Analysis Platform of Rail Transit Vehicle Signal System Based on Data Mining

Chunying Li*, Zhonghua Mu
[1]Department of Electronic Engineering of Zhengzhou Railway Vocational & Technical College, Zhengzhou, Henan, 451460, China
E-mail: chunyingli7@126.com, zhonghuamu6@163.com
*Corresponding author

Keywords: rail transit, data mining, detection, maintenance, vehicle signal system

Received: January 26, 2022

According to the increasing demand of interactive information of rail transit on-board signal equipment, a rail transit on-board monitoring and maintenance system based on data mining is proposed. The association rules for operation data acquisition are defined and based on these rules a correlation rules algorithm is proposed to obtain more reliable understanding and operation quality evaluation of train operation information. A new approach for the safety analysis is proposed for the analysis of failure mode of rail braking system. The proposed approach uses Bayes method reasoning to introduce reliability and probability analysis of braking system, hence guiding the maintenance strategy of the system. From a lot of logs, quickly find key issues, applied in the train test and repair field. The simulation experiment results show that after analyzing the simulation data and the curve, the system extraction results have certain error in the manual calculation results, and the error value is between 0.5 and 0.6, but the overall meets the actual work needs, and optimize the invalid data to reduce the error. The reliable operation and maintainability of the system are verified.

Povzetek: Razvit je sistem za spremljanje železniškega prometa s podatkovnim rudarjenjem z uporabo Bayesove metode za zanesljivost.

1 Introduction

With the rapid development of rail transit, the degree of automation and the amount of interactive information of on-board signal equipment are increasing, while the maintenance time is continuously shortening, which puts forward higher requirements for the highly reliable operation and safe maintenance of signal equipment. The traditional way of tracking the operation of software and hardware by separately recording error codes in each subunit [1], obviously cannot meet the requirements, and requires an intelligent, concise and user-friendly human-machine interface, which can grasp the overall situation of the vehicle signal equipment in real time. It can provide the expected operation data for testers and maintenance personnel, and realize rapid commissioning and maintenance of trains and lines. At present, in the field of urban rail transit, the application of on-board monitoring and maintenance systems is still relatively rare. In the process of developing the domestic vehicle signal system, the existing achievements of many parties were used for reference, and the multi-source information and data were graded in the data processing. Converted into meaningful data, resulting in abnormal monitoring to assist debugging and maintenance personnel in analyzing equipment failures.

In this paper, based on data mining, the design of rail transit on-board monitoring and maintenance system (TIU, Transit-vehicle Interface Unit) is designed. The main purpose of the design is to comprehensively record the running status of on-board signal equipment), ATO (Automatic Train Operation) and other equipment operation status real-time display and comprehensive analysis, and provide functions such as log download, format conversion and offline analysis, so as to achieve highly reliable operation and safe maintenance of signal equipment. The traffic vehicle monitoring and maintenance system is shown in Figure 1.

In the field of railroad transport, SARM (Security, availability, reliability and maintainability) is generally used to study the functional unwavering quality of gear. In 1986, Sweden gave the primary SARM record necessity in a delicate for the acquisition of fast trains, expecting providers to focus on their unwavering quality, viability and security, and to guarantee that all pointers meet a predefined esteem in the wake of dispatching. During the 1970s, Shinkansen disappointment information was dissected in Japan and enhancements to the plan worked on train unwavering quality. The trains in France have severe SARM prerequisites at the plan stage and their upkeep costs are diminishing step by step. Jing *et al.* played out an assessment of the difficulty free state of SARM on a high velocity rail line [2]. Ma *et al.* brings SARM into metropolitan rail line train checking framework [3]. Liu *et al.* dissected the critical components of SARM control in the rail line industry [4]. Zhong *et al.* present a Bayesian technique for the development of probabilistic organizations and play out an assessment of the framework's credibility [5]. Zhou *et* *al.* subjectively dissect the shortcoming tree of the pressure driven sponsor in the control arrangement of a business vehicle [6]. Wu *et al.* proposed a Bayesian piecewise investigation of a straight model [7]. By changing the information, a Bayesian neighborhood direct model with prescient and nearby straight circulation of boundaries is gotten. Zhong *et al.* dissected the SARM control measures taken at various phases of the flagging framework [8]. Vega *et al.* clear up how for use SARM to control a rail route flagging framework [9]. Horton *et al.* introduced a SARM examination of a rapid rail route power supply framework [10].

