High-Accuracy Visualization-Based Grouping of MIMO Multipath Waves

Emmanuel T. Trinidad[1,2*] and Lawrence Materum[2,3]
[1] Don Honorio Ventura State University, Bacolor, Philippines
[2] De La Salle University, Manila, Philippines
[3] Tokyo City University, Tokyo, Japan
*Email: emmanuel_trinidad@dlsu.edu.ph

***Abstract*—**Wireless multipath propagation causes different paths taken by the signal due to interacting objects present in the environment producing multipath components (MPCs). Cluster-based channel models characterize the wireless channel, and different approaches are utilized to cluster the MPCs. Data mining requires different techniques such as visualization to extract important information and find patterns and clusters in the data. A Graphical User Interface (GUI) is developed in this paper to aid the visualization and the manual clustering of MPCs using t-distributed Stochastic Neighborhood Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP). The clustering results of Simultaneous Clustering and Model Selection (SCAMS) are used in this paper. The datasets are embedded into low-dimensional projection and are manually re-clustered. The manual clustering was performed visually and interactively, which achieves a higher Jaccard membership index with a low value of 0.3368, a median of 0.4697, and a high value of 0.8884 for all the datasets.

***Index Terms*—**Graphical user interface, multipath clustering, interactive clustering, wireless propagation, dimensionality reduction

I. Introduction

Wireless channel modeling has been an integral part of developing wireless communication systems. The demand for seamless wireless connectivity poses a massive challenge to system designers. The use of Multiple-Input Multiple-Output (MIMO) systems is one of the enabling technologies of new standards such as the Fifth Generation (5G) mobile cellular systems and Wireless Local Area Networks (WLAN). Utilizing MIMO technology opens up benefits for different use cases, namely beamforming, spatial multiplexing, and diversity, where the trade-off between these produces optimum use cases of the wireless channel [1]. Furthermore, MIMO bridges data mining and channel modeling in which machine learning can be applied to various challenges in channel modeling [2]. Multipath components (MPCs) are produced due to interacting objects in the propagation environment known as scatterers. Literature shows that MPCs arrive in clusters, and an accurate clustering allows a simplified analysis in getting the Channel State Information (CSI). Furthermore, studies have attributed clusters to the scatterers present in the environment.

Cluster-based channel models have been developed to analyze the transmission in a geometric approach that makes use of the concept of clusters. The COST 2100 Channel Model (C2CM) is a geometry-based stochastic model that can replicate MIMO channels' behavior [3]. Techniques for clustering the MPCs on different parameters are proposed in the literature [4]–[6]. Traditionally, the manual approach was used but tended to be laborious, especially for large datasets, and is quite subjective [7], [8]. Recently, machine learning techniques have been employed to cluster the multipaths automatically. A framework in [9] is proposed using the Multipath Component Distance (MCD) to suit the angular nature of the measured data. Currently, no clustering technique outperforms the other approaches, which leads to different accuracy results in clustering the MPCs.

The limitation of the human eye to visualize more than 3-Dimensional (3D) plots poses a challenge in seeing all the features of data points. In Fig. 1, an example of projecting three features of one snapshot of a dataset for a semi-urban scenario generated from C2CM is illustrated. The groupings are not easily identified via manual inspection, which can be tedious if done. Also, the elevation and azimuth of departure are not accounted for in the visualization.

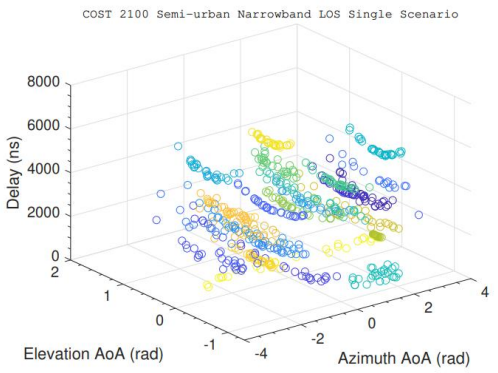


Fig. 1. Visualization of MPC features [10]

Recently, Dimensionality Reduction (DR) techniques have attracted different fields to embed high-dimensional data into low-dimensional plots for visualization. The produced visualization shows cluster tendencies of the data. Visualization aids the manual clustering of MPC and also

Manuscript received May 15, 2022; revised December 13, 2022; accepted January 4, 2023.

validates the automatic clustering performance, establishing a middle-ground approach [11]. Interactive clustering requires a human-in-the-loop process and domain knowledge enables the user to visualize, modify, reject, and or accept the clustering results [12]. Additionally, using a Graphical User Interface (GUI) allows the human-computer interaction, which combines the computing power and domain knowledge of the user resulting in more interpretability and understanding of the clustering results.

This paper proposes a GUI that enables DR techniques to visualize MPCs, refine the cluster, and increase the cluster membership accuracy. The rest of the paper is organized as follows. Section II presents related works in the literature regarding the clustering of MPCs and the use of DR techniques. Section III discusses the techniques used to achieve the goal, followed by Section IV, which presents the methodology. The results are analyzed and discussed in Section V, and Section VI concludes this paper.

