# Clustering-based mmWave Channel Propagation Models for Outdoor Urban Scenarios


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 work investigates the effectiveness of $k$-means and $k$-power-means clustering algorithms in predicting the number of clusters through the use of cluster validity indices (CVIs) and score fusion techniques. Our results show how these two solutions generate an accurate approximation of the mmWave channel model, 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

In this work we are concerned with *multipath* and with the analysis of sorting and grouping the received rays into clusters. A *cluster* is defined as a group of rays with similar attenuation and angular profile. Channel parameters like received power, Time-of-Arrival (ToA), Angle-of-Arrival (AoA) and Angle-of-Departure (AoD) are reported in our simulations by the ray-tracer tool for each arriving ray at the receiver. These parameters are fed into well known *center-based* clustering algorithms, namely, $k$-means [1] and one of its variants, $k$-power-means, in which the input gets partitioned around few centroids or central points. Each MPC is assigned to a specific cluster by calculating the distance to the centroids and choosing the closest one. There is an improvement if we use few channel parameters jointly by replacing the Euclidean distance with the *multipath component distance* (MCD) [2]:

$$MCD_{ij} = \sqrt{||MCD_{AoA,ij}||^2 + ||MCD_{AoD,ij}||^2 + ||MCD_{\tau,ij}||^2}, \quad (1)$$

where $i$ and $j$ are any two estimated MPCs. We use $k$-power-means algorithm with the same distance metric MCD, weighted by the power values $P_l$ of the MPCs: $D = \sum_{l=1}^{L} P_l \cdot MCD(x_l, c_{I_l^{(i)}})$, where index $I_l^{(i)}$ is the cluster number for the $l$-th MPC in the $i$-th iteration.

The preferred clustering solution for a specific algorithm is obtained by finding the value of $K$ (in a certain range) that provides the optimal (min or max) value of a Cluster Validity Index (CVI). The CVIs used in this work are the following: Calinski-Harabasz (CH) [3], Davies-Bouldin (DB) [4], generalized Dunn (GD) [5], Xie-Benie (XB) [6] and PBM [7]. Since no single CVI can capture correctly the validity of any clustering solution (i.e., work well with all data sets), we use a conciliation of multiple CVIs through *score fusion-based* techniques: $SF_a = \frac{1}{M}\sum_{i=1}^{M}\nu_i; SF_g = \left(\prod_{i=1}^{M}\nu_i\right)^{\frac{1}{M}}; SF_h = M\left(\sum_{i=1}^{M}\frac{1}{\nu_i}\right)^{-1}$.

## II. Simulation Results

We simulate 28 GHz communications between transmitter and receiver units using one of the urban scenarios (Rosslyn, VA) delivered with Remcom's ray-tracing tool Wireless InSite (Fig. 1). 22°/15 dBi antennas with maximum transmitted

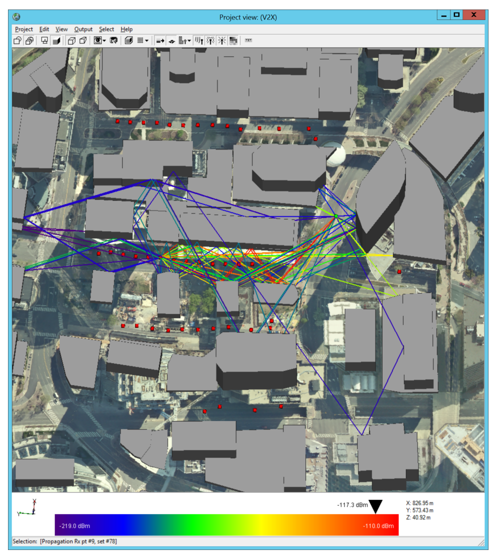


Fig. 1. 44 MPCs at receiver Rx#9.

power of 24 dBm are used at both Tx and Rx locations. We capture the values of the received power, excess delay, angle-of-arrival and angle-of-departure of all MPCs arriving at each randomly placed Rx point. The real part of the complex impulse response (CIR) for this one-time channel realization (Fig. 2) shows the received power levels of all MPCs and their ToA. 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. The results of

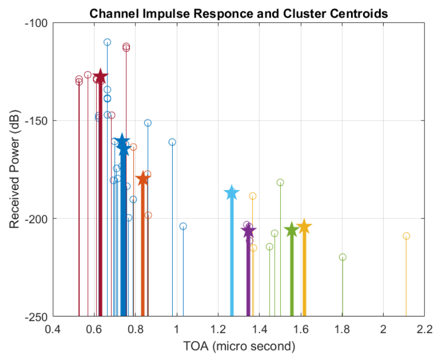


Fig. 2. Clustered CIR at Rx#9 based on $k$-means with MCD.

