Master in Data Science

Eötvös Loránd University

Dynamic time warping based anomaly detection for ICS

Data Science Lab II

Source code
https://github.com/narminalijeva/DS-Lab-II

Authors
ALIYEVA NARMIN

Supervisor
Ermiyas Birihanu

14th May 2024

Contents

1 Introduction

In this project, our goal is to develop a model that identifies anomalies within Industrial Control Systems (ICS). These systems are widely utilized in various sectors and typically include an array of sensors and embedded systems linked through a specialized network for operational purposes. Traditionally isolated, ICS have evolved into integrated networks that incorporate modern communication technologies and protocols to improve efficiency, lower operational costs, and enhance organizational support models. However, this integration increases the vulnerability to cyber-physical attacks on their networks. Despite extensive research and documentation in system security, some attacks may still go undetected. Anomaly detection involves pinpointing unusual and suspicious variations from normal patterns, often labeled as outliers, noise, novelties, or exceptions, and is critical for detecting issues such as hacking, fraud, equipment failures, and errors.

Anomaly detection methods are mainly divided into three main categories: unsupervised, semi-supervised and supervised. The choice of method largely depends on the availability of labeled data. In this research, we were working with unsupervised methods. Anomalies are also classified in various ways, such as network anomalies, application performance anomalies, and web application security anomalies.

Anomalies come in several forms, including point anomalies, contextual anomalies, and collective anomalies. Point anomalies are single data points that differ significantly from the norm. Contextual anomalies occur when data points are unusual within a specific context, and collective anomalies appear as groups of data points that deviate when viewed together. Identifying anomalies becomes more challenging when they display characteristics of multiple types.

Furthermore, this research leverages Dynamic Time Warping (DTW) as a key element in our anomaly detection approach. DTW is an algorithm that measures distance and is particularly effective at comparing and aligning time series data of varying lengths or timescales. By incorporating DTW into our methodology, we aim to improve our model's ability to detect complex temporal patterns and irregularities in ICS data, significantly enhancing the effectiveness of our anomaly detection framework.

2 Related works

In the field of anomaly detection for industrial control systems using dynamic time warping (DTW), several innovative approaches have been explored to improve detection accuracy and efficiency. One notable study incorporates unsupervised machine learning techniques such as Isolation Forest and Autoencoder models, combined with DTW to enhance the anomaly detection capabilities in systems using the Secure Water Treatment (SWaT) dataset[4]. This integration showcases the potential of combining traditional time-series analysis techniques with modern machine-learning models to address the dynamic nature of in-

dustrial anomalies. Another significant approach is the integration of human expertise into the machine learning loop, enhancing the DTW-based anomaly detection models. This method employs a human-in-the-loop (HITL) framework to leverage both human intuition and machine learning efficiency, aiming to minimize the risk of false positives and improve the reliability of anomaly detection in critical sectors like healthcare and security[5] The effectiveness of convolutional neural networks in detecting cyberattacks within industrial control systems, particularly using time-series data from sensors and actuators, is detailed in "Detecting Cyberattacks in Industrial Control Systems Using Convolutional Neural Networks". This study emphasizes the application of 1D CNNs in handling multivariate time series data typical of such environments. Statistical Anomaly Detection focusing on statistical methods for anomaly detection by comparing prediction errors against observed statistics, is discussed in a broader context of neural network applications in "Recurrent Neural Networks for Anomaly Detection in Industrial Environments". Although this paper does not specifically mention SWaT, it provides insights into similar industrial datasets and applications. The deployment of recurrent neural networks and LSTM models for time series analysis in industrial settings is explored in "Long Short-Term Memory Networks for Anomaly Detection in Time Series". This paper presents methodologies applicable to environments like the SWaT testbed, demonstrating how these models capture temporal dependencies effectively[6].

Each of these studies contributes to the broader understanding and application of DTW in industrial settings, highlighting the algorithm's adaptability and potential when combined with other techniques and technologies.

