Road-Side Unit Anomaly Detection

Mohamed-Lamine Benzagouta, Hasnaâ Aniss, Hacène Fouchal, Nour-Eddin El Faouzi

vehicles

MDPI

Article

Road-Side Unit Anomaly Detection

Mohamed-Lamine Benzagouta [1], Hasnaâ Aniss [1], Hacène Fouchal [2,*] and Nour-Eddin El Faouzi [3,4]

[1] ERENA, Département COSYS, Université Gustave Eiffel, 33400 Bordeaux, France; mohamed-lamine.benzagouta@univ-eiffel.fr (M.-L.B.); hasnaa.aniss@univ-eiffel.fr (H.A.)
[2] Lab-I*, Université de Reims Champagne-Ardenne, 51097 Reims, France
[3] ENTPE, Université de Lyon, F-69675 Lyon, France; nour-eddin.elfaouzi@univ-eiffel.fr
[4] LICIT-ECO7, UMR T9401, Université Gustave Eiffel, F-69675 Lyon, France
* Correspondence: hacene.fouchal@univ-reims.fr

Abstract: Actors of the Cooperative Intelligent Transport Systems (C-ITS) generate various amounts of data. Useful information on various issues such as anomalies, failures, road profiles, etc., could be revealed from the analysis of these data. The analysis, could be managed by operators and vehicles, and its output could be very helpful for future decision making. In this study, we collected real data extracted from road operators. We analyzed these streams in order to verify whether abnormal behaviors could be observed in the data. Our main target was a very sensitive C-ITS failure, which is when a road-side unit (RSU) experiences transmission failure. The detection of such failure is to be achieved by end users (vehicles), which in turn would inform road operators which would then recover the failure. The data we analyzed were collected from various roads in Europe (France, Germany, and Italy) with the aim of studying the RSUs' behavior. Our mechanism offers compelling results regarding the early detection of RSU failures. We also proposed a new C-ITS message dedicated to raise alerts to road operators when required.

Keywords: C-ITS; failure detection; road-side unit; data analysis; machine learning

Citation: Benzagouta, M.-L.; Aniss, H.; Fouchal, H.; El Faouzi, N.-E. Road-Side Unit Anomaly Detection. *Vehicles* **2023**, *5*, 1467–1481. https://doi.org/10.3390/vehicles5040080

Academic Editor: Meng Li

Received: 3 August 2023
Revised: 10 October 2023
Accepted: 10 October 2023
Published: 20 October 2023

1. Introduction

C-ITS is a system that allows vehicles to connect with each other. Its development focuses on improving safety, comfort, traffic and energy efficiency. Its main communication strategy is the vehicle-to-everything (V2X), in which data sharing is performed through vehicle-to-vehicle communication (V2V) and vehicle-to-infrastructure communication (V2I). In a C-ITS environment, cooperative awareness messages (CAMs) [1] share real-time information on individual vehicles, whereas the decentralized environmental notification messages (DENMs) [2] include information on events that may have an influence on road safety or traffic conditions (e.g., roadworks, accidents, vehicle breakdowns, and so on), allowing for proactive incident prevention. Nonetheless, given the concept's novelty, the influence of C-ITS services on road networks has yet to be completely recorded and analyzed [3].

Due to the cooperative aspects of the C-ITS, its actors share large amounts of data. The analysis of these data could reveal important information on the performance of the system. Critical information such as system failures, anomalies, or road profiles could be detected by the road and telecommunication operators, as well as end users.

Road-side units (RSUs) are one of the key actors in the C-ITS environment. They are stationary units installed on the side of roads, and act as access points for the vehicles with the infrastructure ensuring V2I communications. They are most useful and become essential in the case of a low-density of C-ITS-equipped vehicles [4]. In addition to the CAM and DENM transmissions, RSUs provide the vehicles with services that are implemented on the infrastructure. These services include [5]:

1. Traffic light maneuver (TLM), which is a service that administrates the generation and sending of Signal Phase and Timing Messages (SPATEMs). Its objective is to control vehicle access to intersections and conflict zones. It sends safety information to the vehicles present in an intersection and informs them of the real-time status of the traffic light and its future status as well as the time period between the two.
2. Road and lane topology (RLT), which is a service that manages the generation and reception of Map Extended Messages (MAPEMs). A MAPEM is a message containing a digital topology map of the geometry of an area. This topology describes lanes, crosswalks, conflict zones, and permitted maneuvers.
3. Infrastructure to vehicle information (IVI) is a road sign service; it uses the Infrastructure to Vehicle Information Message (IVIM) to provide information about physical or virtual road signs such as contextual speeds or road warnings, as well as the presence of roadworks.

Thus, an RSU failure may imply an unavailability (or partial availability) of the infrastructure services in an area of supposed coverage, and it may also imply a degradation of the whole C-ITS system in the case of low-density scenarios. Thus, in order to ensure a proper functioning of the system, RSUs must be monitored and their failures detected and reported promptly.

In this study, we analyze real data that were extracted from road operators and collected through packet capturing using a test vehicle in various roads in Europe (France, Germany, and Italy). We present a specific study objective, which is to detect when an RSU experiences transmission failures. This detection is to be achieved by end users (vehicles), which in turn would inform road operators to recover it.

In this article, the following contributions are proposed:

1. A new methodology that allows vehicles to automatically detect RSU failures.
2. The evaluation of the methodology using a real dataset of CAMs generated in a C-ITS naturalistic driving environment in three countries (France, Germany, and Italy).
3. The proposition of a new C-ITS message to be used to raise alerts regarding detected failures.

The rest of this paper is structured as follows: Section 2 presents the state of the art on RSU failure detection and data analysis techniques and briefly presents the C-ITS in order to understand the role of the communication stack used over the C-ITS. Section 3 presents our solution and the mechanisms used to detect RSU failures. Section 4 is dedicated to describe the alarm message which has to be sent from a vehicle regarding the status of an RSU. Section 5 presents the conclusions and future work.

2. State of the Art

In this section, we discuss various aspects related to our work: the C-ITS, which is the system which we analyzed in this paper; the anomaly detection using relevant datasets; and the clustering algorithms used to classify the data.

2.1. C-ITS Systems

In Europe, the C-ITS uses a rich communication protocol stack which has been defined and standardized by the ETSI standardization institute. Over the *Network layer (defined as geo-networking layer)*, the *Facilities* layer has been designed in order to be an efficient interface with the application layer (close to the driver and the vehicle sensors). Various types of messages are provided by this layer in order to cover a set of use cases (road works warning, traffic jam detection, traffic light control, logistics management). In this study, we base our work mostly on one particular message type, which is the CAM.

