parameters (e.g., power levels, ToA, AoA, AoD) with different values from one MPC to another.

The next step in the clustering process, once the algorithm has partitioned the input data set, is to prove the accuracy of the result. For details about the types of *cluster validation* and of the methods that we implement in our research, we direct the reader to reference [8]. In this study, we pursue internal validation methods by applying the following Cluster Validity Indices (CVIs): Calinski-Harabasz (CH) [9], Davies-Bouldin (DB) [10], generalized Dunn (GD) [11], [12], Xie-Benie (XB) [13] and PBM [14]. Their formulas and the way they measure the cluster size and the cluster separation are described in [8]. As we explained there, each CVI might capture only specific aspects of the clustering solution (i.e., some clusters might not be considered compact just because they have an elongated shape). Therefore, we decide to combine all CVIs in an *ensemble* [15] that provides a more accurate prediction of the clustering quality than any individual CVI. These are the *score fusion-based* techniques where the SF_x scores are defined by the arithmetic, geometric and harmonic mean of the five CVIs [8].

III. Simulation Results

This section describes our ray-tracer simulations and the results of the clustering algorithms applied to the channel parameters estimated with this tool. It also provides the findings on the optimal number K of clusters when we implement both the CVIs and the score fusion techniques mentioned in the previous section. For further details and motivations, the reader is referred to [8]. A wider mmWave spectrum below 100 GHz is analyzed in one of our papers [16]. In the present work, we simulate only 28 GHz communications in an urban scenario (Rosslyn, VA) delivered with the ray-tracing tool (Fig. 1). The transmitter (Tx) is located at a fixed site on a traffic light pole (the green dot) while the receiver (Rx) is in a car that can be at any location marked with a red dot in this picture. The full details of the setup and the way we orient the antennas to receive the transmitted signals are described in [8].

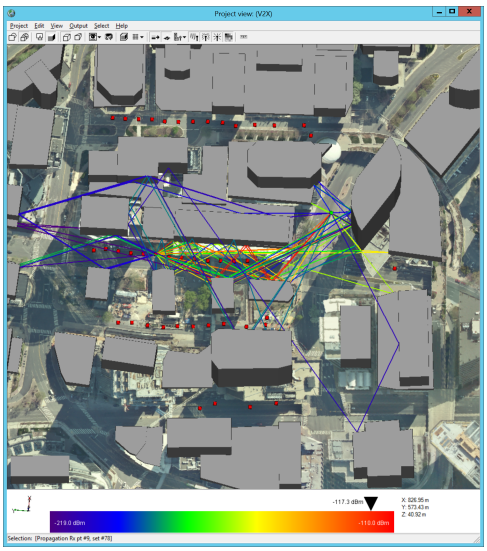

Fig. 1. 44 MPCs at receiver Rx#9.

A. Clustering algorithm results

This section provides the clustering results of the two variants of the k-means algorithm. 44 MPCs are received at receiver Rx#9 on a side street in our urban scenario (Fig. 1). Each MPC is characterized by a power level, AoA, AoD, and a certain excess delay (ToA). For this one-time channel realization, the real part of the complex impulse response (CIR) shows (in different colors) the received power levels and ToA of all clustered MPCs (Fig. 2). A star marks the average power value of each cluster and its average ToA. These values are calculated using the channel parameters of the MPCs in each cluster. The clustering algorithm is k-means with MCD. The 3D results (Fig. 3) show the effect of capturing all five parameters of the MPCs (azimuth & elevation for AoA and AoD, and excess delay) in the clustering process. They allow for a better partition because they correlate the *temporal* and *spatial* characteristics of the radio channel.

Using k-power-means with MCD as another clustering option, we obtain different clustered CIR (Fig. 4) and ToA vs. AoA/AoD (Fig. 5) pictures. The average values of the Rx power and ToA in each cluster (marked with a star in Fig. 4) are very close to each other, even though the MPCs in each cluster are dispersed in time. The 5-dimensional space used for clustering (Fig. 2) is now biased by the received power (Fig. 4), showing that grouping rays around the strongest MPCs is the factor that reduces the number of clusters.