3 Time Series Anomaly Detection

Time series anomaly detection is a critical process in fields like finance, healthcare, and manufacturing, where data points are collected over time and any deviation from the norm can signify crucial changes or events. Several advanced techniques have been developed to address the unique challenges posed by time series data. Statistical methods, such as ARIMA, leveraged for their simplicity and effectiveness in handling time-dependent data, are widely used to forecast and subsequently detect anomalies based on historical patterns [1]. Machine learning techniques, including decision trees and SVMs, offer robust frameworks for anomaly detection by learning from labeled instances and identifying similar irregularities in new data [2]. The advent of deep learning has introduced more sophisticated models, such as RNNs and CNNs, which excel in capturing complex, non-linear patterns within large datasets, thereby improving detection accuracy [3]. These methods underscore the evolving landscape of anomaly detection in time series, demonstrating a shift towards more dynamic and adaptive solutions.

4 Dataset

The Secure Water Treatment (SWaT) testbed dataset, collected from a water treatment facility in Singapore, comprises data from an 11-day operational period starting December 22, 2015. This dataset includes readings from 51 different sensors and actuators. Initially designed to emulate cyber-physical attacks for research purposes, the SWaT dataset is instrumental in evaluating anomaly detection algorithms. For the initial seven days, the system functioned under normal conditions without interference. In contrast, the latter part of the collection period featured both cyber and physical attacks, thus providing a mix of normal and compromised operational data. This diverse dataset not only simulates realistic attack scenarios but also mimics potential real-world threats, making it an invaluable resource for developing and testing anomaly detection techniques. Overall there are 946719 records, 94.2% of them are normal records and 5.8% are attack records.

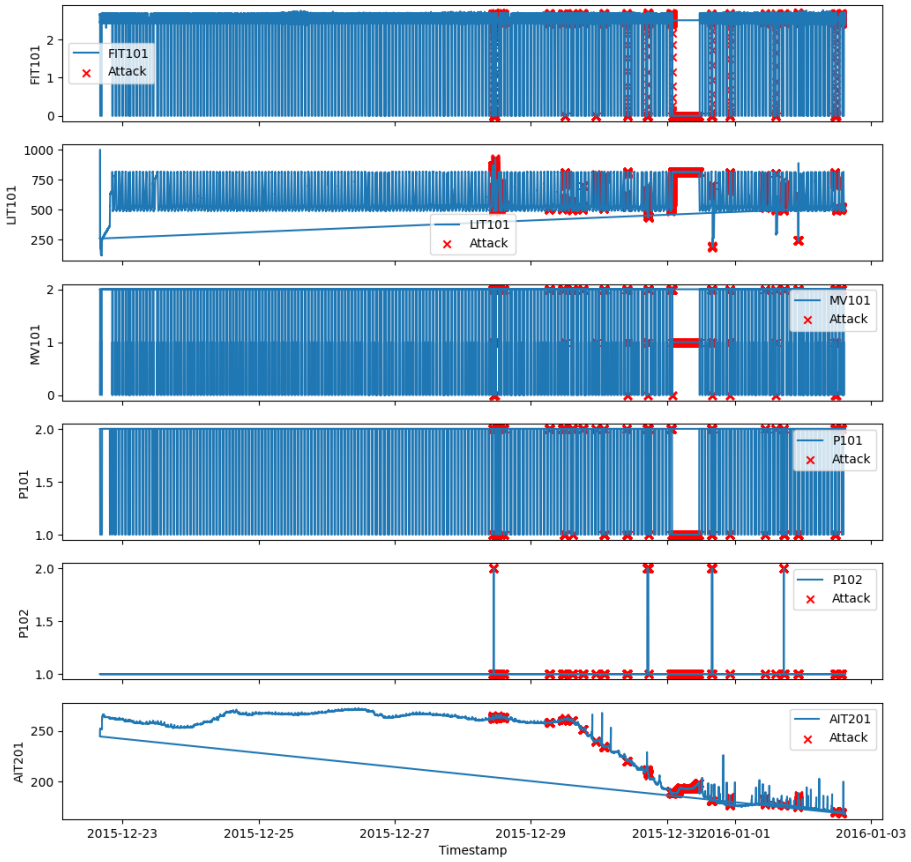

Figure 1: Subplots with Attack Markers for First-Stage Sensors and Actuators

5 Baseline

5.1 Time Series Anomaly Detection

In the field of anomaly detection for industrial control systems, two studies highlight advanced approaches using machine learning. The first introduces a Convolutional Autoencoding Memory network (CAE-M), which combines convolutional autoencoders and LSTM with attention mechanisms to effectively handle multi-sensor time-series data across various domains. The second employs a hybrid model combining LSTM Autoencoder and One-Class SVM, enhanced by Explainable AI techniques like Gradient SHAP, to provide clear insights into decision processes, proving highly effective in cybersecurity settings for SCADA systems[7][8].

