Isolation Forest Based Anomaly Detection: A Systematic Literature Review

Wahid Salman Al Farizi
Department of Electrical and Information Engineering
Universitas Gadjah Mada
Yogyakarta, Indonesia
wahid.salman.af@mail.ugm.ac.id

Indriana Hidayah
Department of Electrical and Information Engineering
Universitas Gadjah Mada
Yogyakarta, Indonesia
indriana.h@ugm.ac.id

Muhammad Nur Rizal
Department of Electrical and Information Engineering
Universitas Gadjah Mada
Yogyakarta, Indonesia
mnrizal@ugm.ac.id

2021 8th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE) | 978-1-6654-3998-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICITACEE53184.2021.9617498

***Abstract*— Anomaly detection using machine learning algorithms is rising lately, especially with increased data volume and velocity. One of the most recent anomaly detection algorithms is Isolation Forest (IF). Despite its simplicity, it excels at dealing with high-dimensional data and excels at speed. However, IF is not without weaknesses, and several researchers have found its weaknesses and at the same time provide solutions. Therefore, to help understanding researches related to IF, this paper will discuss 17 studies related to IF improvement by conducting a systematic literature review that comprehensively discusses IF weaknesses, types of data, and causes of occurrence, as well as dissecting the solutions offered and the fields of research that use IF. From the review, it is known that the main cause of the weakness of IF is the random selections of variables in the data split and the solutions proposed by the researchers are divided into three types: pre-IF, post-IF and method improvement. To our knowledge, there is no literature review related to IF improvement, and we expect this paper to help other researchers in developing anomaly detection based on IF.**

Keywords*— *anomaly detection, isolation forest, systematic literature review

I. INTRODUCTION

Several researchers have developed Anomaly Detection (AD). Many types of AD methods such as density-based, distance-based, neural network, and spectral-based [1]. These algorithms were initially developed for machine learning aimed at classification or clustering and were modified for AD. Until Liu [2] introducing Isolation Forest (IF), an algorithm specially developed for AD.

Domingues et al. [3] evaluated several up-to-date anomaly detection algorithms, including Gaussian Mixture Model (GMM), Kernel Density Estimator (KDE), Mahalanobis Distance, Local Outlier Factor (LOF), One-class Support Vector Machine (OCSVM), and Isolation Forest (IF). The evaluation used 15 different datasets from public data UCI or OpenML and private data. The result has shown that IF outperforms another algorithm in terms of accuracy and time while also demonstrate satisfactory performance in handling high-dimensional datasets. However, several recent studies reported IF's weaknesses and developed anomaly detection algorithm based on IF to overcome these weaknesses.

This paper discusses some of the latest research related to IF's application in various fields and solutions developed to overcome IF's weaknesses. The most prominent problems are the low accuracy in detecting local outliers or conditional outliers [4], [5], time wastage due to the random selection for feature split [6]–[8], and reduced accuracy on a high-dimensional dataset with low-dimensional anomaly data [9], [10]. In addition to the problems in IF, this paper also presents the development of IF for use in various fields such as fraud detection [4], fault detection on manufacture machines [9], [11], [12], attacks on network systems [10], [13], [14]and remote sensing [15], [16].

This paper will conduct a systematic literature review (SLR) based on Kitchenham and Charters' guidelines [17] and divide SLR into three main phases: planning, conducting the review, and reporting. Section. II will introduce the concept of IF, Section. III will discuss planning, Section. IV discussed conducting review and Section. V presented reporting as a result of the review.

II. ISOLATION FOREST

The main concepts of IF utilize two main characteristics of anomaly. i.e., anomalies are few and have different attributes[2]. IF split data into two parts as in a binary tree, splitting is done to limit a certain tree height or until the data can no longer be split. The structure formed is called iTree [2]. Samples isolated at the start of the split or closer to the root of iTree have the potential to become anomalous. In contrast, normal samples are more difficult to isolate and require much slicing until they can be isolated [2]. Notice in Fig. 1, a node on iTree can be an external node with no children or an internal node with two children. The isolated sample is described as an external node, and the number of slices needed to isolate the node is the number of edges calculated from the root node to the external node. The number of edges is then referred to as the path length [18].

