Adv Syst Sci Appl 2023; **03**:27–35
Published online at https://ijassa.ipu.ru.

Deep-Learning-Based Tracing for Satellite Telemetry

Ilya Nachevsky[1], Olga Andrianova[1*], Isaac Chairez[2]

[1] *V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia*
[2] *Tecnologico de Monterrey, Institute of Advanced Materials for Sustainable Manufacturing, Guadalajara, Jalisco, Mexico*

Abstract: The mechanical section of the telemetry data from the scientific satellite "Lomonosov", (such as yaw, pitch, and roll angles of the spacecraft's main axes, along with its programmed and measured velocities) is pre-processed (alignment, reflection, and binarization) and used for anomaly behavior propagation. The main goal of this study is to estimate possible abnormal behavior of the system in the future and to help to restore normal behavior during a limited communication session with a spacecraft. The system model uses a recurrent architecture approach, namely tracing methodology, considering time shifts in the target data sequence. A deep learning strategy is used to model the abnormal behavior using the onboard collected mechanical information as inputs. The results are compared with the onboard anomaly detection system (ARO) data. The reproduction of the obtained information shows better performance compared to traditional estimation techniques, using binary cross-entropy and receiver operating characteristic curve (ROCAUC) as comparison criterion. Future model modifications, which can improve its quality, are discussed at end of the study.

Keywords: trajectory tracing, anomaly propagation, telemetry analysis, recurrent model, time-distributed

1. INTRODUCTION

Telemetry data from the scientific satellite "Lomonosov" could have anomaly behavior, which must be caught during a limited communication session with a spacecraft or propagated based on the previous values of the estimated data, i.e., the preliminary session. Space companies have onboard systems for catching anomalies in data. Still, unfortunately, they are far from ideal and often provide information about a critical anomaly in data after the fact, so it is crucial to develop anomaly propagation systems based on the current data, considering the dynamic capture of the satellite motion.

The problem of restoring dynamic dependencies in the various data types is significant and widely studied in artificial intelligence and data analysis. For example, a recent paper (Tuturov, Andrianova, Sleptsov, Yurkevich, & Kryukova, 2022) has already described an approach to identifying and predicting anomalies, based on the previous experience of the system, using differential neural networks (Poznyak, Oria, & Poznyak, 2019), whose workability is strictly proven by stability theory methods. There also exists a scope of techniques to detect anomalies (Zeng, Jin, Xu, Chen, & Zhang, 2022; Luo & Nagarajan, 2018), that aims to get deviant behavior from a system dynamics. The first approach uses parametric causality, and Double-Criteria Drift Streaming Peaks Over Threshold (DCDSPOT) (Chen, Lu, Fang, et al., 2014) in a space satellite system. In contrast, the second approach employs an autoencoder architecture in a Wireless Sensor Networks (WSN) (DDos

* Corresponding author: andrianovaog@gmail.com

attacs) (Nishio et al., 2022). Another approach uses a CNN-based technique to detect the anomalies of Zero-Day Attacks (Hairab, Elsayed, Jurcut, & Azer, 2022).

This paper aims to propagate deviant or abnormal behavior, which means that the device operates under disturbed conditions or has a defect in the control or mechanical systems. The cases of abnormal system functioning can be described as sophisticated linked moments that can be detected in telemetry data, and the nature of which can be tracked in some patterns of event sequence. In contrast to the papers mentioned above, the main goal is to detect anomalies in the data. This paper intends to propagate target anomaly events, i.e., to estimate the eventual presence of anomaly behavior for future moments.

The main contributions of this study are:

1. Tracing methodology is implemented to propagate anomaly behavior in telemetry data;
2. The developed architecture considers time shifts in target data sequence in the same manner as the onboard anomaly detection system;
3. Propagation results of the new technique show better characteristics than the onboard system.

