Contribution ID: **19** Type: **Poster Presentation**

Deep Learning-Based Log Anomaly Detection

Log data, as the information that records the system running status, is an important part of the system. Anomalies occurring during the system running often need to be searched and rectified with the help of logs. With the increasing scale of large-scale scientific facilities and scientific data centers, the log data has exploded, and the difficulty of log anomaly detection has reached an unprecedented height. Therefore, in order to maintain the stability of increasingly complex systems, it is very important to check log anomalies in time and ensure the safe operation of the system for large scientific facilities and scientific data centers.

Manual based log anomaly detection requires a lot of labor costs. Traditional machine learning log anomaly detection has the characteristics of low hardware dependence and good interpretability, but its ability to extract advanced features or global features is relatively limited, and the effect is not good in practical application scenarios. With the rapid development and gradual maturity of deep learning technology, the application of deep learning technology to log anomaly detection has become a mainstream method in this field. Deep learning uses neural networks to train data that can automatically extract complex features from raw log data, allowing for more accurate identification of abnormal data. However, in the actual environment, the probability of abnormal logs occurring in the system is low, resulting in unbalanced proportion of normal log samples and abnormal log samples, serious imbalance in the training set, and overfitting of the model.

To address these challenges, this paper proposes an improved log anomaly detection method based on Transformer, which incorporates a self-attention mechanism to effectively capture context information in log data and improve model identification. The model processes input data from all locations simultaneously, making parallel computation easier. In addition, the Transformer model is relatively easy to scale to adapt to different log exception detection scenarios.

We validated the proposed method on multiple log datasets. Experimental results indicate that, compared to other log anomaly detection methods, the improved log anomaly detection method based on Transformer can more effectively identify abnormal data.

Primary authors: Ms LIU, Yuanyuan (Computing Center, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, P.R.China; School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, P.R.China); WANG, Jiarong (Institute of High Energy Physics); YAN, Tian (IHEP); QI, Fazhi (Institute of High Energy Physics,CAS)

Presenter: Ms LIU, Yuanyuan (Computing Center, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, P.R.China; School of Nuclear Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, P.R.China)

Track Classification: Track 7: Network, Security, Infrastructure & Operations