II. Related Works

Interfaces can aid in interpreting data, its natural clusters, and the validation of clustering performance. Many interfaces are developed to project and visualize data using PCA, parallel coordinates, scatterplots, and heatmaps [13]–[15]. Multi-dimensional techniques also receive attention in providing the user interactivity and analysis shown in [16]. Gene expression is one of the fields that use cluster analysis. A GUI is developed using Matlab in [16] equipped with external and internal validity indices using a gene expression dataset. A clustering tour is introduced in [17], where t-SNE is used on the dataset and projects an interactive and guided analysis.

In multipath clustering, numerous methods are used to project the features of MPCs, such as angular elevations, azimuths, and the time delay. In [18], the correct number of clusters is evaluated using a factor-inclusion weighing approach where a GUI is presented to visually compare the effective weight for scenarios. Moreover, a tool for visualizing small interacting objects and their location using 3D point cloud data in an urban propagation environment is introduced [19]. Using a GUI in [20], a comparison of four clustering techniques is presented where the user can choose the algorithm and display the accuracy. In projecting the MPCs, some features are not represented due to the limitation of 3D visuals. Visual projections were used to project the MPC in [21], where parallel plot coordinates, heatmap, and t-SNE were used in MPC data, but lacks interactivity. Hence, the aim of this paper is to visually project clustering results using a GUI and modify the cluster membership manually to increase the accuracy further.

III. Dimensionality Reduction Techniques

DR techniques are classified into linear and non-linear methods. This paper uses three DR techniques to project the MPC that assists the manual clustering process through visualization.

A. PCA

PCA is one of the oldest techniques that is still being used today. PCA is a linear DR technique that intends to capture the variance of the dataset and construct a linear combination of the original variables. PCA preserves the global structure of the data and uses orthogonal transformation that captures the maximum variance. PCA can be used to reduce the data to a lower dimension by retaining the principal components. Furthermore, PCA can be used to decorrelate the data and for sphering the data [2], [22], and this approach is used in this study.

B. t-distributed Stochastic Neighborhood Embedding (t-SNE)

t-SNE is classified as a non-linear DR technique that uses the t-distribution, which addresses the crowding problem of SNE. The t-SNE algorithm aims to represent high-dimensional data and project it into low-dimension for visualization [23]. The non-linear structure of the data is captured by t-SNE and seeks to preserve the local structure focusing on the neighboring points. For setting the number of neighboring points in the visualization, t-SNE uses a hyperparameter called perplexity. The **P** and **Q** are the distribution of the pairwise probability in the high and low dimensions in which the Kullback-Leibner divergence is minimized using gradient descent.

C. Uniform Manifold Approximation and Projection (UMAP)

A recent technique named UMAP is also a non-linear or manifold learning DR technique constructed based on algebraic topology [24]. As with t-SNE, the neighboring points can be specified in UMAP using the parameter min points. The UMAP algorithm can be divided into two significant steps: graph construction and layout. The probability distribution that the i^{th} and j^{th} is similar is computed using a distance metric, usually the Euclidean distance. The likelihood graph is then constructed for the similarities **P** and **Q** are the probabilistic similarity in high and low dimensions, respectively. For optimizing the low-dimensional graph, the cost function used is the binary cross-entropy to capture the data's global structure. The Matlab implementation in [25] was used in this study.

D. Multipath Component Distance (MCD)

The MCD is used in different techniques to quantify the separation between the i^{th} and the j^{th} MPC [26]. It was first introduced in automatic clustering algorithms in [27]. The angle of arrival (AOA) and angle of departure (AOD) distances are calculated separately through Eq. 1:

$$\text{MCD}_{\text{AoA/AoD},ij} = \frac{1}{2}\left|\begin{pmatrix}\sin(\theta_i)\cos(\varphi_i)\\ \sin(\theta_i)\sin(\varphi_i)\\ \cos(\theta_i)\end{pmatrix} - \begin{pmatrix}\sin(\theta_j)\cos(\varphi_j)\\ \sin(\theta_j)\sin(\varphi_j)\\ \cos(\theta_j)\end{pmatrix}\right| \quad (1)$$

The distance between the delay parameter τ is obtained by Eq. 2:

$$\text{MCD}_{\tau,ij} = \zeta \cdot \frac{|\tau_i - \tau_j|}{\Delta\tau_{max}} \cdot \frac{\tau_{\text{std}}}{\Delta\tau_{max}} \quad (2)$$

where τ_{std} is the standard deviation of the delays, and ζ represents a scaling factor for the delays. Finally, Eq. (3) calculates the distance between two MPCs:

$$\text{MCD}_{ij} = \sqrt{\left\|\text{MCD}_{\text{AoA},ij}\right\|^2 + \left\|\text{MCD}_{\text{AoD},ij}\right\|^2 + \text{MCD}_{\tau,ij}^2} \quad (3)$$

The t-SNE and UMAP use a distance metric to calculate the pairwise distance of points in high-dimensional space. In this work, the MCD is integrated as a custom distance function for both the DR algorithms. However, the clusters cannot be readily determined by examination after projecting the data with ground-truth values. The PCA was utilized as a pre-processing step where the number of components was preserved and displayed a more optimistic projection. Consequently, the method used in this paper is to apply PCA first before the DR procedures. Fig. 2 illustrates one snapshot per dataset with the ground-truth labels using PCA, t-SNE, and the MCD. The perplexity and number of minimum points of t-SNE and UMAP were set to $\sqrt{N}$, where N is the number of MPC per snapshot. Utilizing UMAP with the same procedure, the results are presented in Fig. 3.