The rest of the paper is organized as follows. In Section 2, telemetry data is described, and the main problem is stated. Section 3 deals with the preprocessing of the measured data. In Section 4, model design, including model structure and training and validation processes, is presented. Simulation results are available in Section 5. Section 6 contains some concluding remarks and ideas for future work.

2. TELEMETRY DATA DESCRIPTION AND PROBLEM STATEMENT

2.1. System under consideration

The system under study is given by telemetry data from the scientific satellite "Lomonosov". From a variety of different technical data. The Spacecraft Orientation and Stabilization (SOS) system has been selected because the functional stability (Korolev, 2018) of the entire spacecraft directly depends on the operation of this system. It can be helpful for anomalies and emergency detection, localization, and further prediction after designing a mathematical model of the considered processes.

The onboard software's existing Automatic RecOnfiguration (ARO) mechanism can be considered a baseline for anomaly detection. To imagine the type of data ARO mechanism provides (see Fig. 2.1), This figure shows three triggers of the anomaly detection system in 160 seconds. In ARO indicator, weight values, expertly assigned to each functional element of all spacecraft systems, are summed up. When the value of the ARO indicator leads to faults in the operation of the spacecraft elements and it falls below the expert-defined value, then the procedure for changing the orientation mode of the spacecraft is automatically started. This procedure aims at stabilizing the spacecraft operation.

The following mechanical data, provided by SOS system, are chosen to form the state vector x for further processing and model design:

$$x = \begin{bmatrix} \alpha \\ \upsilon \\ \vartheta \end{bmatrix}$$

where $\alpha \in \mathbb{R}^3$ stands for the main spacecraft axes angles (yaw, pitch, and roll); $\upsilon \in \mathbb{R}^4$ and $\vartheta \in \mathbb{R}^4$ are the vectors of the measured and programmed angular velocities of the four wheels, respectively. For example, Fig. 2.2 depicts the dynamics of the roll angle within 160 seconds, while Fig. 2.3 shows the programmed angular velocity of the first wheel.

Some problems happened due to the specifics of data reading from sensors, processing, and writing in log files. The troubles contain time shifts and data absents. For example, in this telemetry report, the most extended shift is approximately 1.875 seconds (15 system beat). One system beat is 0.125 seconds.

