# Anomaly detection from TLE data


František Dráček
*Faculty of Mathematics, Physics, and Informatics*
*Comenius University Bratislava*
Bratislava, Slovak Republic
dracek1@fmph.uniba.sk

Jiří Šilha
*Faculty of Mathematics, Physics, and Informatics*
*Comenius University Bratislava*
Bratislava, Slovak Republic
jiri.silha@fmph.uniba.sk

Roman Ďurikovič
*Faculty of Mathematics, Physics, and Informatics*
*Comenius University Bratislava*
Bratislava, Slovak Republic
durikovic@fmph.uniba.sk



***Abstract*—This paper presents a framework for the identification of anomalies in the trajectory of Low Earth Orbit (LEO) satellites. The algorithm can identify a broad class of anomalies from two-line element (TLE) data. The method is based on a combination of time series pre-processing techniques and Long Short Term Memory (LSTM) Autoencoder. The main contributions of this study are utilization of Gaussian derivative filter and demonstrating the ability to detect anomalies on rocket bodies, which have not been previously demonstrated.**



**Keywords**—***TLE, satellites, time series, anomaly detection, LSTM***


## I. Introduction

The number of man-made objects in the vicinity of the Earth has been rapidly growing last few years. The presence of both active and inactive satellites, debris, and other objects in causing the orbit to become congested and the likelihood of collisions of satellites is steadily growing. Satellite operators often do not have complete information about other satellites. For those reasons, the ability to identify anomalies in the trajectory of satellites is quite beneficial for operators. The anomalies in the trajectory of satellites can be caused by orbital manoeuvres, or, in the more interesting case, collisions with other small objects or debris. This case is of special interest to us, as detecting such anomalies could lead to an explanation of unexpected satellite behaviour and eventually could be used as an indicator of upcoming disintegration of the object. Hence, detection of anomalies can lead to efficient prediction of fragmentation events, which are the main source of space debris.

### A. Previous Works

The previous research approaches to the detection of anomalies in satellites' trajectories mainly focused on the detection of orbital maneuvers. Patera developed a data processing method, based on fitting polynomial curves to data in a sliding window [1]. Oltrogge and Alfano showed it is possible to use a low pass filter to identify maneuvers [2]. By fitting two sliding windows with polynomials and looking at differences between endpoints, Kelecy showed that it is sufficient to look at the semimajor axis to detect a broad category of maneuvers [5]. More recently, Li and Chen used the clustering of time series segments to identify maneuvers [9]. The major contribution of this work is usage of the first derivative of Gaussian filter during the preprocessing of the data, which significantly improves the results. Another contribution is demonstration of ability to detect anomalies for rocket bodies, while other studies demonstrated this only for LEO satellites.

## II. Satellite Position Data

This section describes dataset that this study works with.

### A. Orbital Elements

To work with positional data of satellites, first, we need to specify a frame of reference. For the rest of this article, we will assume that all positions are expressed in the True Equator, Mean Equinox (TEME) frame of reference [3], which is a special case of a geocentric frame of reference. The position of the satellite can be expressed by state vectors, i.e., position vector **r** and velocity vector **v**. However, it is hard to intuitively understand the nature of orbital motion just by looking at the state vectors. Therefore, satellite position is conveniently expressed in terms of orbital parameters, sometimes also called osculating elements. Those parameters, shown in Fig. 1., are:

- Eccentricity $e$ - oblateness of orbit
- Semimajor axis $a$ - longest distance from the centre
- Inclination $i$ - tilt angle between the orbital plane and equatorial plane of reference
- Right ascension of ascending node $\Omega$ - the angle between x-axis of TEME reference frame and intersection line (node) of orbital and equatorial planes
- Argument of perigee $\omega$ - the angle between ascending node and periapsis, a point in trajectory of orbiting body which is closes to body it is orbiting around, i.e., closest to the focus of the ellipse
- True anomaly $\theta$ - angle indicating the position of the satellite with respect to periapsis

