Anomaly Diagnosis Method for Electric Vehicle Charging Equipment based on Semantic and Probabilistic Reasoning

Qiuchen Yun[1, a], Haiqi Li[2], Tianlin Yang[1], Chengnan Zhao[2], Ou Bai[1]

[1] STATE GRID SMART INTERNET OF VEHICLES CO., LTD, Beijing, China

[2] State Grid Shandong Electric Power Company Heze Power Supply Company, Heze, China

[a]759704907@qq.com

Abstract

Against the backdrop of the large-scale advancement of new energy vehicle charging networks, anomalies such as charging failures and slow charging have become increasingly prevalent. Conventional diagnostic methodologies are constrained by a single information dimension, fragmented textual data, and inadequate intelligence. To address these challenges, this study proposes an anomaly diagnosis method for electric vehicle charging equipment based on semantic and probabilistic reasoning. The method integrates user review texts, equipment operation monitoring data, and wave recording data, excavates implicit fault clues via semantic models, extracts feature variables through the combination of wavelet packet transform and distance correlation analysis, and constructs a hierarchical Bayesian reasoning model encompassing "charging state-feature variable-fault cause" to achieve precise identification of abnormal states and root cause localization. Validation through practical operation and maintenance cases demonstrates that the proposed method effectively fuses multi-source heterogeneous data, achieving a diagnostic accuracy of over 91% for typical charging anomalies. It significantly enhances the reliability and practicality of charging anomaly diagnosis in complex scenarios, thereby providing robust technical support for the intelligent operation and maintenance of charging equipment.

Keywords

Electric vehicle charging equipment, anomaly diagnosis, semantic reasoning, Bayesian network, wavelet packet transform, multi-source data fusion.

1. Introduction

With the continuous growth of new energy vehicle ownership, the charging network, as a core infrastructure, has entered a stage of large-scale operation. As of October 2025, the number of public charging piles in China has exceeded 3.83 million, but 30% of the equipment has faults such as "charging interruption" and "power fluctuation", and more than 40% of car owners have experienced intermittent charging. High-frequency use and cross-generational technological iteration have led to increasingly prominent problems such as incompatible adaptation and hardware aging of charging equipment, manifested as typical anomalies including failure to start charging, abnormal charging shutdown, and slow charging[1]. These issues not only directly affect user charging efficiency but also exacerbate the "range anxiety" of new energy vehicle users, becoming a key bottleneck restricting the improvement of charging network experience.

Current operation and maintenance management for charging equipment anomalies still relies on traditional models, mainly diagnosing through preset fault code mapping tables. There are significant limitations: first, it can only identify explicit faults and is difficult to cover implicit anomalies reflected by user subjective feedback; second, there is a lack of effective association

between user natural language evaluations and equipment operation data, resulting in the failure to mine potential fault clues; third, it relies on fixed rule engines and lacks the ability of semantic understanding and probabilistic correlation of multi-source data, leading to low accuracy and slow response time in root cause localization in complex scenarios[2][5][10], which cannot meet the "efficient, accurate, and proactive" operation and maintenance needs of large-scale charging networks. After the implementation of the new national standard GB 46519-2025 in 2025, some old charging piles failed to meet energy efficiency standards, and the number of complaints about charging interruptions increased by 22% month-on-month, further highlighting the inadequate adaptation of traditional diagnostic methods. Existing approaches-such as sensor-level battery fault detection [4], joint diagnosis of state-of-charge and thermal anomalies [2,3], or data-driven thermal modeling for station-level anomaly detection [5]-remain largely confined to structured operational signals and lack integration with unstructured user feedback. Meanwhile, broader reviews confirm that current charging infrastructures suffer from fragmented monitoring and insufficient intelligent diagnostics across heterogeneous equipment generations [8][9].

This paper aims to propose an anomaly diagnosis method integrating semantic analysis and probabilistic reasoning, breaking through the information limitations and intelligent bottlenecks of traditional diagnostic technologies, and realizing comprehensive perception and accurate diagnosis of abnormal states of charging equipment. By integrating multi-source heterogeneous data and mining explicit and implicit fault information, the method improves the ability of root cause localization of anomalies in complex scenarios.The approach draws inspiration from differentiated maintenance frameworks in power systems [6] and long-distance O&M architectures based on spatial correlation [7], but extends them through semantic-probabilistic fusion to enable dynamic, user-aware anomaly interpretation. It provides scientific decision support for the operation and maintenance of charging equipment, and has important practical significance for reducing operation and maintenance costs and improving the reliability and user experience of the charging network.