The goal of CAMs is to create a cooperative awareness among vehicles; its use is to provide dynamic information about the vehicle such as its position, speed, heading, etc. A CAM is a periodic message that is sent at a frequency from 1 Hz to 10 Hz depending on the speed of the vehicle. It can be sent using V2V or V2I communications.

The architecture of V2V communications is presented in Figure 1, where each vehicle is supposed to have a pseudonym certificate. The messages sent by the vehicle could reach distant vehicles thanks to multi-hop forwarding.

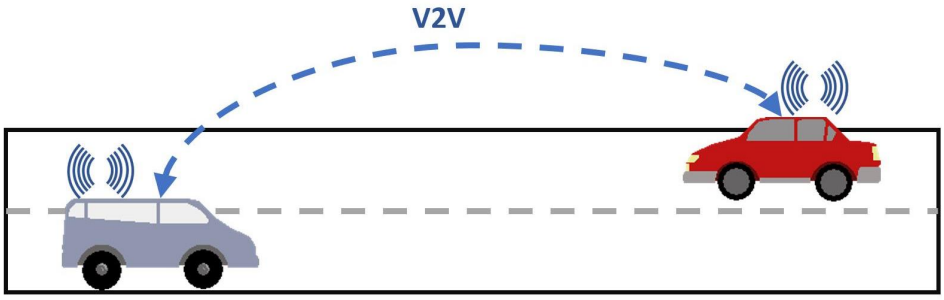

Figure 1. A General Scheme for Vehicle-to-Vehicle Communication.

The general architecture of V2I communications is presented in Figure 2.

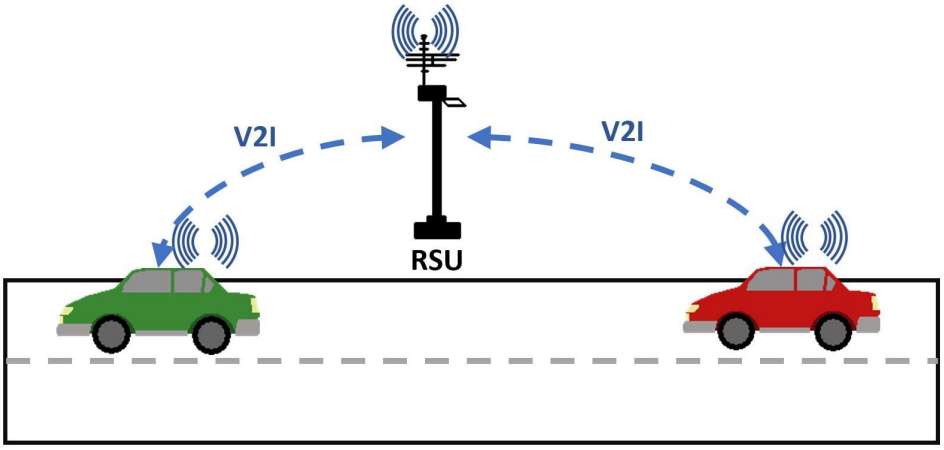

Figure 2. A General Scheme for Vehicle-to-Infrastructure Communication.

An RSU plays the role of the infrastructure as it handles all the received messages from the vehicles and runs the road operator's computations such as traffic management and event recording. It also plays the role of a vehicle for the forwarding aspect and in some cases it disseminates events toward other RSUs within the operator's network if they are in its range.

For the forwarding mechanisms, they are achieved with the geo-networking protocol [6] in the ITS-G5 protocol stack which has been defined by the ETSI. The protocol plays the role of the networking layer. Moreover, in our work, in addition to the CAMs, we also use the geo-networking layer of the packets to access additional information, such as a more precise localization of the concerned stations. Ref. [7] proposes a methodology to check for anomalies on the energy consumption of wireless devices.

2.2. Anomaly Detection

In [8], an anomaly detection approach is the process of defining a region of normal behavior within a set of data and declare whichever data points that do not belong to this region as anomalies. Moreover, anomalies are classified into three categories: (a) point anomalies, which is when a data point is considered aberrant from the rest of the data, (b) contextual anomalies, which is when a data instance is anomalous in a specific context, and (c) collective anomalies, which is when a collection of related data instances is anomalous while the individual data instances may not be anomalous themselves.

Vehicular communication applications are known for their high mobility, which qualifies them as critical applications that require a reliable infrastructure and communication scheme. RSUs are one of the key components of a C-ITS; they mainly ensure the V2I

communication type. An omnidirectional antenna of an RSU is assumed to have equal radio propagation in all directions, but the presence of obstacles such as buildings, tunnels, rivers, or ground elevations can cause signal attenuation in certain directions. Such an RSU is not technically defective. However, a high signal attenuation in some areas of effective coverage could indicate failure.

In [9], antenna failure detection was implemented using a probabilistic model. The model was constructed using real data from field tests, whereas the health assessment of RSUs was performed through a comparison between their behaviors in terms of radio propagation. In [10], a failure detector of VANET systems is proposed, which has some notions of signal attenuation. In [11], numerical simulations were performed in various radio-propagation scenarios, where it was shown that various factors affect the attenuation of a signal. A signal tends to attenuate over large distances in line-of-sight (LOS) contexts, whereas in non-line-of-sight (NLOS) contexts, the presence of a building or obstacle can greatly affect signal intensity (in dBm). Thus, the behavior of radio propagation is highly dependent on the geographical context, be it interurban or urban.

In terms of range, we have observed in our work that, in the interurban context, and with the absence of obstacles, the in-distance (ID) of communication is greater than the out-distance (OD), which are, respectively, the range of the first point where the RSU signal is detected and the range of the last point. This was also observed in [12], where the range of communication between a vehicle and an RSU tends to be greater when the vehicle is moving toward the RSU. Moreover, the authors observed a relationship between vehicle velocity and range and they explain it as a consequence of the Doppler effect. This can also be observed in [13]. The inverse case is, however, observed in [14], where the OD is greater than the ID; however, the authors explain that this is due to the time needed for the mechanism of signal validation in the Wi-Fi access points during a first connection.

DeepADV [15] is an anomaly detection framework for VANETs based on deep neural networks. A threshold is calculated based on the difference between a genuine and an anomalous message and then used to classify them. The algorithm is to be deployed on RSUs and its aim is to detect faulty messages. In [16], an anomaly detection scheme on VANETs using edge computing was proposed. The faults concerned transmission omission and were detected using RSU-based edge network and vehicular edge computing (VEC). These RSUs guard information about a number of vehicles and a number of collected packets, which are then used to determine whether an anomaly (a change in the numbers) has occurred. They tested their anomaly detection strategy on a simulation and found that the strategy was highly effective at high rates of fault ratio (25%). EVAD is a method proposed in [17] to detect anomalies in vehicles using edge computing, wherein a correlation between sensor variables is drawn and used for anomaly detection where, for an observation, if two supposedly correlated variables do not correlate, an anomaly is detected.