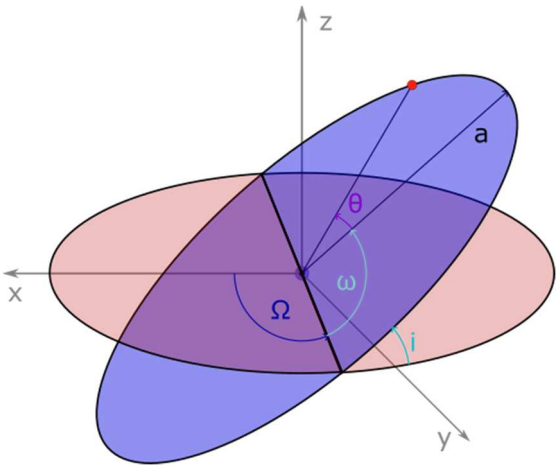


Fig. 1. Orbital parameters. Source: Authors

### B. TLE Dataset

TLE set is a data format, encoding orbital elements of earth-orbiting objects. The TLE database is maintained by the US Space Force, namely the North American Aerospace Defense Command (NORAD). The data are publicly available [10]. The TLE sets are considered as minimal information required for the propagation of Earth-orbiting objects in time. Besides the aforementioned orbital elements, TLE contains


2023 Communication and Information Technologies (KIT) | 979-8-3503-3839-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/KIT59097.2023.10297051

additional information, such as drag term and first and second derivative of mean motion.

However, the TLE orbital elements (mean elements) are not the same as the physical orbital elements (osculating elements). The TLE is averaged and engineered for a specific time propagation algorithm, Simplified General Perturbations model - SGP4. Therefore, the correspondence between mean and osculating elements is not obvious. Fig. 2 shows the semimajor axis of the Topex satellite calculated from TLE data (red) and SGP4 propagation (blue). The propagated semimajor axis, which corresponds to the real physical position, was extrapolated from the first (red) TLE datapoint. The sinusoid behaviour is caused by perturbations in Earth's gravitational potential due to Earth's oblateness. Fig. 3 depicts TLE semimajor axis for a timespan of several years. We can see that scale at which the physical semimajor axis changes, is one order bigger than changes in TLE semimajor axis.

Consequently, the complicated correspondence between mean and oscillating elements presents problems for potential data interpolation and the generation of synthetic data. The TLE data is unevenly sampled. If there was better correspondence between physical and mean orbital parameters, we could use this to resample the TLE time series. However, we showed that this is not the case, therefore we have to use different approach. The only way how we could use SGP4 for resampling, is to use the SPG4 propagator to obtain the state vector at a given time and afterwards iteratively guess the TLE element which would produce the same state vector for zero time propagation. Such an approach is not viable on a large scale. Recently, some of the problems of TLE data were reiterated in [12], where authors note that some of previous studies that used TLE data have failed to consider complicated correspondence between real positions and TLE data.

Another big challenge when working with TLE data lies with the fact that they come without measurement errors, therefore distinguishing between real change in orbital behaviour and noise can be problematic.

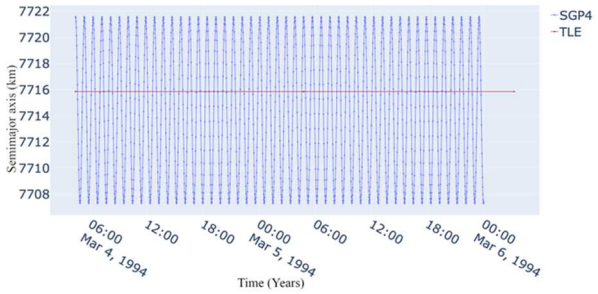


Fig. 2. Topex satellite semimajor axis from TLE data and from SGP4 propagation. Timespan shown is only 2 days, given that SGP4 prediction is accurate only for propagations on order of days. Source: Authors

## III. Anomaly Detection

The main goal of this work is to identify anomalies in the TLE data. Most of the anomalies in trajectory are caused by orbital maneuvers. We do not have a dataset for the maneuvers of satellites, therefore we will use unsupervised learning techniques. In this study, we look at two satellites, namely the Topex satellite, with Norad id 22076. The other object we look at is a rocket body, namely Thor-Burner 2A with Norad id 7412. As we mentioned in the previous section, Kelecy showed that to detect maneuvers, it is sufficient to only look at the semimajor axis [5]. We will adopt this assumption. In some cases, the anomalies affect other orbital parameters, while in others the effect is negligible. So far, we struggle to determine when and which parameters should be used, therefore we will stick with the semimajor axis.