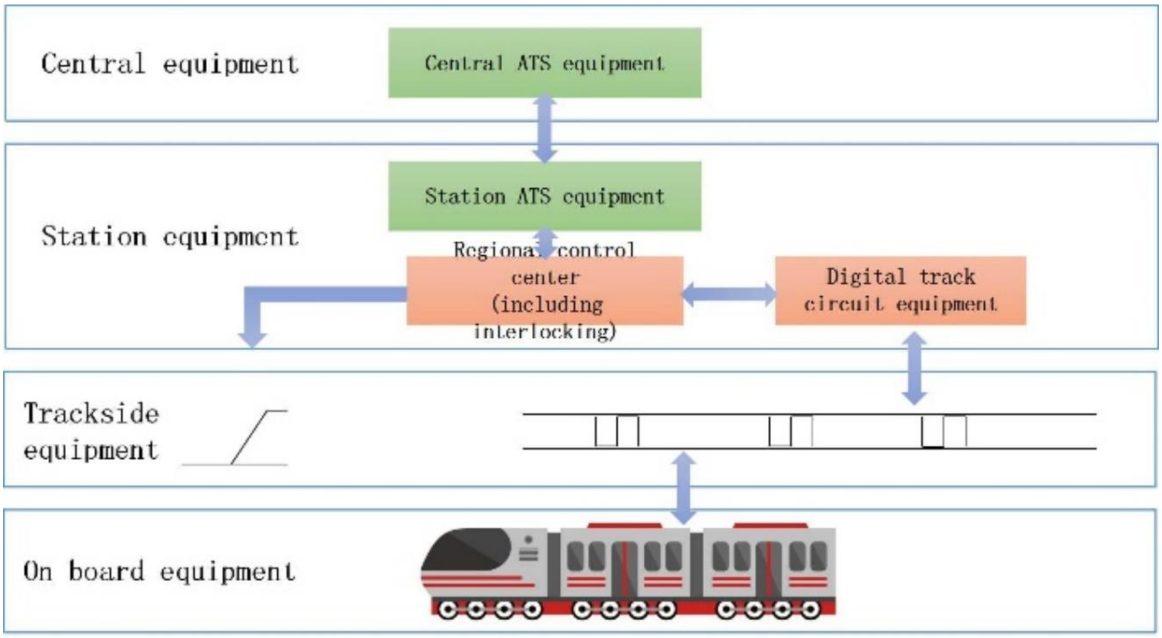

Figure 1: Rail transit vehicle monitoring and maintenance system.

The GO technique is broadly utilized in SARM examination of rail transport frameworks [11]; Zhang *et al.* [12] apply SARM to 9 electromechanical subsystems of travel line 1 of Chengdu Rail, yet there are a few cases effectively utilized in air powered brake frameworks. Numerous specialists utilize the GO-FLOW strategy in rapid rail line security examination [13]. A few specialists apply large information to rail line traffic activity and the board to direct upkeep techniques [14, 15]. The rest of this article is organized as: Section 2 presents the related works in various domains. Section 3 consists of methods comprising the concept. Results and analysis are discussed in Section 4 followed by concluding remarks in section 5.

2 Related work

With the continuous advancement of technology, the technology of video collection data has developed rapidly. Pedestrian movement trajectories can be obtained through video, so as to analyze the characteristics of pedestrian traffic behavior. Zhao *et al.* developed video-based Petrack software for automatic or semi-automatic identification and determination of pedestrian motion positions and trajectories [16]. Wang *et al.* developed a Kinect-based pedestrian trajectory extraction technology for long-term high-precision pedestrian trajectory extraction [17]. Zhigang *et al.* conducted a long-term study on the microscopic traffic behavior of pedestrians. The observed pedestrians need to wear hats of different colors. Using the hats as recognition conditions, the pattern recognition of the captured video images can be used to obtain the motion trajectories of ordinary behaviors and specific behaviors [18]. Chen *et al.* propagandized the pedestrian detection and tracking technology based on Blob analysis, and developed a passenger micro-traffic behavior parameter acquisition system (Ped Trace) to extract traffic behavior characteristic data such as trajectory, passenger speed, pedestrian distance, and acceleration [19]. Ding *et al.* used two-way channel monitoring in urban rail transit stations to manually determine the walking trajectory according to the projected position of the center of gravity of pedestrians' feet, and analyzed the characteristics and laws of pedestrians' overtaking traffic behavior [20]. With the increase of urban population density, the scale of subway construction is also getting larger and larger, and the signal system of rail transit is facing the challenge of more efficient and safer demand. The mobile block train operation control system represented by the Communication Based Train Control System (CBTC) [21] has been popularized to replace the fixed block, and is currently being used for train-to-vehicle communication and fully automatic operation (Fully Automatic Operation, FAO) to further evolve. In order to achieve a shorter running interval of trains in operation, on the one hand, it is necessary to shorten the fault recovery time of the system and make efforts in the direction of immediate maintenance; at the same time, it is also necessary to study the status information of the system and equipment to achieve fault prediction and status repair [22].

Domestic research on PHM (Prognostics Health Management) started later than foreign countries, but relatively speaking, domestic research on PHM in military equipment was earlier [23]. Since the 21st century, the research related to PHM technology has made great progress, mainly reflected in the research of health management and other disciplines. In terms of the top-level design of the PHM system, the aerospace-related research is relatively in-depth [24], and some progress has been made in the PHM research in the fields of machinery, electromechanical, and electronics. The operation safety of the subway has become the research focus of the subway operating companies. In recent years, PHM has gradually been applied to the operation and maintenance management of the subway [25]. However, due to the late start, mature system-level products have not yet appeared. Therefore, in order to improve the operation safety of the subway, improve the maintenance and maintenance efficiency of key equipment, and achieve the purpose of reducing the cost and increasing the efficiency of the subway, the subway system equipment, especially the signal The PHM research of the device is extremely critical [26]. PHM's research in the field of subway is mainly about the health management of high-speed rail equipment, and the application of fault prediction and health management technology to achieve equipment maintenance. For example, algorithms such as VQ and DTW are introduced in the detection of vehicle axle temperature, and the technology originally used in voice signal processing is applied to train health management.

To sum up, based on the current status of maintenance methods of subway equipment at home and abroad and the experience of other industries, combined with the characteristics of domestic rail transit technology, it is proposed to study the maintenance and management technology of rail transit on-board equipment supported by rail transit big data information. development direction [27]. The proposed work can further be extended by using integration approaches of Artificial Intelligence and Machine learning as studied from several studies [28-30]. A technique which considers wavelet frames for micropolar fluid flow is used for high mass transfer [31]. The vibration over the sandwich plates of laminated skew is studied through finite element [32]. The numerical simulation based on space time fractional equation are evaluated [33].

3 System design

In this section, the system structure, functions and association rules based on data mining technique is provided.