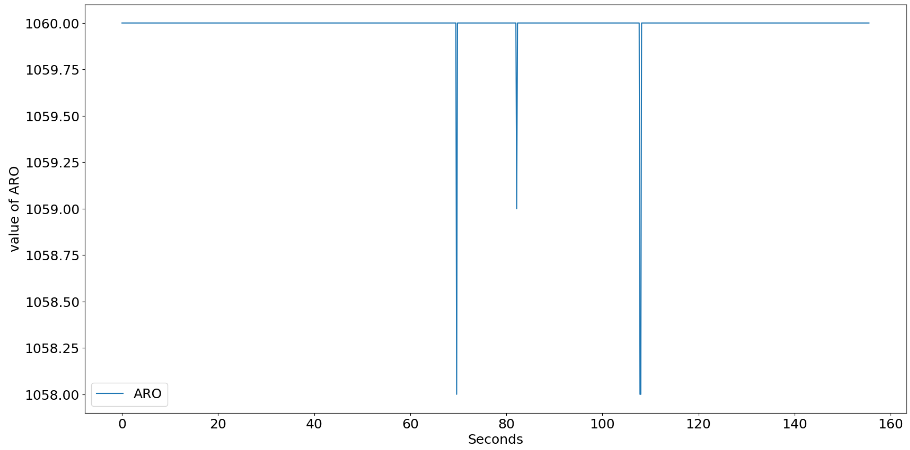

Fig. 2.1. ARO data

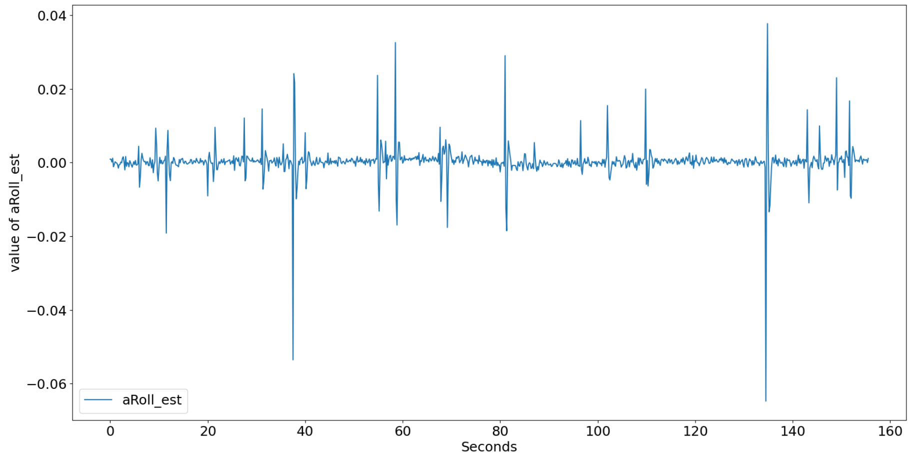

Fig. 2.2. Roll angle

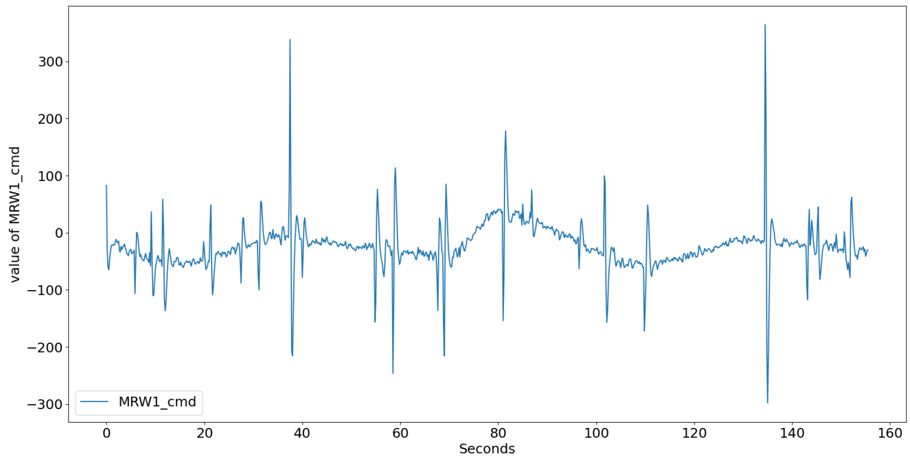

Fig. 2.3. Programmed angular velocity of the first wheel

2.2. Problem statement

The main problem is to propagate anomalies in telemetry data and make anomaly detection better than the existing ARO response. In the presented system, the state vector has nonlinear nature. So, the tracer architecture, which can catch sophisticated connections between vector elements and time dependence due to its recurrent nature (Piech et al., 2015), can be successfully used for model design.

Another problem that needs to be solved in this study is to convert raw telemetry data to a binarized form (where “1” stands for an anomalous value of the element, and “0” indicates the absence of the anomaly at the current time) of each state vector element using statistical methods, such as LOWESS aligner, anomaly detector in the form of quantile barrier (Zeng et al., 2022).

3. PREPROCESSING

In statistics, the term LOWESS refers to “LOcally WEighted Scatterplot Smoothing”. As one of the standard approaches to produce a smooth curve that fits the data points in a scatterplot, this method is applied to the raw telemetry data. The results of its operation for the programmed angular velocity of the first wheel are shown in Fig. 3.4.