Examples of semimajor axis data time series are shown in Figs 3, 4 and 5. Fig. 3 depicts the semimajor axis of the Topex satellite. The main anomaly of the trajectory is highlighted in green. The controlled evolution of the satellite, i.e., when the satellite's semimajor axis was maintained by regular maneuvers, is highlighted in blue. In blue is uncontrolled evolution, decreasing semimajor axis corresponds to the period when the satellite became inactive and the semimajor axis was no longer maintained by regular maneuvers. Fig. 4 is a close up of the Topex semimajor axis restricted to the controlled evolution part. Steep increases in the semimajor axis correspond to orbital maneuvers, which are preformed to maintain the orbit. Fig. 5 depicts the semimajor axis of Thor-Burner, a rocket body. The small anomaly on 17th January (highlighted in green) corresponds to a documented accidental collision with a CZ-4 debris object from a Chinese rocket body[4].

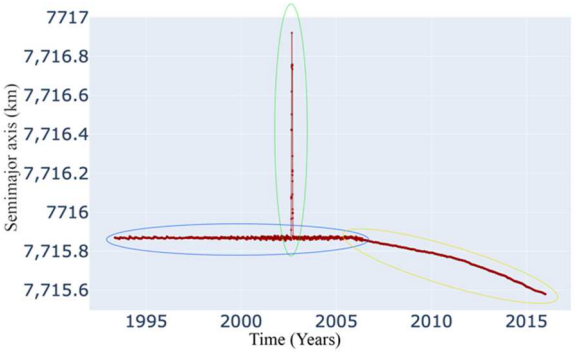


Fig. 3. Topex satellite semimajor axis. Main anomaly is highlighted in green, controlled evolution in blue, uncontrolled in yellow. Source: Authors

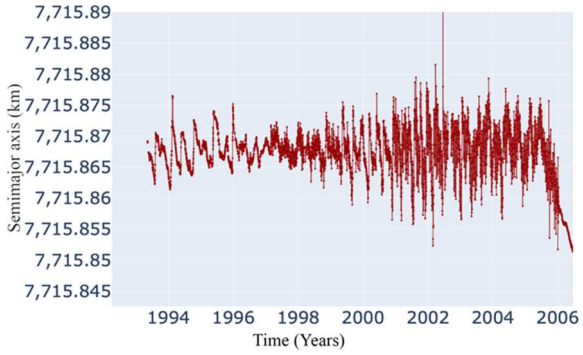


Fig. 4. Topex satellite semimajor axis. Close up snapshot of controlled evolution part of Fig. 3. Source: Authors

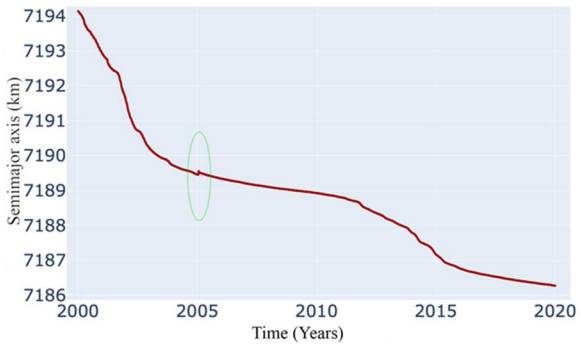


Fig. 5. Thor Burner rocket body semimajor axis. Main anomaly is highlighted in green. Source: Authors

### A. Preprocessing

The time series of semimajor axis parameters are not uniformly sampled, as sampling corresponds to the time of measurement. Moreover, the noisy nature of the data, combined with the fact that data are the result of artificial averaging, makes preprocessing the most crucial aspect of the entire model.

