Cluster-Based Scatterer Identification and Characterization in Vehicular Channels

Laura Bernadó*, Anna Roma*, Nicolai Czink*, Alexander Paier†, and Thomas Zemen*
*FTW Forschungszentrum Telekommunikation Wien, Vienna, Austria
†Institute of Telecommunications, Vienna University of Technology, Vienna, Austria
contact: bernado@ftw.at

***Abstract*—In this paper, we present a new approach for the identification of scattering objects in the delay and Doppler domain. Until now, the identification was done *visually* based on the power delay profile and video material recorded in the measurement campaigns. We propose to use automatic methods based on the local scattering function (LSF), which brings the Doppler domain into play.**

The LSF is a multitaper estimate of the two-dimensional (2D) power spectral density in delay and Doppler. Each peak of the LSF is composed of several multipath components (MPCs) coming from the same scattering object.

Our approach consists of two steps: (i) detection of the relevant peaks, and (ii) assignment of MPCs to the scattering objects using a clustering algorithm. We use a modified a modified version of the *density-based clustering of applications with noise* algorithm, where we use the MPC distance. We apply the method to a set of vehicular radio channel measurements and extract the time-varying cluster parameters.

The clusters have ellipsoidal shape with their longer axis in the Doppler domain. The first detected cluster presents different properties than the rest of the clusters, being larger, constant in time, and more static in the delay-Doppler plane. By properly identifying only the relevant scattering objects, vehicular channel models, such as the geometry-based stochastic channel model, can be simplified significantly.

I. Introduction

In order to design reliable vehicle-to-vehicle (V2V) communication systems, an understanding of realistic V2V propagation channels is required. The vehicular antennas are normally mounted on the roof top of the car, at around 1.5 m height above ground. The radio waves interact with objects located at this height, thus causing multipath components (MPCs) in the received signal. Furthermore, the line-of-sight (LOS) between the transmitter (Tx) and the receiver (Rx) car might be blocked occasionally and intermittently due to bigger vehicles or other objects obstructing the direct LOS link. In these cases, the MPCs become more relevant.

Multipath propagation increases the temporal and directional diversity at the Rx. Hence, appropriate techniques are needed to decode the message under non-LOS conditions. In order to test such techniques, numerical simulations on the computer have to be performed. For that, a mathematical representation of the channel is necessary. A well suited approach for modeling the vehicular channel is the so called *geometry-based stochastic channel modeling* [1], where the scatterers causing the MPCs are randomly placed beside the Tx-Rx link according to a spatial distribution. In [1] the authors provide an accurate parameterized model for LOS situations and its parameters for *highway* and *rural* scenarios. From a receiver point of view, it is not possible to exploit the MPCs, since the received signal is going to be masked by noise. Therefore, by only considering the relevant MPCs reaching the Rx we can use a simpler channel model for numeric simulations. For doing so, one has to identify all relevant MCPs first. Until now, very few studies have been carried out tackling this issue, also because there are not many vehicular radio channel measurements available. Previous publications present some early results obtained by visually analyzing the power delay profile (PDP) and the Doppler spectral density [2].

Contributions of the paper: In this paper we present an efficient and rigorous method for detecting only the relevant MPCs in a set of vehicular radio channel measurements. By using a clustering algorithm we group the MPCs in clusters, each one corresponding to a scattering object. We propose to use the MPC distance in the *density-based spatial clustering of applications with noise* algorithm. Furthermore, we fully characterize the cluster parameters. We take into account the high variability of the environment, and therefore we carry out a time-varying analysis of the measurement data.

Organization of the paper: In Section II, we describe several scatterers identification techniques used in the literature. In Section III, a brief overview of three tested clustering algorithms is given and discussed. The measurement data used in this paper is described in Section IV. Results and conclusions are given in Sections V and VI, respectively.

II. Approaches for Scatterer Identification

It is of great importance to identify which MPCs are truly relevant from the Rx point of view in order to develop simpler mathematical channel models to be used for system performance analysis in numeric simulations. Next, we discuss the previously used technique in the literature, and our new approach.

A. Visual Inspection of the PDP

Previous work that aimed to analyze the scatterer contributions in vehicular communications is based on visual

This research was supported by the projects COCOMINT and NOWIRE funded by Vienna Science and Technology Fund (WWTF), and the EC under the FP7 Network of Excellence projects NEWCOM++. FTW is supported by the Austrian Government and the City of Vienna within the competence center program COMET.

