Self-Attention-Based Time Series Analysis for Anomaly Detection in Vehicle Sensor Data

HARISH B
harishraj@gmail.com
Department of Electronics and Communication Engineering
Sir Issac Newton College of Engineering and Technology Nagapattinam, Tamil Nadu, India

MALATHI P
mala_thi@gmail.com
Department of Electronics and Communication Engineering
Sir Issac Newton College of Engineering and Technology Nagapattinam, Tamil Nadu, India

MIRUTHULA G
mirthu@gmail.com
Department of Electronics and Communication Engineering
Sir Issac Newton College of Engineering and Technology Nagapattinam, Tamil Nadu, India

SANTHIYA R
sandy23@gmail.com
Department of Electronics and Communication Engineering
Sir Issac Newton College of Engineering and Technology Nagapattinam, Tamil Nadu, India

DURGALAKSHMI
Assistant Professor
durgalakshmi@sincet.ac.in
Department of Electronics and Communication Engineering
Sir Issac Newton College of Engineering and Technology Nagapattinam, Tamil Nadu, India

BAHIRUNISHA
Assistant Professor
bahirunisha@sincet.ac.in
Department of Electronics and Communication Engineering
Sir Issac Newton College of Engineering and Technology Nagapattinam, Tamil Nadu, India

Abstract

Modern vehicles are equipped with numerous sensors that continuously generate time-dependent data related to engine performance, temperature, pressure, speed, fuel consumption, battery condition, and other operational parameters. Monitoring these sensor streams is essential for detecting abnormal vehicle behavior and preventing unexpected failures. However, conventional threshold-based and statistical techniques may struggle to identify complex temporal patterns and subtle anomalies in multivariate vehicle sensor data. This paper presents a Self-Attention-Based Time Series Analysis approach for anomaly detection in vehicle sensor data. The proposed method utilizes a self-attention mechanism to learn relationships between different time steps in sequential sensor observations. The model captures both short-term variations and long-range temporal dependencies, enabling improved identification of abnormal operating conditions. The proposed framework consists of data acquisition, preprocessing, normalization, temporal window generation, self-attention-based feature extraction, anomaly scoring, and anomaly classification. Historical vehicle sensor data are used to learn normal operating patterns, while deviations from the learned representation are identified as potential anomalies. Performance can be evaluated using precision, recall, F1-score, accuracy, and area under the ROC curve. The proposed approach provides a flexible framework for intelligent vehicle health monitoring and can support predictive maintenance by identifying abnormal behavior at an early stage.

Keywords: Vehicle Sensor Data, Anomaly Detection, Time Series Analysis, Self-Attention, Machine Learning, Deep Learning, Predictive Maintenance, Intelligent Vehicle Monitoring.

1. Introduction

The development of intelligent and connected vehicles has resulted in the widespread use of electronic sensors and embedded systems. Modern vehicles continuously generate large volumes of sensor data representing different aspects of vehicle operation. Examples include engine temperature, vehicle speed, engine RPM, air pressure, battery voltage, fuel level, coolant temperature, and other parameters. Analysis of these data streams can provide valuable information about vehicle health and operating conditions.

Anomaly detection is an important task in vehicle monitoring because abnormal sensor behavior may indicate component degradation, malfunction, sensor failure, or unsafe operating conditions. Early detection of anomalies can help reduce maintenance costs, prevent unexpected breakdowns, and improve vehicle safety.

Traditional anomaly detection approaches commonly use predefined thresholds, statistical techniques, or conventional machine learning algorithms. Threshold-based approaches are relatively simple but may generate false alarms because vehicle sensor values can vary according to driving conditions, environmental conditions, vehicle load, and operating state. Statistical approaches may also have difficulty modeling highly nonlinear and dynamic relationships between multiple sensors.

Time series analysis provides a more suitable approach because vehicle sensor data are sequential by nature. The value of a sensor at a particular time is often related to previous observations. Therefore, a reliable anomaly detection model should be capable of learning temporal dependencies.

