사물인터넷융복합논문지 Vol. 8, No. 6, pp. 55-63, 2022 https://doi.org/10.20465/KIOTS.2022.8.6.055

Research and Application of Fault Prediction Method for High-speed EMU Based on PHM Technology

Haitao Wang[1], Byung-Won Min[2*]
[1]Ph.D. Student, Division of Information and Communication Convergence Engineering, Mokwon University
[2]Professor, Division of Information and Communication Convergence Engineering, Mokwon University

PHM 기술을 이용한 고속 EMU의 고장 예측 방법 연구 및 적용

왕해도[1], 민병원[2*]
[1]목원대학교 정보통신융합공학부 박사과정, [2]목원대학교 정보통신융합공학부 교수

Abstract In recent years, with the rapid development of large and medium-sized urban rail transit in China, the total operating mileage of high-speed railway and the total number of EMUs(Electric Multiple Units) are rising. The system complexity of high-speed EMU is constantly increasing, which puts forward higher requirements for the safety of equipment and the efficiency of maintenance.At present, the maintenance mode of high-speed EMU in China still adopts the post maintenance method based on planned maintenance and fault maintenance, which leads to insufficient or excessive maintenance, reduces the efficiency of equipment fault handling, and increases the maintenance cost. Based on the intelligent operation and maintenance technology of PHM(prognostics and health management). This thesis builds an integrated PHM platform of "vehicle system-communication system-ground system" by integrating multi-source heterogeneous data of different scenarios of high-speed EMU, and combines the equipment fault mechanism with artificial intelligence algorithms to build a fault prediction model for traction motors of high-speed EMU.Reliable fault prediction and accurate maintenance shall be carried out in advance to ensure safe and efficient operation of high-speed EMU.

Key Words : High speed EMU; Prognostics; Model; Neural Network; Intelligence

요 약 최근 중국에서 중대형 도시철도의 급속한 발전으로 고속철도의 총 운행거리와 총 EMU(Electric Multiple Units) 수가 증가하고 있다. 고속 EMU의 시스템 복잡성은 지속적으로 증가하고 있으며, 이는 장비의 안전성과 유지보수의 효율성에 대한 더 높은 요구사항을 제시한다. 현재 중국의 고속 EMU의 유지보수 모드는 여전히 계획적인 유지보수 및 고장보수에 기반한 사후 유지보수 방식을 채택하고 있어 유지보수가 미흡하거나 과도하게 이루어지며, 장비 고장 처리의 효율성을 떨어뜨리고 유지보수 비용을 증가시킨다. PHM(진단 및 예측관리)의 지능형 운영 및 유지관리 기술을 기반으로 합니다. 본 논문은 고속 EMU의 서로 다른 시나리오의 다중 소스 이기종 데이터를 통합하여 "차량 시스템-통신 시스템-지상 시스템"의 통합 PHM 플랫폼을 구축하고, 장비 고장 메커니즘을 인공지능 알고리즘과 결합하여 고속 EMU의 트랙션 모터에 대한 고장 예측 모델을 구축한다. 고속 EMU의 안전하고 효율적인 작동을 보장하기 위해 고장 예측 및 정확한 유지보수를 사전에 수행해야 한다.

주제어 : 고속 EMU, 고장 예측, 모델, 신경망, 지능형

*교신저자 : 민병원(minfam@mokwon.ac.kr)
접수일 2022년 9월 12일 수정일 2022년 11월 2일 심사완료일 2022년 11월 8일

