Validity Evaluation of Anomaly Detection Using LSTM AutoEncoder for Maneuver Detection

Ryo Kato, Takahito Sakaue, Daiki Mori, Makoto Tanaka
NEC Aerospace Systems, Ltd.
Taichiro Sano, Katsuaki Taya, Masatoshi Ebara, Ryo Nakazawa
Jun Yoshida, Ryosuke Togawa
NEC Corporation

ABSTRACT

Recently, space objects such as GPS satellites, meteorological satellites and telecommunication satellites are one of indispensable infrastructure for many countries. The number of space objects has increased significantly in recent years due to a variety of factors, so there is under threat from emerging risks such as the collision with increasing space debris and threat caused by anti-satellite (ASAT) weapon to safe and stable space environment.
For these risks mitigation, leveraging “safety of flight” capability with the Space Situational Awareness (SSA) system is effective. As a demand, NEC started to develop the unique commercial SSA system prototype in Japan, which we call the NEC Commercial Space Operation Center (NEC ComSpOC). We have collaborated with COMSPOC corporation and have experimentally started SSA operation in-Japan for-Japan. From the operations, we have been able to possess the various SSA information of some space objects such as high accuracy ephemeris, the trend of maneuver and the results of conjunction assessment (CA) uniquely. Fig. 2 shows an example of our system workflow. First, the system gets tracks which have yet to be associated and get relevant ephemeris predictions from database. It is nominally generated by prior orbit determination (OD) run and supplied from Owner/Operator data. Second, the system conducts OD and maneuver processing (MP) to generate pre-fit residuals for each track. Then, the system tags and updates the satellite catalog with those results. Finally, the system assesses the conjunction. However, the system sometimes detects a false maneuver. To improve our system operability, it is necessary to confirm the detection of the maneuvers using a different method is needed to solve these operational problems to improve the capability of Space Situational Awareness (SSA) and Space Domain Awareness (SDA). In this study, we applied the Anomaly detection model based on LSTM AutoEncoder for detecting maneuvers from Two Line Elements (TLEs) of the notable objects of Geostationary Orbit (GEO) to NEC ComSpOC operations and evaluated the validity of the model to apply.
In this study, our model trained with normal values and calculated the anomaly score in the TLEs of GEO objects which contains with high frequency maneuvers, especially those with significant differences from station keeping maneuver, whereas with scarcely maneuvers. As a results, it did partially substantiate the effectiveness of the model NEC developed for GEO objects and the feasibility of maneuver detection of notable objects based on TLEs. Furthermore, our model was partially successful in detecting the maneuver which NEC ComSpOC could not detect. These findings could be used to help SSA/SDA operation. Further research could also be conducted to determine the effectiveness and feasibility of Low Earth Orbit (LEO) objects.

1. INTORODUCTION AND BACKGROUND

In our daily life, space objects such as GPS satellites, meteorological satellites and telecommunication satellites are one of indispensable infrastructure. The number of space objects has increased significantly in recent years due to a variety of factors, so there is under threat from emerging risks such as the collision with increasing space debris and threat caused by anti-satellite (ASAT) weapon to safe and stable space environment.[1]
Fig. 1 shows the monthly number of objects in earth orbit.[2] Starting with the launch of Sputnik 1 in 1957, the number of space objects has continued to grow year by year. Especially since 2006, the number of total objects has been on a significant increase compared to previous years. First of these main causes is the increase of space debris related with ASAT weapon test by China and Russia. In particular, it is said that about 4,000 pieces of debris were come apart by Chinese ASAT weapon test. Second is the satellite collision between the Russian IRIDIUM-33 and the American COSMOS-2251 in 2008. Both satellites collided unintentionally in space, resulting in a lot of debris. Another feature of the increase in space objects over the past few years has been the satellite constellations. Thus,

because of the mix of intentional and unintentional events in modern space, it is important to understand more accurately of the space situation on a regular basis to operate satellites safely.

