SADIS: real-time sound-based anomaly detection for industrial systems

Awaleh Houssein Meraneh, Fabien Autrel, Hélène Le Bouder, Marc-Oliver Pahl

SADIS: real-time sound-based anomaly detection for industrial systems

Awaleh HOUSSEIN MERANEH, Fabien Autrel, Hélène Le Bouder, and Marc-Oliver Pahl

IMT Atlantique, IRISA

Abstract. Industrial cyber-physical systems are critical infrastructures vulnerable to cyber-attacks. Anomaly and intrusion detection are widely used approaches to enhance the security of these systems. Existing detection methods can be categorized into two families. The first family detects only known attacks. The second family overcomes this limitation but often has a high false positive rate and a long detection time. This paper investigates the second family using side-channel leakages, particularly sound, for high-accuracy detection of intrusions and anomalies in various industrial systems. Despite sound signal's advantages, such as low-cost equipment, minimal computational requirements, and non-invasive measurement. Current sound-based anomaly detection (SAD) methods face challenges such as sensitivity to background noise, unbalanced sound data, computational costs, and detection accuracy. To tackle these issues, we introduce robot-arm sound dataset (RASD) and present a real-time sound-based anomaly detection for industrial systems (SADIS) approach that uses a simple and efficient method to fingerprint expected sound data with reduced dimensions. It employs an autoencoder (AE) for data classification and utilizes the Mahalanobis distance (MD) as an anomaly-scoring function, enhancing detection performance. Our experiments demonstrate that the SADIS approach achieves an average attack detection rate of over 96%, with a detection time of less than 1 second and low computational costs.

Keywords: Industrial systems, anomaly detection, side-channel leakage, sound, Autoencoder, Mahalanobis distance, real-time

1 Introduction

Industrial cyber-physical systems (ICPS) face increased vulnerability to cyber-physical attacks due to growing interconnectivity between physical and cyber systems [21, 31]. Attacks on these systems can cause collateral damage, as illustrated by the Stuxnet cyberattack [19], emphasising the need for robust security, including anomaly and intrusion detection. Anomaly, encompassing deviations from normal behaviour, poses challenges in detection due to its unknown. Anomaly and intrusion detection systems (IDS) are security measures for detecting unexpected behaviour, with signature-based intrusion detection systems

(SIDS) identifying known attacks [11] and anomaly-based intrusion detection systems (AIDS) effectively detecting deviations from normal, including zero-day attacks [13]. Current IDS struggle to accurately detect deviations in the physical system of ICPS, resulting in delayed responses and false alarms that risk the manufacturing process. This paper investigates AIDS using side-channel leakages, particularly sound, for high-accuracy detection of intrusions and anomalies in various industrial systems. Therefore, sound-based anomaly detection (SAD) proves to be a practical and widely applicable method in domains such as public video surveillance [8], speech analysis and recognition [5], healthcare [32], predictive maintenance of industrial [10]. To our knowledge, SAD has not yet been applied to detect cyber-physical attacks.

Existing SAD methods face challenges such as sensitivity to background noise, unbalanced sound data and domain shift challenges. Addressing these issues is crucial for enhancing detection performance and adaptability across diverse environments [20, 30]. While many SAD methods are performed offline, real-time detection provides early anomaly detection. In the literature, two real-time approaches exist: continuous online training and detection and offline training with online detection. This work concentrates on the latter approach. Considering the common challenges in anomaly detection, the SAD issues, and recommendations [29] and challenges [1], developing real-time SAD for industrial systems requires collecting high-quality sound data, utilizing low-latency preprocessing, ensuring accurate and fast detection, adapting to domain shifts, and compatibility with various systems.

This paper introduces a real-time SADIS approach for detecting cyber-physical attacks in ICPS. It addresses the challenges of existing methods, including computational costs, detection accuracy, and noise robustness. The SADIS approach efficiently extracts relevant features from sound data using dimension reduction via time-average spectrum (TAS) or principal component analysis (PCA). Additionally, the SADIS employs an autoencoder (AE) for data classification and utilizes the Mahalanobis distance (MD) as an anomaly-scoring function, enhancing detection performance. Moreover, the main contributions of this work are twofold.

1. First, we have generated a sound dataset, robot arm sound dataset (RASD), to address unbalanced sound data challenges[1].
2. Secondly, we have developed the SADIS approach, assessing its robustness using the RASD dataset and enhancing its adaptability through validation on diverse industrial systems.

The paper is structured as follows: Section 2 discusses related studies, Section 3 introduces deep learning concepts and the autoencoder model background, Section 4 presents the SADIS approach, Section 5 details the experimentation setup and results evaluation, and Section 6 concludes with the strengths and limitations of the SADIS approach.

[1] `https://github.com/a23houss/sadis_experimentation_code`

2 Related work

This section presents related studies to real-time anomaly detection of industrial systems using side channels. The comparison of related work is summarized in Table 1. This work compares the SADIS approach to existing methods, evaluating side-channel leakage, real-time performance, parameter sensitivity, robustness against noise, and overall effectiveness in detecting cyber-physical attacks.

