Article

A Digitalization Algorithm Based on the Voltage Waveform of the Multifunction Vehicle Bus

Yangtao Li [1], Haiquan Liang [2,*] and Xiaodong Tan [2]

[1] State Key Laboratory for Traction and Control System of EMU and Locomotive, China Academy of Railway Sciences Corporation Limited, Beijing 100191, China; liyangtao@zemt.cn
[2] Institute of Rail Transit, Tongji University, Shanghai 200092, China; daitong@tongji.edu.cn
* Correspondence: lhq@tongji.edu.cn

Abstract: The MVB is a kind of widely used vehicle-level bus, which is crucial for the normal operation of trains. However, the MVB bus contains device node terminals and the topology is complex, which makes fault location difficult. The voltage waveform of the MVB physical layer can reflect the electrical characteristics and fault characteristics of a network in real time and has the value of assisting the diagnosis of MVB network faults. Based on the characteristics of the MVB physical-layer voltage waveform, this paper studies the optimal sampling rate and feature-extraction algorithm of the MVB voltage waveform and the combination of the data of the data-link layer to assign accurate timestamps to waveform data, and finally presents a design for an MVB dual-mode data-acquisition platform with multiple sampling rates/test conditions. Experimental results show that the proposed algorithm can accurately extract waveform feature information under the optimal sampling rate of 62.5 MHz, which makes the collection of MVB voltage waveform data more reliable and practical.

Keywords: MVB fault diagnosis; feature-extraction algorithm; MVB dual-mode data

1. Introduction

The Train Communication Network (TCN) is crucial for the normal operation of trains. Control information and monitoring information of trains are transmitted through the train communication network, so the TCN is a necessary condition to ensure the normal operation of trains [1–3]. The Multifunction Vehicle Bus (MVB) is the vehicle bus of the TCN, which is used for connecting programmable stations and simple sensors/actors [4–6]. Due to the large amount of on-board equipment, the diverse operating environment and the complex network topology, it is difficult for maintenance personnel to perform timely repairs of MVB network faults, which reduces the efficiency of train operation and maintenance, leading to an increase in operation and maintenance costs.

The communication data of the MVB exist in two main forms during transmission: physical-layer data and link-layer data. MVB devices generate digital link-layer data, and analog voltage waveform signals (i.e., physical-layer data) are generated through digital-to-analog conversion circuits [7,8]. Since the MVB waveforms are analog quantities generated by the conversion of digital data, the information carried by the physical-layer voltage waveform is much richer than the information carried by the link-layer data: the information in the physical-layer voltage waveform not only includes communication data, but also contains the physical features of the waveform itself [9]. These features can characterize different MVB fault modes and whether the device status is healthy or not, and this is exactly the information needed by an online diagnosis system [10,11]. Yang et al. [12] used stacked denoising autoencoders to automatically extract waveform features of the MVB physical layer and used these features to train deep neural networks to achieve MVB network fault-mode classification. Wang et al. [13] proposed a method for diagnosing and locating MVB network faults based on a weighted support vector machine. They

extracted waveform features from the physical layer to characterize different MVB faults and used the WSVM sample weight distribution method of multi-hop edge approaching to deal with the different influences of samples located at different positions in the feature space on a support vector machine (SVM) hyperplane. Li et al. [14] proposed an MVB fault diagnosis method based on physical waveform features and ensemble pruning. They classified different MVB faults with a random-forest classifier and proposed an ensemble pruning method based on diversity indexes and a k-mean algorithm to reduce the number of decision trees and improve the ensemble performance. Wang et al. [15] proposed an MVB fault diagnostic method based on a support vector machine. By extracting the waveform features, fault diagnosis is transformed into a pattern-recognition problem, and an SVM-based classifier was trained for network fault classification and recognition. Li et al. [16] proposed a diagnostic method for MVB terminal faults based on multi-kernel learning support vector machines (MKLSVMs), which combined physical wave features and MKLSVMs to complete terminating fault diagnosis.

The above studies mainly extracted data features from the physical layer or the data-link layer of MVBs for fault diagnosis, and their methods perform well. The results show that waveform features are related to MVB faults and are crucial for fault diagnosis. However, these studies focused more on the theoretical research of MVB fault mechanisms and diagnosis rather than practical applications. Therefore, these studies need data of high quality. Taking the latest study based on MKLSVMs as an example, this study employs a sampling rate of 100 MHz to collect waveform data, which are then analyzed to achieve waveform feature extraction from the data. However, online MVB fault diagnosis requires short diagnostic cycles and real-time data analysis. The data generated by high-speed sampling will increase computational pressure and raise costs. Thus, it is important to acquire a proper sampling rate for applications.

This research is based on MVB physical-layer voltage waveform data, and an MVB voltage waveform digitization algorithm was written to achieve frame data extraction and waveform feature extraction of MVB communication data, simplifying the effective information. By setting different sampling rates, a proper sampling rate for applications has been determined in experiments and the algorithm has been proved to be effective.

2. The MVB Voltage Waveform Digitalization Algorithm

2.1. The MVB Bus and Its Data Characteristics

The MVB bus is a vehicle-level bus of the TCN, and the equipment mounted on the bus includes a traction system, an air conditioning system, a door system, etc. Therefore, the MVB bus is responsible for information transmission and communication between important devices in a car. The MVB bus is equipped with redundant lines, which are divided into supervisory lines and trust lines, both of which transmit the same information at the same time for the receiver to verify the reliability of communication data. The MVB bus has a transmission rate of 1.5 Mbit/s and transmits content using Manchester encoding.

The MVB bus topology is shown in Figure 1. The media of the MVB network are mainly selected according to the communication transmission distance, including ESD (<20 m), EMD (<200 m) and OGF (<2000 m). This research chose an MVB based on EMD media. Since the EMD is the mainstream transmission medium, it is the most representative.

