Theory and Practice of Science and Technology
2025, VOL. 6, NO. 8, 124-128
DOI: 10.47297/taposatWSP2633-456923.20250608

Research on Network Transmission Anomaly Detection and Optimization Based on Signal Decomposition and LSTM

Ning Pan

Big Data Center, Hubei University, Wuhan, Hubei, 430062, China

ABSTRACT
In high-performance network transmission, traffic presents significant nonlinear and nonsmooth characteristics, which brings challenges to anomaly detection. In order to achieve accurate identification and trend warning of network anomalies, a network transmission anomaly detection and optimization method based on ensemble empirical modal decomposition and attention mechanism two-way long short-term memory network is proposed. The study adopts the ensemble empirical modal decomposition to extract multi-scale features, combines with Hurst index to locate anomalous segments, and realizes high-precision network traffic prediction and early warning through the bi-directional long- and short-term memory network with the introduction of the attention mechanism. Experimental results show that the method achieves 0.96, 0.95, 0.95, 0.95, and 0.95 in accuracy, precision, recall, and F1, respectively; the detection latency is stabilized at about 280ms, and the occupancy rate of the central processor does not exceed 30%. The study shows that the method can improve the detection accuracy and advance warning ability of anomalous traffic in complex encrypted network environment, and provide efficient and intelligent technical support for network security protection.

KEYWORDS
Network transmission anomaly detection; Cognitive intelligence; Traffic prediction optimization; Ensemble empirical modal decomposition; Attention mechanism

1 Introduction

Against the background of rising demand for high-speed network communication and data transmission, traditional transmission control protocols have gradually exposed bottlenecks in terms of latency, connection establishment speed, and multipath transmission capability. As a new generation of transmission mechanism based on user datagram protocol, Fast User Datagram Internet Connection Protocols (QUIC) integrates the features of multiplexing, zero round-trip handshaking and end-to-end encryption, which can effectively reduce the connection latency and improve security [1-2]. However, in multi-path, massively concurrent and encrypted traffic scenarios, the nonlinear and non-smooth characteristics of its traffic are more prominent, and traditional static feature detection methods are difficult to adapt [3]. Currently, the combination of signal decomposition techniques and deep learning provides new ideas to solve the above problems. Among them, Ensemble Empirical Mode Decomposition (EEMD) can effectively alleviate the modal aliasing problem by introducing Gaussian white noise and decomposing several times to find the mean value [4]. However, EEMD itself is mainly applicable to feature extraction and lacks the ability to predict future trend-oriented. Bidirectional Long Short-Term Memory (BiLSTM) can simultaneously capture the before and after time series information, while Attention Mechanism (Attention) can further highlight the key feature regions and improve the prediction performance [5]. Therefore, the study proposes a network transmission anomaly detection and optimization method that integrates EEMD and Attention-BiLSTM. The method first utilizes EEMD for multi-scale feature decomposition, combined with Hurst exponent to achieve rapid identification of abnormal segments; Subsequently, BiLSTM was used for bidirectional temporal modeling, and key feature weight allocation was optimized through Attention, aiming to achieve accurate identification and early warning of abnormal traffic in complex network environments, providing new ideas and methods for the application of cognitive intelligence in network anomaly detection. The innovations of the research include combining EEMD with Hurst index for lightweight anomaly detection of encrypted traffic; fusing Attention-BiLSTM to achieve context dependency and key feature focusing; and striking a balance between low latency and low resource consumption in a multi-path QUIC environment. The expected contribution of this method is to provide efficient, accurate, and scalable detection and prediction schemes for active defense of high-performance networks.