3.1 System structure

The system includes a vehicle-mounted unit (lower computer) and a portable maintenance terminal (upper computer), and its hardware and software both adopt a modular design [34]. The on-board unit is placed in the signal cabinet of the train, and has various modes such as network interface, RS-232/422 serial interface, MVB bus interface, etc. It can adapt to the LAN connection method required by ATP, ATO and other equipment, and the required RS-232/422 serial cable connection method, and the MVB bus method used by the vehicle's train management system, etc. The portable maintenance terminal is realized by using a notebook computer. When necessary, it can be connected to the maintenance port of the TIU vehicle-mounted unit by using the RJ-45 network interface. The system structure diagram is shown in Figure 2.

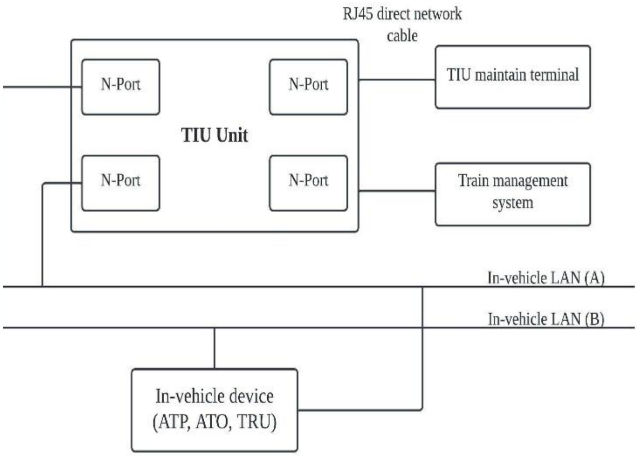

Figure 2: System structure diagram.

A. Hardware structure

CPCI bus technology has the technical characteristics of high openness, high reliability, high versatility, hot swap ability, and good anti-vibration and heat dissipation. Therefore, the CPCI bus architecture is adopted in the hardware design of the vehicle-mounted unit. All TIU boards are installed in a 3U cage. It mainly includes power board, CPU motherboard, analog/digital mixed I/O board, Ethernet card, etc. [35].

The power supply board supplies power for the whole system, selects DC110V/±12V/5V/3.3V standard power module, and sets up galvanic isolation device and filter voltage regulation protection circuit to ensure stable and safe voltage to ATO equipment.

i. The power supply board supplies power for the whole system, selects DC110V/±12V/5V/3.3V standard power module, and sets up galvanic isolation device and filter voltage regulation protection circuit to ensure stable and safe voltage to ATO equipment.

ii. The CPU board is the core of computing and control. The TIU needs to exchange data externally, and it needs to output display and input data for the convenience of debugging. Therefore, the CPU board has high reliability with standard serial port, USB interface, VGA interface and external keyboard and mouse.

equipment.

iii. The analog/digital I/O board is the input and output interface of the TIU, so the analog/digital I/O board has enough input and output channels, including digital input and output, analog output and pulse input.

iv. The Ethernet card is the network interface for the communication between the TIU and other on-board components, and a device with a sufficient number of ports and a firm and reliable port connection is selected. All TIU components are designed according to the wide temperature standard of -40℃～+85℃, and the environmental adaptability and electromagnetic compatibility characteristics conform to the relevant technical standards of rail transit on-board equipment.

B. *Software structure*

Due to the different operating platforms, the software structure includes upper computer and lower computer software. The upper computer software is used for data display and analysis, and is developed using the Windows platform; the lower computer software plays the role of data storage and real-time forwarding, and is developed using a tailored version of Linux. Each software in turn contains multiple modules [36].

The software structure diagram of the lower computer is shown in Figure 3. The software of the lower computer is divided into the bottom general module, the business processing module and the task management module. The low-level general module is closely related to the operating system and hardware drivers. Through the modular programming method, the network port, serial port, file IO, etc. are complicated to set up, and the complex low-level functions are encapsulated into a general module with a simple interface, so as to facilitate the calling of the business processing module. The business processing module separates the processing sub-modules of each data type into a process, and establishes a separate communication channel for it, so as to ensure that the data of different subsystems such as ATP and ATO can be processed in parallel inside the TIU without affecting each other; The task management module is responsible for program startup management, process daemon during program operation, and program exit management to ensure program execution branching and running stability.

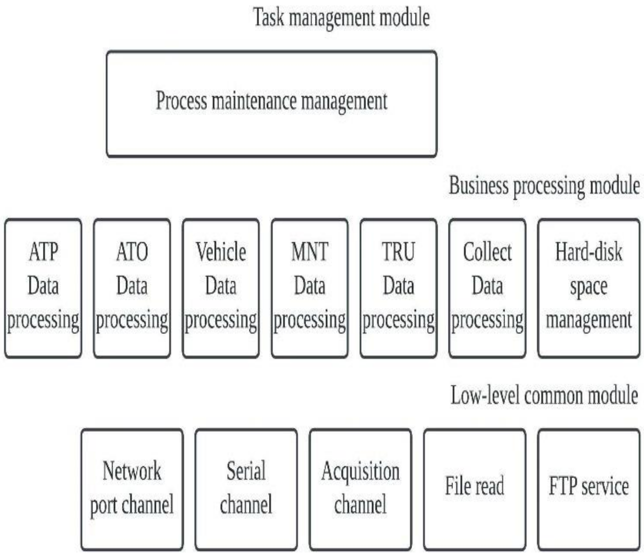

Figure 3: Lower computer software structure diagram.