Based on the EMD, voltage signals are used to transmit information between MVB devices, and link-layer data are used for internal analysis and processing of devices. Data at the data-link layer exist in binary form, while voltage signals at the physical layer transmit information by Manchester encoding. In addition, there are non-data symbols in the transmission process as delimiter bits. Figure 2 shows the encoding format of data and non-data symbols. One of the characteristics of Manchester coding is that there is a transition edge in the middle of the bit cell, which expresses the meaning of the data and is also a clock synchronization signal. With the help of this clock synchronization feature, each data bit can be located in the decoding process, avoiding error accumulation.
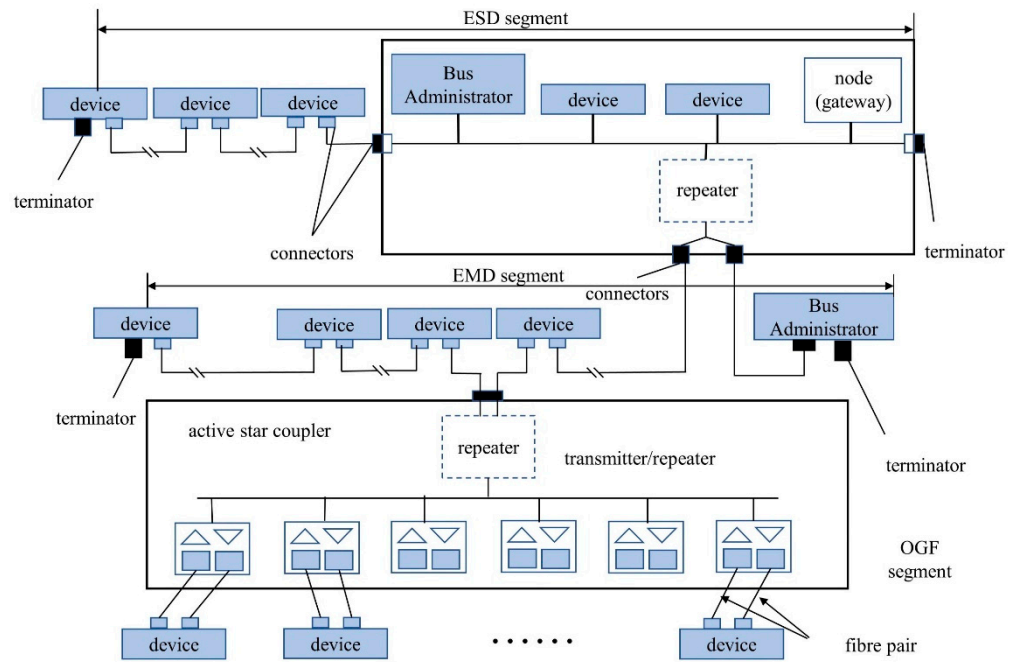
Figure 1. Topology of MVB.

Figure 2. The encoding format of data and non-data symbols: (**a**) non-data symbol encoding; (**b**) "0" and "1" data encoding.

The MVB network polls MVB devices to access different kinds of information. The polling master is the MVB master device and the information sent out is called the master frame, while the polled device is called the slave device and the information sent out is called the slave frame. MVB frame formats are shown in Figure 3.
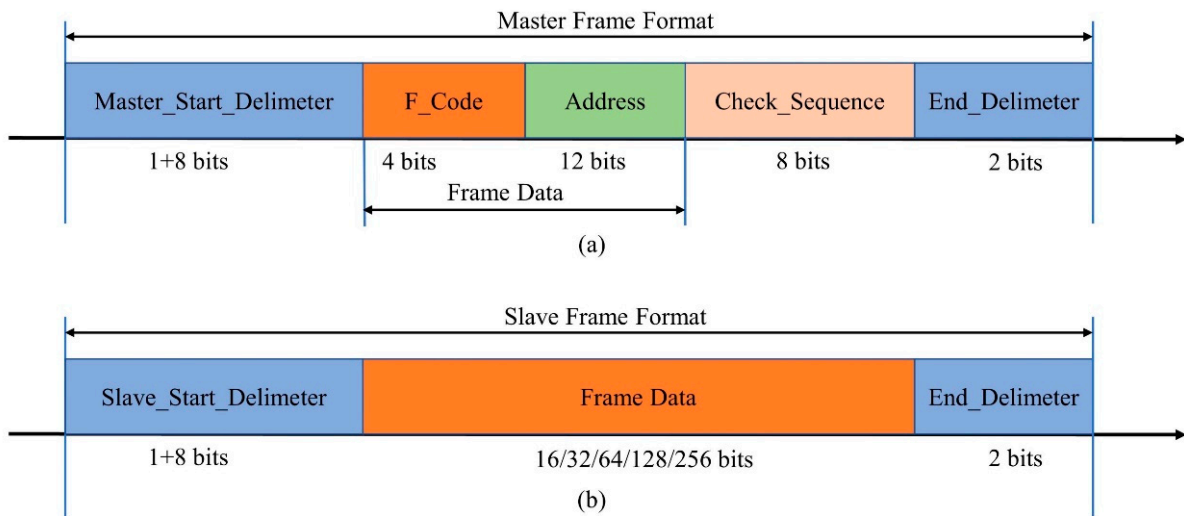

Figure 3. MVB frame formats: (**a**) format of master frame; (**b**) format of slave frame.

In addition, physical-layer waveform data have waveform features; IEC61375-3-1 [4] defines MVB physical-layer voltage waveforms in terms of a variety of features, and a software algorithm is used to extract the following features:

- Steady-state amplitude: The steady-state amplitude of the MVB physical-layer waveform shall be at least ± 1.5 V when connected to the heavy test circuit and at most ± 5.5 V when connected to the light test circuit;
- Overshoot: The overshoot of the waveform, defined as the ratio of the difference between the maximum value of the waveform and the steady-state amplitude to the steady-state amplitude, shall not exceed 10%;
- Positive and negative steady-state amplitude difference: Defined as the difference between the positive steady-state amplitude and the negative steady-state amplitude in two consecutive pulses, it shall not exceed 0.1 V;
- Slew rate: The slew rate of the waveform, defined as the rate of change of the waveform within 100 ns of the zero crossing, shall be more than 15 mV/ns;
- Edge distortion: The edge distortion, defined as the time difference between the idealized and the actual zero crossing, shall not exceed 2% of one bit time.
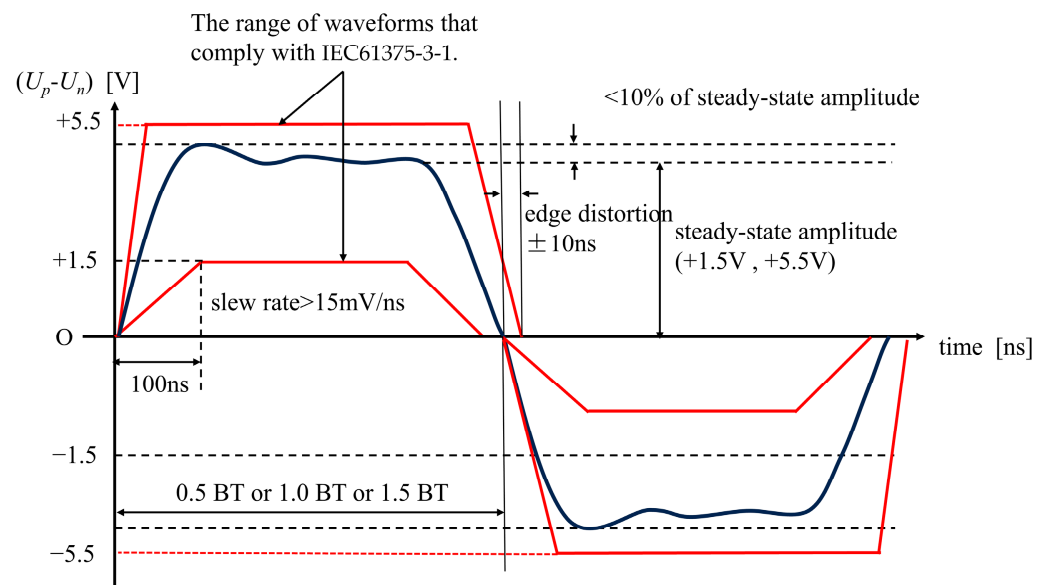- A diagram of waveform features is shown in Figure 4.