B. CVIs and score fusion results

The results of the clustering process are validated and the optimal K value is found by applying the CVIs mentioned

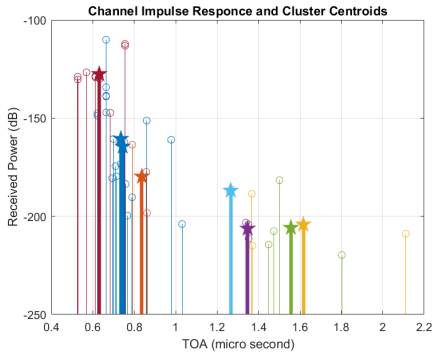

Fig. 2. Clustered CIR at Rx#9 based on k-means with MCD.

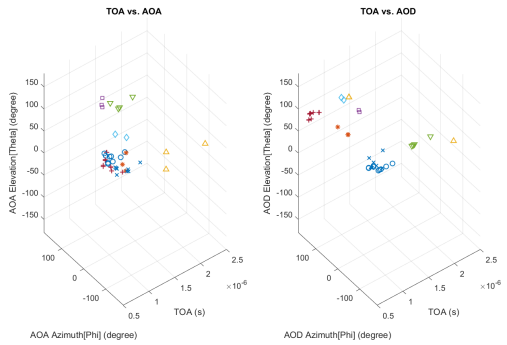

Fig. 3. Clustering via k-means algorithm—ToA vs. AoA, AoD.

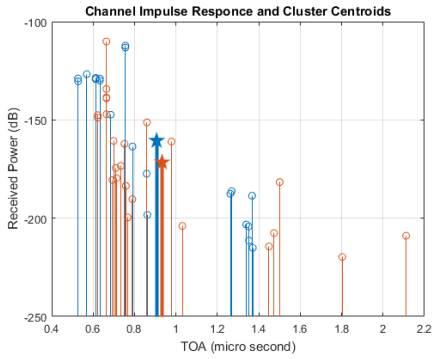

Fig. 4. Clustered CIR at Rx#9 based on k-power-means with MCD.

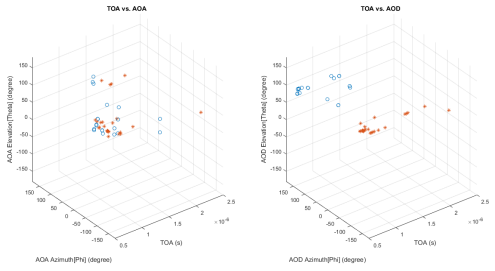

Fig. 5. Clustering via k-power-means algorithm—ToA vs. AoA, AoD.

in Section II. In our scenario, receiver Rx#9 is placed on a side street at approximately 150 m (Euclidean distance) from the transmitter (Fig. 1). The simulation shows only 44 MPCs reaching this receiver. If we estimate 3 rays per cluster, then all MPCs can be grouped in maximum 15 clusters. This guess is always required, to set the initial K input of the clustering algorithm; in our case, the range is $[2, 15]$.

As mentioned, not all CVIs are able to find the optimal K value. When the k-means algorithm is used, indices CH, DB and GD cannot find this number correctly, whereas XB and PBM point to an optimal number of clusters of 6 and 5, respectively. Using the ensemble predictor, we plug the normalized and biased CVI values obtained for Rx#9 (for each input value K) into the score fusion formulas described in [8]. Unfortunately, we cannot predict K using only the maximum value of the CVIs because not all CVIs agree with each other. However, if we use score fusion methods, the SF_g and SF_h scores based on geometric and harmonic mean agree on this value (i.e., $K = 8$ clusters). Moreover, the average of the three scores points to the same value for K. We repeat this study for all 14 receivers installed on the street where Rx#9 is placed. The results show an agreement among the three score fusion values regarding the optimal K value for most of the receivers, except Rx#8, Rx#9, Rx#10 and Rx#14 for which only two scores agree on the same K.