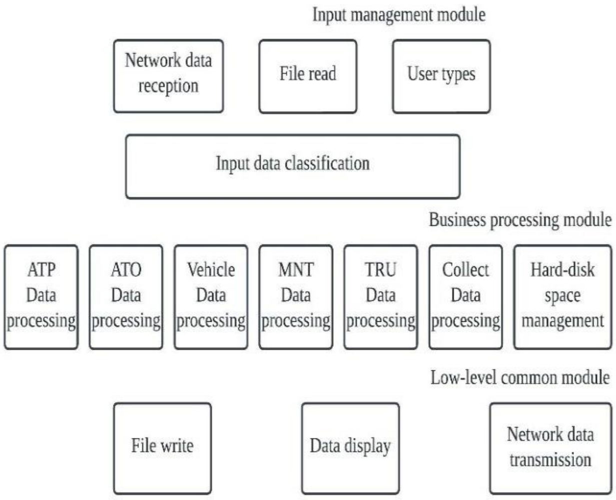

Figure 4: Host computer software structure diagram.

The software structure diagram of the host computer is shown in Figure 4. The upper computer software is divided into input management module, business processing module and output management module. The input management module encapsulates the complex data source processing functions such as network ports, file IO, and user interaction into sub-modules with simple interfaces and convenient calls; the business processing module is responsible for classifying and processing different input data and opening up independent data memory space , establish an independent data processing thread to ensure that the data of different subsystems such as ATP and ATO can be processed synchronously and independently output on the host computer [37]; the output management module is responsible for the module encapsulation of the output function, so as to complete the processing of the business processing module. The data can be classified and displayed, stored and recorded, and forwarded through the network.

3.2 System functions

3.2.1 Lower computer function

The lower computer software is responsible for the acquisition and forwarding of the underlying data, and runs in the on-board TIU host. Real-time forwarding to the maintenance terminal; ① Record the operation data and alarm data of vehicle-mounted signal equipment such as ATP and ATO; ② Relay the operation data of vehicle-mounted signal equipment such as ATP and ATO to the maintenance terminal in real time; ③ transfer the configuration information frame between ATP and MNT; ④ cooperate with MNT to complete TIU's own parameter configuration; ⑤ receive the time synchronization information of ATS and synchronize the local clock; Communication information; ⑦ Collect fan, vehicle power status, and send power-off protection frame to ATO; ⑧ Automatically maintain hard disk space to ensure effective storage of records; ⑨ FTP data service, provide download function of record files.

3.2.2 Host computer configuration

As the window of the system, the host computer software runs in the external maintenance terminal (MNT), the main functions: ① real-time data display, real-time display of the operation and alarm data of ATP, ATO and other equipment; ② offline log display, the display is saved in the machine ③ TIU data download, download the log files stored in the on-board TIU host through FTP; ④ Data format conversion, convert the log file into Excel form; ⑤ System parameter configuration, configuration the system of ATP, ATO, TIU and other systems Parameters; ⑥ Log analysis, assist testers to analyze log data, during the analysis process, the program uses data mining technology to comprehensively analyze logs from different sources and different time periods, so as to select possible useful states or faults [38].

3.3 Application of data mining technology in log analysis

3.3.1 Mining algorithm selection and introduction

The data recorded by TIU has the characteristics of being complex, multi-level and uncertain: ① There are various sources, including the status of on-board equipment such as ATP and ATO, as well as the information of vehicles and trackside equipment; ② There are many states, including door status and speed information , location information, alarm information and other thousands of information states; ③ There are various types of states, and the data types of different information states vary widely, such as Boolean, integer, floating point, string, etc.; ④ The dispersion of log records, The randomness of the time span, etc. According to these characteristics, a one-dimensional simple association rule model that is simple in form, easy to understand, and can effectively capture the relationship between data is adopted in the log analysis algorithm [39].

3.3.2 Principles of association rules

Association rule mining can be formulated as follows: Let I = (i1, i2, ...in) be a set of items, and T = (t1, t2, ...tn) a set of transactions, where each transaction ti is a set of items and satisfies ti. An association rule for ∈I: X→Y, where X∈I, Y∈I, and X∩Y=. Support and confidence are two commonly used indicators to measure the strength of association rules. The support of rule X→Y refers to the percentage of transactions that contain itemset X∪Y in transaction set T, so the support of rule Ts represents the frequency of rule use in transaction set T [40].

$$T_s = \frac{(X \bigcup Y) \cdot count}{n} \tag{1}$$

where n is the total number of transactions in T. Confidence rule, the confidence of X→Y refers to the percentage of transactions that contain both X and Y to all transactions that contain X. It can be regarded as an estimate of the conditional probability P(X|Y). The confidence Hs determines the predictability of the rule [41].

$$H_s = \frac{(X \bigcup Y) \cdot count}{X \cdot count} \tag{2}$$

Association rule mining refers to finding out the association rules in T whose support and confidence are higher than a user-specified minimum support (min-supp) and minimum confidence (minconf) respectively [42].

3.3.3 Algorithms of association rules

The algorithm of association rules, the Apriori algorithm that uses candidate item sets to find frequent item sets, is mainly divided into two steps: 1. Generate all frequent item sets, one frequent itemset is an itemset whose support is higher than minsup; 2. From the frequent item sets Generate all trusted association rules, a trusted association rule is a rule with a confidence greater than minconf. Association rules only need to be generated based on frequent item sets [43]: extracting all association rules from frequent item sets f needs to use all non-empty subsets of f, let a be any non-empty subset of f, then: (f-a) →a, if the confidence Ts satisfies:

$$T_s = \frac{f \cdot count}{(f-a) \cdot count} \geq \min conf \qquad (3)$$

An association rule generation algorithm similar to frequent itemset generation can be used: first generate 1-association rules with only one item of all consequences from the frequent itemset f (k-itemset is a set containing k items), and then use the association rule. The rules generate consequent 2-association rules, which are recursive in turn to generate all frequency sets.