First, we resample data at a frequency of 1 datapoint per day. In the case where we have more than one data point per day, we take the mean of those data points, while in the case of missing datapoints, we use linear extrapolation. We already discussed why it is infeasible to use a physical model to extrapolate and resample TLE data.

The data are smoothed with a weighted moving average filter with distance reach of $d = 6$ neighbours [6]. Let us assume time series $\{t_i, a_i\}$, where $t_i$ is time and $a_i$ is smoothed parameter - in this case semimajor axis. After smoothing $a_i \to \tilde{a}_i$, where:

$$\tilde{a}_i = \frac{\sum_{j=i-d}^{i+d} w_j a_j}{\sum_{j=i-d}^{i+d} w_j}, \tag{1}$$

where:

$$w_j = 1 - \frac{|t_j - t_i|}{t_{max}}, \tag{2}$$

with $t_{max} = max(|t_{i-d} - t_i|, |t_{i+d} - t_i|)$.

Given the fact that we are mainly interested in changes in the semimajor axis, we use a 1D derivative of the Gaussian filter, like one used to detect edges in pictures, to obtain the derivate [7]. To calculate the filter, first we sample Gaussian

$$G = \frac{1}{\sqrt{2\pi}s} \exp(-\frac{x^2}{2s}), \tag{3}$$

at $x = \{-2s, -2s + 1, \ldots, 2s - 1, 2s\}$. The size we use is $s = 5$. Afterwards, we convolve sampled Gaussian with derivative filter

$$h = (G * f), \tag{4}$$

where $f = \{1, 0, -1\}$. Derivative of a given data parameter is then obtained as

$$\tilde{a}_i \to \tilde{d}_i = (\tilde{a} * h)_i, \tag{5}$$

Finally, we use MinMax scaler to normalise data, that is:

$$d_i \to \tilde{d}_i = \frac{d_i - \min d_j}{\max d_j - \min d_i}. \tag{6}$$

### B. LSTM Autoencoder

To detect the anomalous segments from the time series of the Gaussian derivative of the semimajor axis parameter, we divide the time series into subsequences of length $l = 28$. LSTM Autoencoder model is trained to encode and decode those sequences [8]. The encoder decoder scheme is depicted in Fig. 6. The LSTM autoencoder consists of two modules, encoder and decoder. The input sequence is encoded by encoder module with 3 LSTM layers. The hidden state from last layer, $h_3$, is repeated $l$ times and used as an input for decoder module. We use 3 hidden LSTM layers in both the encoder and decoder, with a hidden dimension equal to 3. The reconstructed sequences $X_{reconstructed}$ are compared to inputs $X$ with mean squared error loss function. The sequences with the biggest reconstruction loss are marked as anomalous. The Autoencoder will try to learn the best way to encode sequence segments we will feed into it. Because of the bottleneck in encoding dimensionality, the autoencoder struggles to learn anomalous segments in data. To improve performance, the anomalous sequences can be repeatedly removed from the training set and the model retrained.

For the training of the LSTM autoencoder, entire time series of a single object is used for training, which we deem as justifiable given the data availability and given the fact that we use autoencoder to identify anomalies. The trianint set consists from 8628 subseqences for Topex case and 11999 for Thor Burner case.

The practical implementation of model utilises Python PyTorch library [11].