Table 1. Comparison of related works in anomaly detection for industrial systems. The '-' symbol indicates that the research question was not addressed in the study, while the ✓ symbol indicates that the study investigated the research question

Articles	Side-channel leakage	Real-time performance	Parameter sensitivity	Robustness against noise	Comparison with existing approach
Baseline [16]	sound	-	-	-	-
IDNN [27]	sound	-	-	-	✓
Pu et al. [23]	power	✓	✓	-	-
U-net [28]	sound	-	-	-	✓
Bayram et al. [3]	sound	✓	-	✓	✓
Bai et al. [2]	power	✓	-	✓	✓
SADIS	**sound**	✓	✓	✓	✓

2.1 Anomaly detection based on power fingerprinting

Pu *et al.* [23] proposed a low-cost method that collects power traces and performs real-time malware detection using support vector machine (SVM) classifier. The authors study the parameters sensitivity. However, they do not address robustness against noise and don't compare their approach with existing methods. Bai et al. [2] designed a power-based intrusion detection system (PIDS) that uses the physically-induced dependency between a robot's movement and the concurrent power consumption to detect replay attacks. They use power data and focus on real-time detection and robustness against noise. They also compare their approach with existing methods, but don't study parameter sensitivity.

2.2 Sound-based anomly detection (SAD)

Sound-based Anomaly Detection (SAD) proves highly effective in critical infrastructure, enhancing safety, security, quality control, predictive maintenance, and cost reduction. Low-cost equipment like microphones simplifies sound data capture. Several recent studies, such as Koizumi et al. [16], Suefusa et al. [27], and van der et al. [28], propose adequate methods for monitoring industrial machines with SAD. Koizumi et al. [16] introduced an unsupervised approach within the detection and classification of acoustic scenes and events (DCASE) [2] challenge, primarily providing a baseline for malfunctioning industrial machine investigation and inspection (MIMII) dataset evaluation. In contrast, Suefusa et

[2] https://dcase.community/challenge2020/task-unsupervised-detection-of-anomalous-sounds

al. [27] used an interpolation deep neural network (IDNN)-based SAD approach and compared it to existing SAD methods, including Koizumi et al.'s baseline. Additionally, Van der Vel den et al. [28] proposed the fully connected U-net for sound-based anomaly detection, conducting comparisons with existing SAD methods, encompassing Koizumi et al.'s baseline [15] and Suefusa et al.'s [27] work. However, these studies primarily conducted offline experiments and did not explore parameter sensitivity and robustness against noise. Additionally, sound data is suitable for real-time processing with lower computational costs, enabling early anomaly detection, as suggested by Bayram et al. [3]. Notably, while related SAD studies focus on predictive maintenance, we emphasise real-time SAD for cyber-physical attack detection through side-channel leakage. Both applications identify normal/abnormal machine sounds but serve different purposes: predictive maintenance SAD prevents failures by spotting malfunctions. In contrast, cyber-physical attack detection SAD rapidly identifies abnormal sounds tied to attacks, including zero-day attacks.

3 Preliminaries

This section introduces deep learning concepts, the autoencoder model background and anomalies scoring function Mahalanobis distance (MD).

3.1 Deep learning for anomaly detection

Deep learning, a subset of machine learning, outperforms anomaly detection for complex data through layered neural networks [7]. Like the human brain, deep neural networks are powerful non-linear models with input, hidden, and output layers [4]. This paper employs an unsupervised algorithm, using reconstruction-based methods for detecting unknown attacks and addressing unbalanced data.

Autoencoder for anomaly detection The reconstruction error of an autoencoder is used to identify anomalies. A higher reconstruction error indicates a higher likelihood that the input data is an anomaly. The SADIS approach uses a deep autoencoder to detect abnormalities in industrial sounds in real-time. A deep autoencoder is simply an autoencoder with multiple hidden layers [33].

Background of autoencoder An autoencoder (AE) is a neural network designed to reconstruct input data with high similarity using an encoder function ϕ and a decoder function φ, incorporating a non-linear activation function f. The encoder maps the original data x from the input layer to a lower-dimensional latent space (z) at the hidden layer as follows:

$$\begin{aligned} \phi : \mathbb{R}^n &\longrightarrow \mathbb{R}^d \\ x &\longmapsto z = \phi_{\Theta}(x) = f(Wx + b); \end{aligned} \tag{1}$$

where $d < n$, and n is the dimension of the input data x. $\Theta = \{W, b\}$ represents the encoder network parameters, with W as the weight and b as the bias. The

decoder function is defined as follows:

$$\begin{aligned} \varphi : \mathbb{R}^d &\longrightarrow \mathbb{R}^n \\ z &\longmapsto \hat{x} = \varphi_{\Theta'}(z) = f(W^{'} z + b^{'}) \end{aligned} \tag{2}$$

Where $\hat{x}$ is the reconstructed data. $\Theta^{'} = \{W^{'}, b^{'}\}$ represents the decoder network parameters, with $W^{'}$ as the weight and $b^{'}$ as the bias.