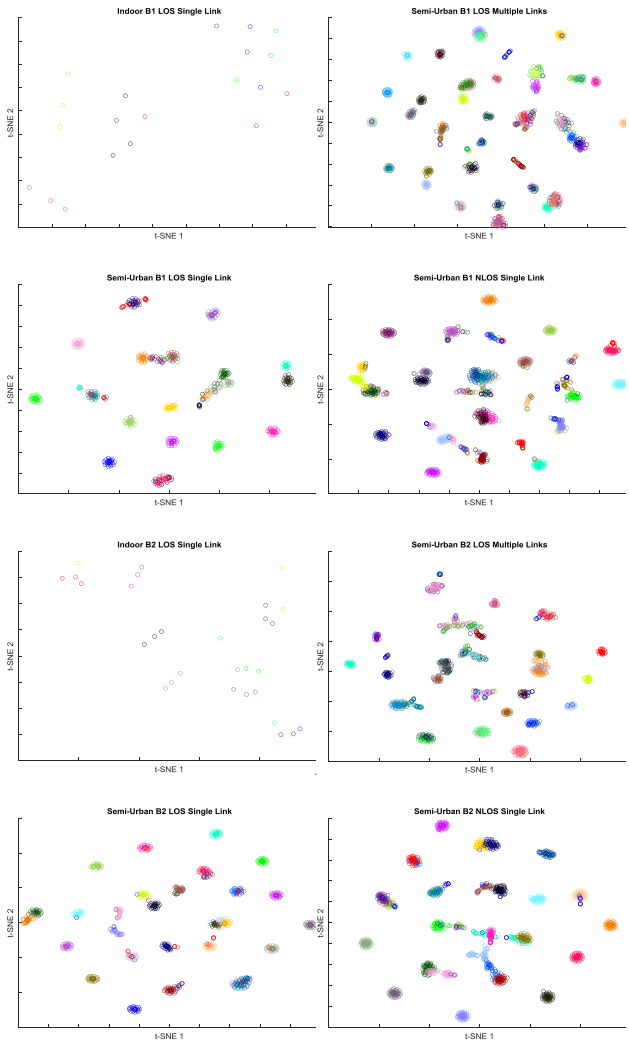


Fig. 2. Ground-truth projection using PCA+ t-SNE and MCD

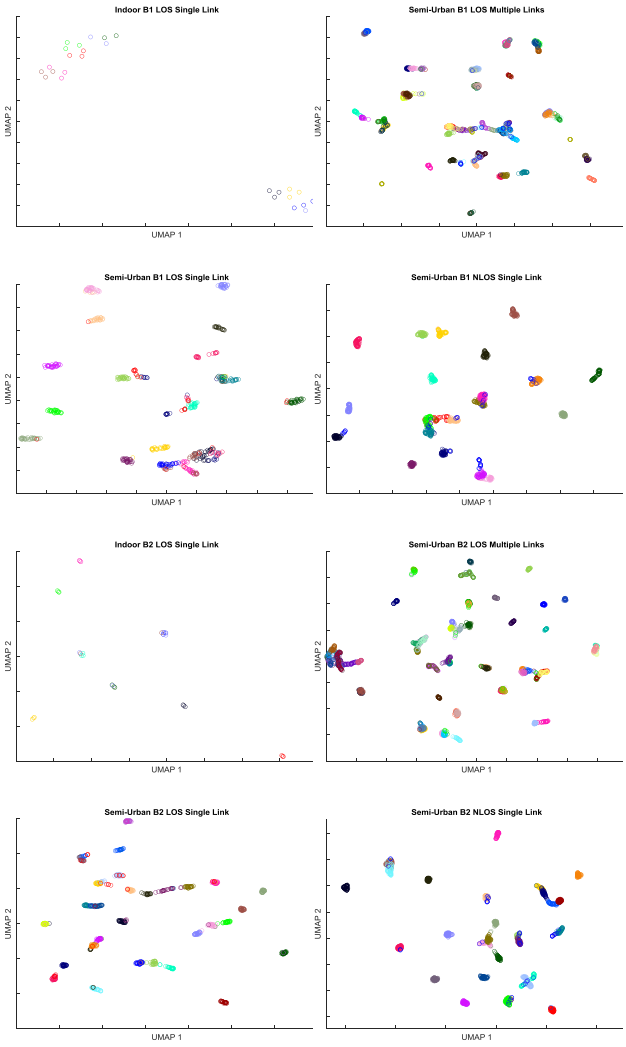


Fig. 3. Ground-truth Projection using PCA+ UMAP and MCD

Fig. 2 and Fig. 3 project the ground truth, showing that the true cluster members are close to one another, which can visually aid the manual refinement process. By visual inspection, the approach of using the DR techniques assists in identifying the membership of each cluster quickly, especially for semi-urban scenarios with large MPCs.