In [18], a CNN is used to extract the spatial, temporal, and spatio-temporal traffic features then used for anomaly detection by the means of employing thresholds. In [19], an anomaly detection approach that takes into account the spatio-temporal features of VANET traffic is proposed. The approach consists of two phases; first, deep learning based on a CNN architecture is used for network traffic estimation; second, a decision-making approach based on reinforcement learning is used to identify the normal and anomalous traffic entries.

When it comes to security anomalies, DAMASCO [20] is a security-based system that aims to detect DoS attacks. A statistical approach was used to detect anomalies in vehicular communications where the MAC layer was addressed and assessed to identify potentially malicious nodes by the number of sent packets and block their activity. In [21], a certification-less authentication method was adopted, wherein the RSUs are only trusted partially and are granted the vehicle's information only partially. Moreover, anomaly detection was performed through clustering; precisely, the agglomerative clustering was used on the traffic data represented as time series using the dynamic time wrap distance [22].

In [23], a methodology was proposed to analyze data collected from agriculture areas during many years and verify the anomalies on the productions and their relationship with weathers conditions. In [24] a Mobile phone Network Data (MNF) based framework to detect anomalies in real time was proposed. The framework is based on two steps, the first one is an offline unsupervised learning done on the MND, the second step is the online real time detection of anomalies.

To the best of our knowledge, only a few works have focused on V2I fault detection and we are not aware of any work that proposes an online method for the fault detection of RSUs.

2.3. Clustering Algorithms

In machine learning, different approaches to handle the data and to extract information exist. A detail of some of them is proposed this section.

2.3.1. Hierarchical-Based Algorithms

The hierarchical clustering algorithm is based on algorithms denoted as **linkage**. The linkage is an algorithm which focuses on merging or splitting within a cluster. Some examples of this algorithm are as follows:

- Single linkage: handles the two closest points in the cluster;
- Complete linkage that handles the two farthest points;
- Average linkage that uses an average fictive point to represent the cluster;
- Centroid linkage that uses the most representative point of the clusters.

Slink [25] and Clink [26] are two algorithms that optimize the basic principle of agglomerative clustering using, respectively, the single and the complete linkage also denoted Agnes (agglomerative nesting) and Diana (divisive analysis) algorithms.

BIRCH (balanced iterative reducing and clustering using hierarchies) [27] is used to find clusters in large datasets using less memory and run-time. It uses the concept of clustering feature (CF), which is a kind of summary of a cluster and CF-Tree. It is known as an algorithm that builds a cluster with only one scan of the dataset but it can improve the result with more small scans.

2.3.2. Partitioning-Based Algorithms

The k-means [28] algorithm is used to select at random k (k is a number defined by the user) data points, to consider them as a cluster and to regroup the other points in the closest cluster until all points are inside a cluster. The center of the cluster is the mean of all the data points of the cluster and it is updated after a new point is added to the cluster. It is light and efficient for many small datasets.

The K-Medoïds [29] algorithm works as the k-means algorithm but does not rely on an average of a cluster center but on the most representative data point of the cluster.

2.3.3. Artificial Neural Network (ANN) Algorithms

Unlike the previous algorithms, the algorithm which uses an artificial neural network cannot handle a dataset immediately but needs to be trained beforehand. Once trained, it can efficiently handle larger amounts of data than the previous algorithms with higher accuracy. SOM (self-organized map) [30] is an artificial neural network trained to produce a low-dimensional and discrete representation of the data. This makes it a strong visualization tool for high-dimensional data.

In our study, the collected data have to be injected into a classification method. We have chosen the K-Medoïds algorithm in order to build clusters. We intend to use the resulting clusters in order to be able to distinguish between proper behavior and improper behavior. To the best of our knowledge, there is no study about the automatic detection of anomalies on C-ITS devices, even though some recent studies focus on collaborative learning of communication routes [31] and selfish behaviors [32].

3. RSU Failure Detection

In this section, we detail our failure detection technique mainly based on verifying whether the coverage which should be ensured by an RSU is effective. For this issue, we propose a new C-ITS message denoted "Alarm Message", which is sent by each vehicle entering a zone covered by an RSU. The idea is to insert in this message some information about the observed signals by the vehicle, mainly in the ID and the OD. The RSU collects all received data and runs an on-the-fly analysis in order to identify whether an anomaly has occurred.

An antenna is a fundamental component of telecommunication systems such as radio, cellular networks, Wi-Fi, and satellite communications. It is a device that is used to transmit and receive signals in the form of electromagnetic waves. When used for transmission, it takes an electrical signal and converts it into electromagnetic waves that propagate through the surrounding space. When used for reception, it captures the electromagnetic waves from the environment and converts them back into electrical signals. Overall, there exist three kinds of antennas based on radiation patterns. An isotropic antenna is a hypothetical one and radiates over all directions in the horizontal plane as well the vertical one. Its radiation diagram resembles that of a sphere and can be seen in Figure 3. An omnidirectional antenna radiates over all directions in the horizontal plane, as can be seen in Figure 4. It is commonly used in applications wherein signals need to be transmitted and received in multiple directions; an example of the omnidirectional antenna is the dipole antenna. Meanwhile, a directional antenna radiates in the form of a concentrated blob over one direction, as can be seen in Figure 5. Examples of directional antennas include parabolic antennas and the phased array antenna.

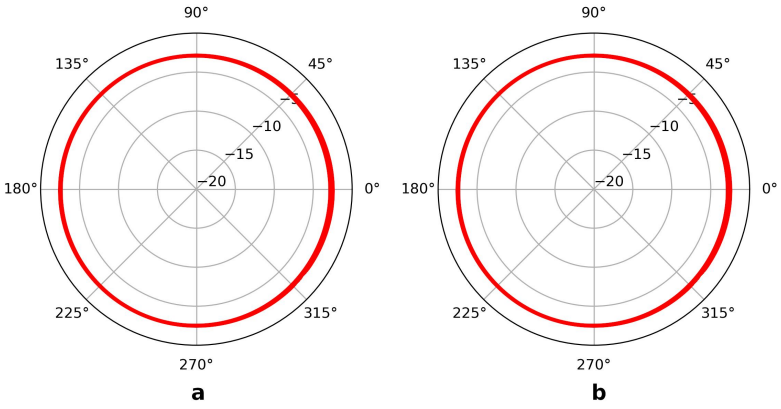

Figure 3. Radiation Diagram for an Isotropic Antenna: (**a**) vertical plane and (**b**) horizontal plane.