Non-linear activations (e.g., sigmoid, ReLU, Tanh) are used to introduce non-linearity and capture complex relationships between the input and output data [26]. SADIS uses ReLU (Rectified Linear Unit) in encoding and decoding, known for efficient and selective neuron activation. Mathematically, ReLU is defined as follows:

$$ReLU(y) = max(0, y) \tag{3}$$

Where $y \in \mathbb{R}$. During training, the autoencoder learns the encoder and decoder functions through backpropagation to minimize the reconstruction error, typically quantified by the mean squared error (MSE) loss function [6].

$$\mathcal{L}(x, \hat{x}) = \frac{1}{n} \sum ||x - \hat{x}||^2; \tag{4}$$

where $||.||$ is a commonly choosen $l2$-norm. Stochastic Gradient Descent (SGD) or Adam optimize Θ and Θ' iteratively, reducing loss and enhancing input reconstruction. Objective: find optimal encoder network and decoder network parameters for minimal loss during training.

3.2 Anomalies scoring function

Anomaly detection relies on scoring functions to differentiate normal from abnormal data. Commonly, the reconstruction error (MSE), denoted as e (Equation (4)), compute the difference between input x with reconstruction $\hat{x}$. Recent research highlights the efficiency of MD as an alternative [9,14], measuring similarity between vector e and distribution Δ. The MD is defined as:

$$MD(e, \Delta) = \sqrt{(e - \mu_\Delta)^T V_\Delta^{-1} (x - \mu_\Delta)}. \tag{5}$$

The MD is computed with Equation (5), where e is an observation vector, μ_Δ and V_Δ^{-1} are respectively mean value and inverse covariance matrix of a distribution. The SADIS approach efficiently detects anomalies using MD as a scoring function.

4 The SADIS approach

This section presents SADIS, the proposed real-time SAD approach. This approach addresses the limitations faced by existing methods in detecting anomalies in ICPS, including high computational costs, low detection accuracy, and

poor robustness to noisy data and domain-shifted conditions. The SADIS approach employs a preprocessing method to extract relevant features from sound data by reducing dimensions using the methods of the time-average spectrum (TAS) or principal component analysis (PCA). Additionally, the SADIS approach employs an unsupervised neural network model, the autoencoder (AE), as a classifier. It uses an anomaly scoring function based on reconstruction loss and MD to set the optimal boundary between normal and abnormal data.

4.1 Overview of SADIS appraoch

As depicted in Figure 1, SADIS is trained on normal sound data using unsupervised learning during the training phase. The training sound data are preprocessed using the equations (7),(8), (9) defined in section 4.2. Then, the preprocessed sound data are fitted to the autoencoder model, which learns relevant features and minimizes reconstruction error using the equation (4) defined in 3.1. The distributtion of normal reconstruction errors sets the threshold for anomaly detection by using equations (5). In online detection 4.4, incoming sound data is preprocessed and compared to the threshold to classify it as abnormal or normal. SADIS is unique in its feature-extracting preprocessing, autoencoder's efficiency, and accurate anomaly scoring for raw sound signals regardless of duration.

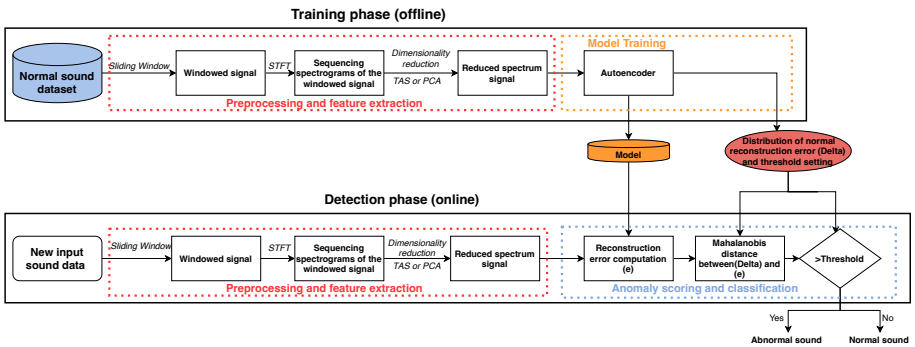

Fig. 1. Overview of the proposed real-time SADIS approach

4.2 Preprocessing step

The sound dataset, denoted by X, comprises N audio signals represented by raw audio signals x_i:

$$X = \{x_1, x_2, \ldots, x_N\}. \tag{6}$$

Each audio signal x_i is composed of n time series samples. In anomaly detection, preprocessing involves dimensionality reduction and feature extraction to

enhance classification algorithms' performance and efficiency. In this paper, a sliding window method is used during the initial data preprocessing phase to detect deviant points by capturing sequential dependencies.

Sliding Window The sliding window algorithm (SW) is an effective preprocessing technique for real-time anomaly detection in time series data. After applying the sliding window algorithm to each audio signal $x_i \in X$, it can be represented as:

$$SW(x_i) = W = \{w_1, \ldots, w_l\}; \tag{7}$$

where $w_j = [j \cdot (k-t), (j+1) \cdot (k-t) + t]$ contains k samples of the raw signal, k corresponds to the window length, and t is the overlap size between windows. The value of l depends on the time window-length k and its overlap t and is such that $l < n$. This windowed representation enhances anomaly detection by preserving temporal dependencies and identifying patterns. This paper explores the impact of window length on the performance of the SADIS approach in terms of detection rate and time.