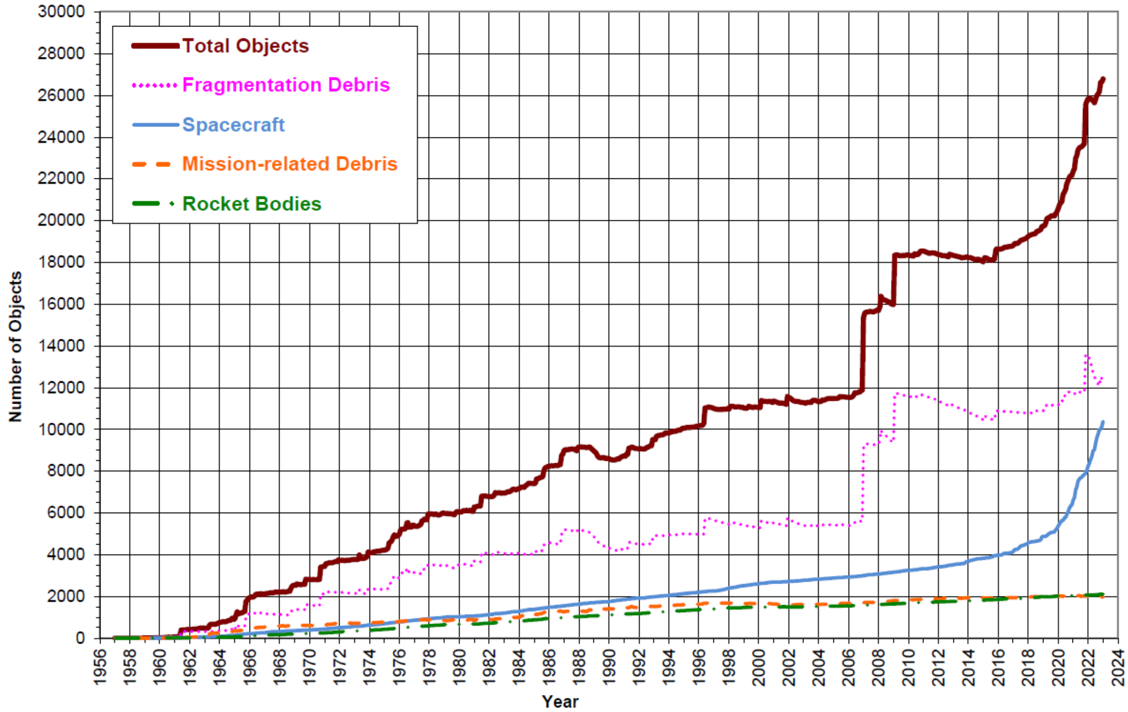

Fig. 1 Monthly number of objects in earth orbit by object type.[2]

For these risks mitigation, leveraging "safety of flight" capability with the Space Situational Awareness (SSA) system is effective. As a demand, NEC started to develop the unique commercial SSA system prototype in Japan, which we call the NEC Commercial Space Operation Center (NEC ComSpOC).[3] We have collaborated with COMSPOC corporation and have experimentally started SSA operation in-Japan for-Japan. From the operations, we have been able to possess the various SSA information of some space objects such as high accuracy ephemeris, the trend of maneuver and the results of conjunction assessment (CA) uniquely. Fig. 2 shows an example of our system workflow. First, the system gets tracks which have yet to be associated and get relevant ephemeris predictions from database. It is nominally generated by prior orbit determination (OD) run and supplied from Owner/Operator data. Second, the system conducts OD and maneuver processing (MP) to generate pre-fit residuals for each track. Then, the system tags and updates the satellite catalog with those results. Finally, the system assesses the conjunction. However, the system sometimes detects a false maneuver. To improve our system operability, it is necessary to confirm the detection of the maneuvers using a different method.