Figure 4. Features of MVB waveforms.

2.2. MVB Voltage Waveform Digitalization Algorithm Extraction Algorithm

The reason why MVB network dual-mode data are extracted is that link-layer data are only used for MVB device communication and do not contain specific network status information, which results in MVB diagnostic devices/software based on data-link-layer data having almost no effect on the diagnosis of MVB physical-layer faults, and the detection and location of specific faults still requires a lot of manpower and resources. However, the physical-layer voltage waveform exists in the transmission media between devices and is affected by the electrical characteristics of the train network and the status of the devices, so the waveforms could be applied to diagnose MVB network faults. The environment for compiling the algorithm is MATLAB.

MVB physical-layer voltage waveform data need to undergo several steps, including decoding, feature extraction, anomaly detection and synchronization with link-layer data. A flow chart of the digitalization algorithm is shown in Figure 5.
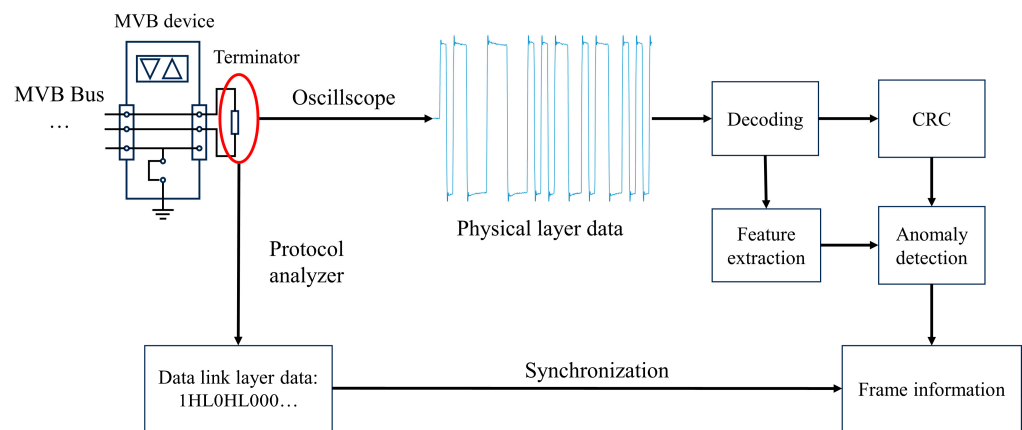
Figure 5. Flow chart of the digitalization algorithm.

2.2.1. Decoding

The electrical characteristics of MVB on-board devices vary, and the working environment is complex, so the voltage waveform of the MVB physical layer is not an ideal square-wave signal, which leads to unnecessary interference. In order to facilitate the decoding of the MVB voltage waveform, preprocessing of the waveform data is first carried out. The transmission rate, f_{mvb}, of the MVB bus is 1.5 Mbit/s, and the sampling rate of the physical-layer waveform is assumed to be f_s. In other words, a single bit in Figure 4 has s_{bit} sampling points in the actual sampling waveform. Here is the equation.

$$s_{bit} = f_s / f_{mvb} \tag{1}$$

After determining the width of the bit, the waveform signal will be adjusted to a square wave that is easy to handle. IEC61375-3-1 mentions that the steady-state amplitude of the high-level voltage of the MVB bus under light load conditions does not exceed ± 5.5 V; the steady-state amplitude of the high-level voltage under heavy load conditions is not less than ± 1.5 V; and the minimum receiving threshold of MVB equipment is (-0.2 V, $+0.2$ V). Therefore, three intervals can be set: ($-\infty$, -0.2 V), (-0.2 V, $+0.2$ V), ($+0.2$ V, $+\infty$), and the waveform signal is converted into a unified amplitude of $-1/0/1$, similar to a digital quantity. The 0 value represents that the MVB bus is in a silent state; when the MVB bus is active, 1 represents a high level and -1 represents a low level.

The preprocessed data become a string of arrays containing only $-1/0/1$, and the mutual transformation of these three numbers represents that there is an edge in the MVB physical-layer voltage waveform. The time index of all the edges can be obtained by using the algorithm for detecting numerical abrupt changes. In addition, at the physical layer, the signal of data transmission starts from 0 V to a high level, lasts for 0.5 bits, and ends with the non-data symbols "NL" and "NH". This indicates that each frame starts with a change from 0 to 1 and ends with a change from 1 to 0, so the parity order in which the array mutates can indicate the direction of a waveform edge.

As the bit time of the MVB frames is about 666.7 ns, it can be used as a basic unit of time intervals. At the same time, MVB frames consist of four kinds of data or symbols, leading to different intervals between edges. For example, the different intervals of the start delimiter are shown in Figure 6. Thus, we can decode waveforms with the information of edges.

Here is the whole process of decoding:

- Data preprocessing, and waveforms are turned into digital quantities;
- Detection of the position of edges;
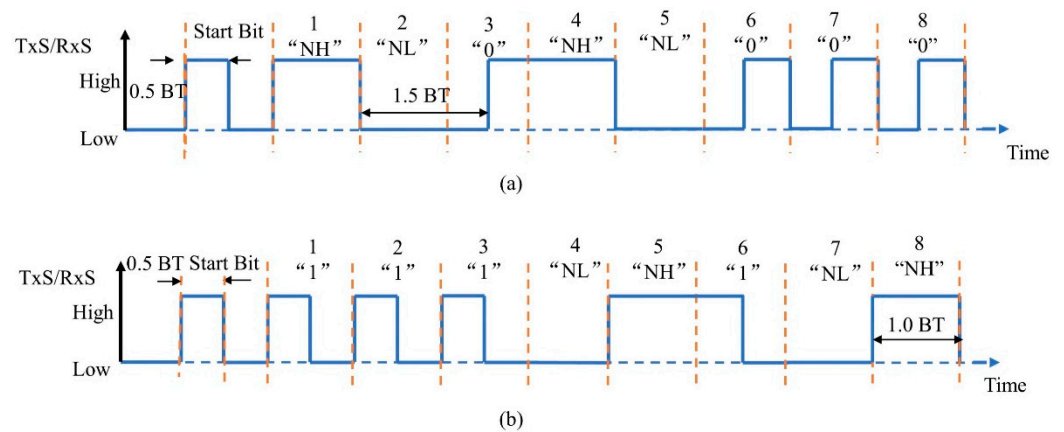- Decoding of the bits according to Table 1.