Deep learning models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs) have been widely used for sequential data. Although these models can capture temporal dependencies, recurrent processing may make it difficult to efficiently model long-range relationships.

The self-attention mechanism provides an alternative approach. Instead of processing observations only sequentially, self-attention calculates relationships between different positions within a sequence. This allows the model to determine which time steps are more important for interpreting the current observation.

This paper proposes a Self-Attention-Based Time Series Analysis framework for anomaly detection in vehicle sensor data. The proposed approach is designed to learn normal temporal patterns and identify observations that deviate significantly from those patterns.

2. Problem Statement

Vehicle sensor data are characterized by high dimensionality, temporal dependency, noise, and varying operating conditions. A single abnormal sensor value does not always represent an actual vehicle fault. Similarly, a combination of individually normal sensor values may indicate an abnormal operating condition.

Conventional threshold-based methods generally analyze sensor values independently and may not capture relationships between multiple sensors and time steps.

Therefore, the main research problem addressed in this work is:

How can self-attention-based time series modeling be used to identify anomalous patterns in multivariate vehicle sensor data while considering temporal dependencies between observations?

3. Objectives

The major objectives of the proposed research are:

1. To analyze multivariate vehicle sensor data as time series sequences.

2. To preprocess and normalize vehicle sensor measurements.

3. To construct temporal observation windows from continuous sensor streams.

4. To develop a self-attention-based deep learning model for temporal feature extraction.

5. To learn normal patterns from historical vehicle sensor data.

6. To calculate anomaly scores for unseen sensor sequences.

7. To classify observations as normal or anomalous.

8. To evaluate the proposed approach using standard anomaly detection metrics.

9. To provide a framework that can support intelligent vehicle health monitoring and predictive maintenance.

4. Literature Review

Vehicle health monitoring has traditionally relied on diagnostic codes, predefined operating limits, and statistical analysis. These approaches can be effective for known failure conditions but may have limited ability to identify previously unseen or complex anomalies.

Machine learning techniques have been increasingly applied to predictive maintenance and anomaly detection. Supervised models can classify known fault categories when labeled training data are available. However, obtaining sufficient labeled failure data from real vehicles can be difficult because abnormal events are relatively rare.

Unsupervised and semi-supervised anomaly detection methods provide an alternative by learning normal behavior and identifying deviations. Autoencoders, clustering algorithms, isolation-based techniques, and probabilistic models have been used for this purpose.

For sequential data, recurrent neural networks have demonstrated the ability to capture temporal dependencies. LSTM and GRU architectures are particularly useful when the relationship between observations extends over multiple time steps.

However, recurrent models process sequences sequentially, which can limit parallelization and make the learning of long-range relationships more challenging. The self-attention mechanism addresses this limitation by directly computing relationships between elements within a sequence.

Self-attention has become an important component of Transformer-based architectures. It allows a model to assign different importance weights to different observations. In vehicle sensor analysis, this capability can help the model focus on time steps that contain meaningful changes or patterns.

The proposed research applies self-attention to multivariate vehicle sensor time series with the goal of identifying anomalous temporal behavior.

5. Proposed System

The proposed system consists of the following major stages:

1. Vehicle Sensor Data Collection
2. Data Preprocessing
3. Feature Selection
4. Time Series Window Generation
5. Self-Attention-Based Representation Learning
6. Anomaly Score Calculation
7. Threshold-Based Anomaly Classification
8. Performance Evaluation

5.1 System Architecture

The overall architecture can be represented as:

Vehicle Sensors

↓

Sensor Data Acquisition

↓

Data Cleaning and Preprocessing

↓

Normalization

↓

Temporal Window Generation

↓

Self-Attention Layer

↓

temporal Feature Representation

↓

Reconstruction / Prediction

↓

Anomaly Score

↓

Normal or Anomalous

↓

Vehicle Health Monitoring

The system receives continuous sensor measurements and transforms them into fixed-length temporal sequences. The self-attention model analyzes these sequences and learns the relationships among different time steps.