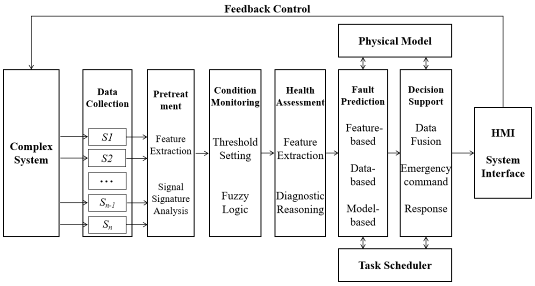
1. Definition of PHM

1.1 Basic Concept and Biackground of PHM

PHM includes two aspects, namely fault prediction and health management. Fault prediction refers to the method of diagnosing and predicting the current and future equipment health status, performance degradation and possible faults according to the historical status and current monitoring data of the system, including determining the remaining life or the length of time for normal operation of components or systems. Health management is the ability to make appropriate planning, decision-making, planning and coordination for tasks, maintenance and support activities according to the information of diagnosis, assessment and prediction, available maintenance resources, equipment use requirements and other knowledge. PHM represents a change of method, a change of maintenance strategy and concept, which realizes the change from traditional sensor based diagnosis to intelligent system based prediction, thus providing a technical basis for accurate and active maintenance activities at the right time and at the right place.

1.2 PHM Technology Architecture

The typical architecture of fault prediction and health management is OSA-CBM(open system architecture for condition-based maintenance) system architecture, which is an important reference in this field. It is a single dimensional seven module functional architecture oriented to general objects. OSA-CBM architecture divides the functions of PHM into seven levels, mainly including data acquisition, feature extraction, condition monitoring, health assessment, fault prediction, maintenance decision-making and human-computer interaction interface, as shown in Fig. 1.



[Fig. 1] PHM technology architecture

The PHM system is composed of seven functional modules. The data flow between each functional module basically follows the above sequence. Any one of the functional modules has the ability to obtain the required data from the other six modules.

2. Problems Encountered in Fault Prediction of High-speed EMU at Present

At present, China has carried out extensive research on fault diagnosis, prediction and health management in the fields of transportation, aerospace, shipbuilding, energy and power. Most of the research subjects are universities and scientific research institutions. The main research content focuses on PHM architecture and intelligent diagnosis and prediction algorithms. Although certain achievements have been made, the overall application research scale and level are still relatively backward, and the basic theoretical research lacks application background support and experimental verification. In particular, the following difficulties still exist in the field of fault prediction and health management of high-speed EMU in China.

2.1 single Detection Method for System and Equipment Status

At present, the state perception of high-speed EMU in China mainly monitors the vehicle state in real time through data acquisition devices such as temperature sensors, vibration sensors, speed sensors, pressure sensors and voltage and current transformers, and measures the state between systems and components with analog quantities and switching quantities. However, there is a lack of detection means within the system or components, which can not accurately locate faults, It is impossible to accurately describe the internal performance degradation trend of the system and components.

2.2 Failure Diagnosis Lacks Advance Prediction Mechanism

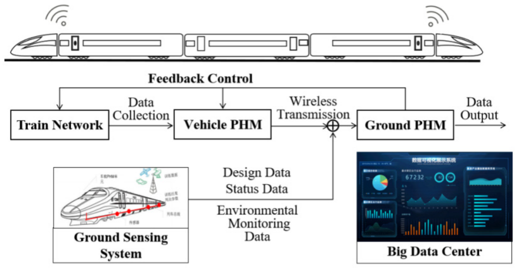
The fault diagnosis of key systems and components of high-speed EMU adopts the post fault processing method, which can not be predicted in advance before the fault occurs, which will cause damage to system components, train delay and other problems affecting the operation order and safety. At the same time, the fault handling after the fault also requires a lot of manpower and material resources to prevent the expansion of the fault range and the deterioration of the fault degree, which will not only increase the maintenance cost, but also significantly affect the safe operation and service quality of EMU.

2.3 Complexity of Multi-source and Heterogeneous Data

The structured, semi-structured and unstructured data generated in the R & D, design, production and manufacturing stages, operation stages and operation, maintenance and repair stages of multiple units are owned by different companies. For example, the production and manufacturing data of the train is in the manufacturing company, and the operation data of the train is in the national operation Department. The detection data of the user's daily maintenance of the subordinate maintenance department of the train is stored in the user's specific system. PHM technology of high-speed EMU is to study the overall state of the whole life cycle of the train. Data in different formats, data from different sources and interface interconnection need to be urgently solved.