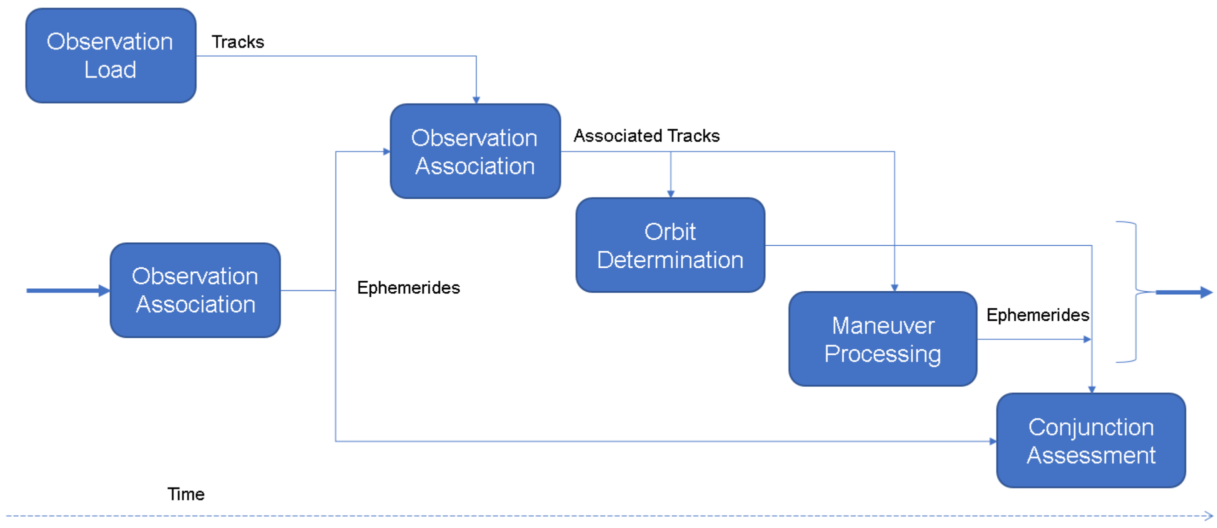

Fig. 2 Typical Catalog Processing Workflow in NEC ComSpOC

In the past few years, a more powerful Deep Neural Network (DNN) [4] has been developed based on machine learning. DNN has shown advantages in learning, time series prediction and other fields, and has achieved great success. In particular, the DNN architecture is very successful in learning feature representation of data, reducing the amount of manual design features. Recurrent Neural Networks (RNN) [5] is a type of DNN. RNN has a hidden layer state as well as fuses historical data with current data. The impact of prediction on time series problems is remarkable, e. g. natural language processing [6], speech recognition [7] etc. Although RNN has a better prediction effect on data with shorter time span. If the time span is large, the gradient reduction problem will occur. To solve this problem, a long short-term memory (LSTM) model [8] is proposed. In this paper, we use LSTM model for anomaly detection in two-line-elements (TLEs) files to maneuver detection while using AutoEncoder [9] to reduce data to improve generalization capabilities of the model.

2. PREDICTIVE MODEL

In this section, we will introduce a framework of DNN for anomaly detection. In Section 2.1, the LSTM is introduced, and reducing the data dimension through AE in Section 2.2.

2.1 LONG SHORT-TERM MEMORY

LSTM is a very effective method to solve the problem of gradient disappearance. It is mostly used for time series a type of deep learning. TLE is a type of time series which is open in public by Space-Track. So, LSTM model can be used to predict anomaly detection by TLE.
Deep learning can approximate complex functions in any form. Look for linear relationships between nonlinear data. You can dig deep into the hidden relationship between data and give full play to the potential of data. Deep learning has advantages that traditional machine learning does not have. In this part we introduce the principle of LSTM.[11]

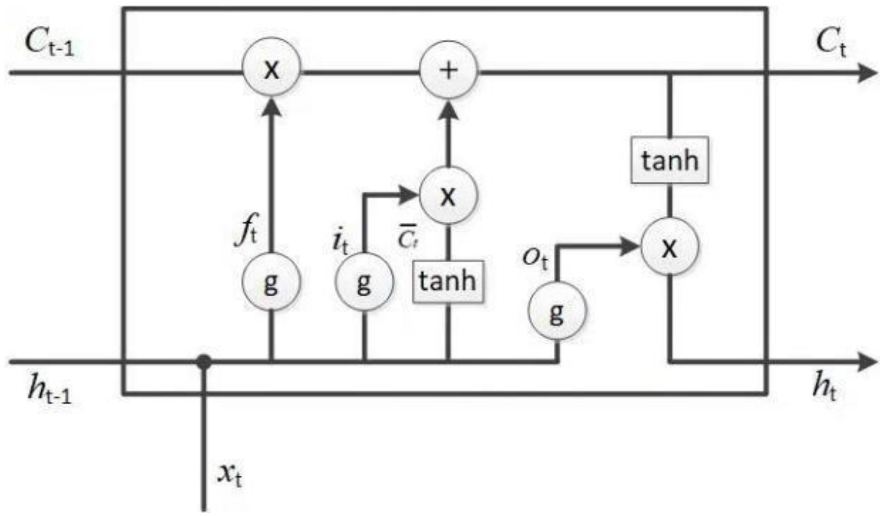

Fig. 3 The cell of LSTM[11]

LSTM avoids long-term dependencies through deliberate design. As can be seen in Fig. 3, LSTM mainly relies on the three middle gates. These three gates are the forgot gate, the input gate, and the output gate. The gate is activated with sigmoid (denoted as g), while the input and unit states are usually converted with tanh LSTM can be defined by the following set of equation:

Forget gate:

$$f_t = g\left(w_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

Input gate:

$$i_t = g(w_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

State update:

$$\overline{C_t} = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \overline{C_t} \tag{4}$$

Output gate:

$$o_t = g(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \cdot tanh(C_t) \tag{6}$$

where W_{ij} represents the weight of neurons i to j, and b represents the bias.[11]

2.2 LSTM AUTOENCODER MODEL

Among the factors in TLEs affecting the anomaly detection, some are highly correlated, and some are weak. Too few features can lead to a lack of information and a full reflection of anomaly changes and too many features can lead to information redundancy and reduced model generalization performance. [11]

Features as a raw material for machine learning systems, the impact on the final model is beyond doubt. The performance of machine learning algorithms depends to a large extent on the choice of data representation or feature representation. When the data can be well characterized as features, satisfactory accuracy can be achieved even with a simple model. Therefore, in the practical application of machine learning algorithms, an important step is how to preprocess the data to get a good feature expression.[11]