6. Vehicle Sensor Dat

The proposed framework can be applied to various vehicle sensor parameters depending on the available dataset.

Potential sensor features include:

Vehicle speed

Engine RPM

Engine temperature

Coolant temperature

Battery voltage

Fuel level

Air pressure

Throttle position

Engine load

Intake air temperature

Brake-related parameters

Acceleration

Gyroscope measurements

Not all parameters are required. The final feature set should depend on the dataset used for experimentation.

A typical input record may have the following srructure:

The actual values should be obtained from the selected data

7. Data Preprocessing

Raw vehicle sensor data may contain missing values, noise, duplicate observations, and inconsistent sampling intervals. Preprocessing is therefore necessary before model training.

7.1 Missing Value Handling

Missing observations can be handled using suitable techniques such as:

Forward filling

Interpolation

Mean or median replacement

Removal of incomplete records

The selected technique should depend on the characteristics of the dataset.

7.2 Noise Reduction

Vehicle sensors may generate short-term fluctuations caused by measurement noise. Filtering techniques can be applied to reduce unwanted variations while preserving meaningful changes.

7.3 Normalization

Because sensor parameters may have different numerical ranges, normalization is performed before model training.

Min-max normalization can be represented as:

$X' = (X - X_{min}) / (X_{max} - X_{min})$

where:

X is the original sensor value,

X_{min} is the minimum value,

X_{max} is the maximum value,

X' is the normalized value.

Normalization prevents parameters with larger numerical ranges from dominating the learning process.

8. Time Series Window Generation

The continuous sensor stream is divided into fixed-length sequences.

For example, if a sequence length of T observations is selected, each model input contains:

$X = [x1, x2, x3, ..., xT]$

where each xt represents a multivariate sensor observation at time t.

For example:

$X \in R^{(T \times F)}$

where:

T = number of time steps

F = number of sensor features

The selection of the window size is an important parameter. A short window may capture only local behavior, while a longer window may provide more temporal context.

The final window size should be determined experimentally

9. Self-Attention Mechanism

Self-attention allows the model to determine the importance of different observations within a temporal sequence.

For an input sequence, three representations are generated:

Query (Q)

Key (K)

Value (V)

The attention operation is commonly represented as:

$$\text{Attention}(Q, K, V) = \text{softmax}(QK^T / \sqrt{dk})V$$

where dk represents the dimensionality of the key vectors.

The attention mechanism calculates relationships between time steps and assigns higher weights to observations that are more relevant.

For vehicle sensor data, this means that the A temperature increase following an RPM change

A battery voltage change associated with vehicle operating state

A speed change associated with engine load

A sequence of sensor changes preceding an abnormal condition

This provides a richer temporal representation than analyzing individual sensor observations

10. Proposed Self-Attention-Based Model

The proposed model can use an encoder-based architecture consisting of:

1. Input Layer
2. Feature Projection Layer
3. Positional Encoding
4. Multi-Head Self-Attention
5. Feed-Forward Network
6. Normalization and Residual Connections

7. Latent Representation

8. Reconstruction or Prediction Layer

9. Anomaly Score Calculation

10.1 Input Layer

The input layer receives normalized multivariate time series windows.

ture Projection

The original sensor features are projected into a higher-dimensional representation suitable for attention processing.

10.3 Positional Encoding

Because attention does not inherently encode sequence order, positional information can be incorporated into the input representation.

This allows the model to distinguish between:

Sensor state at time t1

and

Sensor state at time t2.

10.4 Multi-Head Self-Attention

Multiple attention heads can learn different relationships within the sensor sequence.

One attention head may focus on short-term changes, while another may learn relationships involving longer temporal patterns.

10.5 Feed-Forward Network

The attention output is passed through a feed-forward network to learn nonlinear representations.