2 Research on network transmission anomaly detection and optimization based on EEMD-attention-BiLSTM

2.1 Network anomaly detection model based on EEMD and hurst exponent

In modern network communication environments, especially in high-performance data transmission based on the

* **Corresponding Author:** Ning Pan, panda20087@hubu.edu.cn

QUIC protocol, network traffic often exhibits the complex characteristics of non-linearity and non-smoothness. This makes the traditional methods based on static statistics such as mean and variance suffer from poor adaptability and high leakage rate in anomaly detection.Hurst exponent, as an important measure of the self-similarity of time series, has been widely used to detect the long-range dependence characteristics in network traffic [6-7]. Meanwhile, Empirical Mode Decomposition (EMD) technique has already performed well in non-stationary signal processing by virtue of its good local adaptability. However, the underlying EMD is prone to mode aliasing problems in practical applications, which affects the detection accuracy [8]. To address these challenges, a research proposes an anomaly detection method that integrates cognitive intelligence. By introducing EEMD technology and combining sliding windows with Hurst exponent, a traffic anomaly determination model is constructed. This model uses EEMD to simulate human perception and decomposition ability of complex signals, and then reconstructs the signal in segments according to sliding windows. The Rescaled Range Analysis (R/S) method is used to calculate the Hurst value to determine abnormal segments, thereby achieving rapid identification of sudden abnormal traffic in QUIC networks. This design enables the model to intelligently identify and predict network anomalies. In the EEMD process, white noise is first added to the original signal sequence, decomposed multiple times, and the Intrinsic Mode Function (IMF) is averaged to obtain stable decomposition results. The specific calculation for defining a single denoising decomposition is shown in equation (1).

$$X_i(t) = X(t) + \epsilon_i(t), i = 1,2,\cdots,N \tag{1}$$

In Eq. (1),$X(t)$isthe moment ofthe original signal sequence;$\epsilon_i(t)$(is theGaussian white noiseshowing the first)$i$the second addition;$N$is the number of iterative rounds of the noise addition decomposition of the EEMD. The final IMF result is the average of the results of each decomposition, specificallyas shown in Eq. (2).

$$\mathrm{IMF}_k(t) = \frac{1}{N}\sum_{i=1}^{N}\mathrm{IMF}_k^{(i)}(t) \tag{2}$$

In Eq. (2),$\mathrm{IMF}_k(t)$(is thevalue ofthe first)$k$eigenmode function at time$t$(value [) (9-10)].The resulting low-frequency IMF components are reconstructed to form a new signal. Subsequently, the new signal was segmented by a fixed window length and the Hurst index was calculated for each segment. The Hurst exponent is defined using R/S analysis, specificallyas shown in Eq. (3).

$$H = \frac{\log\left(\frac{R(n)}{S(n)}\right)}{\log(n)} \tag{3}$$

In equation (3), $H$is the Hurst index; $R(n)$ is the range extreme deviation; $S(n)$ is the standard deviation; and$n$ is the number of sample points. If the H-value of a segment is much higher than the normal benchmark, it can be determined that there is an abnormal traffic behavior of self-similarity enhancement in that segment. The structure of EEMD-based network anomaly detection model is shown in Fig. 1.