2. Related Theoretical Foundations

2.1. Wavelet Packet Transform Principle

Wavelet packet transform can effectively extract feature information of different frequency bands by decomposing and reconstructing signals in multiple layers. After performing i-layer wavelet packet decomposition on charging voltage and current signals, the reconstructed signals of each node are obtained using low-pass and high-pass filter coefficients, and the feature is characterized by calculating the energy moment. The formulas are as follows:

$$S \approx \sum_{j=1}^{2^i} S_{ij}$$

$$E_{ij} = \int \left| S_{ij}(t) \right|^2 \times t dt = \sum_k \left| x_{ij}^k \right|^2 \times j$$

Where Sij is the reconstructed signal, xijk is the discrete point value of the reconstructed signal, and E is the energy moment feature value.

2.2. Distance Correlation Analysis

Distance correlation is used to quantify the correlation strength between factor indicators and fault rates. The correlation coefficient is obtained by calculating the ratio of distance covariance to distance variance, and the formula is:

$$\mathrm{dCor}(X,Y) = \frac{\mathrm{dCov}(X,Y)}{\sqrt{\mathrm{dVar}(X)\mathrm{dVar}(Y)}}$$

When the correlation coefficient is ≥ 0.8, the factor is determined as a major fault-influencing factor, providing a basis for feature variable selection.

2.3. Bayesian Reasoning Model

Bayesian reasoning calculates the posterior probability through prior probability and likelihood probability based on Bayes' theorem to realize the probabilistic inference of fault causes. The core formula is:

$$P(C_i \mid F) = \alpha P(C_i) P(F \mid C_i)$$

\[P(C_i|F) = \alpha \cdot P(C_i) \cdot P(F|C_i) \]

Where α is the regularization factor, $P(C_i)$ is the prior probability of the fault cause, $P(F \mid C_i)$ is the conditional probability of the feature variable, and $P(C_i \mid F)$ is the posterior probability of the fault cause.

3. Design of Anomaly Diagnosis Method

3.1. Multi-source Data Collection and Dataset Construction

Multi-dimensional data is obtained through standard interfaces to construct a dataset of charging equipment operation status, including:

(1) Charging order information: transaction serial number, equipment number, vehicle VIN code, charging time period, charging capacity, etc.;

(2) Operation monitoring data: charging voltage, current, power, ambient temperature and humidity, equipment operation years, manufacturer information, operation location, etc.;

(3) Wave recording data: millisecond-level voltage and current wave recording signals;

(4) Text data: user review texts submitted through APPs and customer service hotlines.

According to the actual charging phenomena, the charging states are divided into normal charging, abnormally terminated charging (including BRM message reception timeout, insulation fault, and electrical energy metering anomaly) and slow charging. Combined with fault causes and operation and maintenance guidelines, a typical fault handling knowledge base is constructed.

3.2. Semantic Model-Driven Sample Data Generation

Based on semantic models, entity recognition and relationship extraction are performed on customer service records and user review texts to extract key information such as equipment number, charging time period, and anomaly description. Combined with charging order data of the corresponding time period, fault samples including equipment number, charging start and end time, and voltage and current anomaly features are generated. By optimizing generation parameters and prompt words, the conversion of text information into structured sample data is realized, which is supplemented to the sample database of the probabilistic model, solving the problem of scarcity of implicit anomaly samples.

3.3. Feature Extraction and Selection

Wavelet packet transform processing: Perform 4-layer wavelet packet decomposition and reconstruction on wave recording data to extract fundamental wave and harmonic signals of various orders, and calculate energy moment feature values;

Distance correlation analysis: Quantify the correlation strength between harmonics of various orders and fault types, select features with a correlation coefficient ≥ 0.8, and finally determine key fault features such as current fundamental wave and 5th harmonic amplitude;

Feature set construction: Integrate 13 types of feature quantities including charging voltage, current, power, ambient temperature and humidity, seasonal factors, equipment operation years, and manufacturer type to form a complete feature set.

3.4. Feature Discretization and Sample Database Construction

Continuous features (such as voltage and current) are discretized into normal/abnormal binary variables based on business thresholds; features such as equipment operation years, seasons, and operation locations are discretized according to set rules (e.g., equipment operation years exceeding the threshold are determined as abnormal, and high-temperature-prone seasons are marked as specific categories); vehicle manufacturers and charging pile manufacturers are divided into mainstream manufacturers and other manufacturers. Based on historical fault data, charging order data, and processed feature quantities, a three-dimensional sample database of "feature variable-charging state-fault cause" is constructed, and the conditional probability of each fault cause corresponding to each feature is statistically determined.