IV. Methodology

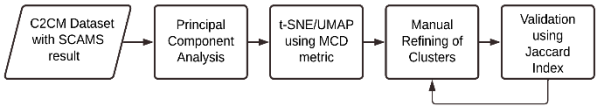


Fig. 4. Methodology of the study

The methods used in this study are discussed in this section. The datasets used, the design of the GUI, and the cluster validation are discussed. Fig. 4 summarizes the methodology used in this paper. The approach of this paper is to incorporate visualization as an additional step in the clustering process which further improves the accuracy of the automatic clustering approaches. The use of the MCD in calculating the distance in high-dimension is incorporated into the DR techniques. With the use of

validity indices, the results of the automatic approach of SCAMS have been increased by manually refining the membership using the projections.

A. Datasets

The datasets used in this study are taken from the Institute of Electrical and Electronics Engineers (IEEE) data port [10]. The C2CM was used to produce MPC data from eight-channel scenarios, each with 30 trials of snapshots listed as follows:

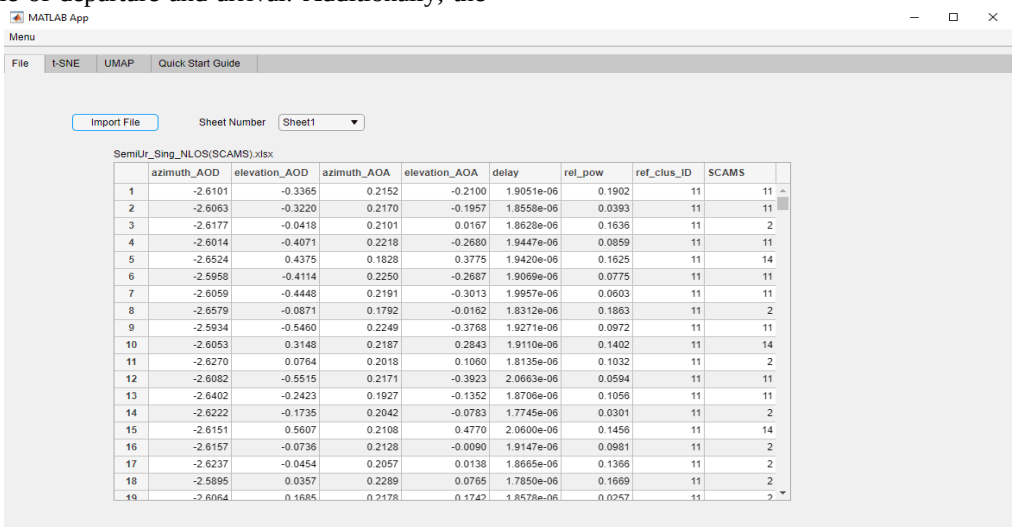
1. B1, indoor, single link, line-of-sight
2. B1, semi-urban, multiple links, line-of-sight
3. B1, semi-urban, single link, line-of-sight
4. B1, semi-urban, single link, non-line-of-sight
5. B2, indoor, single link, line-of-sight
6. B2, semi-urban, multiple links, line-of-sight
7. B2, semi-urban, single link, line-of-sight
8. B2, semi-urban, single link, non-line-of-sight

Each snapshot of the dataset is represented by matrix $\boldsymbol{X} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots \boldsymbol{x_L}\}$ which contains rows of MPC represented by $\boldsymbol{x}_\ell = [\tau_\ell, \theta_{\ell,\mathrm{AOA}}, \varphi_{\ell,\mathrm{AOA}}, \theta_{\ell,\mathrm{AOD}}, \varphi_{\ell,\mathrm{AOD}}]$ where $\boldsymbol{x}_\ell$ represents the ℓ^{th} MPC, τ represents delay, θ is the elevation angle of departure and arrival, and φ is the azimuth angle of departure and arrival. Additionally, the reference cluster membership of each MPC per snapshot is also provided.

The clustering results from SCAMS presented in [28] are used in this study as the clustering results to be modified. The computed cluster membership of the algorithm is manually refined to increase the accuracy further.

B. GUI Design

The GUI is developed using the Matlab App Designer, where components are dragged in a canvas making it more flexible in positioning components. The components are then given some callback functions to perform the necessary outcomes. The GUI is designed with four tabs, one for the file input, two for the t-SNE and UMAP algorithm, and one tab for the quick start guide. Additionally, the Jaccard index is calculated with the given input data clustering results and after merging or modifying the clusters. The File tab shown in Fig. 5 allows the user to input an Excel file with reference and clustering results as the final two columns of the file. The reason for this is to let the interface know the reference and cluster result for the Jaccard computation.