As the second clustering algorithm in our analysis, we use the k-power-means variant. The distance metric and the local and global centroids used in the CVI formulas are all weighted by the power of each MPC. The same five CVIs and three score fusion factors are used to validate the clustering results when the same set of MPCs (at Rx#9) is clustered. Now, the optimal K number is 2. We repeat the study for all 14 receivers located on that street. This time, for each receiver on the street, all three score fusion values report the same optimal K.

C. Cluster-based channel modeling

Channel model analysis implies a *large-scale channel model*, which is an indication of the received power attenuation with distance, including path loss and shadowing, and a *small-scale channel model*, which is related to environment (e.g., power decay rate, path arrival rate, RMS delay spread and angular spread). *Cluster-based channel modeling* applies to both models, and has been lately given more attention, especially in mmWave indoor scenarios [17], [18], [19]. Our paper fills the gap and studies the influence of clustering to the channel model in an outdoor scenario.

Path loss (PL) is the signal attenuation due to a decreased antenna reception when the distance between Tx and Rx increases; it is associated with a *path loss exponent* n that shows how fast path loss increases in various environments. The large-scale propagation model also accounts for *shadowing loss*, which is caused by the absorption of the radiated signal by obstacles and scattering structures. The shadowing factor χ_σ in (1) is typically modeled by a random variable with log-normal distribution with zero mean and standard deviation σ:

$$PL(d)[dB] = PL_{FS}(d_0) + 10n log_{10}\frac{d}{d_0} + \chi_\sigma \qquad (1)$$

where $PL(d_0)$ is the free-space loss at reference distance d_0 given by: $PL_{FS}(d_0)[dB] = 20log_{10}\frac{4\pi d_0}{\lambda}$. In a previous paper [20], we developed the large-scale channel model for the same urban outdoor scenario. Hundreds of random Rx points were generated at each receiver location, every 10 m, on the same street considered for the NLOS scenario, in order to estimate the path loss exponent (n) and the standard deviation (σ) of the shadowing factor used in (1). This time, we considered only one Rx point at each location, but we captured *all* MPCs arriving at this receiver (for clustering purpose). The number of clusters (CLx) estimated with the k-means clustering algorithm at each of the 14 locations on the street and the strongest path in each cluster are summarized in Table I. While this table shows the path loss of the MPC

TABLE I
PATH LOSS FOR STRONGEST MPC [dB]—k-MEANS ALGORITHM.

Rx	CL1	CL2	CL3	CL4	CL5	CL6	CL7	CL8
1	102.5	96.6	91.2	184.1				
2	93.5	183.1	130.6	96.7				
3	134.1	97.97						
4	100.6	129.2	180.5					
5	137.8	102.2	182.8					
6	147.5	121.2	105.5	208.1	141.5			
7	128.3	210.1	126.6	125.1				
8	110.1	126.4	111.5	207.8				
9	112.2	147.6	188.5	203.2	186.3	181.6	163.5	110.1
10	117.2	207.9	109.6	133.4				
11	190.7	129.6	129.3					
12	128.7	134.7	196.1					
13	183.1	136.9	123.5	200.8				
14	198.7	129.6	182.7	139.3				

that represents the centroid of each cluster, we plot in Fig. 6 the average value of the path loss for all clusters at each Rx location, and we compare it with the curve obtained based on equation (1) where a path loss exponent of $n = 4.71$ was found (in our previous study) for this street when antennas with half-power beamwidth (HPBW) of 22° and 15 dBi gain were used and no beam alignment was implemented for the Tx and Rx antennas. Fig. 6 proves that for the studied distance range (70 to 200 m) the average path loss based on clustering (blue diamonds) is well matched by the path loss model (green triangles) given by equation (1). The conclusion is that a *simplified* path loss model based on clustering eliminates the lengthy process required to describe fully equation (1). In the same time, the plot in Fig. 7 shows the path loss of the dominant MPC in all clusters at each Rx point, and how these defining MPCs compare with the average path loss values (i.e., the red dots) taken over the MPCs recorded at each location. The distribution of the cluster centroids shows that these dominant paths can be used to send multiple data streams between the Tx and Rx. Thus, we can say that knowing the number of clusters at each receiver determines the maximum number of independent streams that can be sent