10.6 Reconstruction or Predictio

For unsupervised anomaly detection, the model can reconstruct the input sequence.

The reconstruction error can then be used to determine whether a sequence is normal or anomalous

11. Anomaly Detection Strategy

The proposed framework can use reconstruction error as an anomaly score.

Let:

X = original input sequence

X̂ = reconstructed sequence

The reconstruction error can be calculated using Mean Squared Error:

$MSE = (1/n)\ \Sigma(X_i - \hat{X}_i)^2$

A higher reconstruction error indicates that the observed sequence differs from the patterns learned from normal data.

An anomaly threshold can be determined from the validation data.

For example:

If anomaly score > threshold → Anomaly

If anomaly score ≤ threshold → Normal

The threshold should be determined experimentally rather than arbitrarily selected

12. Training Methodology

The model can be trained primarily using normal vehicle sensor sequences.

The training process consists of:

1. Collect normal sensor data.

2. Clean and normalize the data.

3. Generate temporal windows.

4. Feed the sequences into the self-attention model.

5. Train the model to reconstruct or predict normal patterns.

6. Calculate validation reconstruction errors.

7. Determine an appropriate anomaly threshold.

8. Evaluate the model using test data.

This approach is useful when labeled fault data are limited.

13. Experimental Setup

The proposed approach can be evaluated using a publicly available vehicle sensor dataset or a dataset collected from an actual vehicle.

The experimental setup should include:

Dataset: [Dataset Name]

Number of samples: [Number]

Number of sensor features: [Number]

Sampling frequency: [Value]

Sequence length: [Value]

Training data: [Percentage]%

Validation data: [Percentage]%

Testing data: [Percentage]%

Framework: [PyTorch / TensorFlow / Keras]

Hardware: [CPU/GPU configuration]

These values should be replaced with the actual experimental configuration.

14. Performance Evaluation

The proposed system can be evaluated using standard classification and anomaly detection metrics.

14.1 Accuracy

Accuracy measures the percentage of correctly classified observations.

Accuracy = (TP + TN) / (TP + TN + FP + FN)

14.2 Precision

Precision measures how many observations classified as anomalies are actually anomalous.

Precision = TP / (TP + FP)

14.3 Recall

Recall measures the proportion of actual anomalies correctly detected.

Recall = TP / (TP + FN)

14.4 F1-Score

The F1-score is the harmonic mean of precision and recall.

F1 = 2 × (Precision × Recall) / (Precision + Recall)

14.5 ROC-AUC

The Receiver Operating Characteristic Area Under the Curve can be used to evaluate the model's ability to distinguish between normal and anomalous observations across different thresholds.

15. Results and Discussion

The actual performance of the proposed model should be reported using experimentally obtained results.

The comparison should be performed under the same dataset and evaluation conditions.

Discussion

The expected advantage of the self-attention approach is its ability to model relationships between different positions within a time series. Unlike simple threshold-based approaches, the model can consider multiple sensor values and their temporal context.

For example, a small increase in engine temperature may not be anomalous by itself. However, if it occurs together with an unusual change in RPM, battery voltage, and engine load over a specific time interval, the combined sequence may represent abnormal behavior.

The self-attention mechanism can assign different importance to these observations and learn such temporal relationships.

However, model performance depends on dataset quality, sequence length, sensor selection, training strategy, and anomaly threshold selection

16. Advantages of the Proposed System

The proposed self-attention-based anomaly detection framework offers several advantages:

1. Temporal Dependency Modeling: It captures relationships between different time steps.

2. Multivariate Analysis: Multiple vehicle sensors can be analyzed simultaneously.

3. Long-Range Relationships: Attention can connect distant observations within a sequence.

4. Reduced Dependence on Manual Thresholds: The model can learn normal patterns from data.

5. Early Anomaly Detection: Abnormal behavior can potentially be detected before serious failure.

6. Scalability: Additional sensor features can be incorporated into the framework.

7. Predictive Maintenance Support: Detected anomalies can support maintenance decisions.

8. Flexible Deployment: The approach can be adapted to different vehicle sensor datasets.

17. Limitations

Despite its advantages, the proposed approach has certain limitations.