	azimuth_AOD	elevation_AOD	azimuth_AOA	elevation_AOA	delay	rel_pow	ref_clus_ID	SCAMS
1	-2.6101	-0.3365	0.2152	-0.2100	1.9051e-06	0.1902	11	11
2	-2.6063	-0.3220	0.2170	-0.1957	1.8558e-06	0.0393	11	11
3	-2.6177	-0.0418	0.2101	0.0167	1.8628e-06	0.1636	11	2
4	-2.6014	-0.4071	0.2218	-0.2680	1.9447e-06	0.0859	11	11
5	-2.6524	0.4375	0.1828	0.3775	1.9420e-06	0.1625	11	14
6	-2.5958	-0.4114	0.2250	-0.2687	1.9069e-06	0.0775	11	11
7	-2.6059	-0.4448	0.2191	-0.3013	1.9957e-06	0.0603	11	11
8	-2.6579	-0.0871	0.1792	-0.0162	1.8312e-06	0.1863	11	2
9	-2.5934	-0.5460	0.2249	-0.3768	1.9271e-06	0.0972	11	11
10	-2.6053	0.3148	0.2187	0.2843	1.9110e-06	0.1402	11	14
11	-2.6270	0.0764	0.2018	0.1060	1.8135e-06	0.1032	11	2
12	-2.6082	-0.5515	0.2171	-0.3923	2.0663e-06	0.0594	11	11
13	-2.6402	-0.2423	0.1927	-0.1352	1.8706e-06	0.1056	11	11
14	-2.6222	-0.1735	0.2042	-0.0783	1.7745e-06	0.0301	11	2
15	-2.6151	0.5607	0.2108	0.4770	2.0600e-06	0.1456	11	14
16	-2.6157	-0.0736	0.2128	-0.0090	1.9147e-06	0.0981	11	2
17	-2.6237	-0.0454	0.2057	0.0138	1.8665e-06	0.1366	11	2
18	-2.5895	0.0357	0.2289	0.0765	1.7850e-06	0.1669	11	2
19	-2.6064	0.1685	0.2178	0.1742	1.8578e-06	0.0257	11	2

Fig. 5. File tab of the GUI

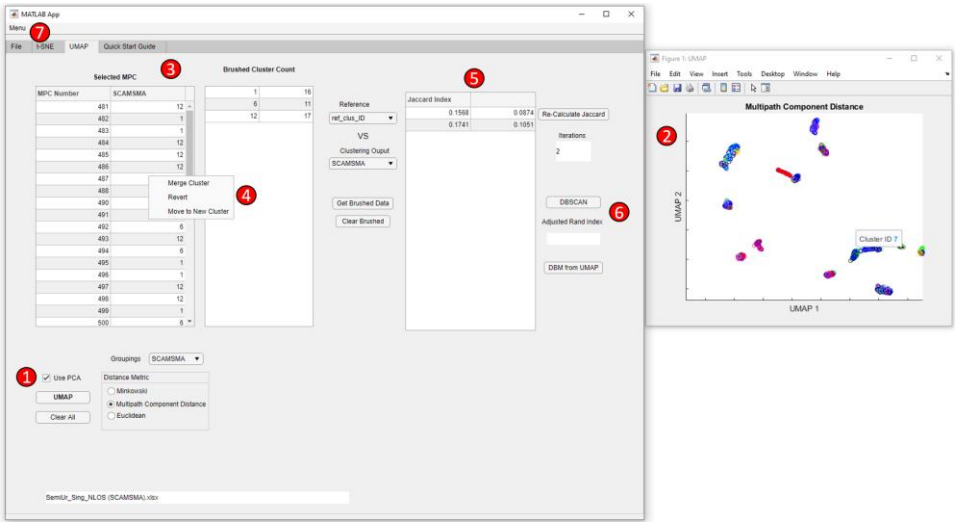


Fig. 6. GUI layout and components

Fig. 6 illustrates the GUI developed with the components that are used interactively. Component ① lets the user select the option to use PCA and the distance metric to be used. The distance metric included are Minkowski, Euclidean, and the MCD. After selecting the parameters, through the t-SNE/UMAP button, the low-dimension representation is computed, and then the results are graphed in a scatterplot represented in component ②. Furthermore, inside this plot, the groups of MPCs are brushed and selected to modify their cluster membership. Component ③ is a table of the brushed MPCs and ④ is a selection menu to merge clusters, revert to their original cluster ID, or assign a new cluster to the selected MPCs. A table is shown calculated inside the GUI ⑤ to view an increase or decrease in the accuracy. Lastly, the component ⑦ allows the user to save the file with the original dataset, the manual cluster IDs, and the transformed variables from t-SNE and UMAP in an Excel sheet. Components of the GUI are summarized in Table I.

TABLE I. SUMMARY OF COMPONENTS IN THE DR TABS

Component	Function
①	Parameter selection with the use of PCA and distance metric to be used
②	Projection of MPC using the selected algorithm (t-SNE or UMAP)
③	Brushed MPC counts table
④	Cluster membership refining pane
⑤	Jaccard index table
⑥	DBSCAN button (future work)
⑦	Menu Bar for saving and DR Tabs

C. Clustering Accuracy

For evaluating the accuracy of the cluster membership of the manually refined clusters, the Jaccard index η is used. The Jaccard index ranges from 0 to 1, where 1 represents a perfect agreement between reference and calculated cluster membership. Jaccard index is computed mathematically in Eq. (4):

$$\eta_{\text{Jac}} = \frac{M_{11}}{M_{11} + M_{10} + M_{01}} \in [0,1] \tag{4}$$

where M_{11} is the number of members that are both present in the calculated and reference cluster, M_{10} represents the members presented in the calculated but not in the reference cluster while M_{01} is the number of members present in the reference but not in the calculated clusters.