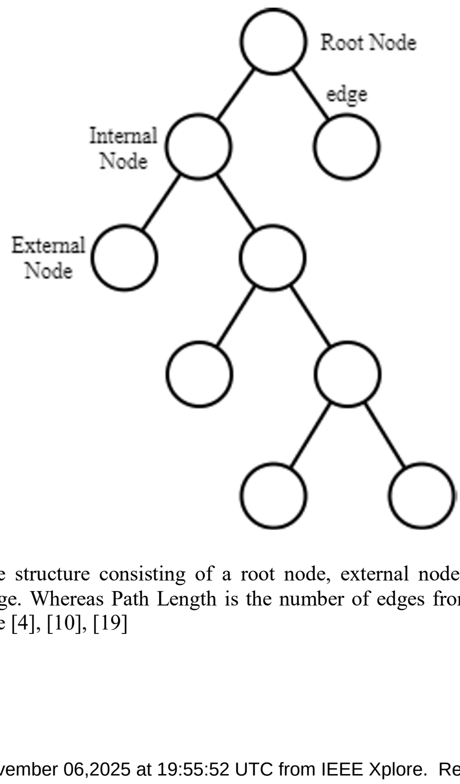

Fig. 1. iTree structure consisting of a root node, external node, internal node and edge. Whereas Path Length is the number of edges from root to external node [4], [10], [19]

III. REVIEW PLANNING

The planning stage consists of explaining why SLR is needed, identifying research questions, and compiling a review protocol.

A. *The reason behind SLR*

IF is the first algorithm explicitly designed for anomaly detection. IF also comes at the right time when data flooding occurs. However, to process various types of data with their various uniqueness, it is necessary to improve the IF algorithm. This paper aims to review strategies undertaken by researchers in utilizing IF in anomaly detection and strategies for overcoming IF weaknesses in various data fields. To achieve those goals, it will answer some Research Questions (RQ).

B. *Research Questions*

To answer the needs and control the direction of the SLR. Research questions should be formulated based on the reasons behind implementing SLR. the questions can be seen in Table I.

C. *Review Protocols*

This study needs protocols for performing SLRs to prevent bias [17]. The protocols in this study are (1) the reason behind the SLR, (2) research questions, (3) literature search strategy, (4) literature selection criteria, (5) result and data discussion, and (6) conclusion or reports. Protocols 1 and 2 have already been implemented above, and the following discussion will discuss the other four protocols.

IV. CONDUCTING REVIEW

A. *Literature Search Strategy*

The difference between SLR and traditional review methods lies in a literature search strategy [17]. SLR has a literature search structured in such a way as to get credible literature sources and avoid bias[17]. In this study, we searched in 4 (four) digital libraries:

- Science Direct (http://www.sciencedirect.com/)
- SpringerLink (http://link.springer.com/)
- IEEE-Xplore (https://ieeexplore.ieee.org/Xplore/home.jsp)
- Sagepub (https://journals.sagepub.com/)

The keywords used are "isolation AND forest", "isolation AND based AND anomaly AND detection", and "isolation AND based AND outlier AND detection". The filter applied is English-language research published in journals in computers and engineering from 2016 to 2020. The results of this search found 2520 researches. Table II displays the search results in detail.

B. *Literature Selection Criteria*

A manual search was carried out to select research related to IF. The research selected must propose an improvement to IF by overcoming IF weaknesses. Research must also use commonly known algorithmic evaluation methods such as AUROC, AUPRC, or F-measure. The results of this manual search found 17 studies as seen in Table III.