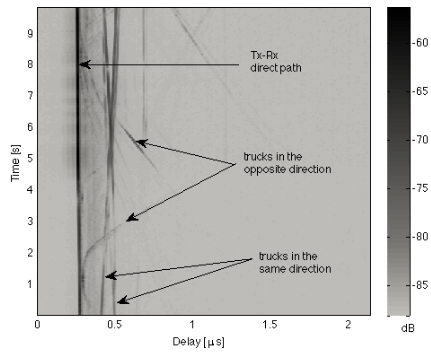


Fig. 1. Scattering identification by visual inspection of the PDP.

inspection of the PDP [2]. By using video material recorded during the measurements, the different MPCs observed in the PDP can be *visually* identified and related to physical objects which interact with the radio waves. Figure 1 shows the MPCs identified with this approach for a vehicular channel measurement. The main MPC contributions correspond to other trucks and cars driving beside the Tx and the Rx car.

There are scenarios where it is clear and easy to distinguish the different contributions and to identify the scatterers causing the MPCs. But in other cases, the task becomes too difficult to be done *visually*, such as in scenarios with rich scattering and strong presence of diffuse components. Furthermore, this method is very time-demanding and the results can be considered to be too *subjective*. Therefore, we should rely on an empirical automated methodology delivering more accurate results.

B. Clustering of MPC using the LSF

In order to overcome the drawbacks of the previous method, we introduce another domain in the analysis, the Doppler shift, which gives additional accuracy in the scattering identification. We do that by analyzing the local scattering function (LSF), which is a time-varying representation of the delay-Doppler power spectrum that is calculated as in [3], [4]. Figure 2 (a) shows the 3D LSF for the same scenario represented in Fig. 1 at $t = 5$ s. Each peak of this plot represents a different scatterer.

First of all, we have to distinguish only the relevant paths from the received signal. There are several peaks with low power not relevant compared to the highest peak. In order to distinguish them, we use a very simple concept, the power threshold criterion [5]: A path can only exist when it has more power than a certain threshold. We choose the threshold to be $-25\,\text{dB}$ from the highest detected peak. Another threshold is required to remove the noise. All components below the noise power plus $15\,\text{dB}$ are set to 0. Figure 3 exemplifies the different thresholds. These thresholds have been crosschecked using the visual inspection approach.

In Fig. 2 (b) the detected paths are shown as red crosses over the 2D view (delay-Doppler shift) of the LSF at $t = 5\,\text{s}$. Applying the peak detection algorithm described previously, three paths can be identified. Figure 2 (b) is a zoomed-in region of subfigure (a). We observe that each one of the scattering contributions is defined by several MPCs, the red crosses. For each scattering object we can group the MPCs in clusters, shown with black ellipses. By using already existing clustering algorithms, we can relate each MPC to one cluster and therefore, to one scatterer. With that, we are able to identify not only the number of relevant scatterers (number or clusters) but also their extension in the delay and Doppler domain.

III. Time-varying Cluster Parameters

Since the environment changes very rapidly due to the high speed of the Tx, Rx, and other moving vehicles, it is logical to expect that the number of clusters as well as their extension also change in time. Therefore, we present the analysis of three different clustering algorithms and describe the time-varying parameters we use for characterizing the clusters.

A. Clustering Algorithms

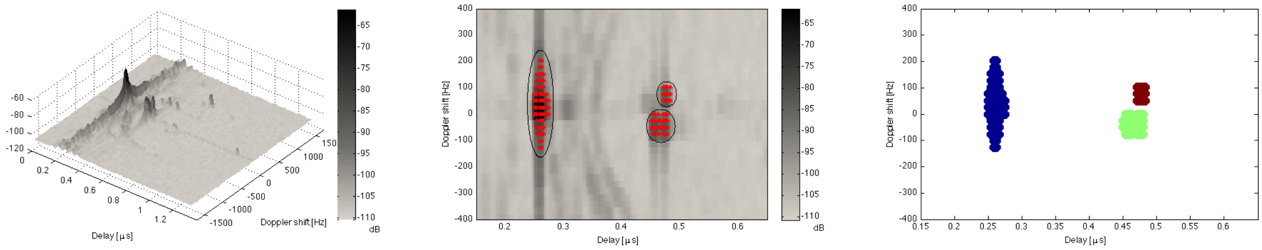
We would like to develop a method for identifying the number of relevant paths, and for each path, its extension in the delay and Doppler domain.