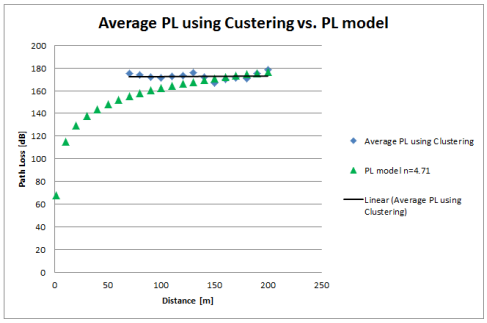

Fig. 6. Average Path Loss using Clustering vs. Path Loss model.

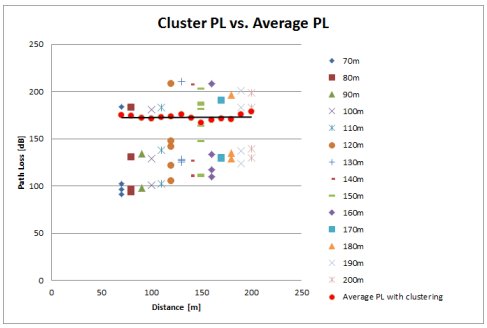

Fig. 7. Clusters Path Loss vs. Average Clustered Path Loss.

in one polarization, and also influences the channel capacity. Also, these centroids provide a possible localization of the reflectors in the dual delay–angle space while the distribution of the MPCs within each cluster describes the dispersion characteristics of these reflectors.

In the second part of our analysis dedicated to the *small-scale channel model* we are interested in two sets of parameters, inter- and intra-cluster parameters that describe the clusters and the rays in each cluster. They matter for a better understanding of the mmWave channel. For example, a good angular dispersion characterization helps design a better control of the beamwidth and directivity of an array antenna. Also, as already mentioned, clustering of MPCs may significantly affect channel capacity; hence, RMS delay spread (RMS DS) is another important factor to us.

We start by analyzing first the *RMS delay spread*, as this parameter is tightly connected with the maximum data rate achievable in the channel. For that reason, we capture the delay spread reported by the ray-tracer at each of the 14 locations on the street (last column in Table II); these numbers are in accordance with outdoor measurement campaigns in an urban canyon [21]. Using the same formulas (2) implemented by the ray-tracer, we calculate the RMS DS values of each cluster based on the partitioning obtained with the k-means algorithm:

$$\sigma_{DS} = \sqrt{\frac{\sum_{i=1}^{L_k}(\tau_i - \bar{\tau})^2 P_i}{P_R}}; \bar{\tau} = \frac{\sum_{i=1}^{L_k} P_i \tau_i}{P_R} \quad (2)$$

where P_i and τ_i are the power and delay of an arriving MPC while P_R and $\bar{\tau}$ are the power of all L_k MPCs in one cluster and their mean arrival time. We repeat this procedure for the

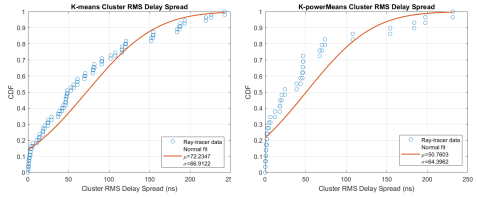

Fig. 8. CDF and truncated normal distribution of clustered RMS DS.

k-power-means algorithm, but for lack of space we omit this table. Nevertheless, the plots of the Cumulative Distribution Function (CDF) of the clustered RMS delay spread for both variants of k-means are captured in Fig. 8 and their fitted distribution parameters are shown in Table III.