TABLE I. RESEARCH QUESTIONS

ID	Research Questions	Objectives
RQ 1	What are the main weaknesses of IF?	Identify the research gaps of each study
RQ 2	What solutions are offered to overcome each of these weaknesses?	Identify the method used with IF
RQ 3	In what areas have researchers applied IF?	Identify the environment and data type suitable for IF

TABLE II. PRELIMINARY SEARCH RESULT

Source	Isolation Forest	Isolation Based Anomaly Detection	Isolation Based Outlier Detection	Quantity
Sciencedirect	584	675	519	1778
springer	260	221	113	594
IEEE	35	23	7	65
Sagepub	18	34	31	83
	897	953	670	2520

TABLE III. MANUAL LITERATURE SEARCH RESULT

Reference	Year	Methodology	Dataset
[9]	2017	Forward Selection Independent Variables (FSIV) and Forward Selection Minimizing the Maximum Reconstruction Error (FSMM)	Primary data in semi-conductor manufacture
[4]	2018	• Expert knowledge • Binary classification	Primary data of Financial data
[13]	2019	SPARK Parallel Computing	A public dataset of IDS UNSW-NB15
[7]	2019	Random Slope Feature Splitting	Public dataset on medical, weather, and satellite
[5]	2019	Correlation Analysis	Public dataset – Aircraft engine data
[10]	2019	PCA based Feature Selection and Feature Extraction	Primary Dataset of Power Control Center
[6]	2019	Weighted Feature-based Feature Selection	Primary data – Docker Container performance data
[8]	2020	Histogram based feature splitting	A public dataset of NASA software defect
[11]	2020	Local Outlier Factor	Public data of concrete (UCI ML Repository)
[20]	2020	Deep learning	Primary data of wind turbine performance
[19]	2020	Gaussian Mixture Model-based clustering	Primary data of wind turbine
[21]	2020	mixture coefficients of a mixture distribution	14 Public ML dataset
[22]	2020	Stack Autoencoder to reduce false positive	Primary dataset: network intrusions data
[23]	2020	Autoencoder based dimension reduction	Public dataset: network intrusions data NSL-KDD
[16]	2020	Multiple feature extraction	Public Hyperspectral data
[15]	2020	Kernel-based	Public Hyperspectral data
[12]	2020	Dual isolation forest	Primary data of Industrial Control System Network

V. RESULT AND DATA DISCUSSION

After reviewing each of the studies selected above, we have obtained SLR results. Each study is grouped according to its suitability for the RQ to present the results for each objective.

A. RQ 1: What are the main weaknesses of IF?

The IF's central concept is to split data randomly until the data are isolated and then calculate the average number of splits needed to isolate the data. The smaller the score obtained, the more likely the data are anomaly[2]. Based on this concept, the researchers found several weaknesses in IF.

1) Low accuracy in the detection of conditional anomalies

The IF split data on randomly selected features and the data's split points are also randomly selected. Conditional anomalies can occur in many types of datasets. The study by Stripling et al. [4] found the problem of AD by IF in a high-dimensional dataset, but only a few features affect the occurrence of anomalies. This weakness was also investigated by Wang et al. [16] in their study stated that in hyperspectral data, certain features generally determine the occurrence of anomalies.

This problem was also found by Zou et al. [6] in their research which stated that each data has a different feature weight, so that random feature selection is not effective. Likewise, as in the research of Khan et al. [5], in datasets with correlated features, hidden anomalies can occur in highly correlated data. Datasets taken under multiple operating conditions can have extreme values that occur in features to hide the anomaly [19].

Conditional anomalies can also lead to false positives or false negatives, such as the imbalance dataset. Imbalance means the number of anomalies is tiny compared to nominal data[8], [23]. Or on hyperspectral data where IF is built based on random selection for each pixel in the entire scene [15].

2) Low accuracy in specific high dimensional data

IF has a good ability in anomaly detection on high dimensional datasets [3], but researchers found that IF can experience difficulties on datasets that have specific properties. Puggini and McLoone's [9] study found that IF experienced an increase in false positives on the high-correlated dataset because slight differences outside the normal value in groups of correlated variables can potentially lead to anomalies rather than significant differences in groups of isolated variables.

A high volume of high dimensional data can also lead to degradation of accuracy. Tao et al. [15] stated that IF can experience a bottleneck on big data if applied to a single computer. The dataset processing capacity in the IF algorithm is limited to the available memory on one machine where the IF-based AD software is installed.

Ahmed et al. [10] examined AD on a very high dimensional dataset with dimensions of up to 1122 dimensions and has a close relationship between dimensions. Standard IF has decreased accuracy because the IF feature splitting process does not involve inter-dimensional relationships. Besides that, processing 1122 dimensions in IF takes much time.