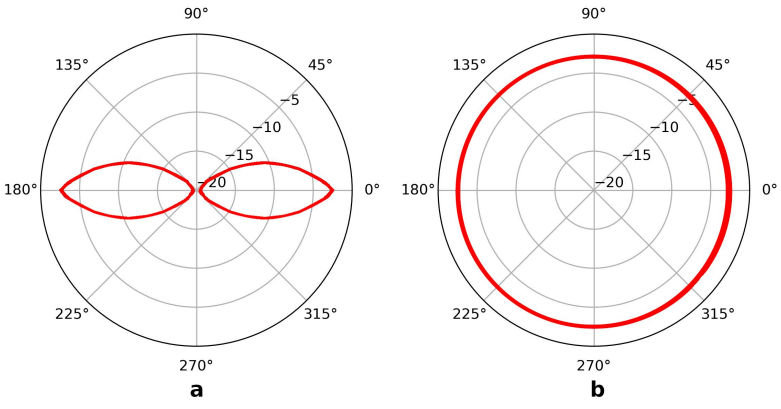

Figure 4. Radiation Diagram for an Omnidirectional Antenna: (**a**) vertical plane and (**b**) horizontal plane.

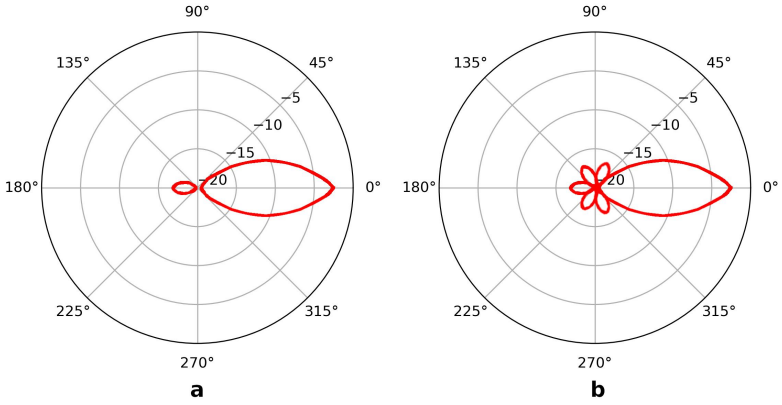

Figure 5. Radiation Diagram for a Directional Antenna: (**a**) vertical plane and (**b**) horizontal plane.

The used RSUs in this study are either omnidirectional or have two directional antennas each. Each antenna is directed toward a side of the road and has a relatively high directivity. As long as the road is relatively parallel to the antenna's lobes, the vehicle will be able to receive messages from the RSU. Our target is a specific failure which is when one of the two antennas fails, and the bidirectional RSU thus behaves as a directional antenna and radiates over only one side of the road.

Within the C-roads and the Indid projects (two European—Connecting Europe Facilities-projects), we collected many datasets through real conditions on more than 300 km of highway roads in Italy, Germany, and France. More than 1 gigabytes of traces were collected in PCAP format (this format is the one used by network analyzers) produced by one driving vehicle and more than 50 road-side units, both in ITS-G5 and C-V2X. For each RSU, more than 20 megabytes were collected in each experiment. This work, however, only focuses on ITS-G5 RSUs, which have shorter ranges than their C-V2X counterparts. More than 10,000 messages were collected by the experimentation vehicle.

In the experiment, each captured packet comes with a radiotap information layer. Radio is a protocol which is the standard for the 802.11 frame injection and reception. It does not belong to the original ITS-G5 protocol stack, but it is created at the moment of capture of a packet. The test vehicle that was used had the ability to generate the radiotap layer concerning the captured packet. The recorded information in the radiotap layer concerns the radio transmission, such as the signal intensity at the moment of capture. If the packet was sent from the vehicle, then the transmission power (in dBm) is stored, and if the packet was received then the signal intensity at the reception moment is stored. We are interested in the latter, since it provides us with the signal intensity of transmissions coming from the RSU.

One way to determine whether the packet was sent or received by the vehicle without looking at the source address is via radiotap. The difference between a sent and received packet is the format of the signal intensity; in the first case, it is denominated transmission power and concerns the signal power that was used to send the packet from the vehicle, whereas in the latter it is denominated antenna signal and concerns the transmission power at the moment of reception of the packet by the test vehicle.

We then need to identify which packets were sent by an RSU in contact with the vehicle. Assuming that each RSU has sent at least one CAM, we search through all CAM instances for the field station-type. The station-type of a vehicle is equal to 5, whereas that of an RSU is equal to 15. We can therefore identify all CAMs that were sent by RSUs and save their MAC addresses. We use the MAC address rather than the CAM's station-ID for identification because the station-ID is supposed to change every 10 min, whilst the MAC address remains static during the experiment.

An RSU could also be identified through its positioning. A vehicle moves, whereas an RSU remains static; therefore, if a station does not change its location in the C-ITS stack in

all records, it is assumed to be an RSU. Concerning the pcap dataset, and so far, we verified that all static position stations have sent at least one CAM with a station-type of 15.

A vehicle is assumed to be under the coverage of an RSU once it starts receiving packets from it until it stops. This method may give false indications as the vehicle could leave the coverage area and enter once again, having thus two coverage periods or more and still be considered as one single coverage. This verification is however not needed since we only consider the packets sent by an RSU in the analysis and the coverage remains irrelevant for the study topic. The comparison between the ID and OD is also taken from the absolute point of view (meaning, as one coverage). Therefore, coverage from the point of view of the vehicle as a metric is dismissed.

To calculate the ID and OD from the pcap files, the Euclidean distance D between two points on the Earth's globe is given by:

$$D = 6378 \times arccos[(sin(lat_1) \times sin(lat_2)) + cos(lat_1) \times cos(lat_2) \times cos(lon_2 - lon_1)]$$

where 6378 is the radius of the Earth in kilometers, lat_1 and lon_1 are the longitude and latitude of the first point, and lat_2, lon_2 are those of the second point.

Each figure from Figures 6–8 represents a set of captured packets by the vehicle from the RSU during the vehicle's travel in the area of coverage. Each packet is represented by the tuple (range, signal), where red dots represent the distance of the packets and blue dots represent their signal intensity. In order to correlate them, both these variables were normalized between zero and one.

Overall, and for the ITS-G5, packets were captured from 23 RSUs from France, Germany, and Italy. These RSUs are located both in urban and interurban locations. Since the presence of buildings and other factors in the urban context causes signal attenuation, the RSU's range can be modified by these factors. Therefore, we chose the interurban context for the comparative study of the RSUs.

When analyzing the data, we have observed several characteristics concerning the properly functioning RSUs. Mainly, that the ID is usually larger than the OD, which is validated by [12]. This particular behavior can be observed in Figures 6 and 7, which concern two proper functioning RSUs both in France; the first one is located in Reims which is in the north eastern part of the country and the second one in Saint Maurice in the suburbs of Paris.