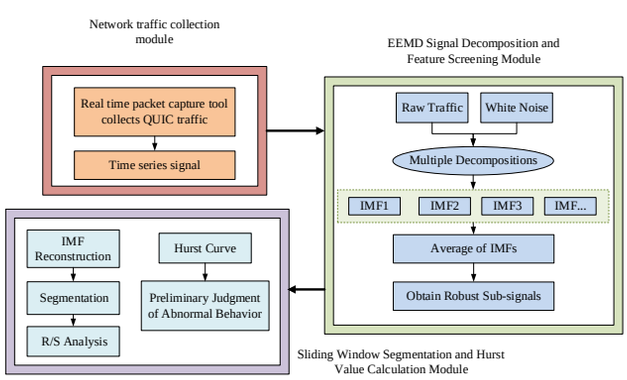


Figure 1 Network anomaly detection structure diagram based on EEMD

As can be seen from Figure 1, the model is divided into three main modules. First, the packets during the transmission of QUIC protocol are captured and grouped in real time by the network traffic acquisition module, and form a continuous time series signal. Subsequently, the EEMD module introduces independent Gaussian white noise into the original signal several times, completes multiple rounds of signal decomposition, and weighted averages each order of the modal function in order to obtain a stable IMF component. Next, the system selects a number of low-frequency IMF components for signal reconstruction, and divides the reconstructed signal into multiple time segments using the sliding window mechanism. For each segment of the signal, its Hurst index is calculated using the R/S statistic to characterize its self-similarity. It is determined whether the Hurst value is significantly higher than the normal benchmark, and if so, it indicates that there may be a network traffic mutation or malicious behavior in that time period [11]. With the features of no need to parse protocol fields, high sensitivity to encrypted traffic, and high adaptability, the model is suitable for lightweight anomaly detection tasks in modern complex high-frequency data transmission networks.

2.2 Optimization model for traffic prediction based on attention BiLSTM

The study proposes an anomaly detection method based on EEMD with Hurst index, which is able to realize rapid localization of abnormal network traffic and better identify the local mutation and volatility enhancement region of traffic. However, the method is mainly oriented to analyze after the fact and lacks advance warning of potential anomaly trends. In practical network protection, sensing potential anomalies in advance is the key to realize proactive defense. Considering the significant time-dependence and nonlinear dynamic characteristics of network traffic, it is difficult to comprehensively portray its evolution law by relying only on statistical indicators [12-13]. Long Short-Term Memory (LSTM), as a classical temporal modeling structure, has good long-term memory capability, which is especially suitable for network traffic, such as periodic and sudden sequence data. Meanwhile, BiLSTM is able to combine before and after contextual information, while Attention helps the model to focus on key time segments, thus improving the accuracy of anomaly prediction [14]. Therefore, the study proposes an anomaly prediction model based on Attention-BiLSTM fusion network on the basis of EEMD feature extraction. Its overall structure is shown in Figure 2.

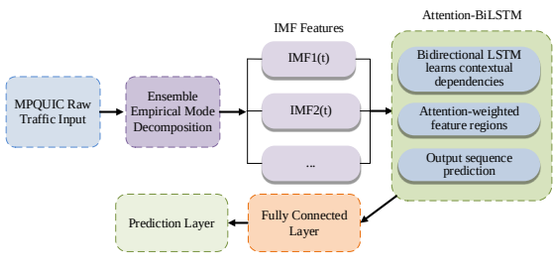


Figure 2 Structure diagram of traffic prediction model based on EEMD attention BiLSTM

Figure 2 shows the complete structure of the EEMD-Attention-BiLSTM prediction model. The original MPQUIC traffic is first decomposed by the EEMD module to obtain multiple IMF feature components and form a feature matrix. Subsequently, these features are input to BiLSTM network, which learns the temporal dependencies of history and future directions respectively; the Attention module assigns attention weights to hidden states to enhance the ability of focusing on anomalous segments. The final fusion information predicts the traffic changes in future moments through the fully connected layer. The system compares the predicted value with the actual flow and marks the anomaly tendency according to the residual size. The model can provide early warning of unknown anomalies and shows better robustness and generalization ability in complex multi-path environment. Among them, the IMF feature matrix obtained by EEMD decomposition is input into the bidirectional LSTM for modeling. BiLSTM produces forward and backward hidden state splicing at each time step as shown in equation (4).

$$h_t = [\overrightarrow{h}_t; \overleftarrow{h}_t] \tag{4}$$

In Eq. (4), $\overrightarrow{h}_t, \overleftarrow{h}_t$ (are the forward and backward hidden states [] (15) (], respectively). The obtained bidirectional hidden state vector not only retains the influence of historical information on the current moment, but also incorporates future information to complement the overall trend, thus providing a more comprehensive and fine-grained representation of the temporal features for the subsequent note Attention. On this basis, the Attention module is able to differentially weight the importance of different time steps to enhance the model's ability to capture key anomalous patterns, the calculation of which is shown in equation (5).

$$\alpha_t = \frac{\exp(V^T \tanh(Wh_t))}{\sum_{t'} \exp(V^T \tanh(Wh_{t'}))} \tag{5}$$

In Eq. (5), W, V are the learnable parameters; α_t is the contribution of the moment t to the overall prediction; T (is the time step [] (16-17)]. The context vector is obtained by weighted summation of the hidden states of all time steps, which is calculated as shown in Eq. (6).