Aminanto [22] found that IF has decreased accuracy by producing a high false-positive rate on imbalanced high-dimensional data. This weakness is very crucial, especially if the cost of false positives is very high.

3) Bias and Performance Issues

Standart IF splits data randomly in parallel with the data dimensions. In their research, Hariri et al. [7] used heat maps to evaluate the AD method on IF and found that this method can cause bias in AD. This bias is the occurrence of ghost areas that have high heat map scores.

Buschjäger et al. [21] evaluating IF fundamentally by finding that the anomaly score assessment using the average path length calculation method has no theoretical explanation based on an understanding of the data distribution. This theoretical explanation can be the basis for improving IF performance.

B. RQ 2: What solutions are offered to overcome these weaknesses?

The researchers proposed several methods to overcome the weaknesses of IF. Researchers not only use statistical methods or other algorithms to strengthen IF but also propose changes to the IF algorithm. They divide the proposed method into three categories. (1) Pre-IF, which is an intervention where data processed by other methods before AD with IF, (2) post-IF, researcher use the output of IF as input for other methods, and (3) method improvement, modification of certain parts of the IF algorithm. Fig.2 shows the proposed methods.

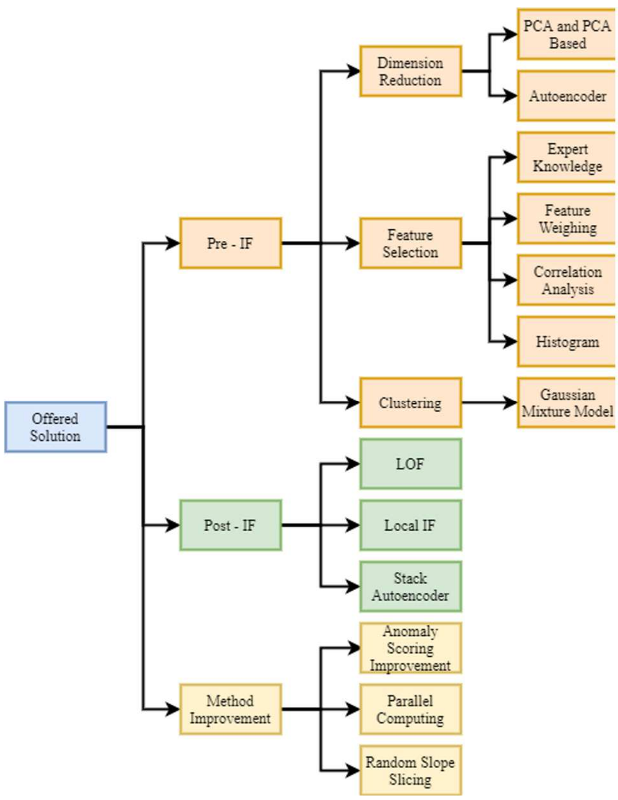

Fig. 2. IF's proposed methods scheme

1) Pre – IF

IF accuracy decreases in high-dimensional data due to the nature of data mentioned in the previous discussion. Therefore, to optimize IF, data pre-processing is required before AD. Data processing can be in the form of reducing dimensions, selecting features, and clustering.

a) Dimension Reduction

Although IF is good at handling high dimensional data, reducing dimension or feature extraction is necessary to maintain the performance and accuracy in large hyper-dimension data. Ahmed et al. [11] used principal component analysis (PCA) to reduce large power system data dimensions. Due to the uniform distribution in the dataset and the Gaussian distribution formed from the sensor or meter measurement noise. Meanwhile, PCA can work well on data with Gaussian distribution. Besides, by using PCA, the researcher can select the number of components representing the data dynamics from the variance criteria. When PCA is suitable for correlated data, it becomes a problem when the anomaly is outside the large cluster of correlated data because PCA will miss the anomaly data. Therefore Puggini and McLoone [9] propose a PCA-based method to find the variables that have the least relationship with the variables selected by the PCA, which is named the Forward Selection Independent Variables (FSIV) method and Forward Selection Minimizing the Maximum Reconstruction Error (FSMM).

