Oil Well Anomaly Detection using Graph Neural Network

Yuting Zhang and Shumeng Zhang

TBSI, Tsinghua University, China

ABSTRACT

- Apply Graph Neural Network-based anomaly detection to multivariate time-series sensor data.
- Pinpoint sensors that cause anomalies.
- Visualize sensor embedding space and sensor relationships as graph edges.

INTRODUCTION

- Oil well operations are extremely costly and abnormal well events can cause production losses and severe accidents.
- **Detection of abnormal events** and **identification of root causes** are essential for engineers to take appropriate actions.

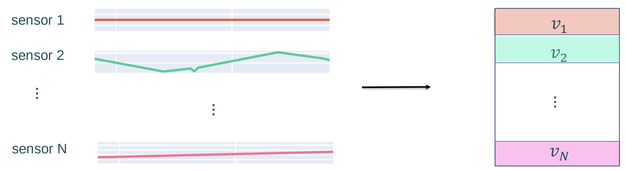
REFERENCES

[1] Deng, Ailin & Hooi, Bryan. (2021). Graph Neural Network-Based Anomaly Detection in Multivariate Time Series.

[2] Vargas, Ricardo; Munaro, Celso; Ciarelli, Patrick; Medeiros, André; Amaral, Bruno; Barrionuevo, Daniel; Araújo, Jean; Ribeiro, Jorge; Magalhães, Lucas (2019), "Data for: A Realistic and Public Dataset with Rare Undesirable Real Events in Oil Wells", Mendeley Data, v1.

METHODOLOGY

1. Sensor Embedding:
captures the characteristics of each sensor.

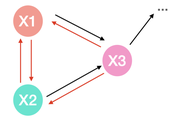


Sensor Data

Sensor Embedding, $v_i \in R^d$

2. **Graph Structure Learning**:
learns relationships between sensors as a directed graphs.

Similarity between embeddings $e_{ji} = \frac{v_i^T v_j}{\| v_i \| \cdot \| v_j \|}$



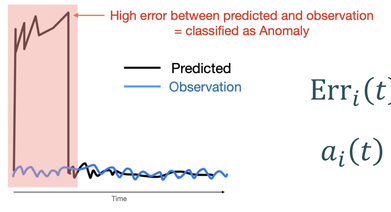
Adjacency matrix $A_{ji} = \mathbb{I}\{j \in \mathrm{TopKEdges}(\{e_{ki}: k \neq i\})\}$

3. **Graph Attention-Based Forecasting**:
predicts values of each sensor based on an attention function over its neighbors.

Attention-Based features $z_i^{(t)} = \mathrm{ReLU}\left(\alpha_{i,i} W x_i^{(t)} + \sum_{j \in N(i)} \alpha_{i,j} W x_j^{(t)}\right)$

Model Forecast $\hat{s}^{(t)} = f_\theta([v_1 \odot z_1^{(t)}, \ldots, v_N \odot z_N^{(t)}])$

4. **Anomaly Detection from the Normalized Error**:

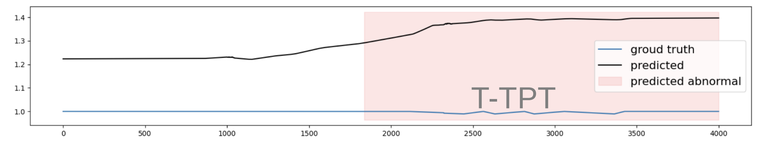


$\mathrm{Err}_i(t) = |s_i^{(t)} - \hat{s}_i^{(t)}|$

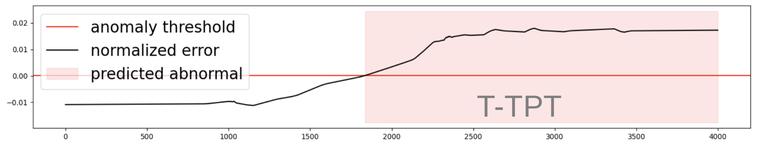
$a_i(t) = \frac{\mathrm{Err}_i(t) - \tilde{\mu}}{\tilde{\sigma}_i}$

RESULT

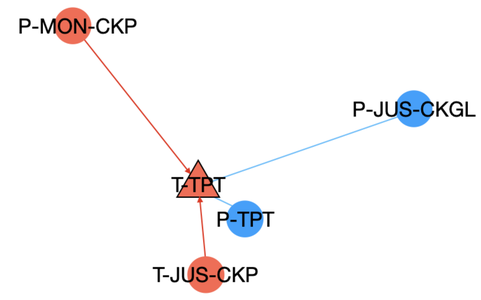
Predicted Vs. Ground Truth, Anomaly Detection in highlighted



Normalized Error and Anomaly Threshold



PCA Plot of Sensor Embeddings



Attention Coefficients to **T-TPT** (k = 3)

T-TPT	0.159
P-MON-CKP	0.043
T-JUS-CKP	0.131

F_1	Precision	Recall
0.73	0.60	0.92

CONCLUSION

- GNN is applied to learn a graph relationships between sensors, predict sensor values and detect anomaly from large deviations from predicted and actual values.
- The result is interpretable: users can understand sources of anomalies and relationships between sensors.