ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

INTEGRATING SYSTEM KNOWLEDGE IN UNSUPERVISED ANOMALY DETECTION ALGORITHMS FOR SIMULATION-BASED FAILURE PREDICTION OF ELECTRONIC CIRCUITS

Felix Waldhauser[1,2,*], Hamza Boukabache[1], Daniel Perrin[1], Stefan Roesler[1], Martin Dazer[2]
[1]CERN, Geneva, Switzerland
[2]Institute of Machine Components, University of Stuttgart, Stuttgart, Germany

Abstract

Machine learning algorithms enable failure prediction of large-scale, distributed systems using historical time-series datasets. Although unsupervised learning algorithms represent a possibility to detect an evolving variety of anomalies, they do not provide links between detected data events and system failures. Additional system knowledge is required for machine learning algorithms to determine the nature of detected anomalies, which may represent either healthy system behavior or failure precursors. However, knowledge on failure behavior is expensive to obtain and might only be available upon pre-selection of anomalous system states using unsupervised algorithms. Moreover, system knowledge obtained from evaluation of system states needs to be appropriately provided to the algorithms to enable performance improvements. In this paper, we will present an approach to efficiently configure the integration of system knowledge into unsupervised anomaly detection algorithms for failure prediction. The methodology is based on simulations of failure modes of electronic circuits. Triggering system failures based on synthetically generated failure behaviors enables analysis of the detectability of failures and generation of different types of datasets containing system knowledge. In this way, the requirements for type and extend of system knowledge from different sources can be determined, and suitable algorithms allowing the integration of additional data can be identified.

INTRODUCTION

Electronic circuits used for safety-critical equipment must be fail-safe systems and meet high reliability requirements by design. Fail-safe radiation monitoring devices used in particle accelerator environments trigger interlocks upon failure detection which cut the particle beam. While this behavior is essential in terms of safety, it also affects the availability of the accelerators and experiments. Continuous condition monitoring and failure prediction allows to perform predictive maintenance and thus to reduce accelerator downtime caused by unexpected failures. This study pursues a data-driven approach allowing online system-level failure prediction of distributed electronic systems with numerous instances. However, characteristics for fault detection at a system level are often complex, diverse and depend on the circuit layout. A common approach to this problem is to use unsupervised anomaly detection algorithms to identify outlying data samples. In this case, anomalies are detected empirically as deviating from the majority of samples but without considering their relation to the system condition or failure states. Hence, the presented approach aims at integrating system knowledge into unsupervised anomaly detection algorithms to establish the link between detected anomalies and true failure states. By quantifying the increase in performance of anomaly detection algorithms, the potential of various types of system knowledge for failure prediction can be assessed. The benchmarking of the algorithms is based on datasets obtained from SPICE simulations of an electronic radiation monitoring device. This novel approach combines failure simulations of electronic circuits and anomaly detection algorithms to select the most suitable type and extend of system knowledge for a failure prediction use case. As a result, the need for resources to generate system knowledge can be minimized, and predictive maintenance algorithms can be targeted at an early stage to reduce unexpected failures and thus increase the availability of large-scale systems.

FRAMEWORK FOR SIMULATION-BASED FAILURE PREDICTION

Selecting the most promising approach for integrating system knowledge into anomaly detection algorithms requires quantifying the associated change in classification performance. Therefore, a fully labeled dataset to benchmark model predictions is essential. However, in most predictive maintenance use cases, labeled faults are scarce as the system failure conditions can be diverse and evolve over time, making it difficult to obtain labels for data events, especially at the transition between normal and abnormal states [1]. Hence, the presented study is based on SPICE simulations of a demonstrator circuit allowing full control of the circuit state on a component-level through netlist modifications. Using simulations as data source provides labeled datasets of both the healthy and faulty system state, and allows to generate various types of system knowledge.

The presented approach combines existing methods for failure simulation and anomaly detection to form a novel methodology allowing efficient configuration of the integration of system knowledge into failure prediction algorithms. The methodology can be subdivided into four main steps, as shown in Fig. 1. Step I requires definition of the failure behavior of the circuit components to generate a benchmarking dataset using simulations of the healthy and faulty state. The detectability study (Step II) aims at evaluating the ef-

* felix.johannes.waldhauser@cern.ch

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

fects of the component degradation on electrical signals. Information on failure detectability as well as common characteristics of simulated faults can be useful for evaluation of the algorithm's performance. In step III, a simulation-based process to generate different types of system knowledge needs to be established. Furthermore, the means of supplying this knowledge to selected algorithms which allow the integration of additional data needs to be defined. In this study, the term system knowledge refers to labeled data samples, i.e., information about the true behavior of the circuit for a specific state. As an example, labels can be assigned to data samples by a system expert reviewing high-ranked anomalies. The last step is to benchmark the approaches by quantifying the increase in performance resulting from integration of the various types of system knowledge.