V. RESULTS AND DISCUSSIONS

The combined PCA, t-SNE/UMAP, and MCD show good visualization, and clusters can be seen by inspection. These parameters are used to modify the SCAMS result and achieve a higher Jaccard membership index. The use of t-SNE and UMAP are used alternatively and sequentially, depending on the nature of the dataset. The process of interaction inside the GUI is illustrated in Fig. 7. Importing one data sheet per scenario and selecting one snapshot is the first step, then selecting the t-SNE or UMAP tab. The usage of PCA and distance metrics is part of the process, whereas the MCD is employed in this study for better visualization. Projecting one snapshot of the data with the SCAMS results inside the plot is displayed in Fig. 8a, which only has a 0.1968 Jaccard membership index. The discrepancy in color labels is apparent because SCAMS has low cluster member accuracy in semi-urban scenarios. This process is repeated until visually noticeable groups of MPCs are processed, as shown in Fig. 8b, which achieves a 0.5051 Jaccard membership index after 39 iterations.

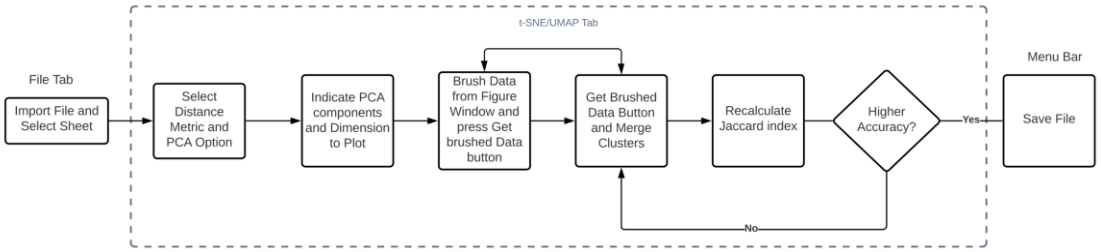
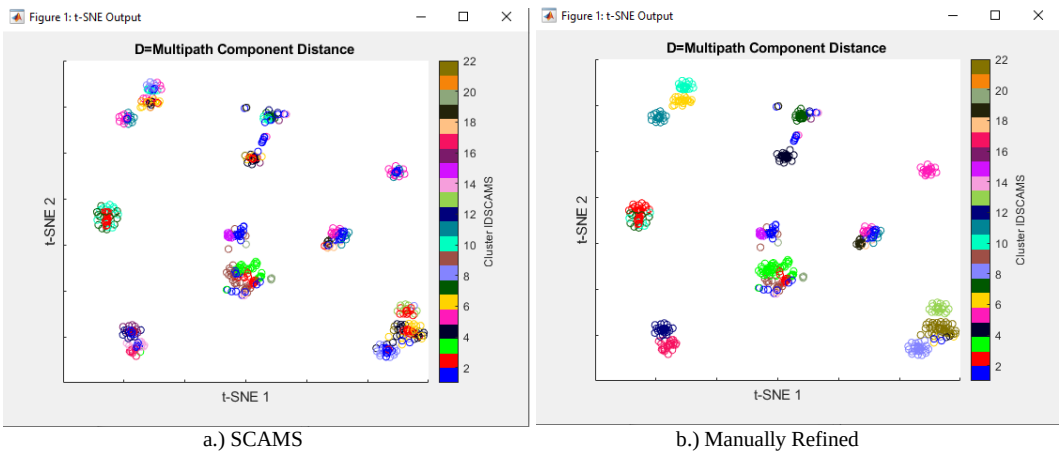


Fig. 7. GUI interaction process



a.) SCAMS b.) Manually Refined

Fig. 8. Projection of MPCs before and after manual refinement

For semi-urban LOS and NLOS, the t-SNE was heavily used, while UMAP was mainly used for indoor and multiple-links scenarios. A total of 227 datasets were manually improved, with indoor datasets with a Jaccard index of one not included for improvement. The mean results of the SCAMS output and this work are summarized in Table II.

TABLE II. RESULTS OF MANUAL CLUSTERING JACCARD MEMBERSHIP

Channel Scenario	SCAMS Output		This Work		Iterations
	Mean	Standard Deviation	Mean	Standard Deviation	Mean
B1 Indoor LOS	0.66984	0.18943	0.86535	0.12856	11.88
B1 Semi-Urban Multiple Links LOS	0.14588	0.01213	0.31511	0.11220	75.37
B1 Semi-Urban Single Link LOS	0.18820	0.02845	0.54599	0.07288	50.63
B1 Semi-Urban Single Link NLOS	0.15975	0.02729	0.45666	0.007728	60.70
B1 Indoor LOS	0.65043	0.185074	0.846619	0.1510	10.19
B1 Semi-Urban Multiple Links LOS	0.14499	0.01401	0.37029	0.04542	89.66
B1 Semi-Urban Single Link LOS	0.18301	0.02120	0.48687	0.06188	48.17
B1 Semi-Urban Single Link NLOS	0.15073	0.02473	0.43651	0.08237	56.30