We studied three different clustering algorithms, all of them belonging to the family of partitioning algorithms. Next, we describe them shortly and explain whether they are useful for our purpose.

KPowerMeans algorithm

The KPowerMeans algorithm [5] takes into account the power of the MPCs. This algorithm iteratively minimizes the total sum of power-weighted distances of each path to its associated cluster centroid. In the following, the single steps of the algorithm are described in more detail.

1) The centroid starting positions $\mu_I^{(0)}$ are chosen randomly from the data set.
2) Every MPC x_l is associated with a cluster centroid such that the function of the total sum of differences, $D = \sum_{l=1}^{L} P_l d(x_l, \mu_{I_l^{(i)}})$, is minimized, with L being the total number of MPCs, and P_l the power of each MPC. We use the MPC distance d [5], which allows to combine parameters that come in different units and reads
$$MCD_{ij} = \sqrt{d_{\tau,ij}^2 + d_{\nu,ij}^2 + d_{power,ij}^2}, \tag{1}$$
where $d_{\tau,ij} = \zeta_\tau d_{\mathrm{E}}$, $d_{\nu,ij} = \zeta_\nu d_{\mathrm{E}}$ and $d_{power,ij} = \zeta_{power} d_{\mathrm{E}}$, and d_{E} denotes the Euclidean distance between the MPC and the centroid at the moment. The index $I_l^{(i)}$ is the cluster number for the ith multipath in the ith iteration step $I_l^{(i)} = argmin[P_l d(x_l, \mu_c^{(i-1)})]$. By including power into the distance function, cluster centroids are pulled to points with strong powers.
3) Clusters are chosen such that they minimize the total distance from their centroids.



(a) LSF at t=5 s. (b) Peak detection on LSF. (c) Cluster identification.

Fig. 2. Peak detection and clustering of MPCs on the LSF.

We discarded this algorithm because we could not find common weighting factors, ζ_τ, ζ_ν, and ζ_{power}, for all analyzed scenarios.

Substractive clustering algorithm

The subtractive clustering method [6] is an extension of the mountain clustering method [7]. It assumes that each data point is a potential cluster center and calculates a measure of the likelihood that each data point would define the cluster center, based on the density of surrounding data points [6]. The algorithm does the following:

1) Select the data point with the highest potential to be the first cluster center.
2) Remove all data points in the vicinity of the first cluster center as determined by *radii* (input parameter), in order to determine the next data cluster and its center location.
3) Iterate on this process until all of the data is within *radii* of a cluster center.

We discarded this algorithm because we could not find a common *radii* for all scenarios. We should note that for measurements with few clusters the algorithm works really well, but when we apply the algorithm to measurements with a large number of clusters we do not obtain reliable results.

Modified density-based spatial clustering of applications with noise (mDBSCAN)

The DBSCAN algorithm [8] is designed to discover clusters of arbitrary shape. It requires only one input parameter and supports the user in determining an appropriate value for it. Moreover, DBSCAN is efficient for a large spatial database.

The key idea is that for each point of a cluster the neighborhood of a given radius (*Eps*) has to contain at least a minimum number of points (*MinPts*), i.e., the density in the neighborhood has to exceed some threshold.

To find a cluster, DBSCAN

1) starts with an arbitrary point p and retrieves all points density-reachable from p with respect to *Eps* and *MinPts*.
2) If p is a core point, this procedure yields a cluster with respect to *Eps* and *MinPts*.
3) If p is a border point, no points are density-reachable from p and DBSCAN visits the next points of the database.

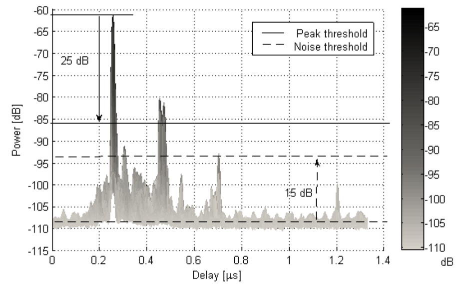


Fig. 3. LSF for a given time instant with power and noise thresholds shown.