Figure 8 represents the behavioral profile of a defective RSU located in Reims. This RSU is bidirectional but behaves as a uni-directional one.

Another characteristic of a functioning RSU in this context is that range is inversely proportional to the intensity of the received signal. We verify this relation by calculating the Pearson's correlation coefficient of all records of range and signal during the coverage period. Thus, well-functioning RSUs should give a negative value below a certain threshold. The relationship between distance and the signal intensity of RSUs that are assumed to function well is not perfectly linear, as the factors that contribute to the signal attenuation are environmental such as air quality and the presence of various obstacles. Thus, the inversely proportional relation does not give a Pearson's correlation coefficient that is strictly equal to -1. We therefore chose a certain threshold to determine the RSUs that are behaving properly and those which are not. We also consider that both distances should be superior to 50 m for interurban RSUs; otherwise, it is considered failure.

Figure 6 addresses the first discussed RSU in the city of Reims. We can observe here that both characteristics are present. The calculated ID is 1229 m, which is the largest distance recorded in the set, and the OD is 671 m. In terms of the relation between distance and signal, we can observe in the figure that the signal tends to increase when the distance becomes shorter. Therefore, the Pearson's coefficient in this set is equal to -0.69.

The RSU in Figure 7 has a lower range than the previous one; thus, fewer packets were captured. But the pattern remains the same and the ID of coverage was calculated to be 407 m, whereas the OD was 300 m. The inverted relation between distance and signal is more visible here, and the correlation coefficient calculated here is equal to -0.83.

For the defective RSU in Figure 8, however, the ID is way smaller than the OD. In fact, this RSU is located on the side of the road that the vehicle was traveling on. The first contact the vehicle had with this RSU was after it had surpassed it on the road by around 297 m, which is the recorded ID, while the OD is at 1572 m. The relationship between signal intensity and range is also not as in the case of previously discussed RSUs. We see in the figure that, in the first 300 packets, the signal and range have a direct relationship, whereas in the last packets, the relationship becomes inverted. The Pearson's correlation coefficient for this set is equal to 0.01. This means that the rule of correlation could not be applied in this context. This RSU is definitely experiencing failure.

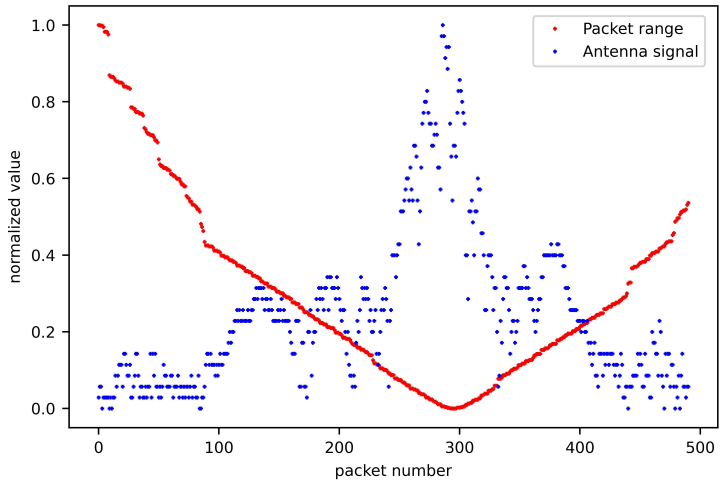

Figure 6. The Correlation between Signal and Distance for a Normal RSU in Reims. Values are normalized between 0 and 1.

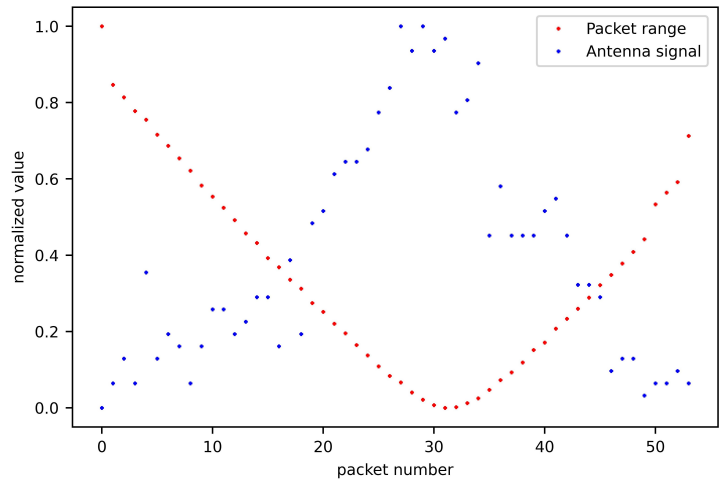

Figure 7. Correlation between Signal and Distance for a Normal RSU in Saint Maurice (Paris). Values are normalized between 0 and 1.

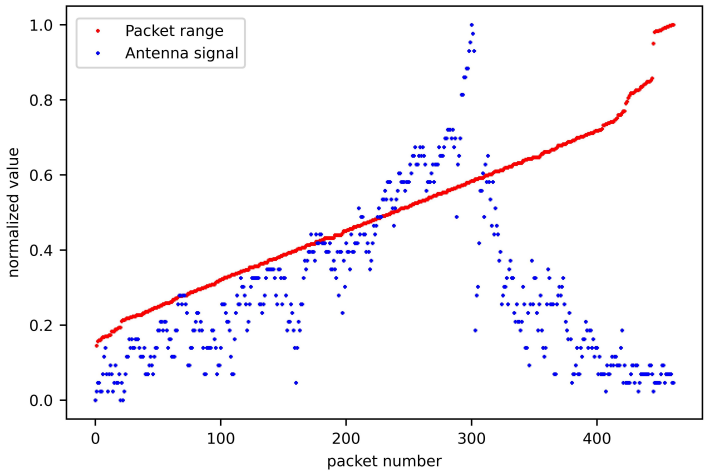

Figure 8. The Correlation between Signal and Distance for a Failing RSU in Reims. Values are normalized between 0 and 1.

The proposed method to determine failure in the case of omnidirectional and bi-directional RSUs requires to verify these properties:

- The ID is greater than the OD.
- The Pearson's correlation coefficient of the coverage data is less than a certain threshold. In our case, we choose it to be less than -0.4. This value was chosen by assessing the whole list of RSUs (Table 1), given that the smallest Pearson coefficient of a functioning RSU in the interurban context is of -0.457 (RSU number 20).
- The coverage should be larger than a threshold, and we chose the threshold to be 50 m. This is a threshold that excludes interurban RSUs in our data and only concerns the urban ones, like RSU numbers 16 and 23, respectively, as seen in Table 1.