4 Results and analysis

Through the design method described in the previous chapter, in the log-assisted analysis, the association rules in data mining can be applied as follows.

The first step is to establish a model, by setting the granularity of conditions and conclusions, artificially setting intervals, selecting single values, setting fuzzy values, and setting the certainty of the rules, mainly in the setting of precise rules and conceptual rules, The train operation data is used as the training data set, and the classification data is tested with the above settings. This process can remove redundant data and irrelevant attributes, and find all high-frequency item groups. For example, if you need to count the data of train stops, you need to set the query time period, a certain platform number, the stop sign, and whether the distance from the stop sign is less than the set value. The above settings are a group of high-frequency items, while in another group in the project group, such as the door opening process, there are also check stop signs and platform numbers.

The second step is to find the support and confidence of the high-frequency item group, set the thresholds for both, and discover the association rules. As shown in the above example, if the stop sign appears in both project groups, the support degree is 0.5, and the platform number also appears in the two projects at the same time. When the stop sign appears, the confidence level is 1. Confidence is always greater than support. If the support degree is greater than the user-set value (temporarily set to 0.5), then these two items are frequent item sets.

The third step is to display and evaluate association rules. Simple one-dimensional association rules are used to calculate frequent item sets, and strong association rules are generated from frequency sets, which must satisfy both minimum support and minimum confidence (temporarily set to 0.5), thus generating two valid values, and so on. The item set is calculated, all frequency sets are generated, and the calculation results are displayed, and finally a subjective screening is carried out. Obvious irrelevant information is excluded. In this way, the association rule mining of a set of information is completed. By mining different information and applying association rules for different purposes, possible errors of programs and possible failures of equipment can be screened out [44]. The abnormal information screened out when the transponder fails is shown in Table 1, which means that the transponder receives the default message at a certain speed and at a certain time. The abnormality will not affect the train running in continuous mode. Under the formula, the abnormality becomes a dominant fault [45].

Table 1: Abnormal information filtered out when the transponder fails

Line number	Time	Speed	Transponder message	Transponder	Screening amount
2711	6:22:54	70.99	default	0x012E004A	1
2712	6:22:55	70.93	default	0x012E004A	1
2713	6:22:55	70.96	default	0x012E004A	1
2714	6:22:56	71.04	default	0x012E004A	1
2715	6:22:57	70.96	default	0x012E004A	1
2716	6:22:57	70.74	default	0x012E004A	1
2717	6:22:58	70.53	default	0x012E004A	1
2718	6:22:58	70.74	default	0x012E004A	1
2719	6:22:59	70.58	default	0x012E004A	1
2720	6:22:59	70.68	default	0x012E004A	1

By monitoring such abnormal signs, it is possible to regularly check the running status of the backup system equipment, and find the fault of the backup system equipment in time. To avoid the situation that when the main system is working, the backup system has failed but has not been repaired in time, and the backup system cannot take over in time after the failure of the main system [46].

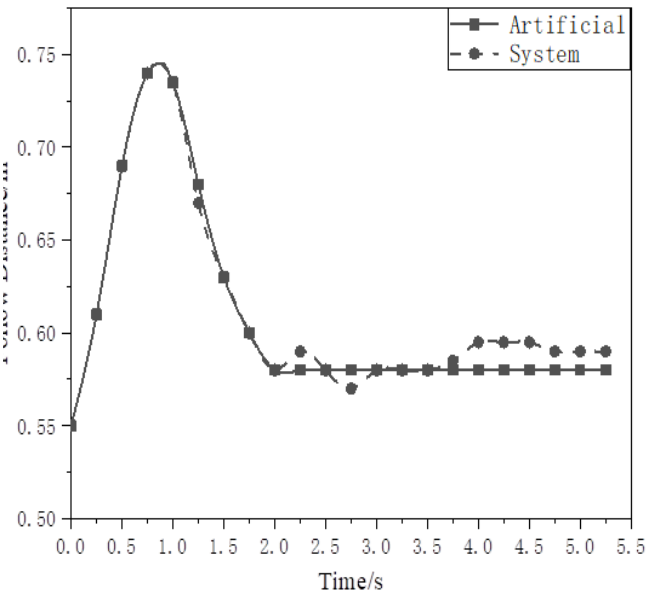

Figure 5: Following distance feature extraction results.

Use the designed rail transit on-board signal system to conduct simulation experiments to extract various behavioral characteristic parameters, and compare the results extracted by the system with the manual calculation results, as shown in Figures 5, 6, and 7 below. After analyzing the simulation data and curves, it can be seen that there is a certain error in the extraction results of the system compared with the manual calculation results, error values are between 0.5 and 0.6, but it generally meets the actual work requirements. At the same time, the system data mining method can be further optimized. Invalid data such as inflection point data are cleaned to reduce errors.

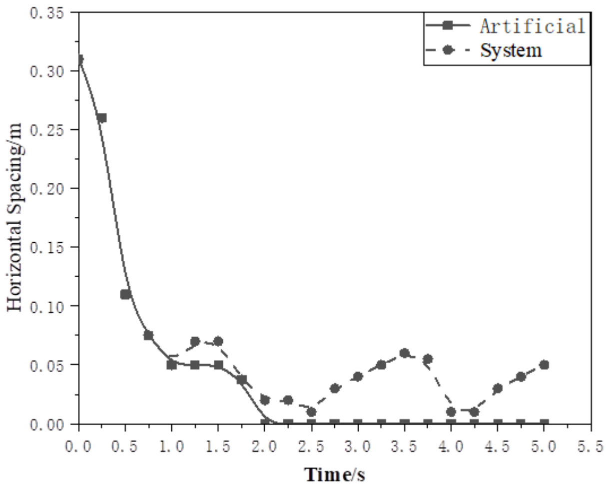

Figure 6: Horizontal spacing feature extraction results.