Feature extraction The spectrogram, obtained by applying the short-time Fourier transform (STFT) to the audio signal waveform, is a time-frequency representation widely used as a feature in various studies [3, 16]. The audio signal is divided into overlapping windows, and the STFT is computed for each window to extract its spectrogram. This results in a sequence of spectrograms, each represented by a two-dimensional array of size $m \times n^*$, where m represents frequency components and n^* represents the number of features. Thus, the windowed audio signal W can be represented as a sequence of spectrograms given by:

$$STFT(W) = \{s_1, \ldots, s_l\}; \tag{8}$$

where s_j is the spectrogram obtained from the j-th window. The windowed signal W has a total size of $l \cdot (m, n^*)$, where l represents the number of windows.

Reduction of dimensions To enhance SAD in industrial systems and reduce dimensions and noise, SADIS employs either TAS or PCA as dimensionality reduction techniques.

Time-Average spectrum (TAS) computes the average of each spectrogram [22] along the spectrogram matrix's m-dimension to minimize machine operation noise. This technique computes the average of each spectrogram along the spectrogram matrix's m-dimensional as shown by the equation:

$$s_j^{TAS} = \frac{1}{m} \sum_{i=1}^{m} s_{i,j}. \tag{9}$$

Here, $s_{i,j}$ corresponds to each spectrogram of the windowed signal, and s_j^{TAS} is the resulting time-averaged spectrogram, which is a vector of size n^*. By applying the TAS method for each spectrogram of the windowed signal W, we

concatenate all the averaged spectrograms into a matrix of dimension (l, n^*) represented by: $S^{TAS} = [s_1^{TAS}, s_2^{TAS}, \ldots, s_l^{TAS}]$.

Principal component analysis PCA identifies the principal components of the dataset by finding eigenvectors of its covariance matrix and projects the data onto a lower-dimensional space. PCA reduces each windowed signal's spectrogram s_j into a smaller vector denoted as s^{PCA}. By applying PCA to each spectrogram in the windowed signal, we obtain the following matrix: $S^{PCA} = [s_1^{PCA}, s_2^{PCA}, \ldots, s_l^{PCA}]$. The resulting signal S^{PCA} is a matrix of dimension (l, n^*), similar to the TAS reduction method described earlier.

The preprocessed signal, denoted as S, results from signal windowing, spectrogram extraction, and dimensionality reduction (TAS or PCA). S is used as the input for the anomaly detection algorithm, representing 128-dimensional spectrograms. We assess their impact on detection rate and time by varying window length and reduction methods (TAS or PCA). Through experiments, we demonstrate the significance of preprocessing methods in obtaining input data with reduced dimensions capturing relevant features of raw audio data.

4.3 Offline training phase

SADIS is trained with only normal sound data during the training phase due to its unsupervised learning. The training sound data are preprocessed using the predefined equations in 4.2. The autoencoder is trained and learns latency and relevancy features of training data while minimising the reconstruction error by using the MSE function defined in equation (4). Let S_i denotes the i-th input signal data, while $\hat{S}_i$ signifies the i-th reconstructed signal data generated by the autoencoder The computed reconstruction errors e during training are defined as follows:

$$e_i = MSE(S_i, \hat{S}_i). \tag{10}$$

From these errors e_i, we define a distribution Δ of normal reconstruction error as follows:

$$\Delta = \begin{bmatrix} e_0 \\ e_1 \\ \vdots \\ e_{n_{tr}} \end{bmatrix}. \tag{11}$$

where n_{tr} denotes the size of training data. The defined distribution of normal reconstruction errors Δ sets the threshold and scores anomalies by using the MD defined in equation (5).

Thresholding The variability and complexity of anomalies make it difficult to set a threshold σ for anomaly detection. This study uses preprocessed normal sound data not used in training to determine the threshold. The trained auto-encoder reconstructs these data as $\hat{S}_j^{\sigma}$ and computes the reconstruction

error e_i^σ using the MSE equation (4). The resulting error set E^σ is then formed as $E^\sigma = \{e_0^\sigma, \ldots, e_{n_\sigma}^\sigma\}$, where n_σ represents the size of normal sound data used for threshold setting. Using equation (5), SADIS calculates the MD between each e_i^σ and a precomputed distribution Δ (defined in equation 11), yielding a list of distance values D^σ represented as $D^\sigma = \{d_0^\sigma, \ldots, d_{n_\sigma}^\sigma\}$. After multiple threshold evaluations, the Gamma distribution percentile is employed. In line with the baseline systems in the DCASE challenge [15, 16], this approach fits a gamma distribution to D^σ scores and utilizes the inverse of the 90th percentile of the cumulative distribution function as the decision threshold σ. After offline training, the trained autoencoder model and the distribution of normal reconstruction errors Δ and the threshold σ are saved. This paper does not evaluate training time, as it occurs offline and involves a substantial amount of normal sound data.