To solve these problems, use AE to reduce the dimensions of the data. AutoEncoder can learn the effective representation of input data through unsupervised learning. This method is called encoding, and its size is smaller than the input data, so that AE can be used to reduce the dimension.[11]

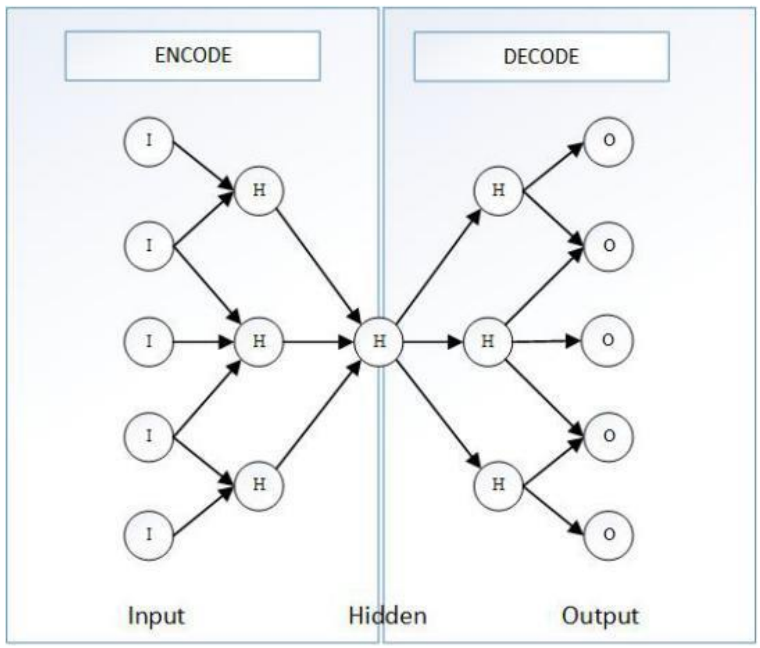

Fig. 4 AutoEncoder.[11]

In Fig. 4, the most left side is the input layer, the middle is the hidden layer, and the far right is the output layer. Suppose the input is X = (x1, x2, x3, ⋯ , xn), and xi ∈ [0,1], AE maps X to an implicit layer, denoted as H = (h1, h2, h3, ⋯ , hs), and hi ∈ [0,1], this process is called Encode.[11]

The output of the hidden layer is expressed as:

$$H = g(w_1 \cdot X + b_1) \tag{7}$$

The output H of the hidden layer is called an implicit variable, and the Z of the output layer is reconstructed by the implicit variable with the same structure as the X of the input layer. This process is called Decode.[11]

The output of the output layer is expressed as:

$$Z = g(w_2 \cdot X + b_2) \tag{8}$$

Z can be seen as a prediction of data X using feature H. Using AE to reduce the dimensionality of the data, the result is H. [11]

The loss function represents the difference between X and Y. The weight between the layers is adjusted by reducing the value of the loss function. This reduces the loss of effective information during the dimension reduction process.[11]

3. MATERIAL AND METHOD

AutoEncoder is one of the major branches of Deep Neural Network that is used to reconstruct the entered data. It is widely used as a noise removal in pictures [12,13]. The anomaly detection algorithm compares the data with the prediction model. If the difference between the data and the prediction model is more than the defined limits, the data can be said as the anomaly data [14]. AutoEncoder can also be used to detect anomalies by training models from normal data. When abnormal data is found, this model cannot be reconstructed into input data. Reconstruction that occurs will be far from the input data which makes the gap between input data and reconstruction data. After calculating the gap value, it will be determined whether the data are included in the anomaly.[12]

One of the methods for reconstructing time series data patterns is LSTM. It is a specialized type of neural network, which falls within the class of Recurrent Neural Network (RNN). Unlike the conventional feed-forward neural network architectures, RNNs employ feedback connections from their output layers back to the input layers, where each of these feedback connections can be used to serve as a time-delay gate [15,16]. Therefore, RNN architecture can represent explicitly the influence of past output values on the computation of the current output, making it ideal to model the auto correlation structure of time series or sequence data.[12]

Anomaly detection model generated by LSTM can predict anomaly score from historical data. Model generation using the LSTM method begins with the accumulation of TLEs historical data.. In this study, we used two different moving types of GEO satellites shown Table. 1. MICHIBIKI-3 is one of Japanese GPS satellite which is called Quasi-Zenith Satellite System (QZSS). It mainly conducted stational keeping maneuver in this training and test span. The tool used in this study was Jupyter notebook (basis python) and library Tensor Flow.

Table. 1 List of satellites TLE spans used for train and test.