The mean percentage Jaccard index increase for bands 1 and 2 is 29.19% and 30.16%, respectively. On the other hand, the lowest increase among the semi-urban scenarios is in the multiple links, with 140.48% for B1 and 153.43% for B2. The increase for the LOS semi-urban link was calculated to be 190.11% for B1, while a 166.04% increase was achieved for B2. Lastly, for the semi-urban NLOS single link, 185.87% for B1 and an increase of 189.60% for B2. The increase in the indoor datasets was lower than the semi-urban datasets because the indoor datasets already had a high Jaccard index from SCAMS. The approach is sensitive to multiple links datasets and has a lower increase due to overlapping clusters, as seen in the ground-truth projections. Additionally, the multiple links have the highest number of iterations due to the cluster overlaps.

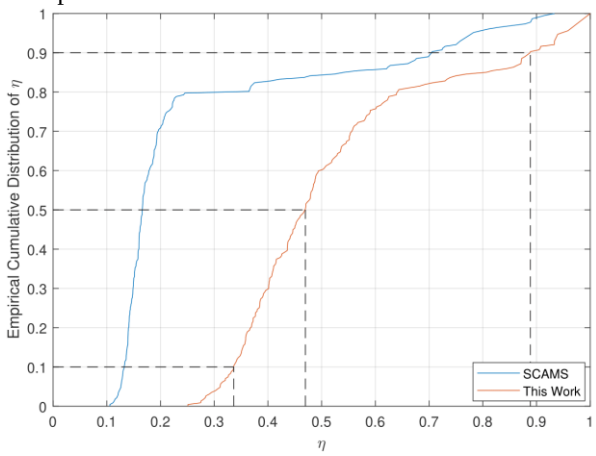


Fig. 9. Comparison of ECDF for all scenarios

In Fig. 9, the empirical cumulative distribution function (ECDF) for the Jaccard indices of all the scenarios is illustrated. The ECDF of SCAMS output and the result of this work shows an increase in the 10^{th} 50^{th} and 90^{th} percentile. For the 10^{th} percentile, the SCAMS has 0.1324, while the result in this study attains 0.3363 with a difference of 0.20392 improvement. The median for SCAMS and this work is 0.1663 and 0.4697, respectively, resulting in a 0.3034 increase. Finally, in the 90^{th} percentile, a 0.1825 difference is achieved from 0.7059 of SCAMS and 0.8884 in this method.

VI. CONCLUSION

The results of using DR techniques in projecting clusters of wireless multipaths are presented in this work. Using PCA, t-SNE, UMAP, and MCD can aid the validation and refining of cluster membership. A GUI is developed and used to modify the SCAMS clustering results of the C2CM datasets and aids the human-in-the-loop process and interactivity in clustering. The results suggest that the strategy of utilizing DR results in a significant increase in Jaccard accuracy. On the other hand, the semi-urban multiple linkages provide an issue in visually distinguishing overlapping clusters, resulting in lesser accuracy than the other semi-urban instances. Although manual refining seems laborious, automatic clustering techniques can be used in the dimensionally reduced data, which is considered for future work.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Emmanuel Trinidad worked on developing the visualization interface, its validation, formal analysis, investigation, data collection, and draft writing. Lawrence Materum made the funding acquisition, research conceptualization, methodology directions, review, editing, draft writing, supervision, and project administration.

ACKNOWLEDGMENT

De La Salle University is acknowledged for supporting this work.

REFERENCES

[1] E. Bonek, "Chapter 2 - MIMO Propagation and Channel Modeling," in *MIMO*, A. Sibille, C. Oestges, and A. Zanella, Eds., Oxford: Academic Press, 2011, pp. 27–54.

[2] S. M. Aldossari and K. C. Chen, "Machine learning for wireless communication channel modeling: An overview,"

Wireless Personal Communications, vol. 106, no. 1, pp. 41–70, 2019.

[3] L. Liu, *et al.*, "The COST 2100 MIMO Channel Model," *IEEE Wireless Communications*, vol. 19, no. 6, pp. 92–99, 2012.

[4] Y. Li, J. Zhang, P. Tang, and L. Tian, "Clustering in the Wireless Channel with a Power Weighted Statistical Mixture Model in Indoor Scenario," *China Communications*, vol. 16, no. 7, pp. 83–95, 2019.

[5] S. Mota, F. Perez-Fontan, and A. Rocha, "Estimation of the Number of Clusters in Multipath Radio Channel Data Sets," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 5, pp. 2879–2883, 2013.

[6] D. D. N. Abinoja and L. Y. Materum, "BIC-Based Optimization of the Identification of Multipath Propagation Clusters in MIMO Wireless Systems," in *2016 International Symposium on Antennas and Propagation (ISAP)*, 2016, pp. 428–429.

[7] J. Laurila, K. Kalliola, M. Toeltsch, K. Hugl, P. Vainikainen, and E. Bonek, "Wideband 3D Characterization of Mobile Radio Channels in Urban Environment," *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 2, pp. 233–243, 2002.