We chose the algorithm parameters as *MinPts* $= 1$ and *Eps* $= 7$. Our modification of the algorithm consists of using the MPC distance to calculate the distance between points. The MPC distance is defined in (1), but for this algorithm, the power term is not necessary. With these values we observed that the algorithm worked as we expected. Since we observed that the clusters are larger in the Doppler domain, we give more importance to it by setting the weighting factors to $\zeta_\nu = 6$ and $\zeta_\tau = 5$. We observed good results with this algorithm for the training and the validation data sets, therefore we decided to carry out our time-varying cluster parameter analysis with the DBSCAN algorithm. Figure 2 (c) shows the result of applying DBSCAN on the data displayed on the other two subfigures. Each one of the MPCs is colored according to the cluster assigned by the algorithm.

B. Extracted Parameters

After choosing the appropriate cluster algorithm we are going to extract the following parameters: number of clusters, and extension of the cluster in the delay and Doppler shift domains.

- The *number of clusters* N_c identified in each experiment indicates the number of relevant scattering objects.
- Each cluster is defined by its central position and its *extension in the delay* S_τ *and Doppler* S_ν *domains*, which are going to be different depending on the studied scenario. The extension is defined as the largest minus the shortest point belonging to the cluster in both domains.

Since the environment in vehicular communications is highly time-varying, the cluster parameters are going to be time-varying as well. Furthermore, we make a very simple classification of the clusters between the first detected cluster (the one with minimum delay) and the rest. This is necessary due to the different characteristics observed between them, as one can already see in Figs. 2 (b)-(c).

For each experiment, we will calculate the temporal expected value over the whole measurement run. Afterwards, we will present the mean value over all measurements per scenario. Based on these results, the conclusions will be drawn.

IV. Measurement Data

The data used in this paper was collected during a measurement campaign named DRIVEWAY'09 [2] conducted in Lund, Sweden, in June 2009. The carrier frequency is $5.6\,\mathrm{GHz}$, and the measured bandwidth covers a total of $240\,\mathrm{MHz}$ in 769 frequency bins with a frequency separation $\Delta f = 312\,\mathrm{kHz}$. A measurement run consists of $S = 32500$ snapshots at a repetition time of $t_{rep} = 307.2\,\mu\mathrm{s}$ resulting in a total measurement time of $10\,\mathrm{s}$.

The Tx and the Rx car are equipped with a channel sounder. Each car has mounted a linear array with four circular patch antennas perpendicular to the driving direction. The antennas had main lobes such that they covered the four main propagation directions [9]. For the analysis performed in this paper, we consider the combined antenna radiation pattern, so that we achieve a $360°$ coverage in the azimuth plane.

Even though the data collected in the DRIVEWAY'09 campaign covers a wide range of scenarios of importance for safety-related ITS applications, in this paper we are going to focus on one of them. We select the *general LOS obstruction* scenario, where the Tx and Rx are driving in the same direction on the highway at around $120\,\mathrm{km/h}$ each. There are big trucks circulating in both directions beside them. During the measurement runs, the LOS between the two cars is intermittently obstructed.

V. Results

We have a total of 12 measurement runs available preformed in the *general LOS obstruction* scenario. In this section we are going to present the detailed result for one of the measurement runs and calculate the temporal mean of the cluster parameters. At the end, we give a summary table of the temporal means for all 12 measurement runs and the average value for the whole scenario.

We calculated the LSFs from the data, detected the peaks, and applied the clustering algorithm. In Fig. 4 the time-varying cluster parameters are shown. Figure 4 (a) plots the number of detected clusters. During the whole measurement run there are 3 to 4 detected clusters, which correspond to the visually identified paths in Fig. 1. During the second half of the measurement, 5 clusters are more often detected. This is due to the fact, that the LOS is not blocked anymore and new MPCs appear. We have not implemented yet a cluster tracking algorithm, this is why the number of clusters oscillates.