In Table 1 can be seen the summary of all calculations made in regards to the RSUs. For each RSU, its context is given, be it interurban or urban. The max range is given, as well as the ID and OD, and the Pearson coefficient. The RSUs in Figures 6–8 are numbered 4, 8, and 2, respectively. We can see that RSUs 2, 9, 10, 13, 14, 17, 18, and 23 have the Pearson's correlation coefficient condition as false, meaning that there is no sufficient inverse linear relationship between range and signal intensity in their schemes. For RSU 2, it is a known failure for it is in the interurban context; as for the rest of the RSUs, their context is urban, and the low correlation may be due to the variation in signal intensity due to the environment or due to a failure. We can also observe that for RSUs 2, 6, 13, 14, 19, and 22, the OD inferior to the ID condition is false; for RSU 2, it is due to a failure, whereas for RSU 6 it is likely due to environmental reasons even though its context is interurban. Another explanation could be as the one discussed in [14], where the ID is smaller than the OD due to the time taken to validate the signal coming from the RSU. This is also applicable to RSU 19; both RSUs have significant correlation between signal and range. Meanwhile, RSUs 13, 14, 19, and 22 are urban RSUs, and thus the condition is false due to environmental reasons.

Moreover, we used the K-Medoïds algorithm in order to classify the behavior of RSUs using the following variables: range, signal strength, vehicle speed, and Pearson's coefficient of the signal-to-range correlation. The latter attribute is repeated for all packets from the same RSU. The classification results are shown in Figure 9, wherein the orange cluster regroups the packets that are transmitted from a failing RSU. The Medoïds of each cluster are represented by the cyan larger dots. This classification will be enhanced in order to give a more precise classification, which will be used in the future on all vehicles in order to be able to detect non-proper RSUs.

Then, an alarm will be sent through a specific message defined in the next section.

Table 1. Recorded communications with the RSUs.

RSU	Context	Max Range	In-Distance	Out-Distance	Pearson's Coef	Coef Condition	ID > OD Cond	Range > 50 Cond
1	interurban	979 m	979 m	393 m	−0.6048	correct	correct	correct
2	interurban	1572 m	297 m	1572 m	0.0108	incorrect	incorrect	correct
3	interurban	1546 m	1546 m	1112 m	−0.5705	correct	correct	correct
4	interurban	1229 m	1229 m	671 m	−0.6921	correct	correct	correct
5	interurban	2871 m	2871 m	1139 m	−0.6213	correct	correct	correct
6	interurban	502 m	447 m	502 m	−0.6198	correct	incorrect	correct
7	interurban	580 m	580 m	398 m	−0.807	correct	correct	correct
8	interurban	407 m	407 m	300 m	−0.8393	correct	correct	correct
9	Urban	501 m	501 m	467 m	−0.1690	incorrect	correct	correct
10	Urban	502 m	502 m	436 m	−0.3441	incorrect	correct	correct
11	Urban	289 m	289 m	225 m	−0.5013	correct	correct	correct
12	Urban	296 m	296 m	220 m	−0.7935	correct	correct	correct
13	Urban	218 m	155 m	197 m	0.8439	incorrect	incorrect	correct
14	Urban	280 m	196 m	280 m	0.1489	incorrect	incorrect	correct
15	Urban	350 m	329 m	241 m	−0.7976	correct	correct	correct
16	Urban	456 m	320 m	15 m	−0.5148	correct	correct	incorrect
17	Urban	1023 m	1023 m	936 m	−0.3466	incorrect	correct	correct
18	Urban	537 m	537 m	503 m	−0.0266	incorrect	correct	correct
19	interurban	1857 m	548 m	1857 m	−0.6720	correct	incorrect	correct
20	interurban	410 m	410 m	100 m	−0.4571	correct	correct	correct
21	Urban	249 m	249 m	214 m	−0.5072	correct	correct	correct
22	Urban	1542 m	202 m	667 m	−0.5235	correct	incorrect	correct
23	Urban	709 m	709 m	48 m	−0.1011	incorrect	correct	incorrect

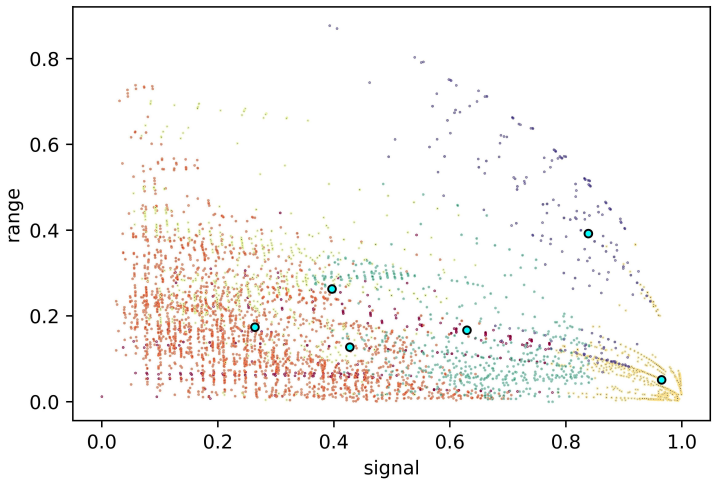

Figure 9. Clustering of Behavior of RSUs.

4. Alarm Message

In this section, we detail the structure of our proposed message, which makes it possible to raise an alarm about RSU failures. This message is composed of the timestamp

and the location when the RSU is met the first time and the timestamp and the location when it is met the last time. The vehicle records the timestamp and location of received messages from the RSU and filters the last one and the first one. As an option, we add the signal strength RSSI. This message has to be sent to the road operator using ITS-G5 or C-V2X channels. Unlike the typical C-ITS messages, this message will target the RSU which has been analyzed. It will be sent through an uni-cast protocol. If ITS-G5 is used, that means the message will be forwarded by vehicles driving in the opposite roads. The structure of the packet is shown in Figure 10.

Message-type	Version	Sender-id	In-timestamp	In-location	Out-delta-time	Out-location	RSU-Id	In-RSSI	Out-RSSI
1 byte	1byte	2 bytes	5 bytes	4 bytes	2 bytes	4 bytes	2 bytes	2 bytes	2 bytes

Figure 10. The structure of the alarm packet.

A vehicle builds this message each time it meets an RSU. At that moment, it collects beacons and CAMs sent by the RSU and keeps the first message and the latest message. When a new message is received, the former one is dropped. When no message is received from the RSU after a time delay, the previous one is considered as the last one.

If an RSU overlaps another RSU, the vehicle filters CAMS and beacons sent from both RSUs in order to be able to alert each RSU accordingly. The algorithm in Algorithm 1 gives the instructions that a vehicle should follow to report an RSU's status.