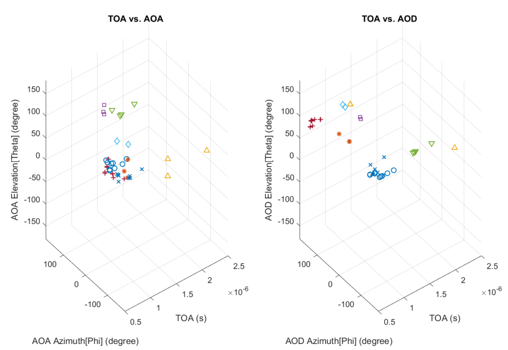


Fig. 3. Clustering via $k$-means algorithm—ToA vs. AoA, AoD.

the clustering process are validated and the optimal $K$ value is found by applying the CVIs and score fusion techniques mentioned in Section I.

We are primarily concerned with the analysis of the root mean square (RMS) delay spread (DS), as this parameter is tightly connected with the maximum data rate achievable in the channel. We capture the delay spread reported by the ray-tracer at each of the $14$ locations on the street, and we compare with the RMS DS values calculated for each cluster based on the partitioning obtained with both variants of $k$-means algorithm (Fig. 4). Plots of the Cumulative Distribution Function (CDF) of the clustered RMS delay spread for both variants of $k$-means are captured in Fig. 5. The statistics extracted from the clustered solutions of the RMS DS represent an accurate approximation of the values estimated by the ray tracer without clustering, thus providing a first order of magnitude estimate of the delay spread and maximum data rate in the channel.

We are also interested in two other sets of parameters, inter- and intra-cluster parameters that describe the clusters and the rays in each cluster (Table I). The *inter-cluster* parameters are the *cluster power decay rate* $\Gamma$ (i.e., the decay rate of the strongest path within each cluster), and the *cluster inter-arrival time* (i.e., the relative delay between two adjacent clusters). The *intra-cluster* parameters are 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.

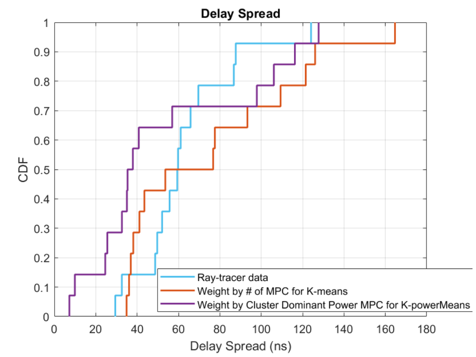


Fig. 4. CDFs of the RMS DS for ray-tracer values and cluster-based values.

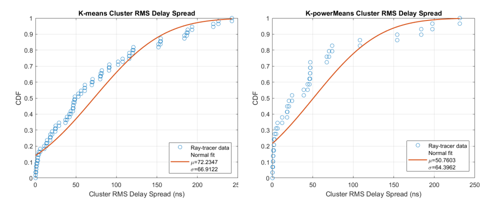


Fig. 5. CDF and truncated normal distribution of clustered RMS DS.

TABLE I
INTER-CLUSTER PARAMETERS AND THEIR DISTRIBUTION PARAMETERS

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

## ACKNOWLEDGMENTS

This work is supported in part by MathWorks under Research Grant “Cross-layer Approach to 5G: Models and Protocols.”

## REFERENCES

[1] 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.
[2] 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.
[3] T. Caliński and J. Harabasz, “A Dendrite Method for Cluster Analysis,” *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, January 1974.
[4] 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.
[5] 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.
[6] 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.
[7] 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.