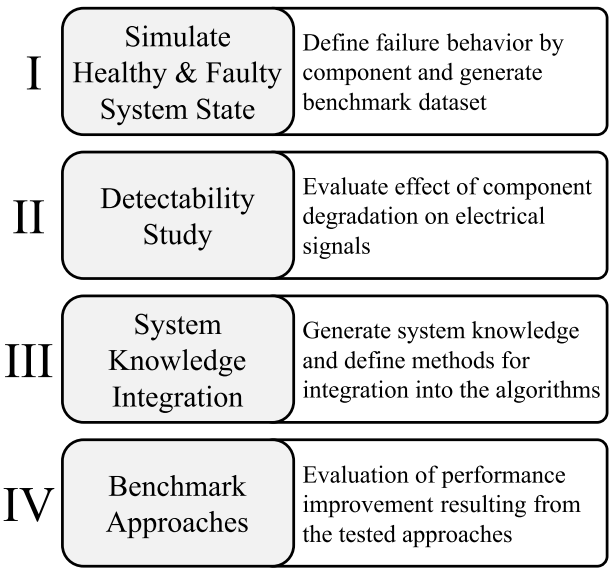

Figure 1: Methodology for benchmarking of simulation-based failure prediction.

Failure behavior and common failure characteristics identified by analysis of the simulations are limited to the component models used in the simulations and are thus not necessarily related to failure characteristics observed in tests of physical devices. Hence, this paper does not intend to study failure characteristics of a specific circuit but rather studies possibilities to enhance failure prediction algorithms by supplying system knowledge.

BACKGROUND AND RELATED WORK

The demonstrator circuit used in this study is one element of the CERN Radiation Monitoring Electronics (CROME) [2]. Mixed-field radiation monitoring systems, such as CROME, are installed along the accelerator chain at CERN to protect people and the environment from unjustified exposure to ionizing radiation. CROME is the newest generation of radiation monitoring devices and is based on a System-on-Chip (SoC) architecture. Being an occupational safety system, CROME devices are connected to the interlock system and cut particle beams if a defined radiation threshold is exceeded. Due to the fail-safe system design, which is imperative for safety, detected internal failures will likewise trigger interlocks and can thus directly impair the availability of CERN experiments. Despite extremely low failure rates and high failure detectability by design, predictive maintenance of CROME devices aims at minimizing unexpected failures by continuous, data-driven failure prediction and condition monitoring.

It has been shown in [3] that unsupervised anomaly detection algorithms can identify outlying data events in CROME operational time-series datasets. However, it remains unclear whether detected anomalies represent failure precursors - which would require maintenance actions - or rare events which are technically anomalous but not related to component degradation or system failures. To establish this link, unsupervised algorithms are extended by system knowledge (and thus become technically semi-supervised algorithms) to improve the classification performance and thus the probability that a detected anomaly is failure-related.

Failure simulations of analog circuits, as shown in this study, are commonly used to benchmark failure detection algorithms [4, 5], to evaluate failure detectability [6], or to make design improvements for higher reliability [7]. As physics-of-failure simulations are often complex, many studies simulate failures based on changes of component properties, i.e., by defining explicit value ranges for fault states deviating from the nominal values [4, 8, 9]. Most system-level failure detection studies consider binary classification tasks, i.e., differentiating between the normal and faulty state [4, 5, 9], whereas [10] train a multi-step algorithm to localize failures.

To evaluate approaches for using system knowledge for failure prediction, this study considers two types of algorithms: Isolation Forest and autoencoder. The Isolation Forest algorithm [11] is an unsupervised algorithm based on isolation trees assuming that anomalies are rare and different from the rest of the data and can thus be isolated earlier than normal samples. The Hybrid Isolation Forest (HIF) [12] is a semi-supervised version of the Isolation Forest providing two extensions: an unsupervised extension calculating the distance to normal data and a supervised extension calculating the distance to labeled anomalies. Hence, it allows the integration of known true anomalies, e.g., labeled by a system expert. Both extensions make use of the decision tree-like data separation in isolation trees: the dataset is split on each node with respect to one feature until each sample reaches its terminal node. The unsupervised extension calculates the centroid of normal data and uses the distance of the terminal node of a sample to this centroid to introduce a score s_c. The supervised extension calculates the centroid of the provided labeled anomalies and uses the distance between the terminal node of a sample and this centroid to calculate s_a. Combining the normalized scores $\tilde{s_c}$ and $\tilde{s_a}$ weighted by α_1 and α_2 with the normalized score of the standard Isolation Forest $\tilde{s}$ results in the HIF anomaly score s_{hif} for a data sample x and sub-sampling size n (Eq. (1)

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

according to [12]).

$$s_{hif}(x,n) = \alpha_2 \cdot (\alpha_1 \cdot \tilde{s}(x,n) + (1-\alpha_1) \cdot \tilde{s_c}(x)) + (1-\alpha_2) \cdot \tilde{s_a}(x) \quad (1)$$

The autoencoder is a deep neural network consisting of an encoder and decoder network which are connected by a low-dimensional middle layer. Samples are projected to the middle layer forming a low-dimensional latent representation of the data. Autoencoders are trained to learn the latent representation of normal data allowing reconstruction with low errors whereas anomalies will result in higher reconstruction errors. Most semi-supervised autoencoder approaches use labeled (i.e., failure-free) normal data to train the autoencoder models (e.g., [13, 14]) in contrast to fully unsupervised approaches using operational datasets which may be contaminated by faulty system states.