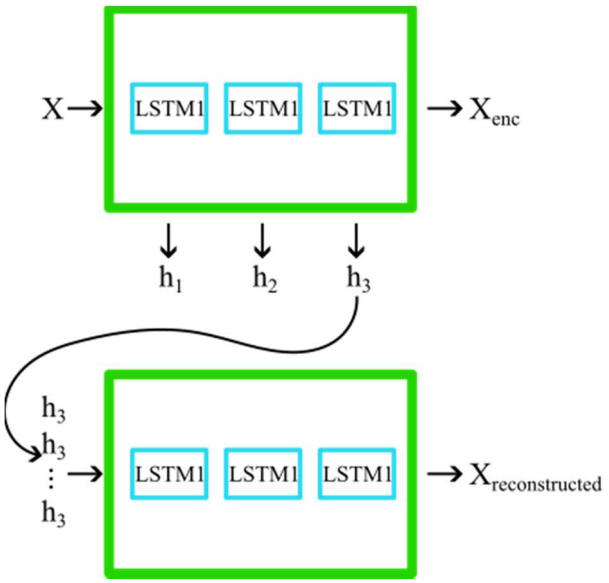


Fig. 6. LSTM autoencoder scheme. X is input sequence, $X_{enc}$ is encoded sequence, $h_i$ are hidden states of LSTM layers and $X_{reconstructed}$ is recnstructed sequence. Source: Authors

## IV. Results

Gaussian derivative of the semimajor axis, combined with the weighted average, works surprisingly well and identifies anomalous changes in the semimajor axis parameter. Fig. 7 depicts semimajor axis of Topex satellite along with the reconstruction loss obtained from LSTM autoencoder. The main anomaly of Topex satellite (green highlight in Fig. 3) trajectory clearly corresponds to the spike in loss function. Therefore, we can conclude that our model is able to successfully identify the main anomaly in the Topex's trajectory. However, this is not the only anomaly present in the Topex's trajectory, there are also numerous anomalies caused by small scale orbital maneuvres, which occurred during the controlled evolution of the satellite. To be able to identify those anomalies, we well partially filter out the main anomaly.

The reconstruction loss for Topex satellite with partially filtered out main anomaly is shown in Fig. 9. In this case, once again, the peaks in the reconstruction loss (grey) clearly correspond to maneuvres.

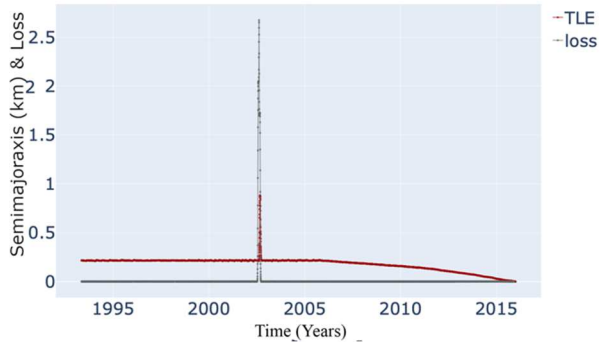


Fig. 7. Topex reconstruction loss and semimajor axis. We observe that the main anomaly (green highlight in Fig. 3) was successfully detected. Source: Authors

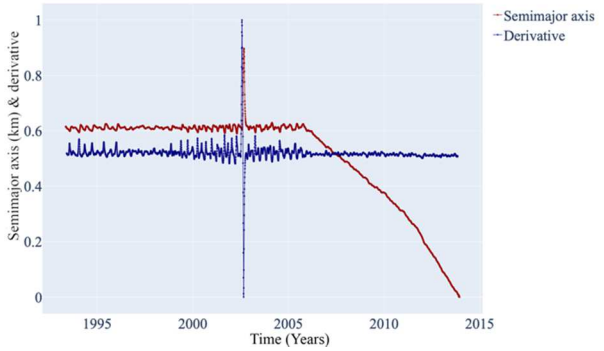


Fig. 8. Topex semimajor axis (red) and its gaussian derivative (blue). The main anomaly was filtered out partially, in order to make the plot more descriptive. Source: Authors