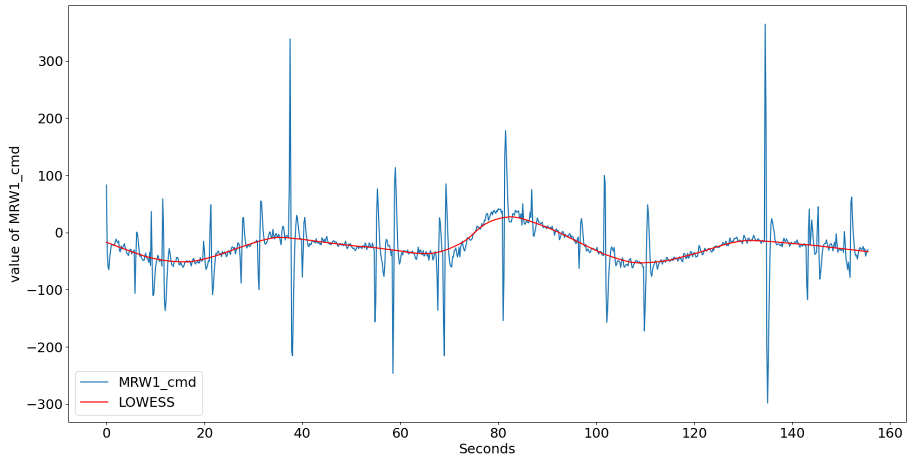

Fig. 3.4. LOWESS algorithm applied to the programmed angular velocity of the first wheel

Further, negative values are reflected in the positive area, as the module of LOWESS stands for the barrier of the quantile barrier function

$$qBar\left(|LOWESS(x)|\right)$$

where

$$qBar(z) = \begin{cases} 1 & \text{if } Q_x(0.95) < z, \\ 0 & \text{if } Q_x(0.95) \geq z, \end{cases}$$

and Q_x is the quantile by sampling the state vector.

Fig. 3.5 demonstrates data processing results, namely alignment, reflection, and binarization for the programmed angular velocity of the first wheel data. The binarized plot has 100 values by ordinate axis for the reasons of clarity. The binarized data for each element of the state vector is recorded with the sequence of labels from 0 to 10 (the spacecraft main axes angles, measured and programmed angular velocities of the wheels). In this form, the dataset is used for further network training.

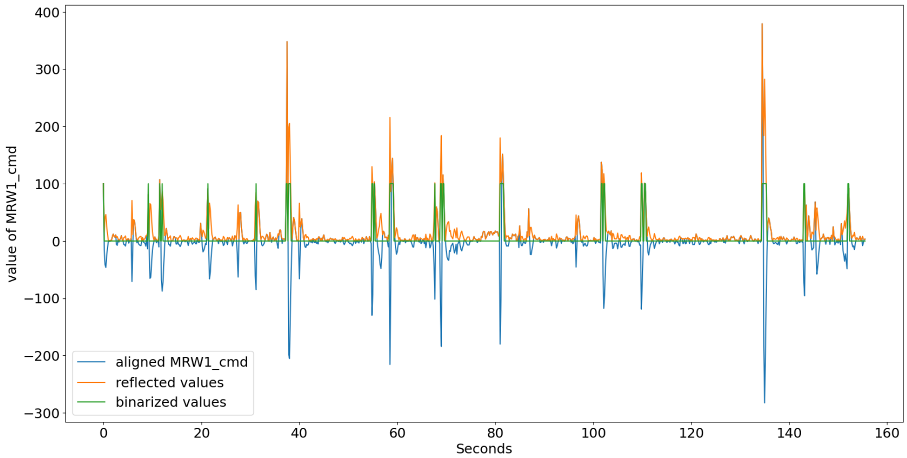

Fig. 3.5. LOWESS algorithm applied to the programmed angular velocity of the first wheel

4. MODEL DESIGN

The recurrent architecture approach has already been used for propagation sequence behavior (Piech et al., 2015). In this work, a similar methodology, namely tracing, is applied. The sequence of events has information about sophisticated linked moments, which are causes of abnormal behavior.