(a)

(b)

Figure 6. Start delimiter of master frame (**a**) and slave frame (**b**).

Table 1. Decoding table according to the information of edges.

Parity Order of Edges	Interval of Edges	Position of Edges	Decoding Result
Even number	0.5 BT	Middle of bit cell	None
Even number	0.5 BT	Edge of bit cell	"1"
Even number	1.0 BT	Middle of bit cell	"1"
Even number	1.0 BT	Edge of bit cell	"NL"
Even number	1.5 BT	Middle of bit cell	"NL"
Even number	1.5 BT	Edge of bit cell	"1" + "NL"
Odd number	0.5 BT	Middle of bit cell	None
Odd number	0.5 BT	Edge of bit cell	"0"
Odd number	1.0 BT	Middle of bit cell	"0"
Odd number	1.0 BT	Edge of bit cell	"NH"
Odd number	1.5 BT	Middle of bit cell	"NH"
Odd number	1.5 BT	Edge of bit cell	"0" + "NH"

2.2.2. Feature Extraction

In Section 2.1, five waveform features of the MVB physical layer are described, which can be associated with the time indexes of the edges. In Section 2.2.1, it was described that all time indexes of numerical abrupt changes will be captured. Thus, the corresponding actual waveform edges will be captured, and they are the reference points used to calculate the steady-state amplitude, slew rate, edge distortion and other waveform features.

- Steady-state amplitude: Actually, the waveform between two adjacent edges is a pulse. With the waveform function, steady-state amplitudes of every pulse can be calculated.
- Overshoot: Within a pulse, the ratio of the maximum amplitude to the steady-state amplitude is the overshoot.
- Positive and negative steady-state amplitude difference: Between three continuous edges, for example, the first bit cell in Figure 6b, there are positive and negative steady-state amplitudes. Their difference value is the target feature.
- Edge distortion: An extraction diagram of edge distortion is shown in Figure 7. The slope of the voltage waveform edge is rapid and the duration is short. Here are the steps to obtain edge distortion:

  1. Obtain the adjacent sampling points (green spots in Figure 7) of the waveform edge;
  2. Fit the sampling points with a linear curve (red line) to replace the actual waveform;
  3. Calculate the actual zero-crossing point (orange spots) of the curve;
  4. Repeat steps 1–3 for the next waveform edge;
  5. The interval can be obtained based on the adjacent estimated zero-crossing points, and the edge distortion is the ratio of the interval to 666.7 ns.

- Slew rate: The algorithmic approach to calculating the slew rate is to find the first point more than 100 ns away from the actual zero-crossing point, divide its amplitude by the time interval (100 ns) and obtain the slew rate.
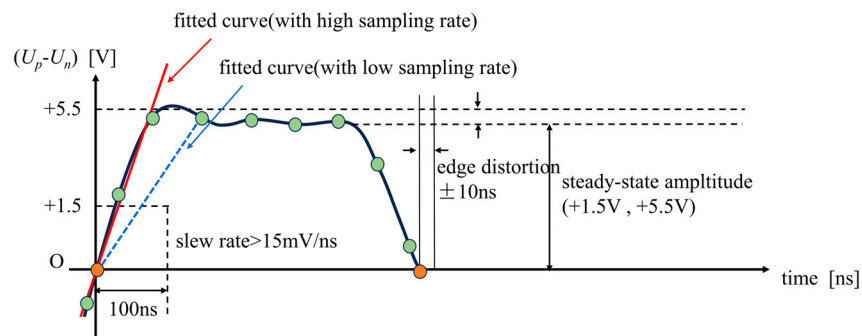

Figure 7. Diagram of the edge distortion and the fitted curve.

2.2.3. Anomaly Detection

As shown in Figure 3, data frames have a fixed format and length. The format of a data frame can be checked during the decoding process:

Any frame that does not contain a master/slave start delimiter at the start of the data frame or does not contain an end delimiter at the end is detected as a data-frame format error.

Before the CRC check, the length of the data frame is checked. Except for five specified lengths, data frames of other lengths are judged as data-frame format errors.

As the data frame contains a CRC check sequence (cyclic redundancy check; Figure 3), the CRC verification algorithm can be used to check whether the decoding is correct and analyze the reliability of the data in fault detection. IEC61375-3-1 provides the CRC check code, which is represented in Equation (2).

$$G(x) = x^7 + x^6 + x^5 + x^2 + 1, \tag{2}$$

The address and F code in the frame are the sequences to be checked. They are divided according to Equation (2) modulo 2 to obtain the remainder. After the remainder is filled up to 7 digits with "0", the parity check code of the sequence to be checked is added at the end, and, finally, the complement is taken as the final CRC sequence.

When the CRC sequence is inconsistent with the original CRC check sequence of the data frame, the algorithm determines it as a CRC check error.

The waveform feature parameters extracted in Section 2.2.2 have a specified range requirement, based on which it can be determined that there is a fault in the electrical parameters of the data frame.

2.2.4. Data Synchronization

The experimental part of this study collected MVB communication data from multiple MVB cards' communication processes, using an MVB protocol analyzer to save link-layer data and an oscilloscope to record MVB physical-layer voltage waveforms. The protocol analyzer records data types as digital quantities, so the data storage pressure is small, and the single recording duration can reach tens to hundreds of seconds. However, the physical-layer voltage waveform consists of analog data with a sampling rate of tens of MHz, the data storage pressure is high, and it can only record tens of ms. The data frames recorded by the protocol analyzer all carry timestamp information, while the data collected by the oscilloscope usually do not. In order to provide accurate timestamp information for the physical-layer voltage waveform, this study matches segments of physical-layer data with link-layer data.

Here are the synchronization procedures:

1. The first data frame of the physical-layer data is compared with the link-layer data to obtain all indexes of the same data frames.
2. Using ones of these indexes as the synchronization starting point, the data frames of the physical layer are compared one by one with the data-link-layer data (data frames that are abnormal in Section 2.2.3 will be skipped).
3. When all data synchronization is completed, it can be considered that dual-modal data synchronization is completed and that an accurate timestamp for the physical-layer data is obtained. Otherwise, procedure 2 is repeated.

2.3. Requirements of the Digitalization Algorithm

The digitalization algorithm requires two inputs: one consists of the decoded data frames obtained by the protocol analyzer, and the other is the original MVB voltage waveform. The duration of voltage waveform acquisition must be shorter than the duration of data acquisition by the protocol analyzer and can only be acquired after the protocol analyzer starts collecting data. Additionally, the waveform should not contain any significant interference or noise.