Algorithm 1: Create Alarm message

Input: $RSUs[RSU\_ID, in\_location, out\_location, in\_RSSI, out\_RSSI, RSU\_Time]$ list of all RSUs with their attributes, $timeThreshold$ timestamp, $currentTime$ timestamp, m last received C-ITS message

```
if m.origin == RSU then
  if m.RSU_ID NOT IN RSUs[RSU_ID] then
    /* First message from this RSU */
    RSU ← createRSU()
    RSU.in_location ← getVehicleLocation();
    RSU.in_RSSI ← m.RSSI
    RSU.RSU_ID ← m.RSU_ID
    RSUs.add(RSU)
  else
    /* Second message from this RSU */
    RSU ← RSUs.get(RSU_ID == m.RSU_ID)
    RSU.out_location ← getVehicleLocation();
    RSU.out_RSSI ← m.RSSI;
    RSU.RSU_Time ← m.time
  end if
end if
for RSU in RSUs do
  if RSU.RSU_Time IS NOT NULL then
    T = currentTime − RSU.RSU_Time
    if T ≥ timeThreshold then
      sendAlarmMessage(RSU);
      RSUs.remove(RSU_ID == m.RSU_ID)
    end if
  end if
end for
```

The road operator has the list of all RSUs in any area, including their geographical context, which could be interurban or urban. Each RSU is associated with a score determining its status. If the score for an RSU is high, this means that the RSU is properly functioning,

otherwise, this means that the RSU is experiencing failure. If the score is around zero, the RSU is either urban or there were not enough alarm messages reported by end users. The algorithm in Algorithm 2 gives the instructions that the road operator should follow in order to treat the upcoming alarm messages. Upon the reception of an alarm message, the geographical context of the concerned RSU is verified; if it is interurban, then the ID and OD are compared, and when the ID is significantly smaller than the OD, then we are in the case of an RSU failure such as in Figure 8 and the score of the concerned RSU is decremented by one; otherwise, it is incremented by one. Lastly, whether the RSU is interurban or urban, its overall range is verified. If it does not exceed a minimal threshold, then the score is also decremented. Once an RSU's score reaches the minimal threshold, then the road operator considers that enough vehicles reported that the RSU is in failure and reports it. A time window is added so that the scores do not become saturated, and at the beginning of each time window the scores are set to zero. The window is shifted over the time and its width can be fixed depending on the whole environment.

Algorithm 2: Alarm message treatment

Input: $RSUs :< RSUID, Score >$ list of RSUs with their scores, $FailureThreshold$, $rangeThreshold$, $maxScore$, $time$ $timeWindow$, dt, $distanceThreshold$
Output: $time$
$m \leftarrow receiveAlarmMessage$;
$in\_distance \leftarrow calculateRange(m.RSUID, m.in\_location)$
$out\_distance \leftarrow calculateRange(m.RSUID, m.out\_location)$
if $m.RSUID$ *is Rural* **then**
/* no obstacles, not urban */
if $in - distance < (out - distance + rangeThreshold)$ **then**
/* reduce the score by one */
if $RSUs[RSUID].Score > -maxScore$ **then**
$RSUs[RSUID].Score- = 1$
end if
else
if $RSUs[RSUID].Score < maxScore$ **then**
/* increment the score by one */
$RSUs[RSUID].Score+ = 1$
end if
end if
end if
if $in - distance < distanceThreshold$ **or** $out - distance < distanceThreshold$ **then**
/* failure */
if $RSUs[RSUID].Score > -maxScore$ **then**
$RSUs[RSUID].Score- = 1$
end if
end if
for R in RSUs **do**
if $R.Score < FailureThreshold$ **then**
Report(R.RSUID);
end if
end for
$time \leftarrow time + dt$
if $time \geq timeWindow$ **then**
for R in RSUs **do**
$R.Score \leftarrow 0$
end for
$time \leftarrow 0$
end if

5. Conclusions

In this paper, the detection of failures in operational RSUs by end users is investigated. This issue is critical and crucial for the C-ITS. We have shown that maintenance of roadside units could be achieved in a simple way without any additional investment or any equipment. Vehicles collect RSU observations during their journeys and analyze them using our failure detection mechanism. This mechanism compares the observed behavior to the known proper one, and when an anomaly is detected, any vehicle raises alerts to the road operators in order to inform them about the suspicious RSU. RSU behavior profiles are studied as well and a difference between urban and interurban is drawn and carefully considered in this study.

We have observed that, in the case of a functioning RSU, there is an acceptable inverse correlation between the range and signal intensity of its transmissions. We have also observed that, in the interurban context with line-of-sight communications, the ID tends to be larger than the OD for a functioning RSU. Both these properties were used for our assessment of RSUs.

We have so far addressed one particular RSU failure, which is when the RSU radiates only over one direction rather than two directions. But there are other kinds of failures, for instance, the RSU could be completely dysfunctional in both directions. One way for the vehicles to detect that is for them to receive an alarm message about an RSU from a single vehicle, and when the informed vehicles find no RSU on the road (because it recently became faulty), they can conclude that there is an RSU in the designated area that went dysfunctional; thus, they can also report it. Software failures are also a possibility, such a malformed packet from an RSU, in which case we could extend our failure mechanism by calculating the ratio of malformed packets from an RSU; if it exceeds a certain threshold, this would mean that the RSU is experiencing failure.

In future works, we will assess other variables from the datasets, such as the vehicle's velocity, in order to enhance our failure detection mechanism. We will also build larger RSU behavioral profiles in order to be much more precise during the data analysis step.

Author Contributions: The authors confirm their contribution to the article as follows: study conception and design: M.-L.B., H.F., H.A. and N.-E.E.F.; analysis and interpretation of results: M.-L.B., H.F. and H.A.; initial manuscript preparation: M.-L.B., H.F., H.A. and N.-E.E.F. All authors have reviewed the results and approved the final version of the manuscript.

Funding: This work was done thanks to CINEA funding for Indid project (Agreement N° INEA/CEF/TRAN/M2018/1788494).

Data Availability Statement: The data on which this study is based were collected within the scope of two European Union projects (C-roads and Indid) on more than 300 km of highway networks in Italy, Germany, and France.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. *ETSI EN 302 637-2*; Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service. ETSI: Valbonne, France, 2014.
2. *ETSI EN 302 637-3*; Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Application; Part 3: Specifications of Decentralized Environmental Notification Basic Service. ETSI: Valbonne, France, 2014
3. Fouchal, H.; Wilhelm, G.; Bourdy, E.; Ayaida, M. A testing framework for Intelligent Transport Systems. In Proceedings of the 2016 IEEE Symposium on Computers and Communication (ISCC), Messina, Italy, 27–30 June 2016.
4. Kchiche, A.; Kamoun, F. Centrality-based Access-Points deployment for vehicular networks. In Proceedings of the 17th International Conference on Telecommunications, Doha, Qatar, 4–7 April 2010.
5. *ETSI TS 103 301*; Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Facilities Layer Protocols and Communication Requirements for Infrastructure Services. ETSI: Valbonne, France, 2020
6. *ETSI EN 302 636-4-1*; Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 4: Geographical Addressing and Forwarding for Point-to-Point and Point-to-Multipoint Communications; Sub-Part 1: Media-Independent Functionality. ETSI: Valbonne, France, 2014.