Satellite	Training		Test	
	Start Date	End Date	Start Date	End Date
High Mobility GEO satellite	1 Jan 2021	31 Dec 2021	1 Jan 2022	20 Dec 2022
MICHIBIKI-3	1 Jan 2021	31 Dec 2021	1 Jan 2022	27 Feb 2023

Table. 2 shown parameters to hypertune the LSTM AutoEncoder model. The “Window size” denotes the time span of segment. “Batch size” denotes the number of segments using in one-time train. “LSTM Encoder hidden dimension” denotes the number of nodes on LSTM Encoder. As the number of nodes increases, the system may handle more complex data, but there is also a risk of overfitting. “LSTM Decoder hidden dimension” denotes the number of nodes on LSTM Decoder.

Table. 2 Hypertuning Parameter derived as a function of the training process.

Hypertuning Parameter	Optimal Tuning Option
Window size	2
Batch size	32
LSTM Encoder hidden dimension	64
LSTM Decoder hidden dimension	64

4. RESULT AND DISCUSSION

Fig. 5 shows anomaly score of high mobility GEO satellite calculated by the result of LSTM AutoEncoder learning and test using it's TLEs. Blue lines mean anomaly score and orange lines mean maneuvers that were detected in NEC ComSpOC system. Orange bold line means maneuvers that were calculated as recovery results in case of OD failure in our system, so we take no account of them in this study. In our system, high anomaly scores were observed at locations where maneuvers were successfully detected. On the other hand, elevated anomaly scores also exist in some instances where maneuvers where not detected.

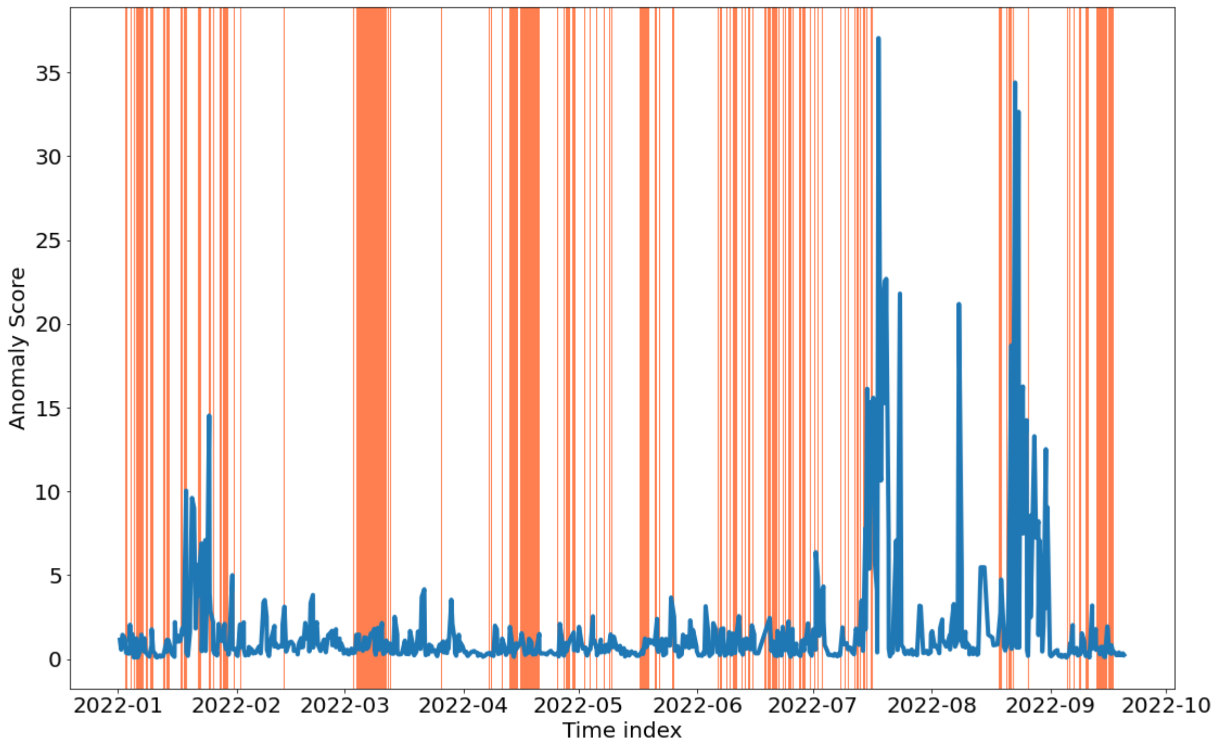

Fig. 5 Anomaly score of high mobility GEO satellite.