3. PHM Platform Design of High-speed EMU

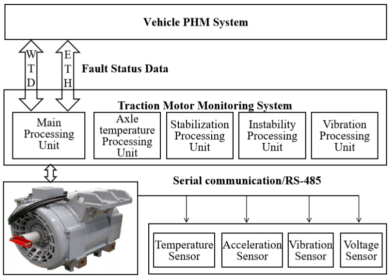
The fault abnormal state of high-speed EMU often involves multiple levels such as train cluster, system cluster and component cluster. The state characteristics of each level are interrelated, which makes fault prediction and location extremely complex. Therefore, to carry out the research on intelligent diagnosis and fault prediction of high-speed EMU, it is necessary to extract and preprocess the real-time status data, dig deeply into the historical data accumulated by train operation, establish the mathematical model of fault prediction of the system, and monitor and logically deduce the feature data and correlation at the train, system and component levels. The fault prediction and health management platform for high-speed EMU is mainly composed of four elements: On-board PHM system, vehicle ground data transmission and communication system, ground information perception system and ground PHM system. The system architecture is shown in Fig. 2.



[Fig. 2] PHM system architecture

The on-board PHM system is used to collect vehicle data in real time and identify the abnormal state of the vehicle in advance through the on-board fault warning model. The system preprocesses the real-time vehicle data, places the early warning model with high real-time requirements and small calculation amount on the vehicle side, and transmits the calculation results to the ground server, so as to improve the real-time calculation efficiency of vehicle data. The on-board PHM system is composed of component level PHM and train level PHM. By adding a component level PHM unit in the control host of traction, braking, axle temperature and other subsystems, the collected internal data can be used for fault warning, prediction and self diagnosis of system status, so as to complete the component level PHM analysis; And use the train network to send the original data, status information and fault information of the components to the on-board PHM host, so as to complete the self diagnosis of the train. The train intelligent display screen displays the diagnosis information and transmits the data to the ground PHM system through the on-board wireless transmission device WTD(wireless transmission device).

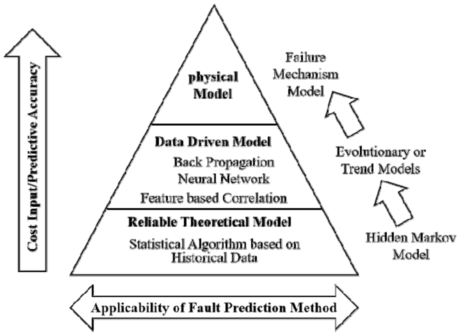
Taking the status monitoring of train traction motor as an example, the structure of on-board PHM system is shown in Fig. 3. The on-board PHM system monitors the operation data of the traction motor in the bogie through the train's own network system, performs real-time calculation and Analysis on the component level characteristic data such as axle temperature, vibration frequency, voltage, current, acceleration and vehicle level data of the traction motor, and transmits the processed warning information and status characteristics to the main processing unit of the train network, and then transmits them to the WTD of the train. Finally, it is sent to the big data platform receiving server of the ground PHM system through the vehicle ground data transmission communication system.



[Fig. 3] PHM model of traction motor

4. Research on Fault Prediction Method of High-speed EMU

The primary task of fault prediction technology is to track and monitor the fault evolution law and degradation state of equipment, use intelligent algorithms to process the information in the whole life cycle of equipment, and achieve the goal of accurate prediction according to the fault evolution law and failure mechanism. The intelligent diagnosis and fault prediction model of the whole vehicle, subsystem or component of high-speed EMU is a prediction model established by abstracting the mathematical method according to its own inherent attribute parameters, logic and functions. As for the current mainstream fault prediction methods, as shown in Fig. 4. This thesis studies the accuracy of prediction, capital cost investment and application universality.