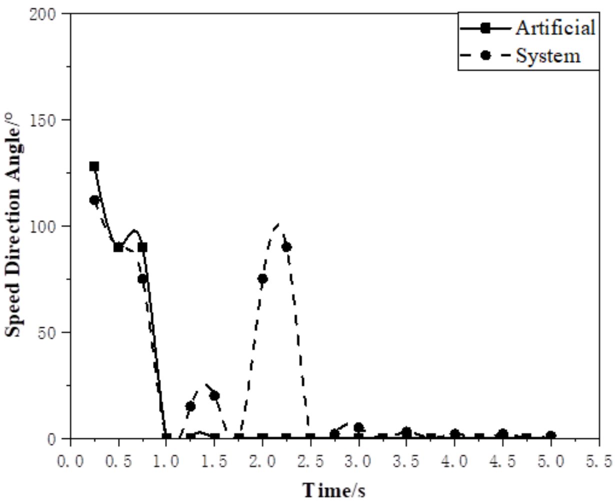

Figure 7: Feature extraction results of included angle in velocity direction.

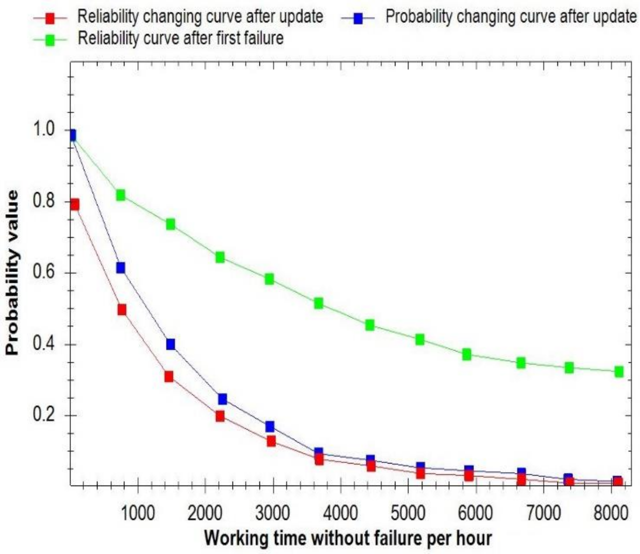

Figure 8: Results through simulation.

Figure 8 depicts the simulation outcomes of the proposed system. The reliability graph before the primary failure is equivalent to the typical probability graph, however it varies from one another after the updating of failure and maintenance. The pattern of the typical probability graph after the framework update is more awful than before the primary failure. This is on the grounds that a few old units stay in their ordinary states in the refreshed framework, making it simpler for the framework to go into a failure state. The measured reliability is additionally lower. After the framework is refreshed, the dependability is more awful than the first framework, albeit the framework can work typically. Over the long haul, the probability of typical activity moves toward the dependability of the framework. It ought to be noticed that unwavering quality is the premise of security, while safety mirrors the continuous condition of dependability.

5 Conclusions

By designing a rail transit on-board system platform based on data mining, this paper analyzes the system structure and functions required in the process of rail transit on-board monitoring and maintenance, and briefly introduces the data mining algorithm used in the log-assisted analysis process. Finally, the reliability and maintainability of the system are verified by simulation experiments. The application of this system not only realizes the centralized recording of the operation status and alarm of the vehicle signal system, but also provides a powerful tool for the debugging personnel to monitor the operation of the equipment and analyze the program loopholes, and greatly improve the efficiency of the maintenance personnel to analyze the logs. And make fault early warning a possibility to prevent problems before they happen.

References

[1] Huang, C., & Huang, Y. (2021). Urban rail transit signal and control based on Internet of Things. *Journal of High Speed Networks*, (Preprint), 1-14. *10.3233/JHS-210664*

[2] Jing, G., Ding, D., & Liu, X. (2019). High-speed railway ballast flight mechanism analysis and risk management–A literature review. *Construction and Building Materials*, *223*, 629-642. https://doi.org/10.1016/j.conbuildmat.2019.06.194

[3] Ma, J., Su, C., Yang, Y., Wu, M., & Jiang, B. (2016). The field test for influence of ram-compacted piles with bearing base on settlement of embankments in China Beijing-Shanghai high-speed railway on deep soft soil. *Japanese Geotechnical Society Special Publication*, *2*(3), 217-220. https://doi.org/10.3208/jgssp.CHN-49

[4] Liu, Z., Wang, L., Li, C., & Han, Z. (2017). A high-precision loose strands diagnosis approach for isoelectric line in high-speed railway. *IEEE Transactions on Industrial Informatics*, *14*(3), 1067-1077. 10.1109/TII.2017.2774242

[5] Zhong, J., Liu, Z., Han, Z., Han, Y., & Zhang, W. (2018). A CNN-based defect inspection method for catenary split pins in high-speed railway. *IEEE Transactions on Instrumentation and Measurement*, *68*(8), 2849-2860. 10.1109/TIM.2018.2871353

[6] Zhou, L., Tong, L. C., Chen, J., Tang, J., & Zhou, X. (2017). Joint optimization of high-speed train timetables and speed profiles: A unified modeling approach using space-time-speed grid networks. *Transportation Research Part B: Methodological*, *97*, 157-181. https://doi.org/10.1016/j.trb.2017.01.002