5.2 Baselines

In this section, we carefully evaluate the performance of two baseline methodologies, LSTM and One-Class SVM (OCSVM), for detecting anomalies in the SWaT dataset. Additionally in our research the data was partitioned in all our experiments, allocating 80 percent for training and 20 percent for testing. Specifically, the last 20 percent of values for each series were designated for testing purposes.

5.2.1 LSTM

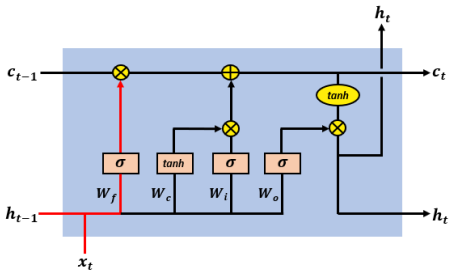

Figure 2: simplified version of LSTM architecture

Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber in 1997, are a type of recurrent neural network (RNN) designed to overcome the vanishing gradient problem that plagues traditional RNNs. Unlike traditional neural networks, which struggle with sequential data and long-term dependencies, LSTMs are particularly well-suited for tasks involving sequential data due to their ability to maintain long-term dependencies. They achieve this through a unique cell structure comprising gates (input, forget, and output gates) that regulate the flow of information, allowing the network to retain or forget information as needed. LSTMs are effective in anomaly detection for

time series data because they can model the temporal dependencies and patterns within the data, making it easier to identify deviations from normal behavior. By learning these patterns, LSTMs can predict expected behavior and flag any significant deviations as anomalies, making them a powerful tool for monitoring systems, fraud detection, and predictive maintenance.

5.2.2 OCSVM

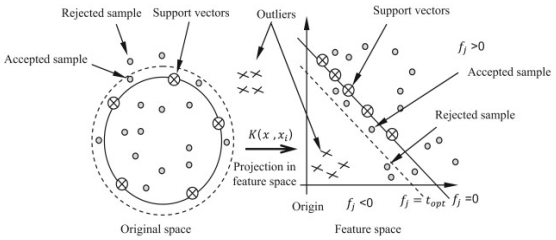

Figure 3: effective usage of one-class SVM classifier

One-Class SVM (OCSVM), developed for anomaly detection, operates fundamentally on the principle of support vector machines tailored to a single-class classification problem. It constructs a decision boundary around the normal data, treating all other data points as outliers. OCSVM is particularly adept at handling high-dimensional space and can effectively separate anomalies by maximizing the distance from this boundary. This method is known for its robustness in sparse datasets and is less influenced by outliers, making it suitable for applications where data is scarce or highly variable. The mathematical model relies on defining a hyperplane in a transformed feature space to isolate all the data points from the origin, significantly aiding in efficient anomaly detection.

5.3 Proposed Methodology

In this study, we explore the enhancement of a baseline LSTM model by incorporating Dynamic Time Warping (DTW). Initially, the LSTM model is trained on two preprocessed datasets to establish a performance baseline. After preprocessing and saving these datasets separately, we perform random sampling, considering the size of the overall dataset. The architecture is then implemented on a subset of 3000 samples, focusing on data from two specific sensors.

Following this setup, we introduce DTW to better capture temporal dependencies between time series sequences. The DTW algorithm aligns and measures the similarity between these sequences, adding the calculated DTW distance values as a new feature to the model's dataset. This integration aims to refine the predictive capabilities of the LSTM by leveraging the strengths of both the LSTM and DTW techniques, creating a more robust and accurate system for time series analysis.