[Fig. 4] Algorithms of fault prediction

4.1 Model based fault prediction technology

In this thesis, the advantages and disadvantages of hidden Markov model / hidden semi Markov model are mainly studied in the model-based fault prediction method. Hidden Markov model has a rigorous mathematical structure. Through statistical learning, the relationship between monitoring data and fault status can be mined, and the fault prediction and health status evaluation of equipment can be carried out. HMM has great advantages in equipment health state recognition, but hidden Markov model has certain limitations when it predicts the state of complex equipment because the state time probability is exponentially distributed. Hidden semi Markov model (HSMM) is an extension of HMM, which increases the dwell time of each state and overcomes the limitation caused by the exponential distribution of the probability of dwell time. HSMM model has stronger modeling ability and higher recognition accuracy for health status of complex equipment.

4.2 Data Driven Fault Prediction Technology

For complex equipment, it is difficult to establish physical model due to its complex structure, many constituent systems and various failure forms. However, it is easy to collect information in the whole life cycle through monitoring system and sensors. Therefore, data-driven method is usually used to predict the fault of complex equipment. By collecting and processing the information in the whole life cycle of the equipment, the method uses machine learning to build the mapping relationship between data input and output to achieve the fault prediction of the equipment. Typical data-driven fault prediction methods include artificial neural network(ANN), time series, decision tree(DT), support vector machine(SVM) and other intelligent prediction methods.

Compared with the regression analysis and time series analysis methods in the traditional statistical category, neural network is one of the most widely used methods in fault prediction methods and application research. It can learn from samples and try to capture the intrinsic functional relationship between sample data. In particular, BP neural network has strong nonlinear mapping ability, which can better perform nonlinear classification. Moreover, BP neural network has high speed in processing information, high fault tolerance rate, and can associate and remember external information. It is a very effective fault prediction algorithm. This is also an important reason why this thesis chooses BP neural network as the fault prediction of traction motor of high-speed EMU.

In the process of fault prediction, we use sigmoid activation function to complete the nonlinear transformation of data and solve the problem of insufficient expression and classification ability of linear models. The formula is as follows.

$$S(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

In order to optimize the parameters of the neural network model, we use the loss function. The function of the loss function is to calculate the difference between the forward calculation result of each iteration of the neural network and the real value, so as to guide the next training in the right direction. The smaller the value of the loss function, the closer the predicted value of the model is to the true value. The loss function is obtained in the forward propagation calculation and is also the starting point of the backward propagation. The formula is as follows.

$$E = \frac{1}{2N}(T-Y)^2 = \frac{1}{2N}\sum_{i=1}^{N}(t_i - y_i)^2 \qquad (2)$$

The matrix can be represented by capital letters, where T represents the real label, Y represents the network output, i represents the i-th data, and N represents the number of training samples.

4.3 Prediction Method Based on Traditional Reliability Theory

The train operation environment and working conditions are extremely complex, and the parameters collected by some systems and components are limited, so data-driven and mechanism analysis can not meet the requirements of modeling. Therefore, it is necessary to analyze and mine the evolution laws of these faults using mathematical statistics, and predict the reliability of equipment through the density function, probability and other indicators of fault occurrence. Typical fault prediction methods based on statistical reliability include dempster-shafer(DS) evidence theory, fuzzy logic(FL), fault tree analysis (FTA) and Bayesian network.

5. Fault Warning and Temperature Prediction Model of Traction Motor

The traction motor of high-speed EMU is the core component of the power system of EMU. It will generate heat loss during operation. If the temperature of the motor rises sharply and exceeds the maximum temperature it can bear due to changes in internal resistance and poor heat dissipation, it will cause equipment failure and lead to train delay or stop. This thesis takes the traction motor of CR400AF high-speed EMU as an example to illustrate the idea of building the fault warning and temperature prediction model of key components of the EMU.