3.5. Construction and Diagnosis Implementation of Bayesian Reasoning Model

Table 1. Typical Fault Classification and Causes

Number	Fault Classification	Fault Cause
1	BRM reception timeout	Abnormal timing of vehicle communication protocol
2		Abnormal timing of charging equipment communication protocol
3		Others
4	Insulation monitoring alarm	Vehicle failure to feedback detection signal in a timely manner
5		Vehicle insulation anomaly
6		Insulation anomaly of charging equipment interface
7		Insulation anomaly of charging equipment gun line
8		Other causes
9	Electrical energy metering anomaly	Fault of metering module
10		Communication fault
11	Slow charging	Mismatch of asset files
12		Dynamic power distribution
13		Lack/fault of charging module
14		Active power reduction of charging equipment
15		Other causes

Model structure design: Based on the hierarchical logic of "charging state-feature variable-fault cause", a directed acyclic graph (DAG) of the Bayesian network is constructed. The charging state classification set Y = {normal charging, BRM message reception timeout, insulation fault, electrical energy metering anomaly, slow charging}, and the fault cause classification covers 15 core faults (as shown in Table 1);

Probability parameter learning: Use sample database data to statistically determine the conditional probability of feature variables and the prior probability of fault causes;

Anomaly diagnosis reasoning: Input real-time feature variable evidence, calculate the posterior probability of each fault cause through the Bayesian formula, and determine the optimal diagnosis result based on the maximum a posteriori criterion;

Model dynamic update: Combine new detection data and annotation conclusions to automatically adjust prior probability parameters to ensure the continuous optimization of diagnostic accuracy. (Table .1)

3.6. Generation of Operation and Maintenance Strategies

According to the fault causes obtained from diagnosis, the typical fault handling knowledge base is associated to generate targeted operation and maintenance strategies, including standardized guidelines such as fault handling steps, required tools, and precautions, which are output to the operation and maintenance management system to provide precise support for on-site maintenance.

4. Verification and Analysis of Actual Cases

To verify the effectiveness of the proposed method, 50 charging stations (including 23 hub stations, 17 community stations, and 10 remote stations) in 3 cities under a provincial charging operator were selected as test objects, covering 210 DC fast charging piles and 185 AC slow charging piles with an operation period of 1-8 years. The test period was from March 2025 to September 2025, and a total of 326 valid abnormal events were collected. Typical cases are analyzed in detail as follows.

4.1. Case 1: BRM Message Reception Timeout Anomaly

4.1.1. Anomaly Phenomenon

A DC fast charging pile (No. CZ-202305) at a hub station (Level 1) triggered 12 charging interruptions during June 12-June 15, 2024. User review texts showed "interrupted 3 times after 10 minutes of charging" and "normal charging after changing piles". Charging order data showed that the voltage was 380V and the current was 52A when charging terminated, with no obvious overrun values.

4.1.2. Diagnosis Process

✓ Data input: Extract user review texts ("interrupted 3 times after 10 minutes of charging", "normal after changing piles"), equipment operation data (equipment No. CZ-202305, operation period of 4 years, manufacturer A), and wave recording data (current fluctuation amplitude reached 8A during charging);

✓ Semantic analysis: Extract core features "charging interruption + frequent occurrence + recovery after changing piles" through the semantic model, and map them to candidate samples of "BRM message reception timeout-suspected protocol incompatibility";

✓ Feature extraction: Perform wavelet packet transform on wave recording data to extract the 5th harmonic amplitude feature, calculate a distance correlation coefficient of 0.86, and determine it as a key feature;

✓ Probabilistic reasoning: After inputting feature variables, the Bayesian model calculates the posterior probabilities of each fault cause: abnormal timing of vehicle communication protocol (42%) > abnormal timing of charging equipment communication protocol (38%) > others (20%).

4.1.3. Verification Result

According to the diagnostic recommendations, operation and maintenance personnel first checked the communication protocol version of the charging pile and found that the equipment still used the old version of the protocol from 2019, which was not compatible with the ISO 15118 protocol standard of new models such as BYD Seal 06 and Zeekr 001. After updating the

communication protocol through firmware upgrade, the equipment did not trigger any charging interruptions for 30 consecutive days, with a fault resolution rate of 100%. Traditional diagnostic methods require troubleshooting hardware faults one by one, with an average time of 45 minutes, while this method only takes 12 minutes to complete localization, improving diagnostic efficiency by 73.3%.

4.2. Case 2: Insulation Monitoring Alarm Anomaly

4.2.1. Anomaly Phenomenon

A DC fast charging pile (No. CZ-202008) at a remote station (Level 3) triggered an insulation monitoring alarm on August 10, 2024, and charging failed to start. The equipment log showed "insulation resistance below 10MΩ", and the user reported "immediate power failure after plug insertion".