4.1. Model structure

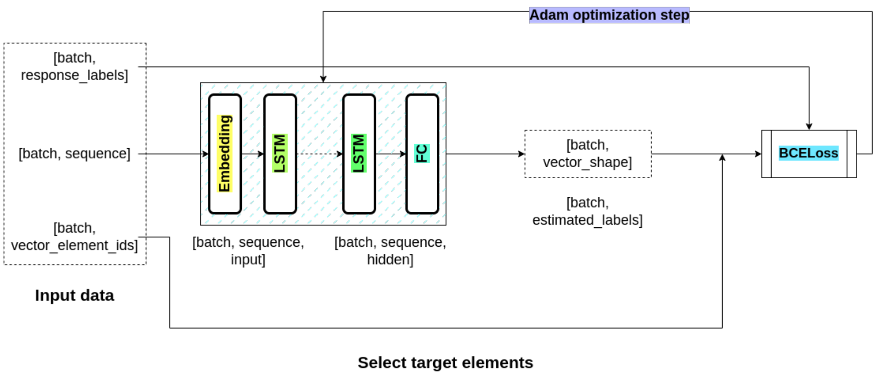

Fig. 4.6. Model architecture

Fig. 4.6 depicts the first part of information coded on the embedding layer (Khrulkov, Hrinchuk, Mirvakhabova, & Oseledets, 2019). The approach helps to represent data in the new form and catch knowledge of relation from parts of the sequence. The second one is the recurrent neural network (RNN)-based layer considering a Long-Short Term Memory (LSTM) implementation. The layer can remember information and gather it through all query items. This approach helps to get specific dependencies between events or even a chain of events. After getting a response from the layer, we need to make a solution about the nature of the sequence by compressing the sparse answer space. The third part is a fully connected layer, which compresses the LSTM response and helps predict the sequence's nature by items. By the way, we make a propagation about the next step of the sequence with analysis from different sources (telemetry from each sensor). The learning strategy is based on the optimization by Adam method (Kingma & Ba, 2017) through back-propagation through-time mechanism (Rumelhart, Hinton, & Williams, 1986).

4.2. *Training and validation process*

The above-described architecture is trained with satellite telemetry data. Fig. 4.7 and Fig. 4.8 demonstrate training and validation processes. ROCAUC is used as a metric such as in related work (Piech et al., 2015), which gives clearer quality in case of unbalanced classes. Binary cross-entropy (BCE loss function) is chosen as a loss function, which shows a dissimilarity measure between two distributions and is based on Kullback-Leibler Divergence (*Kullback-Leibler Divergence*, n.d.). Thus, minimizing the loss function means that the loss value between the target labels and the estimated labels becomes smaller:

$$\begin{gathered} J(x^*, \hat{x}^*) = \{j_1, \ldots, j_N\}^\top, \\ j_k = -\left[\hat{x}_k^* \cdot \log x_k^* + (1 - \hat{x}_k^*) \cdot \log(1 - x_k^*)\right], \quad k \in \{1, \ldots, N\}, \\ J(x^*, \hat{x}^*) \to min \quad \sim \quad ROCAUC(x^*, \hat{x}^*) \to max \end{gathered}$$

where J stands for the loss function, x^* is a target value, and $\hat{x}^*$ is an estimated value.