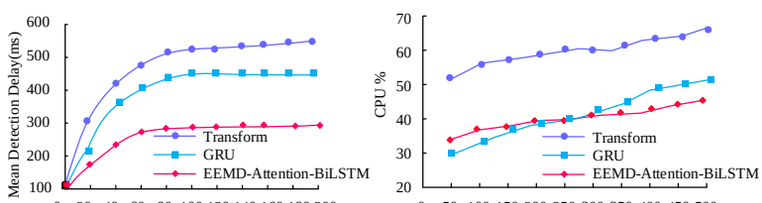
$$c = \sum_t \alpha_t h_t \tag{6}$$

In Eq. (6), c is the context vector. The vector contains global timing dependency and key feature information, which is finally input into the fully connected layer to generate the prediction value and determine whether there is a potential anomaly according to the prediction residuals, and at the same time, the adaptive discrimination under different network states is realized by setting the dynamic threshold mechanism; this process combined with the sliding window updating strategy, which enables the model to maintain real-time tracking of potential anomalies in the continuous traffic change and the Early warning capability.

3 Performance analysis of network transmission anomaly detection and prediction models

To verify the effectiveness of the proposed EEMD-Attention-BiLSTM-based network transmission anomaly detection

and optimization method, the experiments are completed on a unified hardware and software platform. In terms of hardware configuration, a workstation equipped with Intel Core i9-13900K processor with a main frequency of 3.0 GHz, a memory capacity of 64 GB, and NVIDIA RTX 4090 GPU (24 GB of graphics memory) is used; in terms of software configuration, the operating system is Ubuntu 22.04 LTS, and the deep learning framework is selected as PyTorch 2.2.0. Signal decomposition and data preprocessing use Python 3.10 and its scientific computing libraries such as NumPy, Pandas, PyEMD, etc., and the network protocol stack simulation relies on the ns-3.39 and QUIC protocol extension module to complete. The data source of the experiment is the publicly available A Large One-Month QUIC Network Traffic Dataset from Backbone Lines (CESNET-QUIC22), which consists of one-month QUIC network traffic data from the Jie CESNET2 education and research network, covering a period of one month. QUIC Network Traffic Dataset from Backbone Lines, CESNET-QUIC22, which consists of one month of QUIC network traffic data from the Jet CESNET2 education and research network covering about 500, 000 users and more than 500 large organizations. In order to evaluate the performance difference between different models in terms of detection latency and Central Processing Unit (CPU) occupancy, comparative experiments were conducted on Transformer neural network and Gated Recurrent Unit (GRU), and the results are shown in Fig. 3.



(a) Comparison results of training time for different algorithms

(b) Comparative analysis of computing resource utilization of different models

Figure 3 Comparative analysis of different algorithms in terms of training time and computational resource utilization

Figures 3(a) and 3(b) show the variation curves of detection delay with the number of training iterations and computational resource utilization with the number of samples for different models, respectively. As can be seen from Figure 3(a), EEMD-Attention-BiLSTM exhibits the lowest average detection latency in all iteration stages, which is stable at about 280ms, while GRU and Transformer are stable at about 430ms and 520ms, respectively. The main reason for the delay difference is that the EEMD decomposition filters out the redundant noise components in advance, which reduces the redundant computation of the model in the temporal feature extraction stage, while the Attention mechanism shortens the backpropagation path after weighting the key features. The higher delay of the Transformer is mainly affected by the computational complexity of the multi-head self-attention, and its computational overhead is significantly increased in the case of larger sequence lengths . From Fig. 3(b), it can be seen that the CPU occupancy of EEMD-Attention-BiLSTM always stays below 30% and rises slowly with the increase of the number of samples, which shows high computational efficiency; the CPU occupancy of GRU peaks at about 50% after the number of samples exceeds 300, while that of Transformer increases from about 50% to nearly 65%. gradually rising to nearly 65%. In order to further verify the contribution of each module to the overall model performance, an ablation experiment based on EEMD-Attention-BiLSTM is designed, and its prediction performance is shown in Table 1.

