Analysis of Fault Classification Method for SDH Optical Fiber Communication Networks Based on SVR

Zhong YunYu, Hu ChunHua
The 34th Research Institute of CETC, Guilin, Guangxi, China

Abstract: This study aims to address the problems of low accuracy, poor adaptability to complex faults, and slow response in traditional fault classification methods for Synchronous Digital Hierarchy (SDH) optical fiber communication networks, and to improve the efficiency and reliability of SDH network fault diagnosis. The research adopts the Support Vector Regression (SVR) algorithm as the core classification tool, combined with fault feature extraction technology and parameter optimization methods. First, a comprehensive SDH network fault dataset is constructed, which includes feature parameters of common fault types such as signal loss (LOS), bit error rate (BER) exceeding the standard, frame alignment error (FAE), and path mismatch, and the dataset is preprocessed through normalization and outlier removal to eliminate interference factors. Second, the SVR model's key parameters (including kernel function type, penalty factor C, and gamma coefficient) are optimized using the grid search method combined with 5-fold cross-validation to determine the optimal parameter combination that balances classification accuracy and generalization ability. Finally, the preprocessed fault feature data are input into the optimized SVR model for training and testing, and the model's performance is compared with traditional fault classification methods such as BP neural network and decision tree based on evaluation indicators including accuracy, recall, F1-score, and processing time. The results show that the SVR-based fault classification method achieves an average accuracy of over 96.5%, which is 8.2% and 11.7% higher than that of BP neural network and decision tree respectively; its recall rate for complex faults reaches 95.3%, and the average processing time per sample is reduced by 0.32s compared with traditional methods. This method can effectively identify various faults in SDH optical fiber communication networks, providing a reliable technical support for rapid fault location and maintenance of the network.

Keywords: SVR; SDH Optical Fiber Communication Network; Fault Classification; Parameter Optimization; Classification Performance

1. Introduction

1.1 Research Background and Significance

Synchronous Digital Hierarchy (SDH) optical fiber communication networks serve as the core infrastructure for high-speed data transmission in 5G networks, data centers, and national backbone communication systems, demanding ultra-high reliability and real-time fault handling. However, traditional fault classification methods for SDH networks face critical limitations: rule-based methods rely on manual experience and fail to adapt to complex fault scenarios (e.g., overlapping fault features between bit error rate (BER) anomalies and frame alignment errors (FAE)); shallow machine learning methods such as BP neural networks easily fall into local optima, leading to low classification accuracy for rare faults; and most methods lack efficient feature integration, resulting in prolonged processing time that cannot meet the real-time maintenance needs of large-scale SDH networks. Against this backdrop, exploring a fault classification method based on Support Vector Regression (SVR) addresses the technical bottlenecks of traditional approaches. This study holds both theoretical and practical value: theoretically, it enriches the application of SVR in nonlinear fault pattern recognition, providing a new perspective for the intersection of machine learning and communication network maintenance; practically, it improves the accuracy and speed of SDH network fault diagnosis, reducing maintenance costs caused by misdiagnosis or delayed fault handling, and ensuring stable operation of critical communication services.

1.2 Review of Domestic and Foreign Research Status

Foreign research on SDH network fault classification has focused on machine learning applications. Early studies adopted decision trees and k-nearest neighbors (KNN) for fault identification, but these methods show poor generalization when facing high-dimensional fault features. Recent works have attempted to use deep learning models such as CNNs, but they require massive labeled data and complex computing resources, which are difficult to deploy in practical SDH network maintenance systems. Additionally, foreign research rarely optimizes model parameters for SDH-specific fault characteristics, leading to suboptimal classification performance.

Domestic research has made progress in SDH fault mechanism analysis and feature extraction. Scholars have identified key fault indicators (e.g., signal power fluctuation, frame loss rate) and developed feature selection algorithms to reduce data dimensionality. However, most domestic studies use SVR in its original form without targeted parameter tuning, resulting in low accuracy for complex faults (e.g., concurrent LOS and path mismatch). Existing literature also lacks systematic comparison between SVR and mainstream methods under unified evaluation criteria, making it hard to verify the advantages of SVR in SDH fault classification.

2. Relevant Theories and Technical Foundations

2.1 Principles and Fault Types of SDH Optical Fiber Communication Networks

SDH networks adopt a standardized frame structure (STM-N) with a fixed frame period of 125μs, enabling synchronous transmission of multiple services. Core technical features include hierarchical multiplexing, built-in error monitoring (via B1/B2/B3 bytes), and self-healing rings, which support rapid service recovery but increase fault feature complexity.

Common SDH network fault types are categorized based on their causes and manifestations:

Loss of Signal (LOS): Caused by fiber breakage or severe signal attenuation, characterized by sudden signal amplitude drop to zero.

BER Exceeding Standard: Triggered by noise interference or equipment aging, with BER values exceeding the threshold of 10^{-9}.

Frame Alignment Error (FAE): Resulting from frame header mismatch, leading to intermittent service interruption.

Path Mismatch: Arising from incorrect service mapping, manifested by inconsistent path identifiers between network nodes.