Sadaf and Sultana [23] use an autoencoder (AE) to add features in the form of "attack" and "normal" labels to high-dimensional network traffic data so that the IF algorithm can improve detection accuracy as it focuses on two features provided by AE.

b) Feature Selection

In low correlated data, feature selection can be used to select the most important features [10]. Stripling et al. [4], in their study, found that some high-dimensional data have more significant features than others. On nominal data, expert knowledge can quickly identify these features.

Using expert experience to select nominal feature data and feed it to the IF algorithm, [4] can find previously undetected conditional anomalies.

Zou et al. [6] conducted AD research on a dataset where each feature's influence level can vary in each sample data. In this condition, the accuracy of IF decreases. They introduced the weighting method. This method is applied to docker container data by monitoring the most dominant resource usage for each container. Selecting features based on resource weight can improve the accuracy of IF.

Khan et al. [5] use the correlation analysis method for the feature selection approach. The correlation analysis is a method that measures the relationship between variables and then classifies them based on the correlation coefficient. This method can perform AD more effectively because it isolated the group of variables causing the anomaly.

Ding and Xing [8] also use the feature selection approach in overcoming the weaknesses of IF. In their research, they convert the feature's distribution value into a histogram, then splitting is done on the value that lies at the distribution boundary or the value with low frequency. This method achieved convergence faster than the random splitting point by IF standard.

c) Clustering

The data obtained under different operating conditions can hide anomalies if the data are not separated based on conditions. Operating conditions can be behavioral attributes or contextual attributes. When dealing with such data, it is necessary to pre-process the data before doing AD with IF. In their research, Chen et al. [23] proposed data separation based on operational conditions using the Gaussian Mixture Model (GMM). They choose this method because the data used by [24] are complex data with nonparametric distribution.

2) Post-IF

Another approach to increasing the accuracy of IF in conditional outlier detection is to use the IF output as input for other algorithms. Instead of making the IF output the final result, Alsini et al. [11] made the IF output an outlier candidate reprocessed using the LOF and increased the accuracy by the sliding window method used in selecting candidates from the IF results.

In their research, Li et al. [15] found that random selection of data points on hyperspectral data causes many false alarms. Therefore, they propose using twice IF, the first to find global anomaly maps, the second to refining global anomaly maps to find the local anomaly. Elnour et al. [12] also use two Ifs. However, their research strengthened the second IF using PCA.

Aminanto et al. [22]. In their study, investigate a high false-positive rate on imbalanced high-dimensional data. A stacked autoencoder (SAE) is used as a data processing algorithm after IF. SAE is a set of autoencoders (AE) arranged sequentially. The AE output becomes the next AE input to produce a reconstruction error (RE). AE defines the anomaly as a variable with a high RE score.

3) Method Improvement

In addition to using other algorithms to help IF overcome its weaknesses, researchers also modify the IF algorithm to cover existing weaknesses. Buschjäger et al. [21] dissect the basic theory of IF and suggest that scoring anomaly using average path length is the same as calculating the mixture components' estimated coefficient. Therefore, to optimize IF, [21] proposes an estimate of mixture coefficients of mixture distribution as an anomaly scoring method.

Another improvement was also made by Tao et al. [13] by utilizing the independent nature of iTree to modify IF and implement it on a parallel computer. iTree construction on bigdata processing requires extensive memory resources so that a bottleneck can occur when using one machine, so Tao et al. [13], with the construction of iTree on a parallel computer, has succeeded in overcoming the limited memory resources for the IF process on one machine.

Another study by Hariri [7] looked at the nature of data splitting on IF, which was always horizontal or vertical. This study then proposed an IF improvement by making the data splitting direction randomly. The method succeeded in eliminating the bias and ghost area formed from horizontal and vertical splitting.