5.1 Establish Real-time Temperature Warning Sub Model

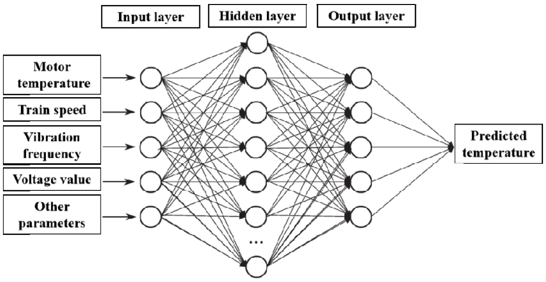
First of all, starting from the working principle of traction motor, the core judging index of motor state - motor stator temperature and its influencing factors are analyzed, and the range of characteristic variables is defined. The judgment index shall cover not only the number of characteristic variables directly related to temperature, such as motor current, train speed, traction level, weather conditions (temperature, rain and snow), but also indirect influence variables, such as the ramp and curve radius of the line section, which will affect the traction power of the motor and thus the motor temperature.

The second step is to extract the operation data of each EMU for two days in a natural month by sampling, and to cover different lines as much as possible. Arrange the sampling data in time order and hide the train number and car information. 1000 groups of data were randomly selected, and the correlation between each variable and motor stator temperature was verified by Pearson analysis through SPSS software. Select data indicators with correlation greater than 0.2 and significance (two sides) less than 0.01 to build the model. See Table 1 for eligible data indicators.

〈Table 1〉 Description of data

Motor stator temperature	Speed	Rotor Frequency	DC Voltage	Tap Position	Constant Speed state	Ambient Temperature	Route Slope
Pearson Correlation Coefficient	0.584	0.584	0.494	0.362	0.362	0.275	0.238
Significance (two-sided test)	0	0	0	0	0	0	0

The third step is to screen out the historical fault information of 7 traction motors of CR400AF EMU, all of which are common grounding faults caused by poor insulation of motor winding. According to the design principle and data analysis, the motor temperature changes greatly before and after such faults occur. Select the motor temperature within 0.5h before the fault and the related data in Table 1 as the fault samples, and select the normal historical data of 50 multiple units for a total of 100h as the normal samples to form a sample library with the fault samples. The current data collection cycle is 10s, so the sample library contains 37260 groups of data. Every 6 groups of data are recorded as a sample, and 6210 samples are generated, including 6000 normal samples and 210 fault samples. The sample data is randomly scrambled and divided into training set and test set according to the ratio of 7:3. There are 4347 samples in the training set, including 4200 normal samples and 147 fault samples; The test set includes 1863 samples, including 1800 normal samples and 63 fault samples. A three-layer traction motor temperature early warning neural network model is established by using BP neural network learning algorithm. The input neurons include the data index variables in Table 1, and the output neurons of the model are traction motor temperature early warning information. The schematic diagram of the model is shown in Fig. 5.



[Fig. 5] Schematic diagram of traction motor temperature BP neural network model

Finally, the model is trained with the training set sample data at a learning rate of 0.1, and the target relative error is 1×10-2, and the maximum number of iterations is 5000. After the training, the generalization ability is detected by using the test set data. When the generalization ability is no longer improved, the training is terminated to complete the construction of the real-time temperature warning model of the traction motor based on the neural network algorithm. The prediction results of the model test set are shown in Table 2. The accuracy rate of the model is 99.14%, and the recall rate is 93.65%.

〈Table 2〉 Description of data

Predictive Target	Forecast Results
TP(True Positive)	59
FN(False Negative)	4
FP(False Positive)	12
TN(True Negative)	1799
Accuracy=(TP+TN)/(TP+FN+FP+TN)	99.14%
Precision=TP/(TP+FP)	83.10%
Recall=TP/(TP+FN)	93.65%