FAILURE SIMULATION USING SPICE

The demonstrator circuit used in this study is one stage of the power supply circuit of the CROME system. The 24 V input voltage is filtered and fed to a DC/DC converter (Fig. 2) which supplies two voltage regulators with 5 V which in turn generate output voltages at 3.3 V and 2.5 V.

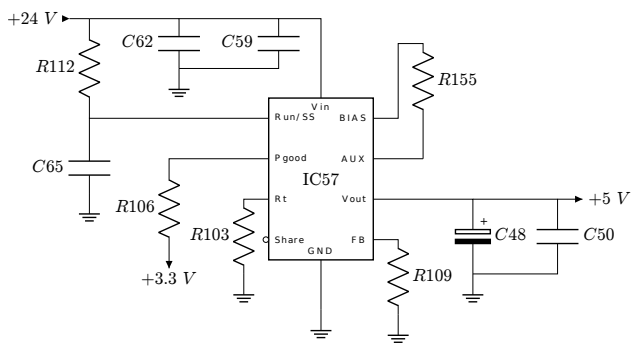

Figure 2: Snippet of the schematics of the demonstrator circuit showing the DC/DC converter.

Simulations of the healthy system state allow variations of the component values within the respective tolerance range. Failures considered in this study are limited to soft faults, i.e., faults caused by component degradation leading to changes of the component's properties. Degradation is therefore simulated using value ranges for component parameters which cover the interval between one end of the tolerance range and a defined end-of-life (EOL) criterion. The example in Fig. 3 refers to the capacitor C50 (see Fig. 2) with a nominal value of 100 nF and 10 % tolerance range while a reduction of the capacitance by 20 % is considered as EOL. Failure simulations are limited to one failing component at a time. Simulations are therefore classified as "healthy" if the values of all components are within the tolerance ranges. If the value of one component is outside of the tolerance range, i.e., in the degradation range, the simulation is classified as "failure".

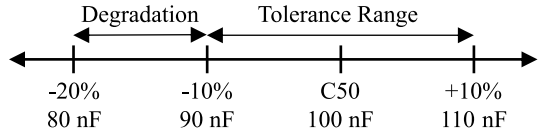

Figure 3: Value range for capacitance degradation of C50.

The EOL critera are defined based on the specific type of each component. Capacitors are the most critical components of electronic circuits in terms of reliability [15] while changes in capacitance or series resistance are usually good indicators for the capacitor's condition [15–17]. This study considers capacitor failures only but could be extended to other component types, e.g., resistor or inductor degradation, if a specific value range for degradation is provided.

The capacitors of the demonstrator circuit can be grouped into two categories: Multilayer Ceramic Capacitors (MLCCs) and Tantalum electrolytic capacitors (Ta capacitors). Typical EOL criteria are 10 % reduction of the capacitance C for MLCCs [16, 17] and 20 % reduction for electrolytic capacitors [18]. Changes of the series resistance R_{ESR} are usually neglected for MLCCs, while a typical EOL criterion for electrolytic capacitors is an increase of R_{ESR} by a factor of 2 [17]. However, as the tolerance range is ±10 % for most MLCCs of the demonstrator circuit and ±20 % for Ta capacitors, the EOL criteria are shifted towards more advanced degradation, as shown in Table 1.

Table 1: Selected EOL Criteria for the Simulation of Component Degradation

Component Type	End-of-Life (EOL) Criterion
MLCC	$C/C_0 = 80\,\%$
Ta capacitor	$C/C_0 = 70\,\%$ $R_{ESR}/R_{ESR,0} = 2$

Component states for healthy and faulty simulations are controlled by manipulating the SPICE netlist of the circuit using a custom netlist-manipulator script handling both the tolerance and the degradation simulations. Simulations can be divided into runs and iterations, where one run consists of 6 iterations. For simulations of the healthy state, the component values are varied within the tolerance range at each iteration. For fault simulations, the degradation of the failing component is divided into steps, each step corresponding to one iteration, ultimately covering the entire degradation range in one run. At each step of the failure simulations, the parameters of the remaining components are varied within the tolerance range. This ensures large variation of the simulation results and allows attributing characteristics to component degradation while mitigating other interfering effects. The distribution of component values within the tolerance range is assumed to be normal which is implemented as a truncated normal distribution covering 3σ (99.73 %) based on [19].