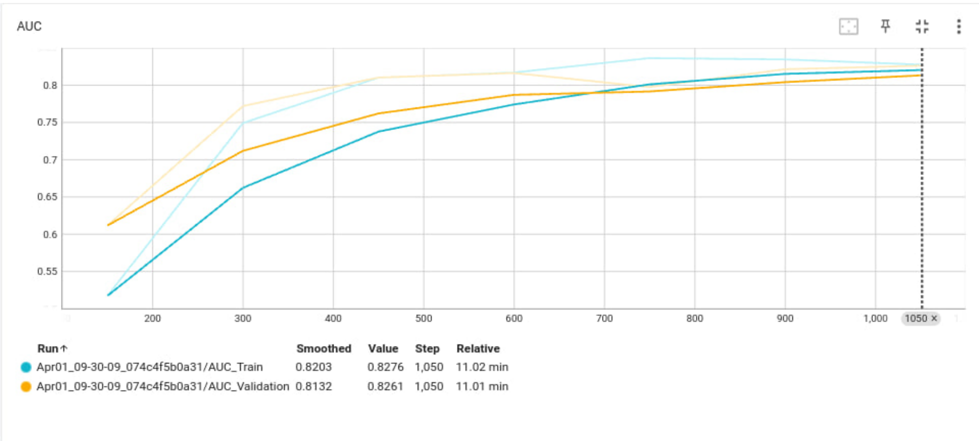

Fig. 4.7. ROCAUC train/test curves

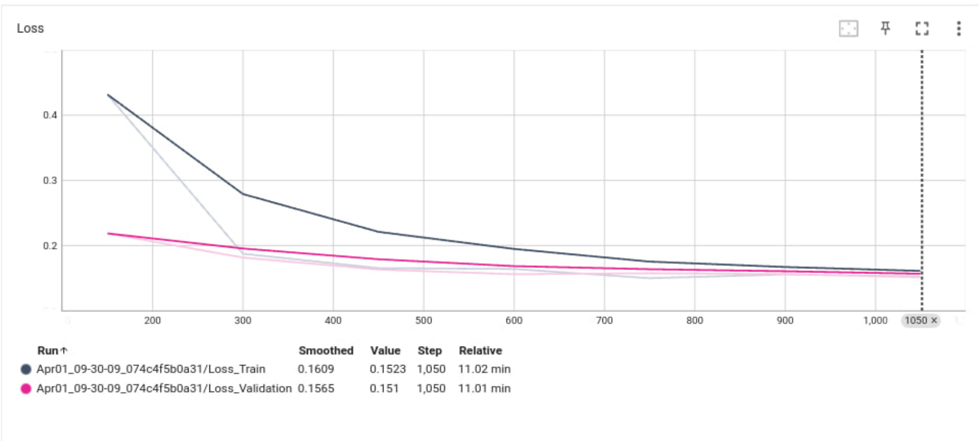

Fig. 4.8. BCE loss train/test curves

5. SIMULATION RESULTS

Fig. 5.9 demonstrates results (smoothed values) of tracing, reflecting the potential appearance of anomaly events. Each time point has information about the previous sequence, and the new time point has a prediction for each info source (sensors). The approach is robust concerning time shifts and information disappearing because the data batches have been shuffled in a training period, and the neural network can predict such events based on only one element in sequence (dictionary of known keys - axes, measured and programmed wheel velocities).

The quality of the onboard ARO system is 54% by AUC metrics, while the presented neural network in a pre-trained condition has a quality of about 82% by AUC metrics. This value of above the results presented before considering the implementation of traditional regressors and static neural networks. Intending to simplify access to the proposed model, the authors of this manuscript prepared a digital repository where the software used to develop the presented results is included. These programs can be found in github repository.