3. Experimental Results

3.1. Experimental Platform

In this study, a CPCI chassis equipped with multiple MVB cards was used as the MVB network simulation platform, and then an MVB protocol analyzer and multi-channel oscilloscope were used to capture the data of the MVB link layer and the voltage waveform data of the physical layer. A picture of the MVB network simulation platform is shown in Figure 8.
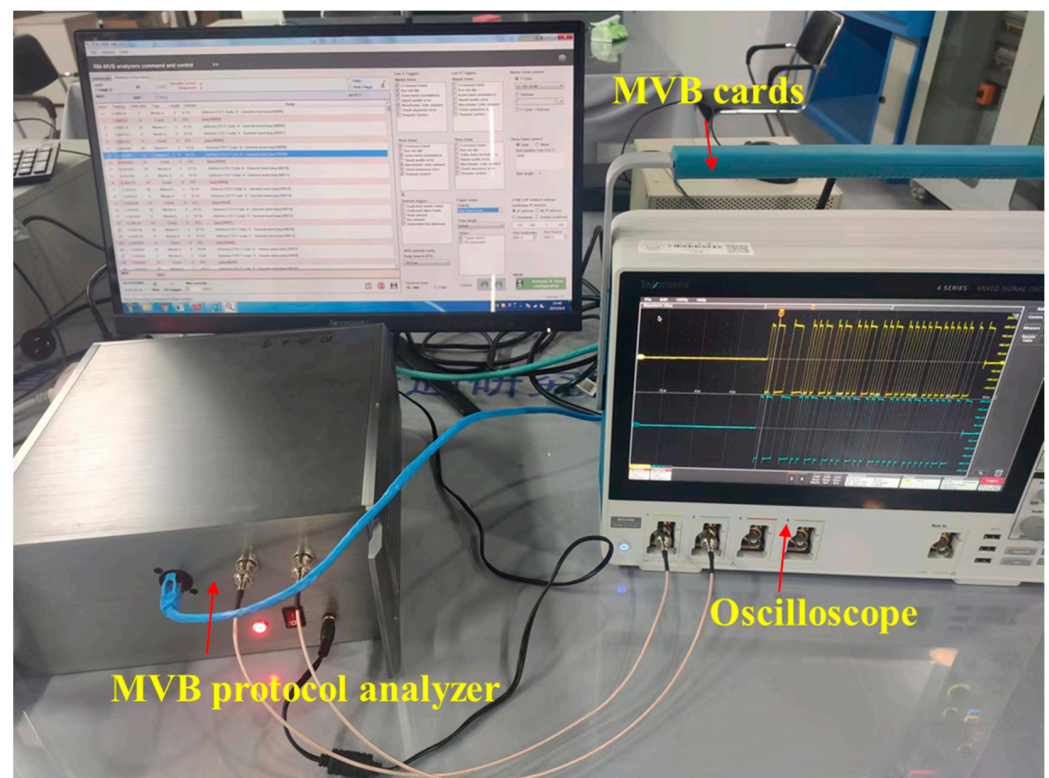

Figure 8. MVB network simulation platform.

Here are the technical parameters of the equipment shown in Figure 8. Three CPCI MVB cards (class 4 devices according to IEC 61375-3-1 2012) are inserted in the CPCI chassis, all of which possess bus administrator (BA) capability, enabling them to independently

manage MVB communication activities. The MSO44 Tektronix oscilloscope, equipped with a maximum sampling rate of 6.25 GHz and an input bandwidth of 200 MHz, achieves voltage measurements on MVB dual lines and electrical isolation of the input through two PINTECH N1008B differential probes. The bus analyzer uses the MVB protocol analyzer RB-MVB/AN02 produced by AMIT (assembled together with the differential probes). It can intercept and analyze MVB communication messages without affecting the normal communication activities. The protocol analyzer also transmits the analysis results to a PC via Ethernet.

Since the sampling rate of the physical-layer voltage waveform is limited by both data accuracy and storage analysis cost, it is necessary to optimize the sampling rate of the physical-layer voltage waveform. The sampling rate should be reduced as much as possible without significantly affecting the data content. In the experiment, multiple sampling rates were used to capture MVB physical-layer voltage waveform data, and the ideal sampling rate was determined according to the data quality.

The MVB network simulation platform has normal and fault conditions. The latter sequentially injects impedance mismatch, short-circuit and open-circuit faults. The structure of the MVB terminator using the EMD medium is shown in Figure 9. The role of the terminal resistor is to absorb energy at the end of the MVB network, reduce the degree of voltage signal reflection and reduce communication interference. In addition, the impedance characteristics of the terminal resistor are specific values. Changing the resistance value of the terminal resistor may cause different degrees of signal reflection, causing distortion in the MVB waveform. Based on determining the sampling rate of the MVB physical-layer voltage waveform, dual-modal data under different conditions of the MVB are collected and processed using a feature-extraction algorithm.
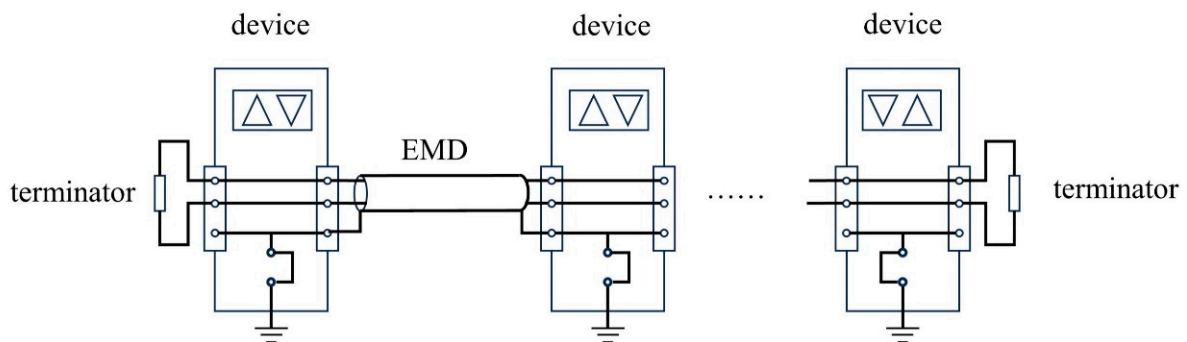

Figure 9. Diagram of terminator using EMD.

3.2. Sampling Data at Different Sampling Rates

The first stage of the experiment was the analysis and comparison of sample numbers under multiple sampling rates. Due to the specifications of the oscilloscope, the sampling rates of the physical-layer voltage waveform involved in the comparison were 6.25 MHz, 12.5 MHz, 31.25 MHz, 62.5 MHz and 125 MHz. Figure 10 and Table 2 present the waveforms and information extracted at different sampling rates.

Table 2. Waveform features of tested sampling rates.