Manipulated netlists are fed to the SPICE engine and simulations are executed on CERN OpenStack VMs ded-

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

icated for compute intensive workloads. Simulations are performed in transient mode for a duration of 10 ms and data of a selection of current and voltage signals is collected and stored for further processing. Datasets used throughout this study consist of statistical measures describing the characteristics of the raw electrical signals logged during simulations. Those measures are in the following referred to as features. Calculated features are common statistical measures for signal classification: mean, standard deviation, Root Mean Square (RMS), number of zero crossings, signal line length, kurtosis and skewness.

FAILURE DETECTABILITY

Restricting the failure simulations to one failing component at a time allows to evaluate the effect of the degradation on the electrical signals for each component and failure type. As a result, it can be assessed whether and how the degradation of a component can be observed by analysis of the output signals. However, this analysis only relates to direct effects while hidden characteristics may still be detected by deep learning algorithms.

Analysis of the datasets resulting from simulations of the healthy and faulty states reveals that many component failures have little or no effect on feature values when comparing them to the healthy state. This can be seen in Fig. 4 showing the values of one feature (*V(tp15)_line_length*, the line length of the voltage signal at node *tp15*) against all simulated failure states and the healthy state (no failing component). The naming of the failure states follows the pattern *component:type* whereas the type *value* refers to degrading capacitance and *R_ESR* to increasing series resistance.

In Fig. 4, the feature values of the failure states C125:value, C126:value, and C48:R_ESR partially deviate from the range of values of the healthy state, which may enable identification of those failures. However, for the remaining failure states, feature values are largely overlapping with those of the healthy state. This behavior applies to most calculated features with deviations always relating to a limited group of components. Common characteristics of this group of failures are exceptionally high signal line length values. Nevertheless, most failures are likely to have little or no direct impact on the characteristics of voltage or current signals and may thus be masked as healthy making them hard to detect.

INTEGRATION OF SYSTEM KNOWLEDGE

To evaluate the potential improvement in classification performance resulting from supplying different types of system knowledge to anomaly detection algorithms, various approaches are implemented and compared against. An overview of algorithms and datasets is shown in Fig. 5. The schematic illustrates the generation of datasets as well as methods to integrate system knowledge into the algorithms, i.e., unsupervised and semi-supervised versions of autoencoder and Isolation Forest algorithms. Table 2 provides

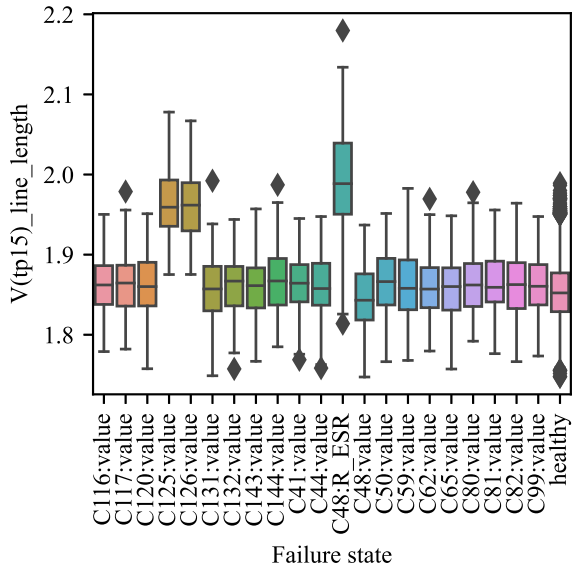

Figure 4: Feature values of V(tp15)_line_length for all simulated failure states and the healthy state.

an overview of the selected algorithms and the respective sources of supplied system knowledge. Synthetic tests refer to the simulation-based generation of samples for healthy or faulty system states. Expert review relates to assigning labels to selected samples. Comparison of the classification performance is done by calculating the Area Under Curve (AUC) score. AUC is selected as measure for quantifying the classification performance of all algorithms as it is well-suited for evaluation of anomaly detection algorithms and invariant to the classification threshold [20].

Table 2: Summary of Algorithms. AE = Autoencoder, IF = Isolation Forest, HIF = Hybrid Isolation Forest

Algorithm	System Knowledge Source
AE_pseudo_healthy	None (unsupervised)
AE_healthytest	Synthetic tests of healthy devices
IF_standard	None (unsupervised)
HIF_random	Expert review, random selection
HIF_score	Expert review, score-based selection
HIF_failuretest	Synthetic tests triggering failures

Datasets used in this study are obtained from simulations of the healthy and faulty states of the demonstrator circuit and consist of statistical measures quantifying the signal characteristics. The contamination of the initial dataset *full_data*, i.e., the percentage of samples that originate from failure simulations, is set to 5 %.