Fig. 6 shows longitude of high mobility GEO satellite which is calculated in NEC ComSpOC system using tracks data in the same span as Fig. 5. From this figure, high mobility GEO satellite would move from longitude 60 deg to -20 deg in July to September 2022. Compare Fig. 5 to Fig. 6, anomaly score increases when the high mobility GEO satellite starts and stops moving.

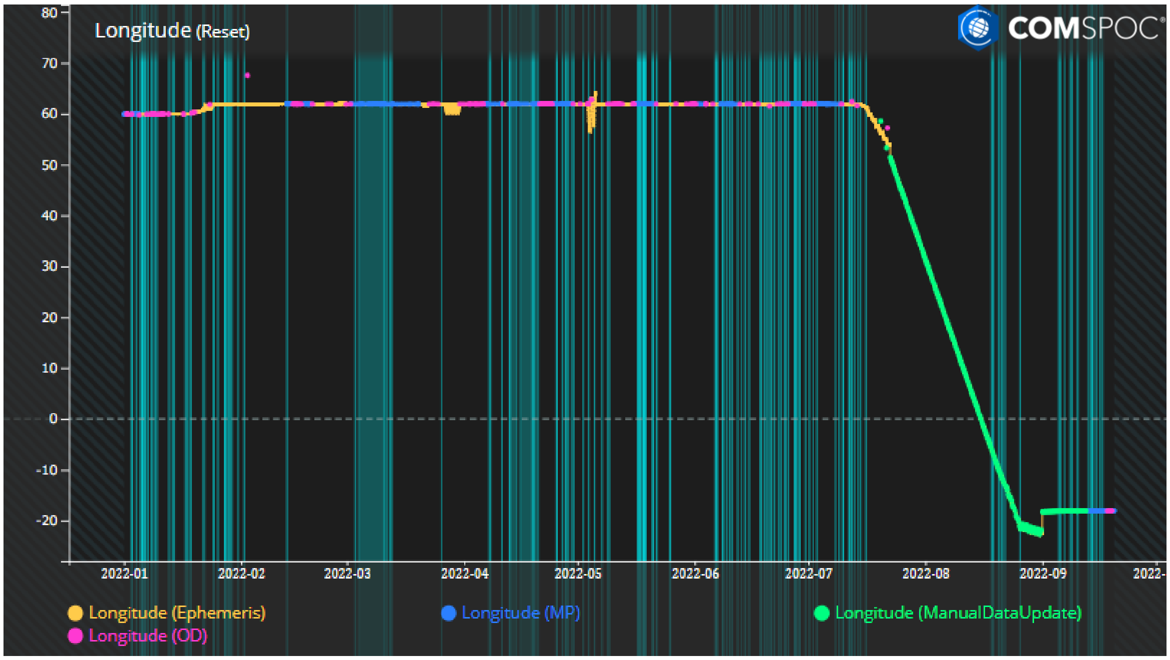

Fig. 6 Longitude history of high mobility GEO satellite.

Fig. 7 shows anomaly score of MICHIBIKI-3 calculated by the result of LSTM AutoEncoder learning and test using it's TLEs. MICHIBIKI-3 has basically conducted stational keeping maneuvers for this test span, thereby orange lines are mostly lined at regular intervals. From this figure, the anomaly score mostly increases when maneuver exists. However, it increases score when maneuver does not exist, especially April 2022.