Each fault type exhibits distinct feature patterns, providing a basis for subsequent feature extraction and classification.

2.2 Basic Principles of SVR Algorithm and Kernel Function Selection

SVR is a supervised learning algorithm that maps input data to a high-dimensional feature space via kernel functions, constructing an optimal hyperplane to minimize regression errors. Unlike traditional regression models, SVR introduces an ε-insensitive loss function to tolerate minor errors, enhancing robustness to noise.

Kernel functions determine SVR’s ability to handle nonlinear problems, with three common types:

Linear Kernel: Suitable for linearly separable data, with low computational complexity but poor performance for nonlinear faults.

Polynomial Kernel: Captures moderate nonlinear relationships but requires tuning of multiple parameters, increasing model complexity.

Radial Basis Function (RBF) Kernel: Maps data to infinite-dimensional space, effectively fitting complex nonlinear fault features with only two key parameters (penalty factor C and gamma coefficient), making it the optimal choice for SDH fault classification.

2.3 Fault Feature Extraction Technology for SDH Networks

Fault feature extraction converts raw SDH network data into discriminative feature vectors, laying the foundation for SVR classification. Key technologies include:

Time-Domain Feature Extraction: Extracts signal amplitude, pulse width, and amplitude fluctuation frequency from real-time signal waveforms, capturing transient features of LOS and FAE.

Frequency-Domain Feature Extraction: Uses Fast Fourier Transform (FFT) to obtain spectral peaks and bandwidth, identifying BER anomalies caused by frequency-specific noise.

Statistical Feature Extraction: Calculates mean, variance, and skewness of monitoring metrics (e.g., B1 byte error count), quantifying long-term fault trends and distinguishing path mismatch from other faults.

3. Design of SVR-Based Fault Classification Method for SDH Optical Fiber Communication Networks

3.1 Construction and Preprocessing of SDH Network Fault Dataset

The dataset is constructed by combining simulated and real operational data to ensure diversity. Simulated data is generated via OPNET software, simulating 10,000 samples of four fault types (LOS, BER, FAE, path mismatch) with 12 feature dimensions (including signal amplitude, BER, frame alignment success rate). Real data is collected from a provincial SDH backbone network, adding 5,000 samples of rare concurrent faults to enhance generalization.

Preprocessing steps eliminate interference and standardize data:

Outlier Removal: Uses Z-score method to filter samples with feature values outside the range of [μ-3σ, μ+3σ], removing 4.8% of abnormal data caused by sensor errors.

Normalization: Applies Min-Max scaling to map all features to [0,1], avoiding dominance of high-magnitude features (e.g., signal power) over low-magnitude features (e.g., error count).

3.2 Design of SVR Model Parameter Optimization Strategy

Parameter optimization targets two key parameters (C and gamma) to balance classification accuracy and generalization. The strategy adopts grid search combined with 5-fold cross-validation:

Parameter Range Setting: C is set to [0.1, 1, 5, 10] to adjust the penalty for misclassification; gamma is set to [0.01, 0.1, 1, 10] to control the influence of individual samples.

Cross-Validation Process: The dataset is divided into 5 subsets, with 4 subsets used for training and 1 for validation in each iteration. This process repeats 5 times to calculate the average accuracy for each parameter combination.

Optimal Parameter Determination: The combination with the highest average accuracy (C=5, gamma=0.1) is selected, avoiding overfitting caused by excessive C or poor generalization caused by inappropriate gamma.

3.3 Construction of SVR-Based Fault Classification Process

The classification process forms a closed loop from data input to result output, with clear logical links:

Data Acquisition: Collect real-time monitoring data (e.g., signal waveform, error byte count) from SDH network nodes via SNMP protocol.

Feature Extraction: Apply time-domain, frequency-domain, and statistical extraction methods to generate 12-dimensional feature vectors.

Data Preprocessing: Execute outlier removal and normalization to obtain standardized feature data.

Model Training: Input preprocessed data into the optimized SVR model for training, generating a fault classification model.

Fault Classification: Input real-time test data into the trained model, outputting fault type and confidence level (e.g., 98.5% confidence for LOS).

4. Experimental Verification and Performance Analysis

4.1 Experimental Environment Construction and Parameter Setting

The experimental environment is configured to ensure reproducibility:

Hardware: Intel Core i7-12700K CPU (3.6GHz), NVIDIA RTX 3060 GPU (12GB), 32GB DDR4 RAM, ensuring efficient model training and data processing.

Software: Python 3.9, Scikit-learn 1.2.2 (for SVR implementation and parameter optimization), Matplotlib 3.7.1 (for result visualization).

Parameter Setting: The dataset is split into training set (70%) and test set (30%); the SVR model uses RBF kernel with optimized parameters (C=5, gamma=0.1); comparison models (BP neural network, decision tree) use default parameters from Scikit-learn for fair comparison.

4.2 Experimental Data Source and Preprocessing Implementation

Experimental data combines simulated and real data:

Simulated Data: Generated via OPNET, covering 4 single fault types with 10,000 samples, each with 12 feature dimensions.

Real Data: Collected from a provincial SDH backbone network, including 5,000 samples (3,000 single faults, 2,000 concurrent faults), ensuring consistency with actual operational scenarios.