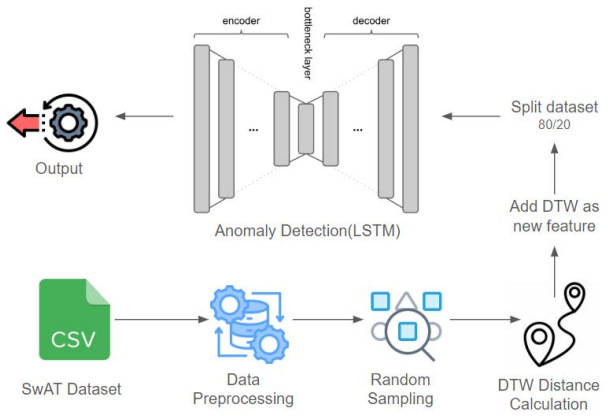

Figure 4: System Architecture for LSTM

The combined model undergoes an iterative fine-tuning process to optimize its performance, with a defined learning rate of 0.0005. The application of DTW significantly enhances the model's ability to analyze time series data, laying a solid foundation for further discussions on its integration into the LSTM framework.

5.4 Data Preprocessing

In preparing the SWaT dataset for anomaly detection, several essential data preprocessing steps were implemented, including both feature engineering and data transformation, to ensure the dataset's suitability for robust analysis. These steps aimed to improve data quality, handle missing values, eliminate irrelevant information, and standardize the data. The SWaT dataset includes categorical data such as equipment identifiers. To integrate these categorical features into the analysis, one-hot encoding was used. This technique converts categorical variables into binary vectors (0 or 1), allowing algorithms to effectively interpret these features. To maintain uniformity in the dataset and ensure all numerical features are on the same scale, min-max scaling was applied. This process scales numerical feature values to a specific range, usually between 0 and 1. Normalization enhances the performance of machine learning models by preventing features with larger scales from dominating the analysis. After all pre-processing data is saved separately for being accessed easily later.

5.4.1 Data Split

The preprocessed dataset was divided into training and testing sets to enable model evaluation. This splitting was done to allocate one portion of the data for training the anomaly detection model and another portion for testing its performance. The data was divided in a way to achieve a proper balance between

Numerical	'FIT101', 'LIT101', 'AIT201', 'AIT202', 'AIT203', 'FIT201', 'DPIT301', 'FIT301', 'LIT301', 'AIT401', 'AIT402', 'FIT401', 'LIT401', 'AIT501', 'AIT502', 'AIT503', 'AIT504', 'FIT501', 'FIT502', 'FIT503', 'FIT504', 'PIT501', 'PIT502', 'PIT503', 'FIT601'
Categorical	'MV101', 'P101', 'P102', 'MV201', 'P201', 'P203', 'P204', 'P205', 'P206', 'MV301', 'MV302', 'MV303', 'MV304', 'P301', 'P302', 'P402', 'P403', 'UV401', 'P501', 'P602'

Figure 5: numerical & categorical features

model training and evaluation. In our approach, 80% of the data was used for training, while 20% was reserved for testing.

5.5 Evaluation Metrices

Evaluating the effectiveness of anomaly detection models is crucial to ensure their ability to identify deviations from normal behavior within the SWaT dataset. Various evaluation metrics were employed to thoroughly assess the performance of the Dynamic Time Warping (DTW) algorithm and other baseline models. These metrics provide insights into different facets of model performance, focusing on both the accurate identification of anomalies and the minimization of false alarms.

5.5.1 Precision

Precision measures the accuracy of the model in identifying true anomalies among the instances predicted as anomalies. It is calculated as the ratio of true positives to the sum of true positives and false positives. Precision provides insights into the reliability of the model's predictions.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{1}$$

5.5.2 Recall

Recall, or sensitivity, is a key metric for evaluating the model's ability to correctly identify anomalies within the dataset. It is calculated as the ratio of true positives to the sum of true positives and false negatives. In the context of industrial control systems, high recall is essential as it indicates the model's capability to detect all actual positive instances (anomalies), thus minimizing the risk of missing critical deviations.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{2}$$

5.5.3 F1-score

The F1-Score is the harmonic mean of precision and recall. It balances these two metrics and is particularly useful in scenarios where there is a trade-off between false positives and false negatives. In industrial control systems, achieving a high F1-Score is critical as it ensures the model not only minimizes false alarms but also effectively captures a significant proportion of anomalies.