Table 1 Classification and prediction performance of ablation experiments

Model	Accuracy	Precision	Recall	F1	AUC
EEMD-Attention-BiLSTM	0.96	0.95	0.95	0.95	0.98
w/o Attention	0.94	0.94	0.93	0.93	0.97
w/o BiLSTM	0.94	0.936	0.92	0.93	0.96
w/o EEMD	0.93	0.92	0.91	0.91	0.95
w/o EEMD-Attention	0.91	0.90	0.89	0.90	0.94
Only LSTM	0.90	0.89	0.88	0.88	0.93

As can be seen from Table 1, the complete EEMD-Attention-BiLSTM model in terms of accuracy, precision, recall, F1 and Area Under the Receiver Operating Characteristic Curve (AUC) metrics all achieved the best performance, reaching 0.96, 0.95, 0.95, 0.95, 0.95 and 0.98, respectively, indicating that the synergistic effect of the EEMD decomposition and the Attention mechanism can significantly improve the feature extraction precision and the timing-dependent modeling capability. When the Attention module is removed, the accuracy and F1 drop to 0.94 and the AUC drops to 0.97, indicating that the Attention mechanism contributes positively to the prediction results in highlighting key features and suppressing redundant information. When BiLSTM is removed, accuracy drops to 0.936 and F1 to 0.93, indicating that the bidirectional loop structure is advantageous in capturing the pre and post timing dependencies. The effect of removing EEMD is more pronounced, with recall dropping to 0.91 and F1 to 0.91, due to enhanced modal aliasing and noise interference resulting from the undecomposed signal. If EEMD and Attention are removed at the same time, the performance decreases further

and F1 is only 0.90. The lowest of all metrics, accuracy 0.90 and AUC 0.93, is achieved when only LSTM is used, indicating that a single recurrent structure without the introduction of multi-scale decomposition and Attention weight assignment has limited ability to model the transmission characteristics of a complex network.

4 Conclusion

In order to achieve efficient prediction and anomaly advance warning of encrypted and multipath traffic in complex network environment, the study constructed a network transmission anomaly detection and optimization model based on the fusion of EEMD and Attention-BiLSTM. The method achieves fast localization of anomalous segments through EEMD multiscale decomposition combined with Hurst index, and then uses BiLSTM to capture bidirectional timing dependencies and highlights the key features through Attention, thus realizing high-precision prediction of anomalous trends. The experimental results show that, in terms of detection latency and computational resource consumption, the average detection latency of the model in each iteration stage is stabilized at about 280 ms, which is significantly lower than that of the GRU at about 430 ms and that of the Transformer at about 520 ms; at the same time, the CPU occupancy rate is always lower than 30%, whereas it reaches the maximum of about 50% and 65% for the GRU and the Transformer, respectively. The study shows that the multi-scale decomposition of EEMD and the feature weighting mechanism of Attention effectively improve the timing modeling accuracy and significantly enhance the ability to sense potential anomalies in advance. However, despite this, the study still has room for improvement. Since the experiments are mainly based on a single public QUIC dataset, the generalization performance of the model under different network environments and protocol types needs further verification, and the response speed of extreme sudden anomalies is not optimized enough. In the future, the adaptability and real-time performance of the model can be further improved by introducing cross-domain multi-source traffic datasets, fusing lightweight deep models and online learning mechanisms.

About the Author

Ning Pan, February 1972.2, Male, MBA, Senior Experimentalist, Research direction: Computer network and application, smart campus construction management.

References

[1] Wang X Y, Zhang H X, Wang B, Abnormalities of efficiency in resting state functional brain network in first-episode paranoid schizophrenia[J]. Chin. J. Behav. Med. Brain Sci., 2021, 30(03): 219-225.