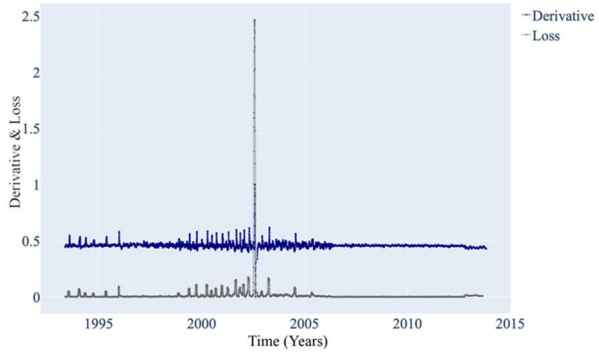


Fig. 9. Topex reconstruction loss (grey) and derivative of semimajor axis (blue). The main anomaly was partially filtered out from data - to amplify anomalies in the part of data corresponding to controlled evolution. Source: Authors

Finally, the reconstruction loss for Thor Burner rocket, along with normalized semimajor axis is shown in Fig. 10. The highest peak of reconstruction loss (in grey) successfully identifies the main anomaly of Thor Burner rocket. The other peaks in reconstruction loss correspond to sudden changes in rate of decrease (i.e., derivative) of semimajor axis. This is probably a result of controlled change in acceleration of the rocket. In this sense, those are still anomalies, however, not very interesting ones. Fig. 10 also demonstrates why usage of LSTM autoencoder is justified in this application. From the gaussian derivative (blue) the identification of anomalous segments is not straightforward, while they are much more pronounced in the reconstruction loss.

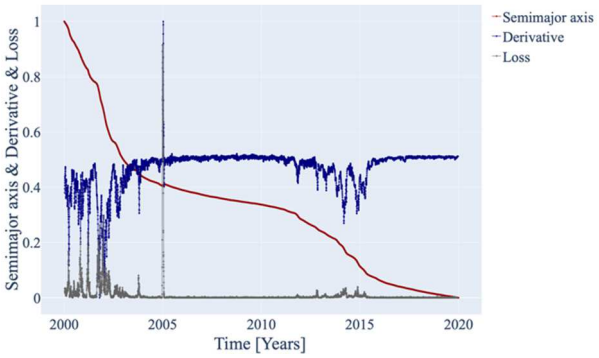


Fig. 10. Thor-Burner reconstruction loss (grey), derivative of semimajor axis(blue) and semimajor axis (red). We observe that the main anomaly (green highlight in Fig. 5) was successfully detected, also with points where the rate of change of the semimajor axis changed. Source: Authors

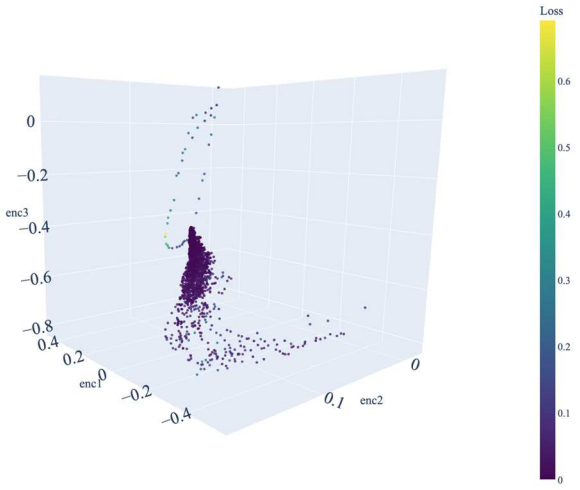


Fig. 11. Thor-Burner LSTM autoencoder hidden representation. The axes $enc_1$ , $enc_2$ , $enc_3$ correpond to bottleneck dimensions of encoded inputs. Source: Authors

## V. Discussion

The Gaussian derivative filter works surprisingly well when used with TLE data, however, one could argue, that after we obtain derivative, it is sufficient to use thresholding to identify anomalies, instead of using LSTM autoencoder. However, the LSTM Autoencoder approach is more robust, in the sense that it is not sensitive to user defined constants. It also seems that LSTM autoencoder performs better for cases with significant presence of noise in the derivative, even after the smoothing. In the case of Topex satellite, LSTM autoencoder basically predicts constant line, which is not superior when compared to derivative thresholding. For Topex, derivative of semimajor axis is same for entire evolution apart from the anomalies. However, this might not be true for all satellites. We expect LSTM autoencoder to perform superiorly in cases where derivative has two modes. This is partially visible in the case of Thor Burner. The usage of LSTM autoencoder is also justified by offering us pathways for future research, mainly the possibility of further clustering the hidden encoding of LSTM autoencoder. Such clustering could lead to identification of various modes of satellite behaviour. Plot of hidden representation features is shown in Fig. 11. Colour of datapoints corresponds to reconstruction loss. Clearly, the anomalous points are outside the main cluster. However, so far, we have not been able to identify different behaviour modes. The LSTM approach also allows us to conveniently incorporate the rest of the orbital parameters into our model.