7. Bennis, I.; Zytoune, O.; Driss Aboutajdine, D.; Fouchal, H. Low energy geographical routing protocol for wireless multimedia sensor networks. In Proceedings of the 2013 9th International Wireless Communications and Mobile Computing Conference, IWCMC 2013, Sardinia, Italy, 1–5 July 2013; pp. 585–589.
8. Chandola, V.; Banerjee, A.; Kumar, V. Anomaly Detection: A Survey. *ACM Comput. Surv.* **2009**, *41*, 58. [CrossRef]
9. Worrall, S.; Agamennoni, G.; Ward, J.; Nebot, E. Fault detection for vehicular ad-hoc wireless networks. In Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV), Gold Coast, QLD, Australia, 23–26 June 2013.
10. Liu, J.; Ding, F.; Zhang, D. A Hierarchical Failure Detector Based on Architecture in VANETs. *IEEE Access* **2019**, *7*, 152813–152820. [CrossRef]
11. Lytaev, M.; Borisov, E.; Vladyko, A. V2I Propagation Loss Predictions in Simplified Urban Environment: A Two-Way Parabolic Equation Approach. *Electronics* **2020**, *9*, 2011. [CrossRef]
12. Demmel, S.; Lambert, A.; Gruyer, D.; Larue, G.; Rakotonirainy, A. IEEE 802.11p Empirical Performance Model from Evaluations on Test Tracks. *J. Netw.* **2014**, *9*, 1485–1495. [CrossRef]
13. Halili, R.; BniLam, N.; Yusuf, M.; Tanghe, E.; Joseph, W.; Weyn, M.; Berkvens, R. Vehicle Localization Using Doppler Shift and Time of Arrival Measurements in a Tunnel Environment. *Sensors* **2022**, *22*, 847. [CrossRef]
14. Ammoun, S.; Nashashibi, F. Design and efficiency measurement of cooperative driver assistance system based on wireless communication devices. *Transp. Res. Part Emerg. Technol.* **2010**, *18*, 408–428. [CrossRef]
15. Tejasvi, A.; Bhavya, G.; Ayush, A.; Vinay, C.; Fei Richard, Y. DeepADV: A Deep Neural Network Framework for Anomaly Detection in VANETs. *IEEE Trans. Veh. Technol.* **2021**, *70*, 12013–12023.
16. Weber, J.S.; Ferreto, T.; Zincir-Heywood, N. Exploring Anomaly Detection Techniques for Enhancing VANET Availability. In Proceedings of the 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring), Florence, Italy, 20–23 June 2023.
17. Guo, F.; Wang, Z.; Du, S.; Li, H.; Zhu, H.; Pei, Q.; Cao, Z.; Zhao, J. Detecting Vehicle Anomaly in the Edge via Sensor Consistency and Frequency Characteristic. *IEEE Trans. Veh. Technol.* **2019**, *68*, 5618–5628. [CrossRef]
18. Nie, L.; Li, Y.; Kong, X. Spatio-Temporal Network Traffic Estimation and Anomaly Detection Based on Convolutional Neural Network in Vehicular Ad-Hoc Networks. *IEEE Access* **2018**, *6*, 40168–40176. [CrossRef]
19. Nie, L.; Wang, H.; Gong, S.; Ning, Z.; Obaidat, M.S.; Hsiao, K.F. Anomaly Detection Based on Spatio-Temporal and Sparse Features of Network Traffic in VANETs. *IEEE Access* **2019**, *7*, 177954–177964. [CrossRef]
20. Valentini, E.P.; Filho, G.P.R.; De Grande, R.E.; Ranieri, C.M.; Júnior, L.A.P.; Meneguette, R.I. A Novel Mechanism for Misbehavior Detection in Vehicular Networks. *IEEE Access* **2023**, *11*, 68113–68126. [CrossRef]
21. Tan, H.; Gui, Z.; Chung, I. A Secure and Efficient Certificateless Authentication Scheme with Unsupervised Anomaly Detection in VANETs. *IEEE Access* **2018**, *6*, 74260–74276. [CrossRef]
22. Bourdy, E.; Piamrat, K.; Herbin, M.; Fouchal, H. New Method for Selecting Exemplars Application to Roadway Experimentation. In Proceedings of the International Conference on Innovations for Community Services, Žilina, Slovakia, 18–20 June 2018.
23. Moso, J.C.; Cormier, S.; de Runz, C.; Fouchal, H.; Wandeto, J.M. Anomaly detection on data streams for smart agriculture. *Agriculture* **2021**, *11*, 1083. [CrossRef]
24. Akopyan, E.; Furno, A.; El Faouzi, N.-E.; Gaume, E. Unsupervised Real-time Anomaly Detection for Multivariate Mobile Phone Traffic Series. In Proceedings of the ESANN European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Virtual, 6–8 October 2021; pp. 469–474. [CrossRef]
25. Sibson, R. SLINK: An optimally efficient algorithm for the single-link cluster method. *Comput. J.* **1973**, *16*, 30–34. [CrossRef]
26. Defays, D. CLink: An efficient algorithm for a complete link method. *Comput. J.* **1977**, *20*, 364–366. [CrossRef]
27. Zhang, T.; Ramakrishnan, R.; Livny, M. BIRCH: An Efficient Data Clustering Method for Very Large Databases. In Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, Montreal, QC, Canada, 1 June 1996.
28. MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1967; Volume 1, pp. 281–297.
29. Kaufman, L.; Rousseeuw, P. Clustering by means of medoids. In Proceedings of the Statistical Data Analysis Based on the L1 Norm Conference, Neuchatel, Switzerland, 31 August 1987; pp. 405–416.
30. Kohonen, T. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **1982**, *43*, 59–69. [CrossRef]
31. Shan, A.; Fan, X.; Wu, C.; Zhang, X.; Men, R. Dynamic Selfish Node Detection with Link Quality Consideration in Vehicular Networks. *IEEE Trans. Veh. Technol.* **2023**, *72*, 8827–8843. [CrossRef]
32. Wu, C.; Liu, Z.; Liu, F.; Yoshinaga, T.; Ji, Y.; Li, J. Collaborative Learning of Communication Routes in Edge-Enabled Multi-Access Vehicular Environment. *IEEE Trans. Cogn. Commun. Netw.* **2020**, *6*, 1155–1165. [CrossRef]