4.4 Online detection phase

New sound data x^{new} is preprocessed using the equations from 4.2. The resulting preprocessed sound data S^{new} is given as input to the trained model. Its reconstruction error e^{new} is computed as the mean squared error (MSE) between S^{test} and the reconstructed data ($\hat{S}^{new}$). Additionally, the distance metric (MD) between e^{new} and the previously computed distribution Δ is calculated using Equation (5). If this distance value surpasses the threshold σ, the new sound data is classified as abnormal; otherwise, it is considered normal.

5 The SADIS experimentation and results

This paper addresses anomaly detection in industrial robot arms, critical components in production chains. Timely detection of anomalies is crucial for ensuring safety and security in the complex operations of robot arms and the evolving smart manufacturing technology [25].

5.1 Experimental setup

Hardware and software specifications The computational resources employed in this research for generating the model and conducting experiments consist of an *AMD EPYC 7552 48-Core* Processor CPU and *Nvidia Tesla T4* for GPU. However, extensive computational power wasn't necessary for conducting our experiments. To extract the spectrograms features, we use the *Librosa* library,while the deep autoencoder models are implemented and applied using *Keras/TensorFlow*.

5.2 Dataset

We conduct experiments on one sound public dataset MIMII and our collected sound dataset from a robot arm called RASD.

MIMII dataset The MIMII dataset [24] is widely used in SAD research for identifying malfunctions in industrial machinery. It consists of recordings from four machine types (valves, pumps, fans, and slide rails). Each 10-second recording contains both operational machine sounds and ambient noise, serving as a benchmark to assess our approach and compare it with existing methods.

Robot arm sound dataset (RASD): our sound dataset We have created an experimental platform to assess the efficacy of our methodology, generating a novel dataset of industrial robotic arm sounds using the *TinkerKit Braccio*, operated by an *Arduino Uno*. This robot arm is a real-industrial production chain, with six axes controlled by servo motors and versatile configurations for different tasks[3]. The dataset stands out for its diverse real attack vectors and reproducibility, ensuring authentic sound data generation. Our experimental platform setup and the sound dataset are described in the following paragraphs.

Normal sound behavior of RASD The robot arm moves a valuable part safely and precisely from point A to point B. Normal sound data is generated as the arm follows a predetermined trajectory at a calibrated speed, reflecting typical operation in manufacturing settings. This behavior represents a typical operation of a robot arm in real-world manufacturing settings. The recorded normal sound data accurately reflects the scenario of a functioning robot arm.

Anomalous sound generation and vectors description Generating anomalous sound data is challenging and costly due to the rarity of abnormal patterns in real life [3,10]. To address this, attack vectors from literature are used to create cycles of abnormal behavior for the robot arm, focusing on modifying speed parameters and trajectory angles [18]. Speed modification attacks, similar to those in Stuxnet [12, 19], can be hard to detect as they resemble normal behavior. By altering speed and trajectory angles, malicious actors can disrupt robotic arm operations, leading to accidents, production errors, and compromised quality. Though these attack vectors do not cover all possible abnormal behaviors, they represent primary types of attacks in the context of robotic arms. The abnormal sounds were created using different anomaly vectors. *Anomaly Vector 1* involves a slight speed increase in the base axis, resembling normal behavior with low risk of damage. *Anomaly Vector 2* has a higher acceleration in both the base axis and gripper axis, posing material risks like piece falls and affecting machine longevity. *Anomaly Vector 3* entails a speed increase across all axes, posing significant risks to facility safety, accelerating wear on the piece, and reducing machine longevity.

[3] For more details: https://store.arduino.cc/products/tinkerkit-braccio-robot

RASD structure and size The RASD dataset contains over 4260 normal recordings and over 1900 anomalous recordings. Each recording has a duration of 23 seconds, and the signals are sampled at a frequency of $48KHz$.

5.3 Evaluation criteria

Area under ROC curve The receiver operator characteristic (ROC) curve visually assesses binary classification performance by plotting true positive rate (TPR) against false positive rate (FPR). A higher AUC value indicates better differentiation between normal and abnormal data within models [17].

Detection time computation Real-time anomaly detection prioritizes the time taken to detect anomalies as a crucial metric. To compute detection time (DT) accurately according to section 4.4, we track the execution times of each step, including preprocessing time (PT), model time (MT), as well as computation of error reconstruction and Mahalanobis distance (t_{score}). Therefore, the detection time (DT) can be expressed as the sum of these three-time intervals as follows:

$$DT = PT + MT + t_{score}. \tag{12}$$

Reducing detection time ensures efficient real-time anomaly detection.

5.4 Results: the SADIS detection performance

This section evaluates the SADIS detection performance using TAS or PCA reduction methods. We assess its effectiveness on the MIMII dataset and analyze its robustness using RASD. The AUC value serves as the detection rate metric stated in 5.3.