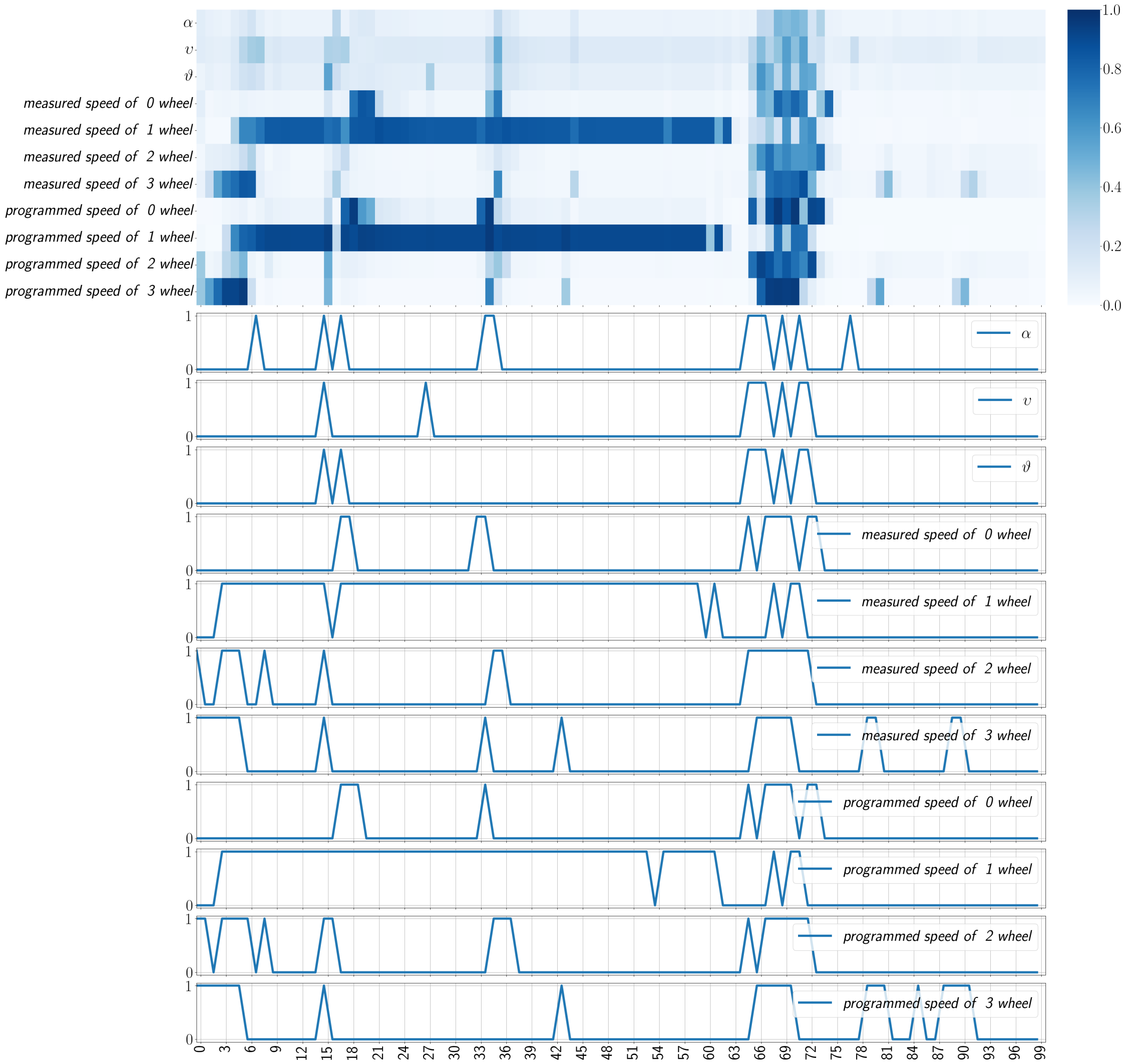

Fig. 5.9. Tracing result with comparison to strict data

6. DISCUSSION

This study describes the propagation of anomalies in satellite telemetry data. The suggested and tested method to improve deviant system behavior propagation can be further used for onboard implementation due to its lightweight architecture. The system can be fine-tuned using new information. In this case, the system with the tracing neural network core is able to predict with actual information about the target satellite device. It could be implemented by fine-tuning the control center and pushing new weights to the satellite board (the model weight is about 700Kb).

Another approach - Time-distributed CRNN (Luna-Álvarez, Mújica-Vargas, Matuz-Cruz, Vianney Kinani, & Ramos-Díaz, 2020) - can also be applied to make a solution to this problem. The idea regarding applying convolution layers to sum the information can get a better result than the current architecture approach (tracing with the embedding layer). However, in this way, a change of the 2d convolutions to 1d is suggested, or the information should be represented in a bit new form to feed the data in a recurrent layer.