[8] K. Yu, Q. Li, D. Cheung, and C. Prettie, "On the Tap and Cluster Angular Spreads of Indoor WLAN Channels," in *Proc. IEEE 59th Vehicular Technology Conference. VTC 2004-Spring (IEEE Cat. No.04CH37514)*, vol. 1, pp. 218-222, 2004.

[9] N. Czink, P. Cera, J. Salo, E. Bonek, J. Nuutinen, and J. Ylitalo, "A framework for automatic clustering of parametric MIMO channel data including path powers," in *IEEE Vehicular Technology Conference*, 2006, pp. 1–5.

[10] J. F. Blanza, A. T. Teologo, and L. Materum, "Datasets for multipath clustering at 285 MHz and 5.3 GHz bands based on COST 2100 MIMO channel model," in *Proc. International Symposium on Multimedia and Communication Technology (ISMAC)*, 2019, pp. 1–5.

[11] L. Materum, J. Takada, I. Ida, and Y. Oishi, "Mobile station spatio-temporal multipath clustering of an estimated wideband MIMO double-directional channel of a small urban 4.5 GHz macrocell," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, pp. 1–16, 2009.

[12] J. Bae, T. Helldin, M. Riveiro, S. Nowaczyk, M. R. Bouguelia, and G. Falkman, "Interactive clustering: A comprehensive review," *ACM Computing Surveys (CSUR)*, vol. 53, no. 1, pp. 1–39, 2020.

[13] T. Metsalu and J. Vilo, "ClustVis: A web tool for visualizing clustering of multivariate data using principal component analysis and heatmap," *Nucleic acids research*, vol. 43, no. W1, pp. W566–W570, 2015.

[14] M. Charrad, N. Ghazzali, V. Boiteau, A. Niknafs, and M. M. Charrad, "Package 'nbclust,'" *Journal of Statistical Software*, vol. 61, pp. 1–36, 2014.

[15] H. Lee, J. Kihm, J. Choo, J. Stasko, and H. Park, "iVisClustering: An interactive visual document clustering via topic modeling," *Computer Graphics Forum*, vol. 31, pp. 1155–1164, 2012.

[16] B. C. Kwon, *et al.*, "Clustervision: Visual supervision of unsupervised clustering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 142–151, 2018.

[17] M. Cavallo and Ç. Demiralp, "Clustrophile 2: Guided visual clustering analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 267–276, 2019.

[18] M. Roque, "Interface for the factor-inclusion weighting approach in determining the number of multipath propagation clusters," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 6, pp. 1768–1776, 2019.

[19] D. M. Diallo, J. Takada, and K. Saito, "Visualization tool of the urban microcell radio propagation paths," *IEICE Communications Express*, vol. 10, no. 11, pp. 834–839, 2021.

[20] J. Teologo A. T., J. F. Blanza, and L. Materum, "Human-Computer Interface for wireless multipath clustering performance," *Journal of Engineering Science and Technology*, vol. 16, pp. 33–45, 2021.

[21] A. D. M. Africa, E. T. Trinidad, and L. Materum, "Projection of wireless multipath clusters using multi-dimensional visualization techniques," *Advances in Science, Technology and Engineering Systems Journal*, vol. 5, no. 6, pp. 1064–1070, 2020.

[22] Y. Li, J. Zhang, and Z. Ma, "Clustering in wireless propagation channel with a statistics-based framework," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1–6.

[23] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of Machine Learning Research*, vol. 9, no. 11, pp. 2579–2605, 2008.

[24] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," arXiv preprint arXiv:1802.03426, 2018.

[25] Connor Meehan Jonathan Ebrahimian, Wayne Moore, S. M. (2022). U. M. Approximation, and M. C. F. E. Projection (UMAP) Uniform Manifold Approximation and Projection. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/71902

[26] M. Steinbauer, H. Ozcelik, H. Hofstetter, C. F. Mecklenbrauker, and E. Bonek, "How to quantify multipath separation," *IEICE Transactions on Electronics*, vol. 85, no. 3, pp. 552–557, 2002.

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

[28] J. Blanza, "Joint identification of the clustering and cardinality of wireless propagation multipaths," *International Journal of Emerging Trends in Engineering Research*, vol. 7, pp. 763–767, 2019.

Emmanuel T. Trinidad received the B.Sc. degree in Electronics Engineering (ECE) from Angeles University Foundation, Philippines, in 2016. Currently, he is taking up the M.Sc. ECE degree at De La Salle University, Philippines. He is currently an instructor at Don Honorio Ventura State University. His research interests are wireless multipath clustering, dimensionality reduction, and TV whitespace.
Email: emmanuel_trinidad@dlsu.edu.ph.

Lawrence Materum received the B.Sc. degree in electronics and communications engineering with honors from Mapúa Institute of Technology (now Mapúa University), the M.Sc. degree in electrical engineering major in computers and communications from the University of the Philippines Diliman through an Analog Devices Fellowship, and the Ph.D. degree in international development engineering (major in electrical/electronics engineering) from Tokyo Institute of Technology through a MEXT Scholarship. Email: materuml@dlsu.edu.ph