The SADIS evaluation on MIMII dataset This experiment compares the SADIS approach with existing SAD methods, presented in section 2, using the MIMII dataset. The results show the average detection rate of TAS and PCA reduction techniques, measured by AUC values. Table 2 provides a summary of the comparison. Prior preliminary tests are conducted to select an optimal window length for the MIMII dataset. The results of this experiment show that

Table 2. The SADIS evaluation on MIMII dataset (AUC values (%)

Machine Type	Baseline [16]	Unet [28]	IDNN [27]	**SADIS**
Fan	66	80	71	**92**
Pump	73	**85**	75	**77**
Slider	85	90	90	**91**
Valve	66	84	**90**	**85**

the SADIS approach outperforms existing methods in two machine types (fan and slider) and is comparable in the remaining two. These results indicate its potential for anomaly detection across various industrial systems. These findings

also highlight the adaptability of the SADIS approach to different conditions and equipment used for sound data collection.

Robustness of the SADIS approach In both offline training and testing, we conduct experiments using RASD to assess the SADIS approach's robustness to parameter sensitivity and noises. RASD has a longer time length compared to MIMII, with 10 seconds for MIMII and 23 seconds for RASD. This preliminary experiment aims to analyze the impact of window-length variation and noise factors. By varying the window length and introducing background noise, we evaluate their effects on the SADIS approach's detection rate, time, and overall robustness against background noise.

Paramters sensitivity and selection Parameter sensitivity analysis is conducted to assess the impact of different parameters on the detection rate and time performance of the approach. Experimental results are summarized in Table 3 and Table 4. The window length is varied within a range of 0.2 to 1.25 as shown in Fig.2 considering the frequent trajectory changes and sound capturing uncertainty [2]. Table 3 and Table 4 outline the performance results and detection time results of SADIS approach using TAS and PCA reduction method respectively.

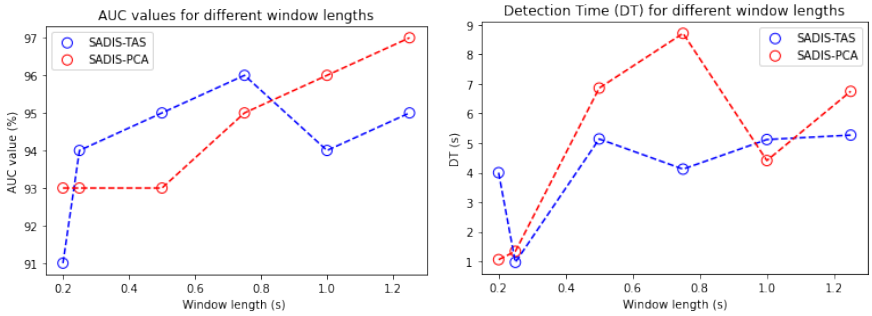

Fig. 2. Left figure depicts the AUC values (detection rate) for various window lengths of TAS and PCA, while the right figure illustrates the detection time (DT) for different window lengths of SADIS-TAS and SADIS-PCA

Table 3. paramter sentivity results of SADIS using TAS as reduction method

Window length (s)	**Performance**	**Time Taken**		
	AUC (%)	**DT(s)**	**PT(s)**	**MT(s)**
0.2	91	3.54	3.1	0.43
0.25	94	**0.97**	**0.95**	**0.01**
0.5	95	5.14	4.33	0.8
0.75	**96**	4.12	3.5	0.61
1	94	5.14	4.81	0.67
1.25	95	5.27	4.65	0.62

Table 4. paramter sentivity results of SADIS using PCA as reduction method

Window length (s)	Performance	Time Taken		
	AUC (%)	DT(s)	PT(s)	MT(s)
0.2	92	**1.06**	**1.05**	**0.01**
0.25	92.5	1.34	1.21	0.13
0.5	93	6.86	6.24	0.62
0.75	95	8.71	8.05	0.65
1	96	4.41	4.24	0.17
1.25	**97**	6.75	6.53	0.22

The results show that varying the window length has minimal impact on the detection rate (AUC values) of SADIS-TAS and SADIS-PCA, with an average TPR of 99‰ and FPR of 2%. However, the window length variation significantly affects the detection time (DT), while the t_{score} is negligible. The PT strongly influences the DT, as shown in the result Table 3 and Table 4, and the MT averages at 0.4 seconds. Based on these findings, window lengths of 0.25 seconds and 0.2 seconds are selected for TAS and PCA, respectively, using the RASD dataset, while different datasets may require varying window lengths.

Robustness of SADIS against background noise We assess SADIS's robustness to background noise through experiments using artificially synthetic test data. Introducing a noise factor (ranging from 0.1 to 0.5) to a subset of normal and abnormal test data allows us to evaluate SADIS's sensitivity to varying levels of background noise. Given that sound parameters are sensitive to background noise, and real-world industrial systems often operate in noisy environments, these experiments provide valuable insights. E.g a noise factor of 0.1 indicates relatively low background noise, while a factor of 0.5 represents higher background noise. The results of this analysis are presented in Table 5 and Table 6 and Figure 3.