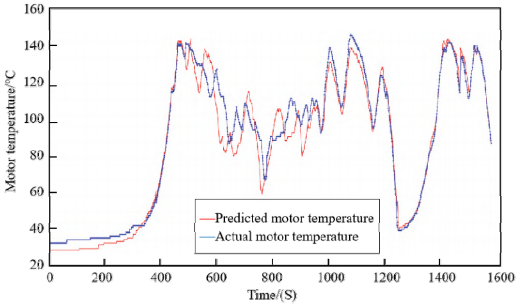
5.2 Establish Temperature Trend Prediction Sub Model

The departure and arrival time of the same train and the maximum running speed of the line are completely the same, so the factors affecting the temperature of the traction motor such as the operation of the multiple units in the same section and the line slope are basically the same. Only the ambient temperature belongs to the unstable variable, which can be obtained from the weather forecast data along the line. Based on this, we can realize the temperature prediction of traction motor.

First of all, from the normal operation data of CR400AF multiple unit in the latest year, the 8 data index variables in Table 1 are sampled according to different seasons and lines, and the stator temperature of the traction motor is regressed to build a motor temperature prediction model.

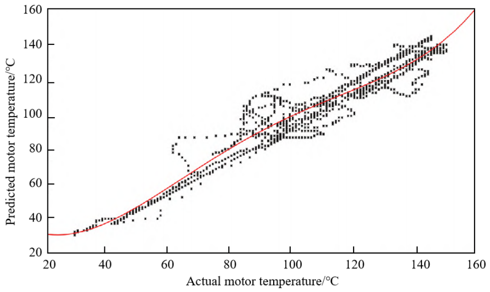
Then, the latest normal operation state data of the EMU in the subsequent sections, including speed, rotor frequency, DC voltage, gear, constant speed state, line slope and electric current, are obtained through the correlation between the operation plan of the day, real-time GPS positioning information and historical data, and the weather forecast ambient temperature data along the line are obtained through the Internet. The acquisition cycle of all data association is 30min, and the temperature change trend of each traction motor in the following 30min under the normal working condition is predicted by the temperature prediction model.

In order to reflect the predicted motor temperature, the actual motor temperature of a certain EMU on a certain day is compared with the predicted motor temperature on that day, and the results are shown in Fig. 6. From the data analysis results, the prediction effect is very good when the driving conditions of the EMU are relatively simple (such as in the state of constant speed and constant gear). When the driving conditions of multiple units are complex (such as repeated speed regulation and frequent traction braking switching), the prediction deviation increases. The reason for this is that under complex working conditions, the gear, motor current, voltage and other indicators related to the prediction model change dramatically, and the WTD vehicle ground data transmission cycle is long (10s), so it is impossible to completely record its change process.



[Fig. 6] Comparison between actual traction motor temperature and model predicted temperature

To further evaluate the prediction effect, the predicted value of the model is fitted with the actual value for analysis, as shown in Fig. 7. It can be seen that the actual temperature is in good agreement with the predicted temperature trend, and the R2 value of the fitting coefficient is 0.97, which basically meets the expected requirements.



[Fig. 7] Fitting analysis between predicted values and actual values of traction motor stator temperature prediction sub model

So far, the stator temperature prediction model of the traction motor has been built. Its functions include two parts: one is to determine whether the traction motor works normally, so as to identify abnormalities and remind the emergency commander to intervene in time. The second is to obtain the temperature change trend information within 30 minutes after the motor is in normal working condition, and compare it with the real-time temperature after the motor fails, so that the emergency commander can intuitively understand the abnormal degree of the motor and provide data support for subsequent analysis and decision-making.

6. Conclusions

On the basis of fault prediction and health management technology, this thesis analyzes the current situation of fault prediction of high-speed EMU, and introduces the architecture of PHM platform of high-speed EMU. In the process of prediction method research, three commonly used prediction technologies are compared and analyzed, and a method based on the combination of component failure mechanism and BP neural network algorithm is used to explore and build a traction motor failure warning and temperature prediction model with the traction motor of high-speed EMU as an application case. It provides a theoretical basis and research ideas for the fault prediction technology of high-speed EMU, and has a good reference significance. The next step will be to study the application of big data technology and intelligent diagnosis technology in PHM and establish a new high-speed EMU Operation and maintenance system with multi-source data fusion, algorithm optimization and intelligent diagnosis.