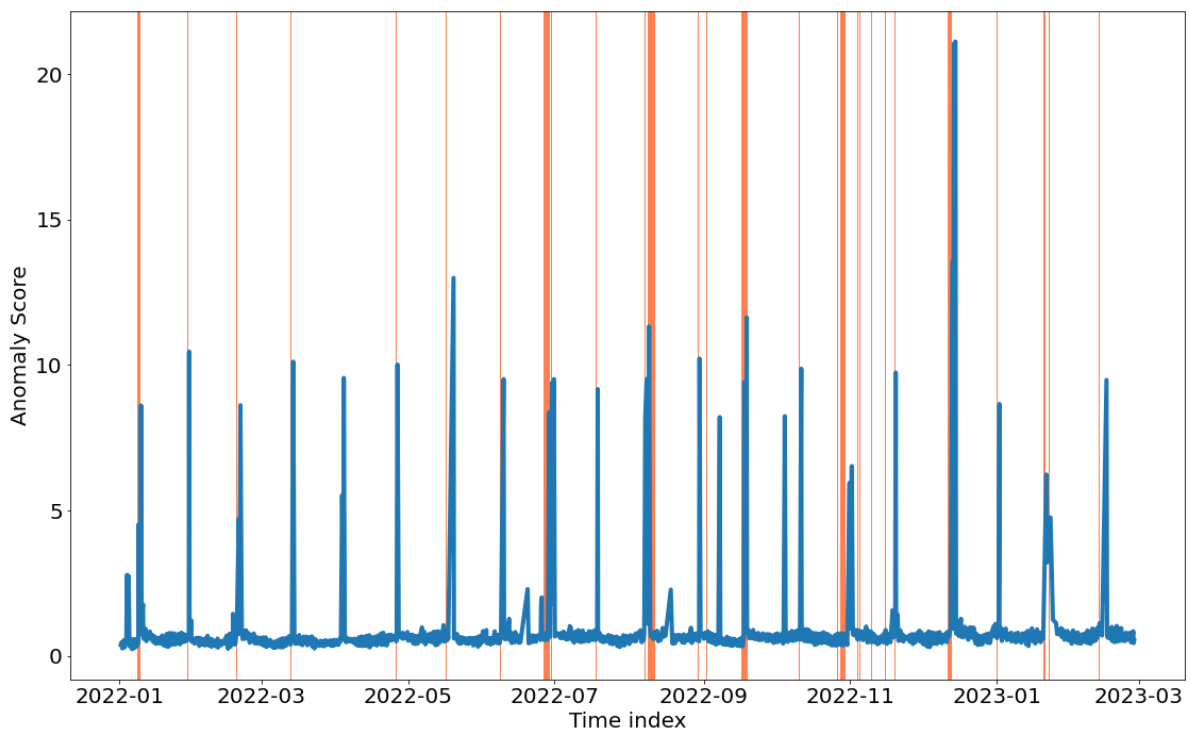

Fig. 7 Anomaly score of MICHIBIKI-3.

Fig. 8 shows longitude of MICHIBIKI-3 which is calculated in NEC ComSpOC system using tracks data in the same span as Fig. 7. From this figure, MICHIBIKI-3 has maneuvered to keep stationary GEO repeatedly for this test span, hence we have concluded that MICHIBIKI-3 maneuvered in April 2022.

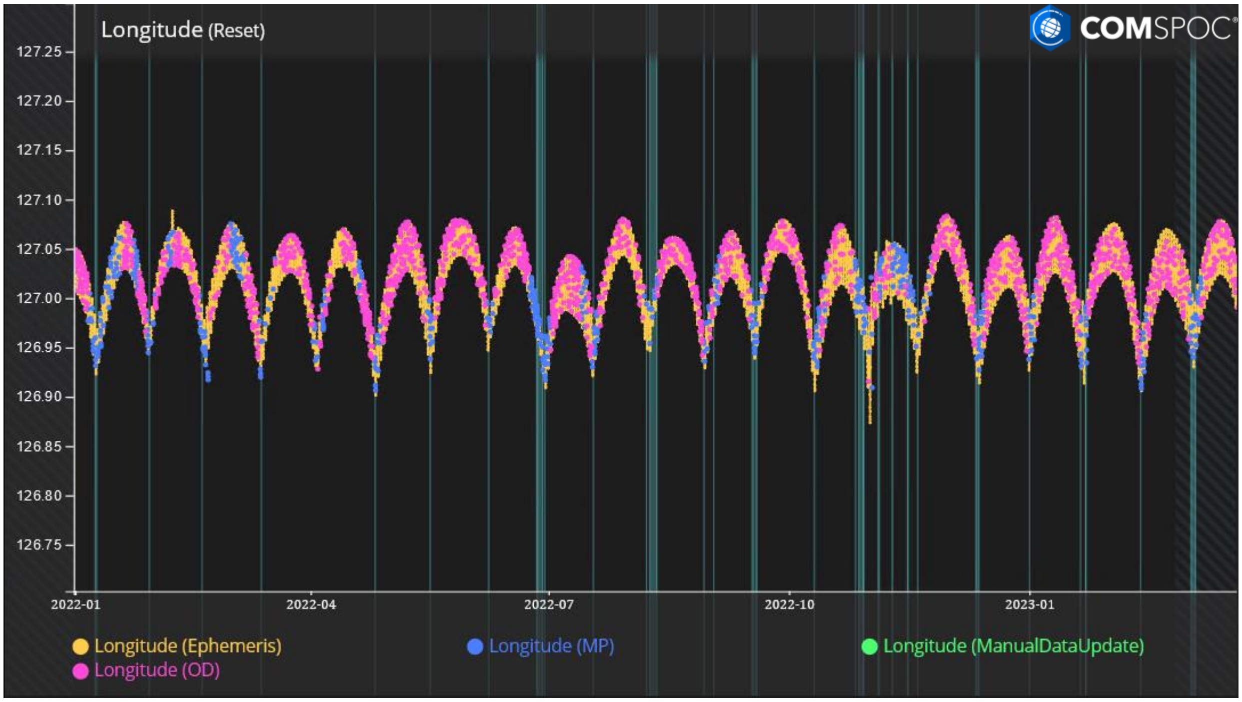

Fig. 8 Longitude history of MICHIBIKI-3.