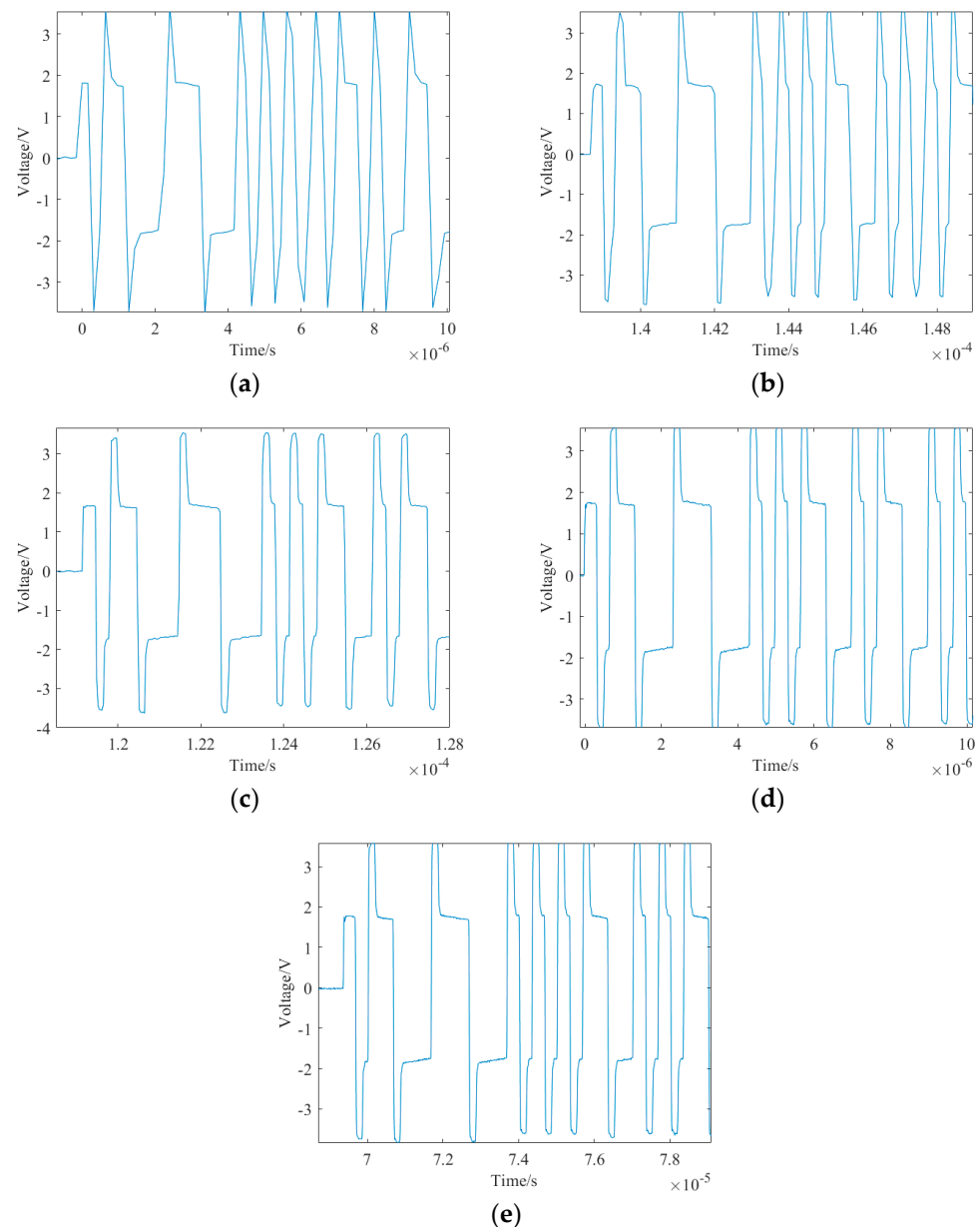| Sampling Rate/MHz	Package Loss Rate/%	Average State Amplitude/V	Average Slew Rate/(mV·ns^{-1})	Average Overshoot/%	Average Edge Distortion/%	Time Stamp	Score
12.5	0	1.81	22.98	95	1.06	4507	70.9
31.25	0	1.795	28.43	91.68	0.4	4307	81.6
62.5	0	1.84	32.5	94.4	0.2	4735	98.3
125	0	1.845	34.48	94.9	0.2	2689	100

Figure 10. (**a**) Waveform at 6.25 MHz. (**b**) Waveform at 12.5 MHz. (**c**) Waveform at 31.25 MHz. (**d**) Waveform at 62.5 MHz. (**e**) Waveform at 125 MHz.

We obtained several indicators from Table 2:

- "Package loss": This is obtained by calculating he percentage of data frames during the sampling period for which correct decoding information could not be obtained, measured against all data frames during the sampling period. It serves as a metric to assess the reliability of information transmission.
- "Timestamp" represents the starting point of the waveform collected. The numbers represent the sequence of data frames collected at the link layer.
- "Scores": These represent the similarity between waveform features at the corresponding sampling rate and the waveform at a 125 MHz sampling rate. They are obtained by calculating the ratios of the waveform features at the corresponding sampling rate to those at a 125 MHz sampling rate (with the "Average edge distortion" inverted).
- The description and the calculation procedures for the remaining variables can be found in Section 2.2.2 on feature extraction.

3.3. Sampling Data under Different Working Conditions

This stage of the experiment included the normal condition and terminator short-circuit/open-circuit/abnormal resistance faults. According to IEC 61375-3-1, the resistance value of the terminator was 120 Ω. The abnormal resistance value was 90 Ω. The sampling rate was 62.5 MHz. Figure 11 presents how these conditions are generated. Figure 12 and Table 3 present the waveforms and information extracted under different conditions.
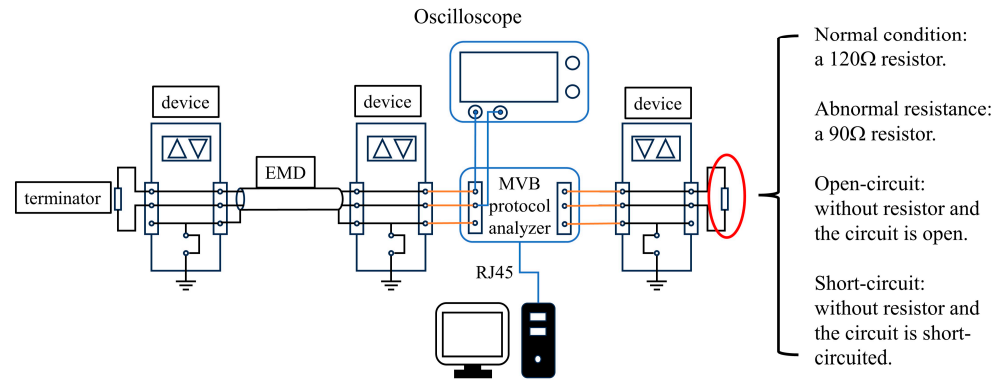