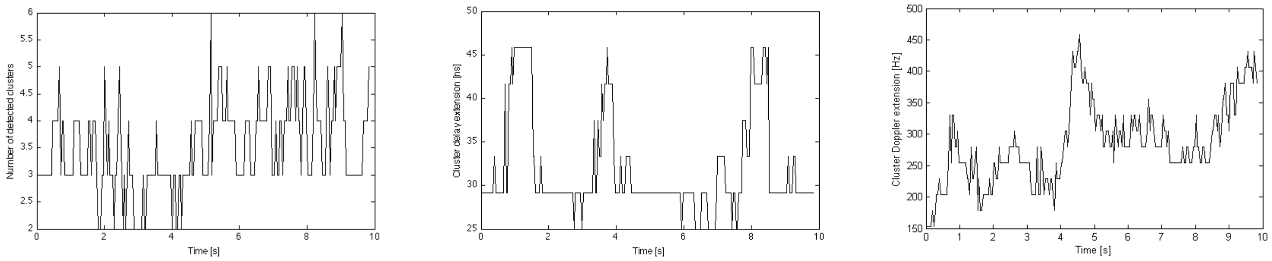
The extension in the delay and Doppler domain for the first detected cluster is depicted in Figs. 4 (b) and (c). The delay extension remains basically constant during the total measurement time of $10\,\mathrm{s}$. However, there are three time intervals where it increases considerably, at around 1, 4, and 8 s. At these time instances one of the objects driving in the opposite direction is precisely placed between Tx and Rx. The MPCs coming from this scatterer are very close to the ones of the LOS and the clustering algorithm is not able to separate them. The detected cluster consists of two *merged* clusters and therefore its extension is larger. A similar phenomena happens for the extension in the Doppler domain, nevertheless not that accentuated. Table I lists the temporal mean, maximum, and minimum value for the cluster parameters, as well as the standard deviation for this measurement throughout the $10\,\mathrm{s}$ run. While for the first detected cluster, the mean, maximum, and minimum values for the extension in the delay and Doppler domain are close to each other, the same values for the other detected clusters are far apart. The minimum values correspond to clusters with a single MPC. The extension happens to be equal to the delay or Doppler resolution, $4.17\,\mathrm{ns}$ and $25.43\,\mathrm{Hz}$ respectively. When two or more clusters are very close to each other, the used algorithm cannot distinguish them and it treats them as a single cluster with larger extension, giving as a result a maximum extension of $104.17\,\mathrm{ns}$ in the delay domain, and $534.06\,\mathrm{Hz}$ in the Doppler domain. Since this does not happen in a general basis, we provide the standard deviation values, which we consider can be useful for developing a channel model based on this cluster-based approach.

TABLE I
Mean values of the time-varying channel parameters for measurement 1.

Parameters	N_c	$S_{\tau,1st}$[ns]	S_τ[ns]	$S_{\nu,1st}$[Hz]	S_ν[Hz]
Mean	4	31.92	14.81	276.69	83.70
Std	0.83	5.81	8.01	70.66	40.51
Max	6	45.83	104.17	508.63	534.06
Min	2	25.00	4.17	127.16	25.43

We performed the same analysis for the 12 measurement runs. Table II presents their temporal mean values for the number of clusters, delay and Doppler extension for the first, and for the rest of the detected clusters. At the end of the table, there is the average among the whole set of measurements, and also the average of the standard deviation calculated for each measurement run. The number of detected clusters oscillates between 1 and 7. However, one cluster was detected in only one occasion, in measurement 6. During this measurement, there is only one strong component, which intermittently disappears, this is why there are no values for the spreads for the other clusters. The results regarding cluster spreads show more agreement among the 12 measurements. Clusters show an ellipsoidal shape, more pronounced in the first one. The first detected cluster presents different characteristics than the rest:

- Its extension in both, the delay and Doppler domain is larger than for the others, and with also larger standard deviation.



(a) Number of detected clusters. (b) Extension in delay of the first detected cluster. (c) Extension in Doppler of the first detected cluster.

Fig. 4. Time-varying cluster parameters for a single measurement run.