First, self-attention-based models can require significant computational resources, particularly for long sequences. Second, the quality of anomaly detection depends strongly on the quality and representativeness of the training data.

If the training dataset contains abnormal behavior that is incorrectly treated as normal, the model may learn those abnormal patterns. Similarly, changes in vehicle type, driving behavior, weather, road conditions, or sensor configuration may affect model performance.

Another limitation is that anomaly detection does not automatically identify the exact mechanical cause of an anomaly. Additional diagnostic models may be required to determine the specific faulty component

18. Future Scope

The proposed research can be extended in several directions.

Future work can investigate Transformer-based architectures specifically designed for long-term time series analysis. Multiscale attention mechanisms can be explored to capture both short-term and long-term vehicle behavior.

Federated learning can also be investigated to train models using data from multiple vehicles without directly sharing raw sensor data.

The system can further be integrated with vehicle diagnostic systems to provide real-time alerts. Explainable AI techniques can be used to identify which sensors and time intervals contributed most strongly to an anomaly decision.

Another promising direction is the development of a hybrid model combining self-attention with LSTM or convolutional layers. Such a model may capture both local temporal patterns and long-range dependencies.

Finally, real-world testing across different vehicle models, driving conditions, and environmental conditions would help evaluate the generalization capability of the proposed approach.

19. Conclusion

This paper presents a Self-Attention-Based Time Series Analysis framework for anomaly detection in vehicle sensor data. The proposed approach processes multivariate sensor sequences and uses self-attention to learn temporal relationships between observations.

Unlike conventional threshold-based techniques, the proposed framework considers the broader temporal context of vehicle sensor measurements. By learning normal operating patterns and calculating deviations from those patterns, the system can identify potentially anomalous vehicle behavior.

The proposed framework consists of data preprocessing, temporal window generation, self-attention-based feature learning, reconstruction or prediction, anomaly scoring, and classification. Its performance can be evaluated using accuracy, precision, recall, F1-score, and ROC-AUC.

The approach has potential applications in intelligent vehicle health monitoring, predictive maintenance, fleet management, and automotive fault detection. However, actual effectiveness must be validated using a representative vehicle sensor dataset and experimentally measured performance.

Future research can focus on larger datasets, advanced Transformer architectures, explainable anomaly detection, real-time deployment, and integration with automotive diagnostic systems.

20. References

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I., "Attention Is All You Need," Advances in Neural Information Processing Systems, vol. 30, 2017.

2. Hochreiter, S., and Schmidhuber, J., "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

3. Cho, K., Van Merriënboer, B., Gulcehre, C., et al., "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," Proceedings of EMNLP, pp. 1724–1734, 2014.

4. Chandola, V., Banerjee, A., and Kumar, V., "Anomaly Detection: A Survey," ACM Computing Surveys, vol. 41, no. 3, 2009.

5. Ruff, L., Vandermeulen, R., Görnitz, N., et al., "Deep One-Class Classification," Proceedings of the 35th International Conference on Machine Learning, pp. 4393–4402, 2018.

6. Malhotra, P., Vig, L., Shroff, G., and Agarwal, P., "Long Short Term Memory Networks for Anomaly Detection in Time Series," Proceedings of the European Symposium on artificial Neural Networks, 2015.

7. Kingma, D. P., and Ba, J., “Adam: A Method for Stochastic Optimization,” International Conference on Learning Representations, 2015.

8. Goodfellow, I., Bengio, Y., and Courville, A., Deep Learning, MIT Press, 2016.

9. Géron, A., Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly Media.

10. Breunig, M. M., Kriegel, H. P., Ng, R. T., and Sander, J., “LOF: Identifying Density-Based Local Outliers,” ACM SIGMOD International Conference on Management of Data, pp. 93–104, 2000.