The quality of the problem solution can also be improved when transformer architecture is used (Vaswani et al., 2017). The transformer architecture can represent the latent space more complexly in comparison to the approach from this paper, which, consequently, will lead to the ability to separate distributions and to make the high dissimilarity between them. It's worth mentioning that the loss function can be picked the same, while the data representation is needed to be changed in a straight way of queries, keys, and values matrices.

ACKNOWLEDGEMENTS

The paper was prepared with the partial financial support of the Tecnologico de Monterrey, Institute of Advanced Materials for Sustainable Manufacturing under the grant Challenge-Based Research Funding Program 2022 number I006-IAMSM004-C4-T2-T.

REFERENCES

Chen, B., Lu, G., Fang, H. Z., et al. (2014). Method of satellite anomaly detection based on least squares support vector machine [j]. *Computer Measurement & Control*, *2022*(3), 690–696.

Hairab, B. I., Elsayed, M. S., Jurcut, A. D., & Azer, M. A. (2022). Anomaly detection based on cnn and regularization techniques against zero-day attacks in iot networks. *IEEE Access*, *10*, 98427–98440.

Khrulkov, V., Hrinchuk, O., Mirvakhabova, L., & Oseledets, I. V. (2019). Tensorized embedding layers for efficient model compression. *CoRR*, *abs/1901.10787*. Retrieved from http://arxiv.org/abs/1901.10787

Kingma, D. P., & Ba, J. (2017). *Adam: A method for stochastic optimization.*

Korolev, A. (2018). Functional stability of navigation information systems. *Izvestiâ vysših učebnyh zavedenij. Priborostroenie*, *61*, 559-565.

Kullback-leibler divergence. (n.d.). http://hanj.cs.illinois.edu/cs412/bk3/KL-divergence.pdf. (Accessed: 2023-04-02)

Luna-Álvarez, A., Mújica-Vargas, D., Matuz-Cruz, M., Vianney Kinani, J. M., & Ramos-Díaz, E. (2020). Self-driving through a time-distributed convolutional recurrent neural network. In *2020 17th international conference on electrical engineering, computing science and automatic control (cce)* (p. 1-6). doi: 10.1109/CCE50788.2020.9299181

Luo, T., & Nagarajan, S. G. (2018). Distributed anomaly detection using autoencoder neural networks in wsn for iot. *IEEE ICC*. doi: arXiv:1506.05908v1

Nishio, T., Nakahara, M., Okui, N., Kubota, A., Kobayashi, Y., Sugiyama, K., & Shinkuma, R. (2022). Anomaly traffic detection with federated learning toward network-based malware detection in iot. In *Globecom 2022-2022 ieee global communications conference* (pp. 299–304).

Piech, C., Spencer, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L., & Sohl-Dickstein, J. (2015, 06). Deep knowledge tracing.

Poznyak, A. S., Oria, I. C., & Poznyak, T. (2019). A survey on artificial neural networks application for identification and control in environmental engineering: Biological and chemical systems with uncertain models. *Annu. Rev. Control.*, *48*, 250-272.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*, 533-536.

Tuturov, A. A., Andrianova, O. G., Sleptsov, E. S., Yurkevich, E. V., & Kryukova, L. N. (2022). Spacecraft emergency situations automated scenarios design within the framework of fdir concept and dynamical neural networks technique. *Advances in Systems Science and Applications*, *4*, 19-30.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention is all you need. *CoRR*, *abs/1706.03762*. Retrieved from http://arxiv.org/abs/1706.03762

Zeng, Z., Jin, G., Xu, C., Chen, S., & Zhang, L. (2022). Spacecraft telemetry anomaly detection based on parametric causality and double-criteria drift streaming peaks over threshold. *Applied Sciences*, *12*(4). Retrieved from https://www.mdpi.com/2076-3417/12/4/1803 doi: 10.3390/app12041803