[2] Yang D, Wang H, He R J, Wind Speed Prediction Based on Improved Empirical Mode Decomposition and Hybrid Deep Learning Models[J]. Smart Power, 2024, 52(01): 1-7.

[3] Xu P L, Yang W G, Improvement of Bird's Nest Detection Network Structure of Transmission Lines Based on YOLO[J]. Smart Power, 2024, 52(04): 54-61.

[4] Zhou F, Wang S P, Zhang K P, A Short-term Traffic Flow Prediction Model for Urban Road Network Based on Graph Transformer Network[J]. Sci. Technol. Eng., 2024, 24(10): 4307-4316.

[5] Du W L, Niu X C, Wang H C, A fault diagnosis method for key transmission components of rotating machinery based on SAM-1DCNN-BiLSTM temporal and spatial feature extraction[J]. Proc. Inst. Mech. Eng. Part C-J. Mech. Eng. Sci., 2025, 239(14): 5810-5822.

[6] Xu K X, Long K P, Lu Y, Joint Secure Transmission and Trajectory Optimization for Reconfigurable Intelligent Surface-aided Non-Terrestrial Networks[J]. J. Electron. Inf. Technol., 2025, 47(02): 296-304.

[7] Sheng H, Xi Z, Dengji L, Construction of a knowledge graph for reusable spinning processes based on bidirectional long short-term memory networks[J]. J. Silk, 2024, 61(12): 52-60.

[8] Huang X Y, Lin Z A, Tang B, Minimum Deployment Method for Wireless Sensor Network Base Stations in Transmission Line Inspection[J]. Chin. J. Sens. Actuators, 2024, 37(12): 2125-2130.

[9] Feng L Z, Liu F R, Wang Y W, Detecting Rumor Based on Graph Convolution Network and Attention Mechanism[J]. Data Anal. Knowl. Discov., 2024, 8(04): 125-136.

[10] Li M L, Zhu Y C, Shen C N, RIS-Assisted ISAC with Non-orthogonal Multiple Access Transmission and Resource Allocation Optimization in Vehicular Networks[J]. J. Electron. Inf. Technol., 2025, 47(04): 1043-1051.

[11] Ma Y Y, Xu M, Jin H B, Multi-Channel Adaptive Feature Fusion for Urban Road Network Traffic Prediction[J]. Comput. Eng. Appl., 2025, 61(07): 334-341.

[12] He H Q, Zhou F Y, Chen M, Fruit Tree Canopy Segmentation by Unmanned Aerial Vehicle Photogrammetry Coupled on Convolutional Neural Network and Attention Mechanism[J]. J. Geo-Inf. Sci., 2023, 25(12): 2387-2401.

[13] Duan Q W, He X Z, Chao Z, Short-term Load Forecasting Model Based on Ensemble Empirical Mode Decomposition and Q Learning Strategy [J]. Mod. Electr. Power, 2025, 42(02): 360-368.

[14] Wang H T, An Energy Efficient Data Transmission Scheme for Wireless Sensor Network[J]. Chin. J. Sens. Actuators, 2025, 38(04): 724-731.

[15] Lu Z, Shen M X, Liu L S, Study on feeding behavior recognition method of fattening pigs based on lightweight networks and attention mechanism[J]. J. Nanjing Agric. Univ., 2023, 46(04): 802-812.

[16] Ma R H, Yu C Y, Tong Y X, Image Denoising Based on Residual Dense Networks and Attention Mechanism[J]. Sci. Technol. Eng., 2025, 25(09): 3795-3805.

[17] Yin L S, Liu P, Sun S C, Traffic Flow Combined Prediction Model Based on Complementary Ensemble Empirical Mode Decomposition and Bidirectional Gated Recurrent Unit Optimized by Improved Sparrow Search Algorithm[J]. J. Electron. Inf. Technol., 2023, 45(12): 4499-4508.