The unsupervised version of the autoencoder (*AE_pseudo_healthy*) is trained on data that is assumed to represent the healthy state of the system, but may nonetheless be contaminated with faults. This pseudo healthy training data is obtained by selecting the least

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

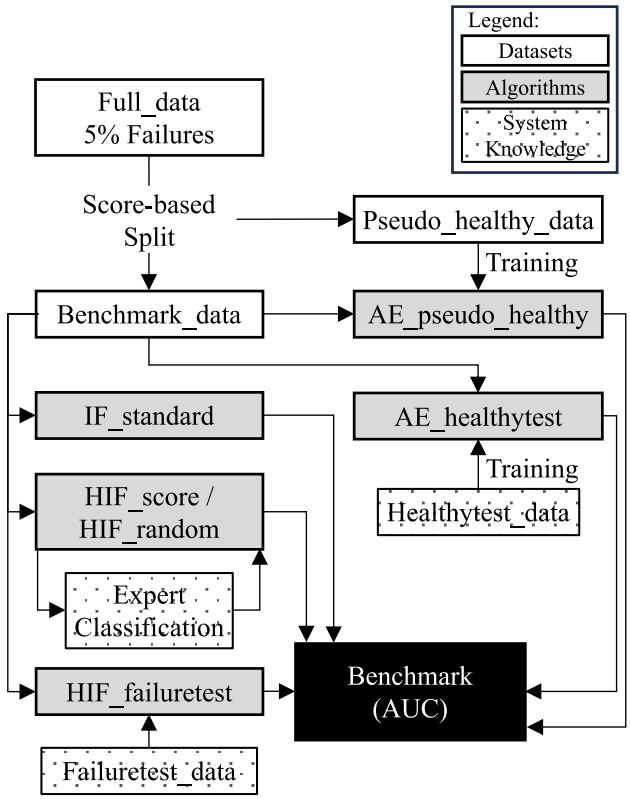

Figure 5: Schematic process for dataset generation, integration of system knowledge, and benchmarking of the algorithms.

anomalous samples from *full_data*. More precisely, a standard unsupervised Isolation Forest algorithm is applied to *full_data* and the samples with the lowest anomaly score are extracted to *pseudo_healthy_data*. The remaining samples are used as main benchmark dataset (*benchmark_data*) for all algorithms. This split is required to allow unsupervised training of the *AE_pseudo_healthy* autoencoder while testing all algorithms on the same benchmark dataset.

AE_healthytest is a semi-supervised autoencoder algorithm which is trained on system knowledge data collected from simulated tests of virtual healthy devices (*healthytest_data*). Here, virtual instances of the demonstrator circuit are created based on the component tolerances. Simulations of these instances represent hardware tests of healthy physical instances of the circuit used to generate data about the normal, i.e., healthy, state. Both autencoders (*AE_pseudo_healthy* and *AE_healthytest*) have the same configuration with 4 hidden layers for both the encoder and decoder network, 50 input dimensions, and a two-dimensional latent space.

Besides autoencoder models, four Isolation Forest-based algorithms are evaluated: the standard unsupervised Isolation Forest and three versions of the HIF with labeled faults provided from different sources. Since the parameters α_1 and α_2 (see Eq. (1)) affect the anomaly scores of the HIF algorithm and thus the classification performance, the configuration of α_1 and α_2 leading to the maximum AUC score is selected.

The unsupervised *IF_standard* is applied directly to the *benchmark_data* dataset to assess its classification performance. All HIF algorithms are likewise applied to *benchmark_data* but are in addition provided with labeled fault samples. Faults supplied to the *HIF_score* algorithm are selected based on the anomaly scores of *IF_standard*. This represents a manual review process with a system expert reviewing samples with high anomaly score. Labeled faults are then supplied back to the algorithm. The *HIF_random* follows a similar approach, but instead of using an anomaly score, labeled faults are selected randomly from *benchmark_data*. *HIF_failuretest* is supplied with labeled faults from simulated tests of faulty devices (*failuretest_data*). This is done by creating virtual instances of the demonstrator circuit based on the defined tolerances, triggering failure modes during simulation and simulating the degradation using the EOL criteria defined in Table 1. These simulations represent reliability tests of physical instances which aim at triggering specific failure modes to collect data about the system failure behavior. Here, samples with highest degradation are supplied first to *HIF_failuretest*.

BENCHMARKING RESULTS

Analyzing the classification performance of the algorithms shown in Fig. 5 allows to detect potential improvements and to evaluate which algorithms responds best to the integration of system knowledge. As generation of system knowledge might be expensive and costs increase with more data supplied to the algorithms, the performance is analyzed with respect to the size of the dataset supplied as system knowledge.

For the semi-supervised autoencoder algorithm, this size corresponds to the number of simulation runs of virtual healthy devices (*n_healthytest*) used to train the *AE_healthytest* model. Both autoencoder algorithms show overall low performance. Nevertheless, the semi-supervised *AE_healthytest* outperforms the unsupervised *AE_pseudo_healthy* with a mean AUC difference of 0.04 (Fig. 6). The performance of *AE_healthytest* does not increase with increasing *n_healthytest*. *AE_pseudo_healthytest* is naturally not related to this parameter but is tested on the same dataset for each analysis to allow comparison.