Table 5. SADIS-TAS detection performance with noise factor

Noise factor	AUC (%)	DT (s)
0.1	71.6	0.42
0.2	70.8	0.42
0.3	69.2	0.42
0.4	67.0	0.41
0.5	58.0	0.42

Table 6. SADIS-PCA detection performance with noise factor

Noise factor	AUC (%)	DT (s)
0.1	82.5	0.48
0.2	82.2	0.49
0.3	81.8	0.48
0.4	81.5	0.49
0.5	81.0	0.49

PCA is less affected by background noise compared to TAS, although it has a longer computation time. These results partially address the domain shift challenge discussed in the introduction section, as the presence of background noise represents a target domain (test data) different from the source domain (train data).

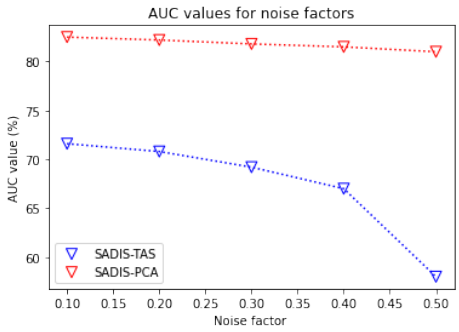

Fig. 3. SADIS performance with noised data

Real-time SADIS performance under different anomalies vectors The real-time performance of the SADIS approach is evaluated in this section, with results presented in Table 7. Three anomaly vectors are generated and described in Section 5.2. These experiments assess SADIS's applicability in real-world scenarios, detecting anomalies with background noise and domain shift conditions.

Table 7. Real-time SADIS performance under different anomalies vectors

Anomaly vector	**SADIS-TAS**				**SADIS-PCA**			
	FPR(%)	TPR(%)	AUC(%)	DT (s)	FPR(%)	TPR(%)	AUC(%)	DT (s)
1	1.2	98.4	95	1.86	2.5	95	91	1.4
2	1	97	94	2.15	3	92	85	1.35
3	**0.02**	**99.8**	**99.9**	1.90	**0**	**100**	**100**	**0.92**

The effectiveness of the SADIS approach is demonstrated in the results. The SADIS approach performs best at detecting anomaly vector 3, which is considered the most dangerous, with a detection rate of 100% for PCA and 99.9% for TAS. Additionally, the SADIS approach achieves an extremely low detection time of 1.90 seconds for TAS and 0.92 seconds for PCA. Anomaly vector 2 had slightly lower performance compared to anomaly vector 3, but still outperformed anomaly vector 1. However, it should be noted that anomaly vectors 1 and 2 are very similar to the nominal behaviour. Despite this similarity, the SADIS approach is able to detect anomalies with an average detection rate of 92% and an average detection time of 1.7 seconds. This demonstrates that the SADIS is capable of detecting abnormal behaviour of the system, even those with slight deviations. This highlights the SADIS approach's ability to detect even minor deviations, addressing the domain shift challenge and successfully detecting anomalies in the presence of background noise and other factors.

6 Conclusion

In this paper, we introduced RASD and presented SADIS, a novel real-time sound-based anomaly detection approach that effectively detects anomalies in industrial systems. With a high detection rate above 96% and a detection time of less than 1 second on average, SADIS can identify abnormal behaviours, even those with slight deviations. The SADIS approach is compatible with various industrial systems, as demonstrated by its high detection rate on the MIMII dataset. Additionally, the method is less sensitive to background noise. While the SADIS approach shows promising real-time anomaly detection performance, there are areas for improvement. Although the rate of false positives is low (less than 1%), further efforts should be made to minimize the need for human intervention in handling these false positives. Future research can explore using multiple parameters to enhance the robustness of side-channel fingerprinting for improved anomaly detection.

Acknowledge This research is part of the chair CyberCNI.fr with support of the FEDER development fund of the Brittany region.

References

1. Ahmed, C.M., MR, G.R., Mathur, A.P.: Challenges in machine learning based approaches for real-time anomaly detection in industrial control systems. In: Proceedings of the 6th ACM on cyber-physical system security workshop (2020)
2. Bai, Y., Park, J., Tehranipoor, M., Forte, D.: Real-time instruction-level verification of remote iot/cps devices via side channels. Discover Internet of Things (2022)
3. Bayram, B., Duman, T.B., Ince, G.: Real time detection of acoustic anomalies in industrial processes using sequential autoencoders. Expert Systems (2021)
4. Bishop, C.M.: Neural networks and their applications. Review of scientific instruments (1994)
5. Borges, N., Meyer, G.G.: Unsupervised distributional anomaly detection for a self-diagnostic speech activity detector. In: 2008 IEEE 42nd Annual Conference on Information Sciences and Systems
6. Bottou, L., et al.: Stochastic gradient learning in neural networks. Proceedings of Neuro-Nımes (1991)
7. Chalapathy, R., Chawla, S.: Deep learning for anomaly detection: A survey. arXiv preprint arXiv:1901.03407 (2019)
8. Foggia, P., Petkov, N., Saggese, A., Strisciuglio, N., Vento, M.: Audio surveillance of roads: A system for detecting anomalous sounds. IEEE transactions on intelligent transportation systems (2015)
9. Gjorgiev, L., Gievska, S.: Time series anomaly detection with variational autoencoder using mahalanobis distance. In: ICT Innovations 2020. Machine Learning and Applications: Proceedings 12. Springer
10. Henze, D., Gorishti, K., Bruegge, B., Simen, J.P.: Audioforesight: A process model for audio predictive maintenance in industrial environments. In: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)