The weak point of our approach remains the selection of hyperparameters of our model. As of now, they are the result of educated guessing and trial-error approach.

The detection of main anomaly of Thor Burner rocket is a significant result, as the identified anomaly correspond to documented collision with second object. To our best knowledge, the detection of anomalies for rocket bodies have not been demonstrated previously, as previous research efforts focused on LEO satellites. As already mentioned, detecting such anomalies could serve as an indicator of disintegration of orbiting body and enable us to predict fragmentation events. Such capability is very desirable, especially in context of orbit congestion.

## VI. Conclusion

While it was already established in previous works that it is possible to use TLE data to identify maneuvers of Satellites, our results show that it is also possible to detect anomalies of rocket bodies from TLE data. This is significant result, as this could lead to improvements in space safety management. The fact that we work with unlabeled dataset presents some limitations to this study.

In future work we would like to explore the utilization of other orbital paraments. It is also quite natural to expect that improvements to our algorithm could be made with the help of labeled dataset.

## Acknowlegment

This work was supported from UK grant UK/327/2023.

## References

[1] R. Patera, “Space event detection method,” Journal of Spacecraft and Rockets vol. 45, pp. 554-559, 2006.

[2] D. L. Oltrogge and S. Alfano, “Determination of orbit cross-tag events and maneuvers with orbit detective,” in Proceedings of the AAS/AIAA Astrodynamics Specialist Conference, Girdwood, AK, USA, 31 July–4 August 2011.

[3] F. R. Hoots, R.L. Roehrich, T. Kelso, Spacetrack Report No. 3, Project Spacetrack Reports, Vol. 80914, pp. 14, 1980.

[4] T.S. Kelso, “Validation of SGP4 and IS-GPS-200D against GPS ephemerides,” 17th American Astronautical Society/American Institute of Aeronautics and Astronautics Space Flight Mechanics Conference, vol. 127, Part 1, pp. 425-440, Sedona, AZ, 2007.

[5] T. Kelecy and K. M. Hamada, “Satellite maneuver detection using Two-line Element ( TLE ) data,” in Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, Maui Economic Development Board (MEDB), Maui, HA, USA, pp. 12–15 September 2007.

[6] S. W. Cleveland, “Robust locally weighted regression and smoothing scatterplots,” Journal of the American Statistical Association, vol. 74, pp. 829-836, 1979.

[7] J. F. Canny, “A Computational Approach to Edge Detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8, pp. 679-698, 1986.

[8] P. Malhotra, R. Anusha, A. Gaurangi, V. Lovekesh, A. Puneet and M. S. Gautam, “LSTM-based encoder-decoder for multi-sensor anomaly detection,” ArXiv abs/1607.00148, 2016.

[9] T. Li and C. Lei, “Space event detection method based on cluster analysis of satellite historical orbital data,” Acta Astronautica, Elsevier, vol 160, pp. 414-420, 2019.

[10] Spacetrack team, Spacetrek, https://www.space-track.org

[11] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, “PyTorch: An imperative style, high-performance deep learning library,” Proceedings of the 33rd International Conference on Neural Information Processing Systems, Curran Associates Inc., Red Hook, NY, USA, Article 721, pp. 8026–8037, 2019.

[12] D. P. Shorten, Y. Yang, J. Maclean and M. Roughan, “Wide-scale monitoring of satellite lifetimes: Pitfalls and a benchmark dataset,” Journal of Spacecraft and Rockets, 2023, in press.