In the case of the HIF algorithms, the size of the system knowledge dataset refers to the number of labeled faults supplied to the algorithms (*n_labeled_faults*). *IF_standard* shows the lowest performance. *HIF_random* results in the highest AUC values with a mean AUC difference of 0.05 compared to *IF_standard*, but does not improve with increasing number of labeled faults (Fig. 7). The performance of the *HIF_score* algorithm is in between the performances of *IF_standard* and *HIF_random* but decreases for more than 10 labeled faults and eventually falls below the performance of *IF_standard* for more than 60 labeled faults (not shown in plot). The performance of the *HIF_failuretest* algorithm

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

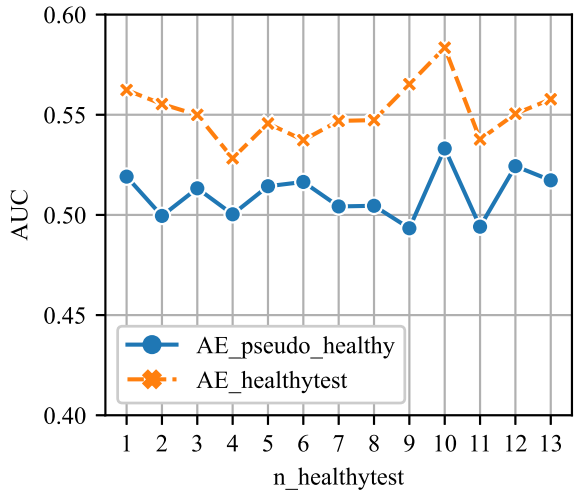

Figure 6: Comparison of the classification performance of autoencoder models with respect to the size of the system knowledge dataset.

is also in between the *IF_standard* and *HIF_random* and increases slightly with an increasing number of supplied labeled faults.

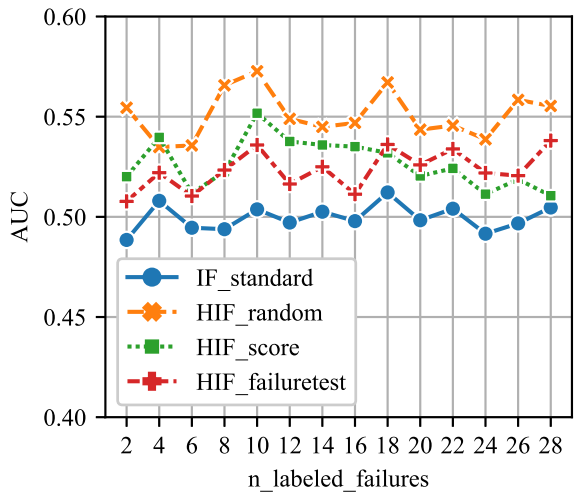

Figure 7: Comparison of the classification performance of Isolation Forest-based algorithms with respect to the size of the system knowledge dataset.

The resulting performances of all algorithms are subject to strong variations (see Fig. 6 and Fig. 7), which can be attributed to the execution process of the analysis. At each step of the analysis, the full benchmarking process as shown in Fig. 5 is executed, which is subject to many random processes, e.g., the selection of fault samples to be included in *full_data*. This can directly affect the classification performance and thus cause strong variations of the results.

DISCUSSION

The overall classification performance is low with AUC values close to 50 % while 50 % corresponds to a random classifier. Detailed analysis of the healthy versus the failure behavior showed that characteristics, i.e., values of the calculated features, of many failures are similar to those of the healthy state (see Fig. 4) making these failures hard to detect. Hence, many failures may be masked as healthy samples and lack characteristics that can be used for detection by the selected algorithms, leading to low classification performance. Furthermore, datasets used in this analysis are small as all samples are generated using computationally intensive simulations. Accordingly, limiting the analysis to detectable failures and samples related to advanced degradation as well as increasing the size of the datasets could lead to better classification performance.

It has been shown that the semi-supervised autoencoder *AE_healthytest* can benefit from labeled healthy data resulting in a higher AUC value for the given task of failure detection in datasets obtained from simulations. However, the performance does not increase if more labeled healthy samples are provided. One possible explanation is the low variance of provided samples resulting from simulations of virtual healthy devices. Here, variation in component values is limited to the initial generation of a device and generated data of one device shows therefore low variance. As a result, few simulation runs of virtual healthy devices already exploit the potential for performance improvement.