The task is to calculate the new values of the RMS DS based on the clustering solution, to check if they are a good representation of the values provided by the ray-tracer tool, or if the channel model can be simplified. Each RMS DS value in the CLx cells in Table II is the result of the MPCs grouped in one cluster for that specific partition. We weight these CLx values with the number of MPCs in each cluster, and the resulted RMS DS values for all 14 receivers are plotted on the same graph (Fig. 9) with the ones given by the ray-tracer (last column in Table II).

TABLE II
CLUSTER-BASED RMS DELAY SPREAD [NS]—k-MEANS ALGORITHM.

Rx	CL1	CL2	CL3	CL4	CL5	CL6	CL7	CL8	RT
1	0.862	0.449	0.755	1.000					29.45
2	12.44	219.5	81.98	18.17					32.61
3	115.7	48.67							48.73
4	61	79.94	115.9						61.06
5	151.8	69.55	154.6						69.63
6	120	187.9	10.93	80.58	48.28				86.76
7	24.44	25	42	46.29					55.73
8	3.025	29.12	56.31	90.8					59.42
9	46.67	102	70.54	5.89	3	4.17	13.5	0.22	52.09
10	108	60.83	2.15	0.968					59.78
11	91.19	18.54	46.81						49.73
12	38.92	154	187						65.85
13	52.93	193	121	242					124.1
14	14.41	43.7	37.1	184					87.69

For the k-power-means algorithm, we use a similar procedure, only that we have only 2 clusters for each receiver (except Rx#1 that has 3), and the weighting function is with the power of the strongest MPC that defines each cluster. These RMS DS values are also plotted in Fig. 9.

As mentioned, one goal is to extract a possible approximation of the channel model when only few MPCs are taken into consideration. When k-power-means is used, almost always, we have only two clusters for all receivers on that street. The power levels of the defining MPCs for the two clusters are in close range. When the k-means version of the algorithm is used, the solution has many more clusters. The difference between the received power levels of the strongest dominant

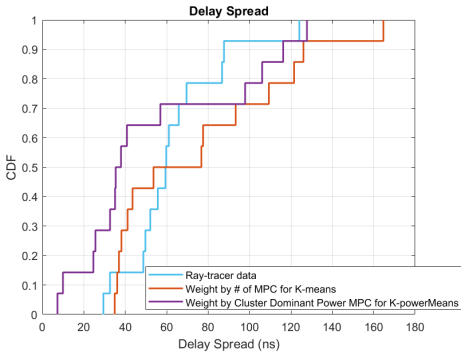

Fig. 9. CDFs of the RMS DS for ray-tracer values and cluster-based values.

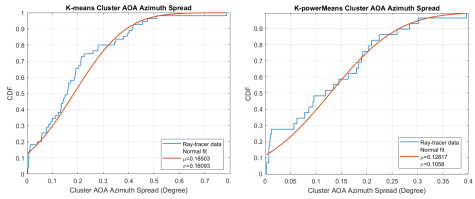

Fig. 10. CDF and normal distribution of the clustered RMS AoA spread.

MPC that defines one cluster and the weakest dominant MPC that defines another cluster is between 61 and 102 dB. Nevertheless, when we consider only the clusters grouped around the two strongest dominant MPCs, the power values of these MPCs are again in close range. That tells us that if nothing else matters, except for the strongest MPCs, we could have a good approximation of the channel model by considering only a limited number of clusters and their associated MPCs.