C. RQ 3: In what areas have researchers applied IF?

The areas of greatest interest are fault detection and diagnosis [5], [6], [8], [9], [11], [20], [24]. Some researchers test the reliability of their proposed improvements to several

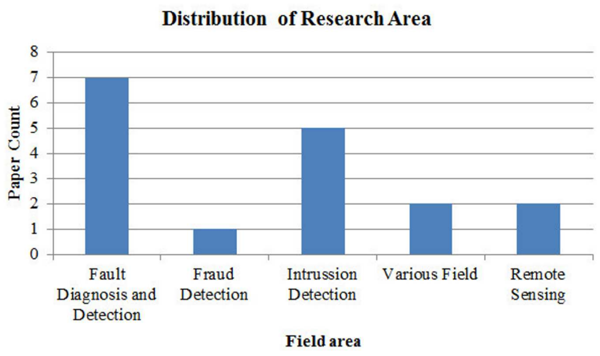

Fig. 3. Distribution of IF research areas

datasets at once [7], [21]. Research like this is grouped into "various fields", which describes the distribution of research areas, while other studies are grouped in their respective fields as shown in Fig. 3.

VI. CONCLUSION

The increasing type and amount of data and the need for fast and accurate data have made AD a rising branch of machine learning in recent years. Isolation forest was introduced in 2008 as one of the AD methods recognized as having good potential and has several weaknesses. Researchers have studied these weaknesses and offer solutions.

This paper reviewed 17 studies that suggest solutions to IF's weaknesses. Most researchers agree that IF's weakness in conditional anomaly detection is due to the random selection of features and the random selection of split points.

However, there is no consensus on the best way to deal with this, and researchers are divided into three methods: Pre-IF intervention, Post-IF intervention, and method improvement.

Research in AD with IF can still be developed in the future. Research topics related to the challenges in applying IF for text-mining and sentiment analysis to the best of our knowledge have never been done. We think this topic is exciting to be further explored with current social media trends.

REFERENCES

[1] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection : A Survey," *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: 10.1145/1541880.1541882.

[2] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-Based Anomaly Detection," *ACM Trans. Knowl. Discov. Data*, vol. 6, no. 1, pp. 1–39, Mar. 2012, doi: 10.1145/2133360.2133363.

[3] R. Domingues, M. Filippone, P. Michiardi, and J. Zouaoui, "A comparative evaluation of outlier detection algorithms: Experiments and analyses," *Pattern Recognit.*, vol. 74, pp. 406–421, Feb. 2018, doi: 10.1016/j.patcog.2017.09.037.

[4] E. Stripling, B. Baesens, B. Chizi, and S. vanden Broucke, "Isolation-based conditional anomaly detection on mixed-attribute data to uncover workers' compensation fraud," *Decis. Support Syst.*, vol. 111, no. April, pp. 13–26, Jul. 2018, doi: 10.1016/j.dss.2018.04.001.

[5] S. Khan, C. F. Liew, T. Yairi, and R. McWilliam, "Unsupervised anomaly detection in unmanned aerial vehicles," *Appl. Soft Comput.*, vol. 83, p. 105650, Oct. 2019, doi: 10.1016/j.asoc.2019.105650.

[6] Z. Zou, Y. Xie, K. Huang, G. Xu, D. Feng, and D. Long, "A Docker Container Anomaly Monitoring System Based on Optimized Isolation Forest," *IEEE Trans. Cloud Comput.*, vol. 7161, no. SEPTEMBER 2018, pp. 1–1, 2019, doi: 10.1109/TCC.2019.2935724.

[7] S. Hariri, M. Carrasco Kind, and R. J. Brunner, "Extended Isolation Forest," *IEEE Trans. Knowl. Data Eng.*, pp. 1–1, 2019, doi: 10.1109/TKDE.2019.2947676.

[8] Z. Ding and L. Xing, "Improved software defect prediction using Pruned Histogram-based isolation forest," *Reliab. Eng. Syst. Saf.*, vol. 204, no. August, p. 107170, Dec. 2020, doi: 10.1016/j.ress.2020.107170.

[9] L. Puggini and S. McLoone, "An enhanced variable selection and Isolation Forest based methodology for anomaly detection with OES data," *Eng. Appl. Artif. Intell.*, vol. 67, no. September 2017, pp. 126–135, Jan. 2018, doi: 10.1016/j.engappai.2017.09.021.