[7] Wu, Y., Qin, Y., Qian, Y., & Guo, F. (2021). Automatic detection of arbitrarily oriented fastener defect in high-speed railway. *Automation in Construction*, *131*, 103913. https://doi.org/10.1016/j.autcon.2021.103913

[8] Zhong, Z. D., Ai, B., Zhu, G., Wu, H., Xiong, L., Wang, F. G., & He, R. S. (2018). *Dedicated mobile communications for high-speed railway* (Vol. 22). Heidelberg: Springer. https://doi.org/10.1007/978-3-662-54860-8

[9] Vega, J., Fraile, A., Alarcon, E., & Hermanns, L. (2012). Dynamic response of underpasses for high-speed train lines. *Journal of Sound and Vibration*, *331*(23), 5125-5140. https://doi.org/10.1016/j.jsv.2012.07.005

[10] Horton, M., Connolly, D. P., & Yu, Z. (2017). Rail trackbed and performance testing of stabilised sub-ballast in normal and high-speed environments. *Procedia engineering*, *189*, 924-931. https://doi.org/10.1016/j.proeng.2017.05.143

[11] Zhan, S., Wang, P., Wong, S. C., & Lo, S. M. (2022). Energy-efficient high-speed train rescheduling during a major disruption. *Transportation Research Part E: Logistics and Transportation Review*, *157*, 102492. https://doi.org/10.1016/j.tre.2021.102492

[12] Zhang, H., Peng, Y., Hou, L., Wang, D., Tian, G., & Li, Z. (2019). Multistage impact energy distribution for whole vehicles in high-speed train collisions: modeling and solution methodology. *IEEE Transactions on Industrial Informatics*, *16*(4), 2486-2499. 10.1109/TII.2019.2936048

[13] Xu, P., Corman, F., Peng, Q., & Luan, X. (2017). A timetable rescheduling approach and transition phases for high-speed railway traffic during disruptions. *Transportation Research Record*, *2607*(1), 82-92. https://doi.org/10.3141/2607-11

[14] Yang, S., Song, K., & Zhu, G. (2019). Stochastic process and simulation of traction load for high speed railways. *IEEE Access*, *7*, 76049-76060. 10.1109/ACCESS.2019.2921093

[15] Mao, Q., Cui, H., Hu, Q., & Ren, X. (2018). A rigorous fastener inspection approach for high-speed railway from structured light sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*, *143*, 249-267. https://doi.org/10.1016/j.isprsjprs.2017.11.007

[16] Zhao, J., Ma, C., Xiao, X., & Jiang, Y. (2022). Research on deformation law of guide rails caused by mine vertical shafts under non-mining action. *Engineering Failure Analysis*, *134*, 106089. *https://doi.org/10.1016/j.engfailanal.2022.106089*

[17] Wang, K., Yan, X., Yuan, Y., Jiang, X., Lodewijks, G., & Negenborn, R. R. (2017, August). Study on route division for ship energy efficiency optimization based on big environment data. In *2017 4th International Conference on Transportation Information and Safety (ICTIS)* (pp. 111-116). IEEE. *10.1109/ICTIS.2017.8047752*

[18] Zhigang, M. (2014, October). Grey Prediction of Urban Rail Transit Machine-Electric Equipment Fault Based on Data Mining. In *2014 7th International Conference on Intelligent Computation Technology and Automation* (pp. 284-287). IEEE. *10.1109/ICICTA.2014.76*

[19] Chen, X., Guo, Y., Li, B., Ge, M., & Xu, C. (2015). Analysis of Dynamic Passenger Flow in Urban Rail Transit Based on Data Mining. In *ICTE 2015* (pp. 2035-2042). *https://doi.org/10.1061/9780784479384.259*

[20] Ding, X., Yang, X., Hu, H., & Liu, Z. (2017). The safety management of urban rail transit based on operation fault log. *Safety science*, *94*, 10-16. *https://doi.org/10.1016/j.ssci.2016.12.015*

[21] Ming, Z. (2015, September). Decision approach of maintenance for urban rail transit based on equipment supervision data mining. In *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)* (Vol. 1, pp. 376-380). IEEE. *10.1109/IDAACS.2015.7340761*

[22] Chen, G. (2018, January). Optimization Design of Passenger Flow in Rail Transit Station in Shanghai Based on Data Mining. In *2017 4th International Conference on Machinery, Materials and Computer (MACMC 2017)* (pp. 583-588). Atlantis Press. *https://doi.org/10.2991/macmc-17.2018.108*

[23] Bai, L., Wang, F. Z., & Zhang, M. (2015). Application of geographic information system (gis) in urban rail transit construction safety and operation monitoring. In *Applied Mechanics and Materials* (Vol. 743, pp. 692-697). Trans Tech Publications Ltd. *https://doi.org/10.4028/www.scientific.net/AMM.743.692*

[24] Liu, X., & Yang, X. (2019, March). Identifying technological innovation capability of high-speed rail industry based on patent analysis. In *2019 8th International Conference on Industrial Technology and Management (ICITM)* (pp. 127-131). IEEE. *10.1109/ICITM.2019.8710710*

[25] Zhang, S. (2012). Data Mining and Analysis of Integrated Circuit Card Data in Subway. In *Future Wireless Networks and Information Systems* (pp. 739-744). Springer, Berlin, Heidelberg. *https://doi.org/10.1007/978-3-642-27326-1_95*

[26] Fu, L., Wang, X., Zhao, H., & Li, M. (2022). Interactions among safety risks in metro deep foundation pit projects: An association rule mining-based modeling framework. *Reliability Engineering & System Safety*, *221*, 108381. *https://doi.org/10.1016/j.ress.2022.108381*