As a first step, we retain the model captured in Table II with the number of clusters for each receiver resulted from clustering with k-means, and the model with only 2 clusters per receiver for the k-power-means variant. Based on Fig. 9, the RMS DS for both clustering solutions provides a good approximation for the RMS delay spread on that street. The k-power-means RMS DS (magenta) plot is more optimistic (i.e., smaller RMS DS) because power is the major weight factor. MPCs with close Rx power levels are many times also close in time; hence a tighter grouping around the dominant ray in each cluster. The k-means RMS DS (brown) plot is a fifty-fifty split around the RMS DS (blue) plot that represents the ray-tracer values because only the 5-dimensional space (ToA, and azimuth and elevation for AoA, AoD) is used for clustering. Another reason for the skew between plots is the very limited number of simulations we considered at each of the 14 receivers. Despite that, the conclusion is that a wireless network architect can use (in the first stage) the fitted truncated normal distributions of the clustered RMS DS (Fig. 8) to get a good estimate of the delay spread and maximum data rate in the channel without using a ray-tracer or measuring channel parameters with expensive, dedicated hardware. A similar approximation is possible for the RMS angular spread (AS) for the azimuth of the AoA (Fig. 10).

In addition to RMS DS and RMS AS, in Table III, we also

summarize other *inter-cluster* parameters part of our analysis. We check the *cluster power decay rate* Γ defined as the decay rate of the strongest path within each cluster, and the *cluster inter-arrival time* defined as the relative delay between two adjacent clusters. The delay of a cluster is taken as the delay of the strongest MPC in the cluster. The inverse of the mean inter-cluster arrival time is the inter-cluster arrival rate Λ. Similar to [4], we prove that the cluster peak power variation could be modeled with a normal distribution in the dB-domain while the cluster inter-arrival time is described by a Poisson process, modeled by an exponential distribution. Besides inter-cluster parameters, there are *intra-cluster* parameters that we only mention here. They can be described in the *time domain* by the average number of rays, ray arrival rate, and ray power decay time, and in the *angular domain* by cluster azimuth and elevation spread.

TABLE III
INTER-CLUSTER PARAMETERS AND THEIR DISTRIBUTION PARAMETERS

Cluster Parameter/Alg	k-means	k-power-means
No of clusters (μ/σ)	3.9286 / 1.3848	-
Power decay (μ/σ)	(-145.06) / 37.184	(-114.13) / 13.677
Arrival rate $1/\Lambda$ [ns]	822	610
RMS DS (μ/σ)	7.22E(-8) / 6.69E(-8)	5.08E(-8) / 6.44E(-8)
RMS AS (μ/σ)	0.18503 / 0.16093	0.12817 / 0.1058

IV. CONCLUSIONS

The paper is concerned with providing accurate channel models for mmWave urban outdoor transmissions at 28 GHz based on clustering the channel parameters estimated with a professional ray-tracer. To cluster data, we compare k-means and k-power-means clustering. Their results are validated using cluster validity indices and score fusion techniques. We tackle both large-scale and small-scale aspects of radio channel models, namely, path loss, RMS delay spread and RMS angle spread. Other inter-cluster parameters like cluster power decay and cluster arrival rate are also investigated. The conclusion is that even for a small set of estimations obtained with the ray-tracer, we still have a good approximation of the channel propagation model for the chosen outdoor scenario. All our observations about the number of clusters, their dominant MPCs and RMS delay spread values emphasize that the clustering nature of the multipath channel has a vital role in the link capacity of the mmWave communication system.

ACKNOWLEDGMENTS

This work is supported in part by MathWorks under Research Grant "Cross-layer Approach to 5G: Models and Protocols." Stefano Basagni was also supported in part by the NSF grant CNS 1925601 "CCRI: Grand: Colosseum: Opening and Expanding the World's Largest Wireless Network Emulator to the Wireless Networking Community."

REFERENCES

[1] G. Hamerly and C. Elkan, "Alternatives to the k-means algorithm that find better clusterings," in *Proceedings of the eleventh International Conference on Information and Knowledge Management (CIKM)*, McLean, VA, November 4–9 2002, pp. 600–607.

[2] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2nd edition, 2016.