[10] S. Ahmed, Y. Lee, S.-H. Hyun, and I. Koo, "Unsupervised Machine Learning-Based Detection of Covert Data Integrity Assault in Smart Grid Networks Utilizing Isolation Forest," *IEEE Trans. Inf. Forensics Secur.*, vol. 14, no. 10, pp. 2765–2777, Oct. 2019, doi: 10.1109/TIFS.2019.2902822.

[11] R. Alsini, A. Almakrab, A. Ibrahim, and X. Ma, "Improving the outlier detection method in concrete mix design by combining the isolation forest and local outlier factor," *Constr. Build. Mater.*, vol. 270, p. 121396, Feb. 2020, doi: 10.1016/j.conbuildmat.2020.121396.

[12] M. Elnour, N. Meskin, K. Khan, and R. Jain, "A dual-isolation-forests-based attack detection framework for industrial control systems," *IEEE Access*, vol. 8, pp. 36639–36651, 2020, doi: 10.1109/ACCESS.2020.2975066.

[13] X. Tao, Y. Peng, F. Zhao, P. Zhao, and Y. Wang, "A parallel algorithm for network traffic anomaly detection based on Isolation Forest," *Int. J. Distrib. Sens. Networks*, vol. 14, no. 11, p. 155014771881447, Nov. 2018, doi: 10.1177/1550147718814471.

[14] M. Kiran, C. Wang, G. Papadimitriou, A. Mandal, and E. Deelman, "Detecting anomalous packets in network transfers: investigations using PCA, autoencoder and isolation forest in TCP," *Mach. Learn.*, vol. 109, no. 5, pp. 1127–1143, May 2020, doi: 10.1007/s10994-020-05870-y.

[15] S. Li, K. Zhang, P. Duan, and X. Kang, "Hyperspectral Anomaly Detection With Kernel Isolation Forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 319–329, Jan. 2020, doi: 10.1109/TGRS.2019.2936308.

[16] R. Wang, F. Nie, Z. Wang, F. He, and X. Li, "Multiple Features and Isolation Forest-Based Fast Anomaly Detector for Hyperspectral Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6664–6676, Sep. 2020, doi: 10.1109/TGRS.2020.2978491.

[17] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.

[18] Y. Chen and W. Wu, "Isolation Forest as an Alternative Data-Driven Mineral Prospectivity Mapping Method with a Higher Data-Processing Efficiency," *Nat. Resour. Res.*, vol. 28, no. 1, pp. 31–46, Jan. 2019, doi: 10.1007/s11053-018-9375-6.

[19] H. Chen, H. Ma, X. Chu, and D. Xue, "Anomaly detection and critical attributes identification for products with multiple operating conditions based on isolation forest," *Adv. Eng. Informatics*, vol. 46, no. March 2019, p. 101139, Oct. 2020, doi: 10.1016/j.aei.2020.101139.

[20] Z. Lin, X. Liu, and M. Collu, "Wind power prediction based on high-frequency SCADA data along with isolation forest and deep learning neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 118, no. September 2019, p. 105835, Jun. 2020, doi: 10.1016/j.ijepes.2020.105835.

[21] S. Buschjäger, P.-J. Honysz, and K. Morik, "Randomized outlier detection with trees," *Int. J. Data Sci. Anal.*, Dec. 2020, doi: 10.1007/s41060-020-00238-w.

[22] M. E. Aminanto, T. Ban, R. Isawa, T. Takahashi, and D. Inoue, "Threat Alert Prioritization Using Isolation Forest and Stacked Auto Encoder With Day-Forward-Chaining Analysis," *IEEE Access*, vol. 8, pp. 217977–217986, 2020, doi: 10.1109/ACCESS.2020.3041837.

[23] K. Sadaf and J. Sultana, "Intrusion Detection Based on Autoencoder and Isolation Forest in Fog Computing," *IEEE Access*, vol. 8, pp. 167059–167068, 2020, doi: 10.1109/ACCESS.2020.3022855.

[24] H. Chen, H. Ma, X. Chu, and D. Xue, "Anomaly detection and critical attributes identification for products with multiple operating conditions based on isolation forest," *Adv. Eng. Informatics*, vol. 46, no. July, p. 101139, Oct. 2020, doi: 10.1016/j.aei.2020.101139.