Preprocessing is implemented via Python:

Outlier removal uses Z-score calculation, filtering 480 simulated samples and 240 real samples.

Normalization applies Min-Max scaling, with code implementing feature-wise scaling to ensure each feature is mapped to [0,1].

4.3 Determination of Classification Performance Evaluation Indicators and Comparative Analysis

Four evaluation indicators are selected to comprehensively assess performance:

Accuracy: Ratio of correctly classified samples to total samples, reflecting overall classification ability.

Recall: Ratio of correctly classified samples of a specific fault type to all actual samples of that type, evaluating recognition ability for rare faults.

F1-Score: Harmonic mean of precision and recall, balancing accuracy and completeness.

Processing Time: Average time to classify a single sample, reflecting real-time performance.

Experimental results show the SVR-based method outperforms BP neural network and decision tree in all indicators.

The SVR-based method achieves higher accuracy and recall due to the RBF kernel's ability to fit nonlinear fault features and optimized parameters that enhance generalization. Its shorter processing time stems from SVR's reliance on support vectors (only 12% of training samples), reducing computational complexity compared to BP neural network's full-sample training.

5. Conclusion

This study achieves three key results: First, it designs a SVR-based fault classification method for SDH optical fiber communication networks, addressing the limitations of traditional methods by integrating targeted feature extraction and parameter optimization. Second, it verifies the method's superiority through experiments—achieving 96.5% accuracy, 95.3% recall for rare faults, and 0.18s per-sample processing time, outperforming BP neural network and decision tree in all indicators. Third, it constructs a standardized dataset and preprocessing workflow, providing a replicable experimental framework for subsequent SDH fault classification research.

This study has two limitations: One is the limited fault type coverage—focusing on single and dual concurrent faults, while lacking analysis of multi-fault fusion scenarios (e.g., simultaneous LOS, BER, and path mismatch) common in large-scale networks. The other is the lack of edge deployment testing—the method is verified in a laboratory environment, and its performance in resource-constrained edge devices (e.g., SDH node controllers) remains untested.

Future research will expand in two directions: On one hand, it will integrate multi-label classification algorithms to handle multi-fault fusion scenarios, enriching the fault type coverage. On the other hand, it will lightweight the SVR model via model compression techniques (e.g., support vector pruning), enabling deployment on edge devices and further improving the practicality of the method in real SDH network maintenance.

References

[1] Li, F. J., & Zhou, X. (2011). SDH Optical Transmission Equipment Commissioning and Maintenance. Beijing: Science Press.

[2] Liu, Z. G. (2021). Analysis and processing of SDH equipment failures. Railway Signaling & Communication, (1), 31.

[3] Sun, J. K. (2019). SDH digital microwave transmission system in radio and television and its fault analysis and processing suggestions. Science and Technology Innovation Herald, 16(28), 132-133.

[4] Lan, H. (2019). Analysis on the reliability of SDH digital microwave transmission in radio and television. China Electronics, (12), 77-78.

[5] Qin, B., & Wang, X. Y. (2020). Fault analysis and processing of smart radio and television bidirectional services. Radio & Television Network, 2(12), 99-100.

[6] Li, M. T., Wang, Y., & Liu, J. (2021). Core ideas of SDH optical fiber communication equipment maintenance and fault handling. Popular Science, (5), 66-67.

[7] Liu, X. M., & Ye, M. (2018). Fault analysis and processing method of SDH transmission network. Scientific Research, (5), 47.

[8] Huang, J. H. (2017). Flexible support vector regression algorithm and its application in

 Copyright @ STEMM Institute Press

fault diagnosis. Journal of Zhejiang University of Technology, 45(6), 629-634.
[9] Wen, C. L., & Lü, F. Y. (2020). Review of deep learning-based fault diagnosis methods. Journal of Electronics & Information Technology, 42(1), 234-248.
[10] Zhang, Z. P., & Zhao, J. S. (2017). Fault diagnosis of chemical processes based on deep belief network. Computers & Chemical Engineering, 107, 39-407.
[11] Li, Y. H., & Wang, H. B. (2021). Research on fault prediction and location of optical transport network in high-speed railway. Journal of Beijing Jiaotong University, 45(3), 321-328.
[12] Chen, X. Y., & Wang, Y. (2024). Machine learning-based alarm classification and correlation in SDH/WDM optical networks. Optical Engineering, 63(2), 026105.
[13] Wang, Y. Q., & Li, X. D. (2025). AI-embedded optical module with millisecond-granularity analysis for autonomous metro transport network. Journal of Lightwave Technology, 43(5), 1234-1242.
[14] Ma, S. Q., & Zhang, Y. L. (2023). Research on fault classification method of SDH optical fiber communication network based on improved SVR. Optical Fiber and Optical Communication Technology, 41(4), 56-62.
[15] Liu, Y. F., & Zhao, G. Q. (2024). Risk early warning and evaluation method for electric power SDH networks based on BP neural network. IEEE Transactions on Power Delivery, 39(2), 876-883.
Copyright @ STEMM Institute Press