TABLE II
TEMPORAL MEAN VALUE OF THE CLUSTER PARAMETERS

Parameters	N_c	$S_{\tau,1st}$[ns]	S_{τ}[ns]	$S_{\nu,1st}$[Hz]	S_{ν}[Hz]
Meas 1	4	31.92	14.81	276.69	83.70
Meas 2	7	45.22	14.76	383.40	84.05
Meas 3	7	49.27	15.10	358.38	84.78
Meas 4	4	34.23	14.73	290.63	83.14
Meas 5	2	38.33	14.69	200.40	83.06
Meas 6	1	17.80	–	89.82	–
Meas 7	6	50.25	14.60	391.64	84.14
Meas 8	4	43.38	14.60	308.74	83.85
Meas 9	3	33.35	14.54	296.22	83.35
Meas 10	6	36.80	13.25	208.03	77.92
Meas 11	3	28.83	14.74	332.54	83.39
Meas 12	2	37.8	14.86	144.65	83.53
Total avg	5	37.26	14.61	273.43	83.20
Std avg	1.58	8.23	6.65	78.81	30.91

- It is present during the whole measurement run (at least for this specific scenario).
- It has the highest power.
- Since Tx and Rx drive in the same direction and at similar speed, it does not change its position in the delay-Doppler plane.

On the other hand, the rest of the detected clusters present a different behaviour:

- There are clusters appearing and disappearing quickly, caused by objects which become only relevant when they are close to the Tx-Rx link.
- There are clusters remaining longer active and moving, caused by objects (mainly trucks and bridges) which have a stronger influence on the radiowave propagation.

VI. CONCLUSIONS

We presented in this paper a new method for scatterers identification based on clustering of multipath components (MPCs). The approach used so far by other researchers is based on visual inspection of the power delay profile. We include a new domain in the analysis by using the local scattering function (LSF). By applying first a peak detection algorithm, we are able to separate only the relevant components from the receiver point of view. Afterwards, a clustering algorithm is used for assigning each one of the MPCs to its cluster, which actually represents a scattering object. The method is run on a computer and there is no need anymore of a person looking through the data and recorded videos during the measurements, which can be highly time-demanding. The algorithm is programmed so that it delivers the number of detected clusters (i.e., detected scatterers), and the extension of the cluster in the delay and Doppler domain. Furthermore, given the time variability observed in vehicular communications, we took into account that the cluster parameters are time-varying. We applied the method to a set of 12 measurement runs performed under *general line-of-sight (LOS) obstruction* conditions, where the LOS between the transmitter and receiver gets obstructed intermittently during the measurement. We observed that the detected clusters have an ellipsoidal shape with its longer extension in the Doppler domain. The properties of the first detected cluster are different than for the rest: the cluster is larger, does not move in the delay-Doppler plane, remains throughout the whole measurement run, and has the highest power. These results, we show the usefulness of a new tool for scatterer identification, which can help researchers in developing simpler but yet accurate vehicular channel models.

REFERENCES

[1] J. Karedal, F. Tufvesson, N. Czink, A. Paier, C. Dumard, T. Zemen, C. Mecklenbräuker, and A. Molisch, "A geometry-based stochastic MIMO model for vehicle-to-vehicle communication," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3646–3657, July 2009.

[2] A. Paier, L. Bernadó, J. Karedal, O. Klemp, and A. Kwoczek, "Overview of Vehicle-to-Vehicle Radio Channel Measurements for Collision Avoidance Applications," in *IEEE 71st Vehicular Technology Conference*, Taipei, Taiwan, May 2010.

[3] G. Matz, "On non-WSSUS Wireless Fading Channels," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 2465–2478, September 2005.

[4] L. Bernadó, T. Zemen, A. Paier, J. Karedal, and B. H. Fleury, "Parametrization of the Local Scattering Function Estimator for Vehicular-to-Vehicular Channels," in *IEEE 70th Vehicular Technology Conference.*, Anchorage, Alaska, USA, September 2009.

[5] N. Czink, "The random-cluster model – a stochastic MIMO channel model for broadband wireless communication systems of the 3rd generation and beyond," *PhD Thesis*, 2007.

[6] D.-W. Kim, K. Y. Lee, D. Lee, and K. H. Lee, "A kernel-based subtractive clustering method." *Pattern Recognition Letters*, vol. 26, no. 7, pp. 879–891, 2005.

[7] R. Yager and D. Filev, "Approximate clustering via the mountain method," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 24, no. 8, pp. 1279 –1284, 1994.

[8] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," *2nd ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, August.

[9] A. Thiel, O. Klemp, A. Paier, L. Bernadó, J. Karedal, and A. Kwoczek, "In-situ vehicular antenna integration and design aspects for vehicle-to-vehicle communications," in *The 4th European Conference on Antennas and Propagation (Eucap 2010)*, Barcelona, Spain, April 2010.