11. Hubballi, N., Suryanarayanan, V.: False alarm minimization techniques in signature-based intrusion detection systems: A survey. Computer Communications (2014)
12. Jin, C.: Cryptographic Solutions for Cyber-Physical System Security. Ph.D. thesis, University of Connecticut (2019)
13. Jyothsna, V., Prasad, R., Prasad, K.M.: A review of anomaly based intrusion detection systems. International Journal of Computer Applications (2011)
14. Kamoi, R., Kobayashi, K.: Why is the mahalanobis distance effective for anomaly detection? arXiv preprint arXiv:2003.00402 (2020)
15. Kawaguchi, Y., Imoto, K., Koizumi, Y., Harada, N., Niizumi, D., Dohi, K., Tanabe, R., Purohit, H., Endo, T.: Description and discussion on dcase 2021 challenge task 2: Unsupervised anomalous sound detection for machine condition monitoring under domain shifted conditions. arXiv preprint arXiv:2106.04492 (2021)
16. Koizumi, Y., Kawaguchi, Y., Imoto, K., Nakamura, T., Nikaido, Y., Tanabe, R., Purohit, H., Suefusa, K., Endo, T., Yasuda, M., et al.: Description and discussion on dcase2020 challenge task2: Unsupervised anomalous sound detection for machine condition monitoring. arXiv preprint arXiv:2006.05822 (2020)
17. Kumar, R., Indrayan, A.: Receiver operating characteristic (roc) curve for medical researchers. Indian pediatrics (2011)
18. Maggi, F., Quarta, D., Pogliani, M., Polino, M., Zanchettin, A.M., Zanero, S.: Rogue robots: Testing the limits of an industrial robot's security. Trend Micro, Politecnico di Milano, Tech. Rep (2017)
19. Matrosov, A., Rodionov, E., Harley, D., Malcho, J.: Stuxnet under the microscope. ESET LLC (September 2010)
20. Mnasri, Z., Rovetta, S., Masulli, F.: Anomalous sound event detection: A survey of machine learning based methods and applications. Multimedia Tools and Applications (2021)
21. Novak, T., Gerstinger, A.: Safety-and security-critical services in building automation and control systems. IEEE Transactions on Industrial Electronics (2009)
22. Park, Y., Yun, I.D.: Fast adaptive rnn encoder–decoder for anomaly detection in smd assembly machine. Sensors (2018)
23. Pu, H., He, L., Zhao, C., Yau, D.K., Cheng, P., Chen, J.: Detecting replay attacks against industrial robots via power fingerprinting. In: Proceedings of the 18th Conference on Embedded Networked Sensor Systems (2020)
24. Purohit, H., Tanabe, R., Ichige, K., Endo, T., Nikaido, Y., Suefusa, K., Kawaguchi, Y.: Mimii dataset: Sound dataset for malfunctioning industrial machine investigation and inspection. arXiv preprint arXiv:1909.09347 (2019)
25. Quarta, D., Pogliani, M., Polino, M., Maggi, F., Zanchettin, A.M., Zanero, S.: An experimental security analysis of an industrial robot controller. In: 2017 IEEE Symposium on Security and Privacy (SP) (2017)
26. Sharma, S., Sharma, S., Athaiya, A.: Activation functions in neural networks. Towards Data Sci (2017)
27. Suefusa, K., Nishida, T., Purohit, H., Tanabe, R., Endo, T., Kawaguchi, Y.: Anomalous sound detection based on interpolation deep neural network. In: ICASSP 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
28. Van Truong, H., Hieu, N.C., Giao, P.N., Phong, N.X.: Unsupervised detection of anomalous sound for machine condition monitoring using fully connected u-net. Journal of ICT Research & Applications (2021)

29. Wolsing, K., Thiemt, L., Sloun, C.v., Wagner, E., Wehrle, K., Henze, M.: Can industrial intrusion detection be simple? In: Computer Security–ESORICS 2022, Proceedings, Part III. Springer
30. Xia, X., Togneri, R., Sohel, F., Zhao, Y., Huang, D.: A survey: neural network-based deep learning for acoustic event detection. Circuits, Systems, and Signal Processing (2019)
31. Yang, J., Zhou, C., Yang, S., Xu, H., Hu, B.: Anomaly detection based on zone partition for security protection of industrial cyber-physical systems. IEEE Transactions on Industrial Electronics (2017)
32. Ye, J., Kobayashi, T., Higuchi, T.: Smart audio sensor on anomaly respiration detection using flac features. In: 2012 IEEE Sensors Applications Symposium Proceedings
33. Zhou, C., Paffenroth, R.C.: Anomaly detection with robust deep autoencoders. In: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining (2017)