[27] Zimbalist, A. (2013). Inequality in intercollegiate athletics: Origins, trends and policies. *Journal of Intercollegiate Sport*, *6*(1), 5-24. *https://doi.org/10.1123/jis.6.1.5*

[28] Wang, H., Hao, L., Sharma, A., & Kukkar, A. (2022). Automatic control of computer application data processing system based on artificial intelligence. *Journal of Intelligent Systems*, *31*(1), 177-192. *https://doi.org/10.1515/jisys-2022-0007*

[29] Sun, L., Gupta, R. K., & Sharma, A. (2022). Review and potential for artificial intelligence in healthcare. *International Journal of System Assurance Engineering and Management*, *13*(1), 54-62. *https://doi.org/10.1007/s13198-021-01221-9*

[30] Cai, Y., & Sharma, A. (2021). Swarm intelligence optimization: an exploration and application of machine learning technology. *Journal of Intelligent Systems*, *30*(1), 460-469. *https://doi.org/10.1515/jisys-2020-0084*

[31] Kumbinarasaiah, S., & Raghunatha, K. R. (2021). A novel approach on micropolar fluid flow in a porous channel with high mass transfer via wavelet frames. *Nonlinear Engineering*, *10*(1), 39-45. *https://doi.org/10.1515/nleng-2021-0004*

[32] Dhotre, P. K., & Srinivasa, C. V. (2021). On free vibration of laminated skew sandwich plates: A finite element analysis. *Nonlinear Engineering*, *10*(1), 66-76. *https://doi.org/10.1515/nleng-2021-0006*

[33] Pedram, L., & Rostamy, D. (2021). Numerical simulations of stochastic conformable space–time fractional Kortewegde Vries and Benjamin–Bona–Mahony equations. *Nonlinear Engineering*, *10*(1), 77-90. *https://doi.org/10.1515/nleng-2021-0007*

[34] Duan, Y. Q., Fan, X. Y., Liu, J. C., & Hou, Q. H. (2020). Operating efficiency-based data mining on intensive land use in smart city. *IEEE Access*, *8*, 17253-17262. *10.1109/ACCESS.2020.2967437*

[35] Wang, D. L., Yao, E. J., Yang, Y., & Zhang, Y. S. (2014). Modeling passenger flow distribution based on disaggregate model for urban rail transit. In *Foundations and Practical Applications of Cognitive Systems and Information Processing* (pp. 715-723). Springer, Berlin, Heidelberg. *https://doi.org/10.1007/978-3-642-37835-5_62*

[36] Wang, F., Xu, T., Tang, T., Zhou, M., & Wang, H. (2016). Bilevel feature extraction-based text mining for fault diagnosis of railway systems. *IEEE transactions on intelligent transportation systems*, *18*(1), 49-58. *10.1109/TITS.2016.2521866*

[37] Massa, S., & Puliafito, P. P. (1999, September). An application of data mining to the problem of the university students' dropout using markov chains. In *European conference on principles of data mining and knowledge discovery* (pp. 51-60). Springer, Berlin, Heidelberg. *https://doi.org/10.1007/978-3-540-48247-5_6*

[38] Dai, X., Qiu, H., & Sun, L. (2021). A Data-Efficient Approach for Evacuation Demand Generation and Dissipation Prediction in Urban Rail Transit System. *Sustainability*, *13*(17), 9692. *https://doi.org/10.3390/su13179692*

[39] Tang, Y. (2021). Risk Chain and Key Hazard Management for Urban Rail Transit System

Operation Based on Big Data Mining. *Discrete Dynamics in Nature and Society*, *2021*. *https://doi.org/10.1155/2021/3692151*

[40] Hu, P., Duan, K., Huo, W., Zhang, Q., Zhou, M., & Ba, Y. (2020, November). Design of state Grid shopping mall heating technology application assistant decision system based on big data analysis. In *2020 International Conference on Computer Engineering and Intelligent Control (ICCEIC)* (pp. 97-100). IEEE. *10.1109/ICCEIC51584.2020.00027*

[41] Ming, Z., Xiaofei, W., & Li, B. (2014, June). The Fault Data Mining of Supervision Equipment of Urban Rail Transit Based on Clustering. In *2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications* (pp. 1045-1048). IEEE. *10.1109/ISDEA.2014.230*

[42] Gusikhin, O., Rychtyckyj, N., & Filev, D. (2007). Intelligent systems in the automotive industry: applications and trends. *Knowledge and Information Systems*, *12*(2), 147-168. *https://doi.org/10.1007/s10115-006-0063-1*

[43] Cheng, X., Huang, K., Qu, L., & Li, L. (2020). A cooperative data mining approach for potential urban rail transit demand using probe vehicle trajectories. *IEEE Access*, *8*, 24847-24861. *10.1109/ACCESS.2020.2970863*

[44] Ma, X. L., Wang, Y. H., Chen, F., & Liu, J. F. (2012). Transit smart card data mining for passenger origin information extraction. *Journal of Zhejiang University Science C*, *13*(10), 750-760. *https://doi.org/10.1631/jzus.C12a0049*

[45] Shuai, L., Limin, J., Yong, Q., Bo, Y., & Yanhui, W. (2014). Research on urban rail train passenger door system fault diagnosis using PCA and rough set. *The Open Mechanical Engineering Journal*, *8*(1). *10.2174/1874155X01408010340*

[46] Du, G., Zhang, X., & Ni, S. (2018, October). Discussion on the application of big data in rail transit organization. In *International Conference on Smart Vehicular Technology, Transportation, Communication and Applications* (pp. 312-318). Springer, Cham. *https://doi.org/10.1007/978-3-030-04582-1_36*