Including MICHIMIBI-3, the services status and delta V information of QZSS is available to the public as NOTICE ADVISORY TO QZSS USERS (NAQU). From these information, MICHIBIKI-3 maneuvered in April 2022. From the learning results of LTSM AutoEncoder model, we could detect maneuver that our system did not detect.

5. SUMMARY AND FUTURE WORK

In this study, we let LSTM AutoEncoder model learn GEO satellite TLEs histories to detect time variation of anomaly score. Then, we compare these results with maneuver histories that calculated in NEC ComSpOC system. As a result, we have identified high anomaly score when the high mobility GEO satellite has started and stopped moving in the direction of longitude. Furthermore, we have detected high anomaly score when our system could not detect maneuver nevertheless it maneuvered in actual.
Meanwhile, there are some high anomaly score that we could not regard time variation of longitude that our system calculated as maneuvered. We did not set anomaly threshold to identify whether the satellite has maneuvered or not. To set anomaly threshold, hypertune learning parameter and use TLEs history longer is presumed to improve this problem. Besides, we will evaluate this model for LEO satellites in the future.

6. REFERENCES

[1] Takao Sato, Daisuke Yamashita, Daiki Mori, Touko Yamada. NEC Aerospace Systems, Ltd., “Application Development with High Definition SSA Information”, In Advanced Maui Optical and Space Surveillance Technologies Conference, September 2019.
[2] Orbital Debris Quarterly News, Volume 27, Issue 1, March 2023, NASA/JSC.
[3] Daiki Mori, Masatoshi Ebara, Takahiro Okada, Hideki Kimura, “Commercial SSA capability in Japan”, In Advanced Maui Optical and Space Surveillance Technologies Conference, September 2019.
[4] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436.

[5] Xiao, Y. S., Wang, W. Q., & Huo, X. P. (2007). Study on the time-series wind speed forecasting of the wind farm based on neural networks. Energy Conservation Technology, 2, 2.
[6] XueFeng, X., & Guo-Dong, Z. (2016). A survey on deep learning for natural language processing. Acta Automatica Sinica, 42(10), 1445-1465.
[7] Deng, L., Hinton, G., & Kingsbury, B. (2013, May). New types of deep neural network learning for speech recognition and related applications: An overview. In Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on (pp. 8599-8603). IEEE.
[8] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.
[9] André Gensler , Janosch Henze, Bernhard Sick, and Nils Raabe, "Deep learning for solar power forecasting - an approach using AutoEncoder and LSTM neural networks," 2016 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2858-2865, 2016.
[10] Herry Pariaman, G M Luciana, M K Wisyaldin, Muhammad Hisjam, "Anomaly Detection Using LSTM-Autoencoder to Predict Coal Pulverizer Condition on Coal-Fired Power Plant" Joint Journal of Novel Carbon Resource Sciences & Green Asia Strategy, Vol. 08, Issue 01, pp89-97, March 2021.
[11] Jinxia Zhang, Xuru Jiang, Xin Chen, Xiaojing Li, Dong Guo, Lixin Cui, "Wind Power Generation Prediction Based on LSTM" Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence, 85-89 (2019).
[12] X. Ye, L. Wang, H. Xing, and L. Huang, "Denoising Hybrid Noises in Image with Stacked Autoencoder", 2015 IEEE International Conference on Information and Automation, 2720-2724 (2015).
[13] H. Chiang, Y. Hsieh, S. Fu, K. Hung, Y. Tsao, and S. Chien, "Noise Reduction in ECG Signals Using Fully Convolutional Denoising Autoencoders", IEEE Access, 7 60806-60813 (2019).
[14] G.D. Nugraha, B. Sudiarto, and K. Ramli, "Machine Learning-based Energy Management System for Prosumer", Evergreen, 7 (2) 309-313 (2020).
[15] Y. Bengio, P. Simard, and P. Frasconi, "Learning Long-Term Dependencies with Gradient Descent Is Difficult", IEEE Transactions on Neural Networks, 5 (2) 157-166 (1994).
[16] S. Hochreiter, and J. Schmidhuber, "Long Short-Term Memory", Neural Computation, 9 (8) 1735-1780 (1997).
[17] Brett Naul, Joshua S Bloom, Fernando P´erez, and St´efan van der Walt. A recurrent neural network for classification of unevenly sampled variable stars. Nature Astronomy, 2(2):151.155, 2018.
[18] Gregory P. Badura, Christopher R. Valenta , Layne Churchill, Douglas A. Hope, "Recurrent Neural Network Autoencoders for Spin Stability Classification of Irregularly Sampled Light Curves", In Advanced Maui Optical and Space Surveillance Technologies Conference, September 2022.