While the overall performance of the Isolation Forest and HIF algorithms is similar to the performance of the autoencoders, all HIF algorithms proved to perform better than the standard IF algorithm, already for a small number of labeled faults. Random selection of labeled faults provided to the algorithm resulted in the best performance. Selecting labeled samples randomly leads to the greatest variety in samples, i.e., the most information provided to the algorithm. For the *HIF_score* algorithm, labeled faults are high-ranked anomalies that correspond in this study mainly to a few specific failure states. Providing multiple samples of only a few failure states can lead to overfitting of the algorithm to those failures while detection of the remaining failures is reduced leading to a reduction in overall performance. Hence, labeling multiple samples that are already considered as anomalies is in this case not suitable to improve the performance. Samples provided as labeled faults to *HIF_failuretest* correspond to a variety of failures triggered by simulations representing reliability tests. Since the performance increases slightly with an increasing number of labeled samples, it can be assumed that the algorithm can benefit from this additional knowledge.

The weak dependency of the performance of the HIF algorithms on the number of labeled faults may be explained by consideration of the mechanisms used to integrate labeled faults into the HIF algorithm. The anomaly score of the HIF algorithm accounts for the distance of a sample from the centroid of the labeled faults. Thus, if a sample is close

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

to a cluster of known faults, it is likely to be faulty as well. However, the cluster of known faults becomes more scattered if the supplied labeled faults have diverse characteristics or if faults masked as healthy are assigned to the same terminal nodes as normal samples. In this case, more labeled faults will result in a greater dispersion of the fault cluster, thus reducing the effectiveness of the distance measure rather than providing additional information. This issue is addressed in [21] by constructing a multi-class HIF algorithm.

In this study, labels assigned to the samples are directly related to the simulation configuration and are therefore error-free. However, in reality, obtaining labeled faults from a manual review, e.g., by system experts, may be difficult and error-prone, as failure characteristics are not necessarily obvious or require an explicit definition of the healthy state for comparison. Especially for randomly selected samples that are in the transition between the healthy and faulty state, characteristics allowing classification may be hidden for manual analysis. Hence, for use cases with complex failure characteristics, obtaining labeled data of the healthy and faulty state from hardware tests is preferable over expert review of data samples. Using data known to be failure-free to train an autoencoder model can help to improve the classification performance. Data of the healthy state may be well-accessible for high-reliability systems with few failures or easy to obtain from hardware tests of healthy instances. Further, generating labeled faulty samples from reliability tests triggering critical failures is suitable to improve the performance of the HIF algorithm. This approach is also more promising than manual labeling of samples that already have a high anomaly score. It has been shown that the size of the dataset provided as system knowledge does not strongly affect the performance in this use case. Hence, providing high-quality data with a variety of samples, e.g., from tests of both healthy and faulty devices, is preferable over providing numerous samples with similar characteristics.

SUMMARY AND OUTLOOK

The presented methodology enables simulation-based benchmarking of failure prediction algorithms to ensure efficient usage of system knowledge and to study the interaction between system knowledge integration and algorithm performance. Defining healthy and faulty conditions of electronic components for the simulations gives full control of the system states and allows generation of several types of system knowledge. It has been shown that the detectability of failures and diversity of failure characteristics are highly relevant for the selection of the best approach. For use cases where characteristics of failures and normal samples are overlapping, making expert review difficult, obtaining system knowledge from hardware tests is the preferred option. Testing healthy devices allows generation of datasets suitable for training of autoencoder models while reliability tests can provide labeled failure samples which can be fed as system knowledge to Hybrid Isolation Forest (HIF) algorithms.

Using the presented methodology to develop data-driven failure prediction algorithms can improve the prediction performance and thus reduce unexpected failures. Therefore, the availability of large-scale systems can be increased while limiting the resources required to generate system knowledge.

In this study, the algorithm performance is measured on the detection of all simulated component failures without considering the effect of the failures on the circuit operability. Hence, the next step is to select only failures which are directly relevant for predictive maintenance, i.e., failures that affect the operability and thus need to be predicted in order to perform maintenance actions. Moreover, only failures that can be detected on system-level by analyzing the output signals should be included. This ensures that the failure prediction algorithm is measured only against detectable failures, which should improve overall classification performance. Information on failures which are critical for system operability but are considered undetectable can be useful for design improvements to increase reliability. In addition, improved feature engineering and optimization of the algorithms, e.g., by performing noise extraction or using a multi-class HIF algorithm, can increase the failure detectability and thus the algorithm performance. Moreover, the empirical observations presented in this study need to be tested for significance based on statistical design of experiments (DOE) methods.

ACKNOWLEDGEMENTS

This work has been sponsored by the Wolfgang Gentner Programme of the German Federal Ministry of Education and Research (grant no. 13E18CHA).