4.2.2. Diagnosis Process

✓ Data input: Equipment log (insulation resistance 8.2MΩ), user text ("immediate power failure after insertion"), operation data (equipment operation period of 5 years, no gun line replacement record, frequent rainfall recently);

✓ Semantic analysis: Extract features "charging start failure + immediate power failure + insulation alarm", and map them to candidate samples of "insulation fault-suspected gun line insulation anomaly";

✓ Feature extraction: Analyze voltage signals through wavelet packet transform to extract zero-sequence current features, with a distance correlation coefficient of 0.91;

✓ Probabilistic reasoning: Output the posterior probabilities of fault causes: insulation anomaly of charging equipment gun line (78%) > insulation anomaly of charging equipment interface (12%) > vehicle insulation anomaly (10%).

4.2.3. Verification Result

Operation and maintenance personnel detected that the gun line's outer skin was aged and damaged due to long-term outdoor use, leading to rainwater infiltration and decreased insulation performance. After replacing the gun line, the insulation resistance was restored to 25MΩ, and the equipment operated normally. Traditional diagnostic methods in this case are prone to misjudging vehicle insulation anomalies, resulting in secondary dispatching. This method directly locates equipment-side faults through multi-source data fusion reasoning, avoiding invalid operation and maintenance and reducing fault handling costs by 60%.

5. Conclusion

The anomaly diagnosis method for electric vehicle charging equipment based on semantic and probabilistic reasoning proposed in this study realizes comprehensive perception and accurate root cause localization of charging anomalies through multi-source data fusion, semantic sample generation, refined feature extraction, and Bayesian reasoning. Verification through actual cases shows that the method achieves a diagnostic accuracy of over 91% for typical charging anomalies, effectively solving the problems of traditional diagnostic technologies such as inability to identify implicit anomalies and low accuracy in complex scenarios, and providing a complete technical solution for the intelligent operation and maintenance of charging equipment.

In the future, the entity recognition accuracy and relationship extraction ability of the semantic model can be further optimized to improve the efficiency of text data conversion; the fault cause classification system can be expanded to include new types of faults such as communication anomalies of digital charging equipment; deep learning algorithms can be combined to optimize the parameter learning process of the Bayesian model, improve the diagnostic performance of

multi-factor coupled faults, and realize the upgrade from "passive diagnosis" to "active prediction", so as to better adapt to the large-scale development and technological iteration needs of the charging network.

Acknowledgement

This work was supported by the State Grid Science and Technology Project “Key Technologies Research and Application of Intelligent Operation & Maintenance and Auxiliary Decision-Making for Electric-Vehicle Charging Equipment” (Project No. 5400-202371588A-3-2-ZN).

References

[1] Xu, X., Gao, X., Li, G., et al. (2022). Analysis on charging safety fault characteristics of electric vehicle charging facilities based on data mining. Electrical & Energy Management Technology, (3), 15–22.

[2] Zhou, J., Gao, T., Wang, Q., et al. (2025). Fault diagnosis of charging piles based on GA-LSTM neural network. Modern Electronics Technique, 48(16), 97–104.

[3] Fu, Z., Zhang, J., Luo, W., et al. (2024). Joint fault diagnosis of battery state of charge and voltage/temperature anomalies. Science Technology and Engineering, 24(8), 3260–3268.

[4] Xiong, R., Yu, Q. Q., Shen, W. X., et al. (2019). A sensor fault diagnosis method for a lithium-ion battery pack in electric vehicles. IEEE Transactions on Power Electronics, 34(10), 9709–9718.

[5] Gómez, P. I., Moreno, A. B., Lin, J., et al. (2022). Data-driven thermal modelling for anomaly detection in electric vehicle charging stations. In *The 2022 IEEE Transportation Electrification Conference & Expo (ITEC)* (pp. 1005–1010). IEEE.

[6] Peng, H., Wen, Y., Wang, L., et al. (2019). Differentiated operation and maintenance technology for power equipment based on a two-layer knowledge architecture. High Voltage Apparatus, 55(7), 221–226. https://doi.org/10.13296/j.1001-1609.

[7] Guo, W., Yin, X., Wang, J., et al. (2018). Design of long-distance operation and maintenance system of charging pile based on Manhattan distance method. Power System Protection and Control, 46(13), 134–140.

[8] Khalid, M., Ahmad, F., Panigrahi, B. K., et al. (2022). A comprehensive review on advanced charging topologies and methodologies for electric vehicle battery. Journal of Energy Storage, 53, 105084.

[9] Lamonaca, S., & Ryan, L. (2022). The state of play in electric vehicle charging services: A review of infrastructure provision, players, and policies. Renewable and Sustainable Energy Reviews, 154, 111733.

[10] Gao, D. X., Wang, Y., Zheng, X. Y., et al. (2021). A fault warning method for electric vehicle charging process based on adaptive deep belief network. World Electric Vehicle Journal, 12(4), 265.