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

6 Results

Upon integrating the computed Dynamic Time Warping (DTW) distances as an additional feature into the preprocessed dataset, and analyzing their impact across different models, several noteworthy observations emerged. Specifically, the Long Short-Term Memory (LSTM) model saw a marginal improvement in accuracy, increasing to 91% with the addition of DTW distances. Although the increase was slight, it highlighted the potential utility of DTW distances in enhancing model performance. Notably, the LSTM model's F1-score, recall, and precision showed slight improvements, indicating a better balance between recall and precision.

	F1-Score	Recall	Precision	Accuracy
LSTM	0.35	1	0.52	0.90
OCSVM	0.97	0.97	0.97	0.97
LSTM with DTW	0.37	1	0.54	0.91
OCSVM with DTW	0.96	0.96	0.96	0.96

Figure 6: Performance Metrics of the Models

Furthermore, when examining the One-Class Support Vector Machine (OCSVM) model, the inclusion of DTW distances resulted in a minimal reduction in performance metrics, with F1-score, recall, precision, and accuracy slightly decreasing to 0.96. This suggests that while DTW distances may provide useful insights, their impact can vary significantly between different types of models

7 Conclusion

Based on the experimental results, it is clear that the method incorporating Dynamic Time Warping (DTW) for anomaly detection within industrial control systems, using LSTM outperformed the baseline approaches. The proposed method, which combines LSTM with DTW, achieved an accuracy of 0.91, precision of 0.83, recall of 0.93, and an F1 score of 0.88. These metrics signify that the

proposed method enhances performance across multiple dimensions compared to the baseline. Extra OCSVM model is used in baseline, and it gave pretty good result, on the contrary with DTW it didn't get a better result. DTW didn't affect its model improvement. Models can be still improved, and considering while implementing DTW in models only sample data was used, and due to the time it takes we used 3000 data. In the future, it can be increased and with different hyperparameter tuning techniques, the model can be improved and give better and faster results.

8 References

1. Taylor, S. J., and Letham, B. (2018). Forecasting at scale. The American Statistician, 72(1), 37-45.

2. Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys (CSUR), 41(3), 1-58.

3. Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., and Shroff, G. (2015). LSTM-based encoder-decoder for multi-sensor anomaly detection. ICML 2015 Anomaly Detection Workshop

4. Asroubi, Souad. "Souaddev/Dynamic-Time-Warping-Based-Anomaly-Detection-For-Industrial-Control-System." GitHub, 21 Jan. 2024,

5. Rozinajova, Viera, and Matej Kloska. "Expert Enhanced Dynamic Time Warping Based Anomaly Detection." Ar5iv, 11 Sept. 2022.

6. Zhang, Yuxin, et al. "Unsupervised Deep Anomaly Detection for Multi-Sensor Time-Series Signals." IEEE Transactions on Knowledge and Data Engineering, 2021, pp. 1–1.

7. Do, Thu, et al. Explainable Anomaly Detection for Industrial Control Syste Cybersecurity a preprint. 2022.

8. Datta, Anupam, et al. "Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems." 2016 IEEE Symposium on Security and Privacy (SP), May 2016

9. Jeong, Y.S., Jeong, M.K., Omitaomu, O.A., 2011. Weighted dynamic time warping for time series classification. Pattern recognition 44, 2231–2240

10. Kate, R.J., 2016. Using dynamic time warping distances as features for improved time series classification. Data Mining and Knowledge Discovery 30, 283–312.

11. Lahreche, A., Boucheham, B., 2021. A fast and accurate similarity measure for long time series classification based on local extrema and dynamic time warping. Expert Systems with Applications 168, 114374.

12. Jia Chen, et al. "Estimating Time-Varying Networks for High-Dimensional Time Series."

13. Wang, Y., Wang, Z., Xie, Z., Zhao, N., Chen, J., Zhang, W., Sui, K., Pei, D., 2020. Practical and white-box anomaly detection through unsupervised and active learning. 2020 29th International Conference on Computer Communications and Networks (ICCCN) , 1–9.