REFERENCES

[1] S. Zhang, F. Ye, B. Wang, and T. G. Habetler, "Semi-supervised learning of bearing anomaly detection via deep variational autoencoders", 2019, Publisher: arXiv Version Number: 2. doi:10.48550/ARXIV.1912.01096

[2] H. Boukabache *et al.*, "Towards a novel modular architecture for cern radiation monitoring", *Radiation Protection Dosimetry*, vol. 173, pp. 240–244, 2017. doi:10.1093/rpd/ncw308

[3] F. Waldhauser, H. Boukabache, D. Perrin, and M. Dazer, "Wavelet-based noise extraction for anomaly detection applied to safety-critical electronics at CERN", in *Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)*, 2022, pp. 1844–1851. doi:10.3850/978-981-18-5183-4_S02-03-080-cd

[4] F. Zhang, Z. Hong, T. Gao, and S. Yin, "A fault detection method for analog circuits based on the wavelet features and one-class KNN", in *2021 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, 2021, pp. 1–6. doi:10.1109/ICSMD53520.2021.9670762

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-TU1BCO02

[5] D. Binu and B. Karivappa, "Support vector neural network and principal component analysis for fault diagnosis of analog circuits", in *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2018, pp. 1152–1157. doi:10.1109/ICOEI.2018.8553786

[6] X. Tang, A. Xu, R. Li, M. Zhu, and J. Dai, "Simulation-based diagnostic model for automatic testability analysis of analog circuits", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 7, pp. 1483–1493, 2018. doi:10.1109/TCAD.2017.2762647

[7] J. Harikumaran *et al.*, "Failure modes and reliability oriented system design for aerospace power electronic converters", *IEEE Open Journal of the Industrial Electronics Society*, vol. 2, pp. 53–64, 2021. doi:10.1109/OJIES.2020.3047201

[8] J. Zhong and Y. He, "Fuzzy diagnosis method based on multi-frequency responses for tolerance analog circuit", in *2009 9th International Conference on Electronic Measurement & Instruments*, 2009, pp. 4–832–4–836. doi:10.1109/ICEMI.2009.5274715

[9] S. Mosin, "Automated simulation of faults in analog circuits based on parallel paradigm", in *2017 IEEE East-West Design & Test Symposium (EWDTS)*, 2017, pp. 1–6. doi:10.1109/EWDTS.2017.8110133

[10] F. Grasso, A. Luchetta, S. Manetti, M. Piccirilli, and A. Reatti, "Single fault diagnosis in analog circuits: A multi-step approach", in *2017 5th IEEE Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, 2017, pp. 1–5. doi:10.1109/AIEEE.2017.8270523

[11] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation Forest", in *2008 Eighth IEEE International Conference on Data Mining*, 2008, pp. 413–422. doi:10.1109/ICDM.2008.17

[12] P.-F. Marteau, S. Soheily-Khah, and N. Béchet, "Hybrid Isolation Forest - Application to intrusion detection", 2017, Publisher: arXiv Version Number: 1. doi:10.48550/ARXIV.1705.03800

[13] M. S. Minhas and J. Zelek, "Semi-supervised anomaly detection using AutoEncoders", 2020, Publisher: arXiv Version Number: 1. doi:10.48550/ARXIV.2001.03674

[14] J. Liu, K. Song, M. Feng, Y. Yan, Z. Tu, and L. Zhu, "Semi-supervised anomaly detection with dual prototypes autoencoder for industrial surface inspection", *Optics and Lasers in Engineering*, vol. 136, p. 106 324, 2021. doi:10.1016/j.optlaseng.2020.106324

[15] M. Ghadrdan, S. Peyghami, H. Mokhtari, H. Wang, and F. Blaabjerg, "Dissipation factor as a degradation indicator for electrolytic capacitors", *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 11, no. 1, pp. 1035–1044, 2023. doi:10.1109/JESTPE.2022.3183837

[16] H. Wang and F. Blaabjerg, "Reliability of capacitors for DC-link applications in power electronic converters—an overview", *IEEE Transactions on Industry Applications*, vol. 50, no. 5, pp. 3569–3578, 2014. doi:10.1109/TIA.2014.2308357

[17] Z. Zhao, P. Davari, W. Lu, H. Wang, and F. Blaabjerg, "An overview of condition monitoring techniques for capacitors in DC-link applications", *IEEE Transactions on Power Electronics*, vol. 36, no. 4, pp. 3692–3716, 2021. doi:10.1109/TPEL.2020.3023469

[18] J. A. Romero, M. H. Azarian, and M. Pecht, "Life model for tantalum electrolytic capacitors with conductive polymers", *Microelectronics Reliability*, vol. 104, p. 113 550, 2020. doi:10.1016/j.microrel.2019.113550

[19] R. Spence and R. S. Soin, *Tolerance design of electronic circuits*. Imperial College Press, 1997, 215 pp.

[20] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection", *ACM Transactions on Knowledge Discovery from Data*, vol. 6, no. 1, pp. 1–39, 2012. doi:10.1145/2133360.2133363

[21] C. Melquiades and F. B. De Lima Neto, "Isolation Forest-based semi-supervised anomaly detection of multiple classes", in *2022 17th Iberian Conference on Information Systems and Technologies (CISTI)*, 2022, pp. 1–6. doi:10.23919/CISTI54924.2022.9820467