REFERENCES

[1] J. Y. Liang, "Research on intelligent diagnosis and fault prediction technology of high-speed trains", Journal of Beijing Jiaotong University, Vol.43, NO.1, pp.63-70, 2019.

[2] B. Liu, J. Shao, H. Lu, Y. Li & M Xie, "Thoughts on the research of fault prediction and health management (PHM) system architecture of EMU", China Railway, pp.1-9, 2022.

[3] Z. Y. Wu, "Application of big data technology in EMU fault warning", Control and information technology, pp.91-96, 2021.

[4] W. T. Hu, "Research on fault prediction and health management of traction motor of multiple units", Lanzhou Jiaotong University, 2021.

[5] D. G. Song and Q. M. Niu, "Research and application of big data PHM system for high-speed EMU", Railway computer application, Vol.27, NO.10, pp.44-48, 2018.

[6] N. Zhe, "Research on fault diagnosis and residual life prediction methods of rolling bearings based on big data technology", Beijing University of chemical technology, 2018.

[7] Z. Y. Wu, "Application of big data technology in EMU fault warning", Control and information technology, pp.91-96, 2021.

[8] K. Zhao, "Design and Implementation of PHM Big Data Architecture for Distributed High speed EMUs", Software, Vol.39, NO.12, pp.90-94, 2018.

[9] C. Z. Chi, S. Y. Feng & J. K. Zhang, "Analysis on development trend of aviation equipment PHM and maintenance support technology", Annual Meeting of China Aviation Industry Technical Equipment Engineering Association 2019, pp.167-171, 2019.

[10] T. Yang, G. K. Yang,, "Vehicle failure rate prediction based on BP neural network", Computer simulation, Vol.26, NO.1, pp.276-271, 2009.

[11] Q. M. Niu, "Research on intelligent analysis model of health status of high-speed railway equipment based on PHM technology", Beijing Jiaotong University, 2019.

[12] Y. G. Chi, "Research and application of fault prediction and health management technology for multiple units", Locomotive electric drive, pp.118-124, 2018.

[13] X. Z. Li, "Fault diagnosis method of high-speed railway signal equipment based on deep learning integration", Journal of Railways, Vol.42, NO.12, pp.97-105, 2020.

[14] X. X. Xie and S. H. Dai. "Research on fault diagnosis of uninsulated track circuit based on deep learning", Journal of Railways, Vol.42, NO.6, pp.79-85, 2020.

[15] Y. Li, "Research on the Application of PHM Technology in the Operation and Maintenance of Beijing Zhangjiakou High speed Railway", Railway Transportation and Economy, Vol.44, NO.9, pp.103-111, 2022.

[16] C. X. Li, "Research on PHM fault model and management of EMU", China Academy of Railway Sciences, 2020.

Wang Haitao[정회원]

▪ He received his bachelor's degree in software engineering from East China Jiaotong University in 2013. He received M.S. degree in computer technology from Ocean University of China in 2016. He has been studying for Ph.D in Fusion of IT professional at Mokwon University, Daejeon, Korea, from 2019. He is currently the director of the Information Center of National Innovation Center of High Speed Train since 2018. His research interests include Information system construction, Big Data and Cloud Computing application. He is recently interested in Big Data Processing and Analysis, Prognostic and Health Management.

민 병 원(Byung-Won Min)[정회원]

▪ He received M.S. degree in computer software from Chungang University, Seoul, Korea in 2005. He worked as a professor in the dept. of computer engineering, Youngdong University, Chungbuk, Korea, from 2005 to 2008. He received Ph.D. degree in the dept. of Information and Communication Engineering, Mokwon University, Daejeon, Korea, in 2010. He is currently a professor of Mokwon University since 2010. His research interests include digital communication systems, information theory and their applications. He is recently interested in multimedia content and Big Data.