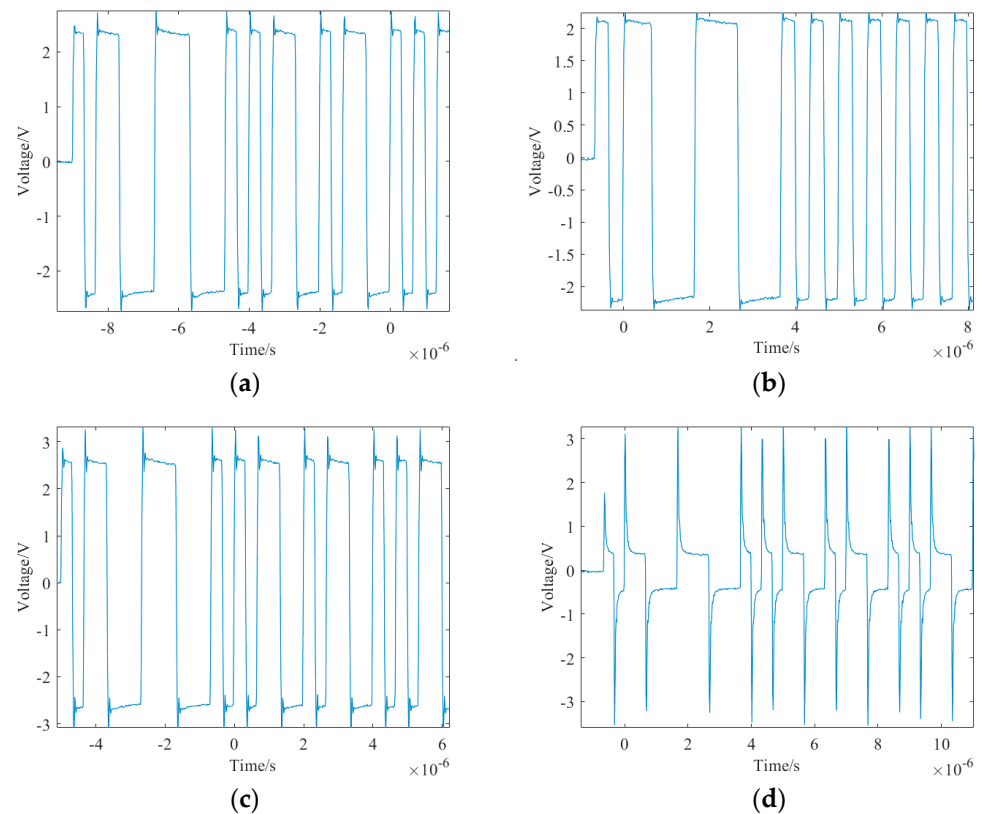**Figure 11.** Sampling data under different working conditions.

Figure 12. (**a**) Waveform without fault. (**b**) Waveform with 90 Ω terminator. (**c**) Waveform in open-circuit condition. (**d**) Waveform in short-circuit condition.

The terminators are resistors installed at the end of the MVB network with a fixed resistance value. Since the signals generated in the bus carry energy, if there is no load to absorb this energy, reflection signals will occur at the terminal, affecting the normal transmission of signals. Similarly, when the terminal resistance value does not match, reflection signals will also occur at the terminal.

Table 3. Waveform features under different conditions.

Features	Normal Condition	Abnormal Resistance	Open Circuit	Short Circuit
Package loss rate/%	0	0	0	18.5
Average State amplitude/V	2.34	2.07	2.54	0.31
Average slew rate/(mV·ns^{-1})	22.30	19.60	24.79	6.07
Average overshoot/%	14.8	8	25	674
Average Edge distortion/%	0.54	0.48	0.53	1.36
Timestamp	1969	2753	3378	none

4. Discussion

4.1. Analysis of Data Sampled at Different Sampling Rates

As shown in Figure 10a, at a sampling rate of 6.25 MHz, the waveform of the MVB physical layer barely exhibits the characteristics of pulse waveforms. The excessively low sampling rate leads to severe distortion of the MVB waveforms. Pulses occupying 0.5 BT in the waveforms do not display a steady state. The severe distortion is not only reflected in the waveform images but also in the CRC results of the MVB frames. Most of the results are wrong, which means the frames are erroneous. However, under normal test conditions, all frames output by the MVB device are error-free, which makes the decoding results at this sampling rate unreliable. Therefore, the sampling rate of 6.25 MHz is not applicable to this study.

As shown in Figure 3, there is a significant difference in the average edge distortion under different sampling rates. The average edge distortion corresponding to 62.5 MHz and 125 MHz is the same, which is more credible, while the average edge distortion under the sampling rates of 12.5 MHz and 31.25 MHz is significantly lower. The definition of edge distortion is the time difference between the idealized and the actual zero crossing The MVB conformance test requires that the edge distortion does not exceed 2%. As shown in Figure 7, the falling edge and the rising edge in the MVB waveform usually last for a short time, and it is extremely difficult to directly collect the actual zero crossing, so the feature-extraction algorithm fits several sampling points along a straight line to approximate the actual zero crossing. However, a lower sampling rate (31.25 MHz and below) is not enough to capture enough data to fit the waveform, and only data within the edge can be used for fitting, causing the estimated actual zero crossing to deviate from the true value, resulting in a difference in the extracted average edge distortion.

Among other feature parameters, there is little difference in the average steady-state amplitude and the average overshoot, while there is a significant difference in the average slew rate. The average slew rate is obtained by calculating the average change speed near 100 ns of the zero crossing. However, just like the actual zero crossing, it is difficult to accurately obtain the value of 100 ns under low sampling, and only sampling points near 100 ns can be selected instead. For example, the sampling interval of a 12.5 MHz sampling rate reaches 80 ns, and the interval to calculate the slew rate may reach 160 ns—far more than 100 ns.

According to the comparison of waveform feature quality at various sampling rates, 62.5 MHz can be regarded as a better sampling rate. Compared with sampling data at lower rates, it can obtain higher-quality waveform features; compared with 125 MHz, the amount of data generated in the same time is halved, reducing data storage costs and computational analysis pressure.

4.2. Performance of the Algorithm under Different Conditions

Under different conditions, there is a significant difference in steady-state amplitude, and the overshoot is also different.

When the terminator is replaced with a 90 Ω resistor and the terminator is removed, the steady-state amplitude changes slightly; after the line impedance characteristics change,

the reflected signal is enhanced. There is also a slight change in overshoot under these two conditions, but it does not affect the decoding and synchronization process of the MVB physical-layer waveform.

However, when the terminator is short-circuited, a strong reflected signal can be seen in Figure 12d, and the steady-state amplitude of the voltage waveform drops rapidly, with an overshoot exceeding 600%. In such extremely harsh conditions, signals with an amplitude less than 1.5 V will not be recognized by MVB devices, causing abnormal MVB network communication. Therefore, the protocol analyzer cannot collect link-layer data, but the voltage waveform digitization algorithm can adjust the threshold to capture frame data.

Through normal operating conditions and fault-condition tests of the MVB voltage waveform digitization algorithm, the experimental results show that this algorithm can achieve waveform feature extraction and data decoding functions.