[3] M. T. Martinez-Ingles, D. P. Gaillot, J. Pascual-Garcia, J. M. Molina Garcia-Pardo, M. Lienard, J. V. Rodríguez, and L. Juan-Llacer, "Impact of clustering at mmW band frequencies," in *Proceedings of the IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, Vancouver, BC, Canada, July 19–24 2015, pp. 1009–1010.

[4] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-Wave Multipath Clustering and Channel Modeling," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 3, pp. 1445–1455, March 2014.

[5] N. Czink, P. Cera, J. Salo, E. Bonek, J. P. Nuutinen, and J. Ylitalo, "A Framework for Automatic Clustering of Parametric MIMO Channel Data Including Path Powers," in *Proceedings of the IEEE 64th Vehicular Technology Conference (VTC Fall)*, Montreal, Quebec, Canada, September 25–28 2006, pp. 1–5.

[6] N. Czink, R. Tian, S. Wyne, F. Tufvesson, J. P. Nuutinen, J. Ylitalo, E. Bonek, and A. F. Molisch, "Tracking Time-Variant Cluster Parameters in MIMO Channel Measurements," in *Proceedings of the Second International Conference on Communications and Networking, CHINACOM*, Shanghai, China, August 22–24 2007, pp. 1147–1151.

[7] N. Czink, P. Cera, J. Salo, E. Bonek, J. P. Nuutinen, and J. Ylitalo, "Improving clustering performance using multipath component distance," *Electronics Letters*, vol. 42, no. 1, pp. 33–45, January 2006.

[8] M. Tehrani Moayyed, B. Antonescu, and S. Basagni, "Clustering algorithms and validation indices for mmWave radio multipath propagation," in *Proceedings of the Wireless Telecommunications Symposium WTS2019*, April 9–12 2019, pp. 1–7.

[9] T. Caliński and J. Harabasz, "A Dendrite Method for Cluster Analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, January 1974.

[10] D. L. Davies and D. W. Bouldin, "A Cluster Separation Measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, April 1979.

[11] J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 28, no. 3, pp. 301–315, June 1998.

[12] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, September 1973.

[13] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 841–847, August 1991.

[14] M. K. Pakhira, S. Bandyopadhyay, and U. Maulik, "Validity index for crisp and fuzzy clusters," *Pattern Recognition*, vol. 37, no. 3, pp. 487–501, March 2004.

[15] K. Kryszczuk and P. Hurley, "Estimation of the Number of Clusters Using Multiple Clustering Validity Indices," in *Proceedings of the 9th International Workshop on Multiple Classifier Systems, Cairo, Egypt*, Springer, Berlin, Heidelberg, April 7–9 2010, pp. 114–123.

[16] B. Antonescu, M. Tehrani Moayyed, and S. Basagni, "Clustering Algorithms and Validation Indices for a Wide mmWave Spectrum," *Information Journal*, vol. 10, no. 9, pp. 1–17, September 2019.

[17] X. Wu, C. X. Wang, J. Sun, J. Huang, R. Feng, Y. Yang, and X. Ge, "60-GHz Millimeter-Wave Channel Measurements and Modeling for Indoor Office Environments," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 4, pp. 1912–1924, April 2017.

[18] Y. Yang, J. Sun, W. Zhang, C. X. Wang, and X. Ge, "Ray Tracing Based 60 GHz Channel Clustering and Analysis in Staircase Environment," in *Proceedings of the IEEE Global Communications Conference GLOBECOM 2017*, Singapore, Singapore, December 4–8 2017.

[19] C. Ling, X. Yin, R. Müller, S. Häfner, D. Dupleich, C. Schneider, J. Luo, H. Yan, and R. Thomä, "Double-Directional Dual-Polarimetric Cluster-Based Characterization of 70–77 GHz Indoor Channels," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 2, pp. 857–870, February 2018.

[20] B. Antonescu, M. Tehrani Moayyed, and S. Basagni, "mmWave Channel Propagation Modeling for V2X Communication Systems," in *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC2017*, Montreal, QC, Canada, October 8–13 2017.

[21] Rappaport, T. et al., "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.