4.3. Comparison with Other Studies

Table 4 presents a comparison with other studies. The sampling rate in this study was much lower than in other studies. Additionally, both physical-layer and link-layer data were collected, and time synchronization was achieved. Therefore, this algorithm can be applied in data-acquisition devices that initiate physical-layer waveform capture through a protocol analyzer, which indirectly increases the density of relevant information. But compared to other studies, the number of waveform features of this algorithm is less, which needs further improvement.

Table 4. Comparison with other studies in terms of data acquisition and processing.

Studies	This Study	Study Based on SVM [13]	Study Based on MKLSVM [16]
Sampling rate/MHZ	62.5	125	100
Data types	Physical-layer waveform and link-layer data	Physical-layer waveform	Physical-layer waveform
Waveform features	4	6	7

5. Conclusions

In order to collect and analyze MVB data more practically, an MVB voltage waveform digitization algorithm was developed, and a sampling rate for physical-layer waveforms was obtained. This algorithm is based on the edge of the voltage signal to realize the decoding of the MVB physical-layer voltage waveform, data-frame verification, and synchronization of physical-layer data and link-layer data. And, according to the position of the edge and the curve fitting method used to determine the actual zero crossing of the waveform so as to obtain four waveform features of the MVB voltage waveform—steady-state amplitude, slew rate, edge distortion, and overshoot—through experiments, the sampling rate of the MVB physical-layer voltage waveform was determined. The experimental results show that 62.5 MHz is the best sampling rate, which can balance lower data volume and higher data quality; whether under normal operating conditions or fault conditions, the MVB voltage waveform digitization algorithm can better realize waveform decoding, anomaly detection and feature-extraction functions.

Compared to previous data-acquisition methods (125 MHz), the sampling rate determined in the study was 50% lower. However, in terms of comprehensive waveform features extracted by the digitalization algorithm, the accuracy was only reduced by 1.7%, and there was almost no difference in waveform features.

However, in extreme fault conditions, such as terminal resistance short circuit, the data-link layer cannot obtain enough valid frame data, which results in the synchronization of this algorithm being unable to handle them.

Additionally, the algorithm's capability to extract waveform characteristics is limited and cannot fully represent the waveform's state.

Further improvement of the algorithm's feature extraction and synchronization strategy is needed. In addition, only limited types of faults could be generated in the experiment. Diverse fault data will be helpful to improve the ability of the MVB voltage waveform digitization algorithm to handle different faults.

Author Contributions: Conceptualization, H.L. and Y.L.; methodology, H.L. and Y.L.; software, X.T.; validation, X.T.; formal analysis, X.T.; investigation, X.T.; resources, X.T.; data curation, X.T.; writing—original draft preparation, X.T.; writing—review and editing, H.L. and X.T.; visualization, X.T.; supervision, Y.L.; project administration, Y.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

References

1. Feng, J.H.; Lu, X.Y.; Yang, W.F.; Tang, J. Survey of Development and Application of Train Communication Network. In Proceedings of the International Conference on Electrical and Information Technologies for Rail Transportation—Electrical Traction, China Electrotechn Soc, Elect Equipment Technol Comm Rail Transportat, Zhuzhou, China, 28–30 August 2015; pp. 843–854.
2. Güldiken, M.C.; Schmidt, E.G.; Schmidt, K.W. Telegram Scheduling for the Multifunction Vehicle Bus (MVB): Algorithms and Evaluation. In Proceedings of the 25th IEEE Symposium on Computers and Communications (ISCC), Rennes, France, 7–10 July 2020; pp. 229–234.
3. Guldiken, M.; Schmidt, E.G.; Schmidt, K.W. Heuristic Algorithms for the Telegram Scheduling on MVB. In Proceedings of the 28th Signal Processing and Communications Applications Conference (SIU), Gaziantep, Turkey, 5–7 October 2020.
4. *IEC 61375-3-1:2012*; Electronic Railway Equipment—Train Communication Network (TCN)—Part 3-1: Multifunction Vehicle Bus (MVB). IEC: Geneva, Switzerland, 2012.
5. Chen, H.R.; Qian, C.Y. Research on Functional Safety of Multifunction Vehicle Bus in Rail Transit. In Proceedings of the IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS), Dalian, China, 20–22 March 2020; pp. 255–259.
6. Xia, M.; Lo, K.M.; Shao, S.J.; Sun, M.A. Formal Modeling and Verification for MVB. *J. Appl. Math.* **2013**, *2013*, 470139. [CrossRef]
7. Li, Z.Z.; Wang, L.D.; Yang, Y.Y.; Du, X.M.; Song, H. Health Evaluation of MVB Based on SVDD and Sample Reduction. *IEEE Access* **2019**, *7*, 35330–35343. [CrossRef]
8. Jiang, Y.; Wang, M.Z.; Su, Z.; Yang, Y.X.; Wang, H.H. Formal Design of Multi-Function Vehicle Bus Controller. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 3880–3889. [CrossRef]
9. Hua, Z.J.; Huang, F.C.; Shen, J.; Chen, Z. An MVB Signal Capturer Based on Microcontrollers for Metro Train On-line Health Monitoring System. In Proceedings of the 6th IEEE International Conference on Software Engineering and Service Science (ICSESS), China Hall Sci & Technol, Beijing, China, 23–25 September 2015; pp. 255–258.
10. Yang, Y.; Wang, L.; Chen, H.; Wang, C. Anomaly Detection Method for MVB Network Based on Variational Autoencoder. *J. China Railw. Soc.* **2022**, *44*, 71–78.
11. Li, Z.; Wang, L.; Yang, Y.; Shen, P. Intermittent Connection Fault Location of MVB. *J. China Railw. Soc.* **2021**, *43*, 107–114.
12. Yang, Y.; Wang, L.; Wang, C.; Wang, H.; Li, Y. Fault Diagnosis Method Based on Deep Active Learning for MVB Network. *J. Southwest Jiaotong Univ.* **2022**, *57*, 1342–1348, 1385.
13. Li, Z.; Wang, L.; Yang, Y. Fault diagnosis of the train communication network based on weighted support vector machine. *IEEJ Trans. Electr. Electron. Eng.* **2020**, *15*, 1077–1088. [CrossRef]
14. Li, Z.; Wang, L.; Shen, P.; Song, H.; Du, X. Fault Diagnosis of MVB Based on Random Forest and Ensemble Pruning. In *Lecture Notes in Electrical Engineering, Proceedings of the 4th International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2019, Qingdao, China, 25–27 October 2019*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 91–100.

15. Wang, J. A Method of Fault Diagnosis and Location for Train MVB Communication Network. *Proc. 16th Annu. Conf. China Electrotech. Soc.* **2022**, *890*, 794–802. [CrossRef]

16. Li, Z.; Wang, L.; Yue, C.; Shen, P. Terminating fault diagnosis of MVB based on MKLSVM. *J. Beijing Jiaotong Univ.* **2019**, *43*, 100–106.