future internet

MDPI

Article

Detection of Man-in-the-Middle (MitM) Cyber-Attacks in Oil and Gas Process Control Networks Using Machine Learning Algorithms

Ugochukwu Onyekachi Obonna [1], Felix Kelechi Opara [1], Christian Chidiebere Mbaocha [1], Jude-Kennedy Chibuzo Obichere [2], Isdore Onyema Akwukwaegbu [1], Miriam Mmesoma Amaefule [3] and Cosmas Ifeanyi Nwakanma [4,*]

[1] Department of Electrical/Electronic Engineering, Federal University of Technology, Owerri 340110, Nigeria; obonnaugochukwu@yahoo.com (U.O.O.); felix.opara@futo.edu.ng (F.K.O.); christian.mbaocha@futo.edu.ng (C.C.M.); isdore.akwukwaegbu@futo.edu.ng (I.O.A.)
[2] Department of Mechatronics Engineering, Federal University of Technology, Owerri 340110, Nigeria; jude.obichere@futo.edu.ng
[3] Department of Mathematics, Federal University of Technology, Owerri 340110, Nigeria; amaefulemiriam51@gmail.com
[4] ICT Convergence Research Center, Kumoh National Institute of Technology, Gumi 39177, Republic of Korea
* Correspondence: cosmas.ifeanyi@kumoh.ac.kr

Abstract: Recently, the process control network (PCN) of oil and gas installation has been subjected to amorphous cyber-attacks. Examples include the denial-of-service (DoS), distributed denial-of-service (DDoS), and man-in-the-middle (MitM) attacks, and this may have largely been caused by the integration of open network to operation technology (OT) as a result of low-cost network expansion. The connection of OT to the internet for firmware updates, third-party support, or the intervention of vendors has exposed the industry to attacks. The inability to detect these unpredictable cyber-attacks exposes the PCN, and a successful attack can lead to devastating effects. This paper reviews the different forms of cyber-attacks in PCN of oil and gas installations while proposing the use of machine learning algorithms to monitor data exchanges between the sensors, controllers, processes, and the final control elements on the network to detect anomalies in such data exchanges. Python 3.0 Libraries, Deep-Learning Toolkit, MATLAB, and Allen Bradley RSLogic 5000 PLC Emulator software were used in simulating the process control. The outcomes of the experiments show the reliability and functionality of the different machine learning algorithms in detecting these anomalies with significant precise attack detections identified using tree algorithms (bagged or coarse) for man-in-the-middle (MitM) attacks while taking note of accuracy-computation complexity trade-offs.

Keywords: amorphous cyber-attacks; process control network; anomaly detection; machine learning; man-in-the-middle attacks; SCADA

Citation: Obonna, U.O.; Opara, F.K.; Mbaocha, C.C.; Obichere, J.-K.C.; Akwukwaegbu, I.O.; Amaefule, M.M.; Nwakanma, C.I. Detection of Man-in-the-Middle (MitM) Cyber-Attacks in Oil and Gas Process Control Networks Using Machine Learning Algorithms. *Future Internet* **2023**, *15*, 280. https://doi.org/10.3390/fi15080280

Academic Editor: Francesco Buccafurri

Received: 5 July 2023
Revised: 4 August 2023
Accepted: 14 August 2023
Published: 21 August 2023

1. Introduction

The oil and gas industry is deemed critical infrastructure due to the fact that it is a major contributor to the world's energy needs, and disruption to its operation could lead to a major impact on the consumers and can lead to devastating effects ranging from catastrophic process safety incidents which may lead to loss of lives, destruction of assets and destruction of the environment, to economic issues to host nations. The choice of standard information technology (IT) open systems, their associated communication protocols, and their preference over proprietary dedicated operational technology (OT) systems has exposed PCN to insecure communications which have given room to cyber-attacks [1]. The May 2021 Darkside Ransomware attack on the Colonial Pipeline in the USA disrupted and stopped the transportation of gasoline and jet fuel when the computerized equipment managing the pipeline was attacked. After gaining access to the company network of the

Colonial Pipeline, Darkside Ransomware was deployed against the company's IT network by intruders [2].

In the last decade, there have been highly sophisticated threats on critical infrastructures with devastating impacts, and not much attention has been given to the oil and gas industry's vulnerability to attacks. For the continual safe operation of an oil and gas facility, there is a need to ensure that all the network components of a process control system are protected from intruders or attackers through effective monitoring and surveillance which may help in the early detection of attacks and proper mitigation of such attacks [3–5]. The process control network (PCN) is the interconnection of real-time network devices used to monitor and control industrial processes. It ensures effective communication between the sensors, controllers, processes, and the final control elements [6]. The process variables in the PCN serve as inputs to the controllers which make real-time decisions on the final control elements to ensure a continuous and safe operation of the plant. A real-time adjustment or modification of the input variables results in the controller affecting the change in the operating conditions of the logic solvers which eventually results in altered outputs to the final control elements. There is a need to ensure secure communication between the field sensors, the controllers, and the final control elements [6].

As with all other sectors of the economy, continuous digital growth has impacted the oil and gas industry. Industrial control systems (ICS) are used to operate in isolation, without bridging over information technology (IT) infrastructures. Industry 4.0 enabled the integration of multiple industrial technologies in ICT, and engineers are now able to monitor operations remotely as well as maintain supervisory control and data acquisition (SCADA) systems in real-time. This digital revolution has exposed once air-gapped OT infrastructures to a myriad of new attack surfaces and vectors [7–9]. With the advancement in the Industrial Internet of Things (IIoT), early identification and prevention of attacks that can lead to PCN disasters can be achieved by continuous monitoring using algorithm-based smart monitoring systems [10–12].

ICS operational technology networks can be penetrated by malicious cyber-attackers. Even though there are intrusion detection systems (IDS), firewalls, demilitarized zones, and data diodes that help in isolating ICS operational technology networks, these security measures cannot be assumed sufficient to stop all malicious penetrations of the air-gapped OT networks. Hackers can access the network through compromised software updates, insider attacks, infected thumb drives, and spear phishing attacks to penetrate heavily isolated and air-gapped OT networks. The Stuxnet malware is a famous example of a worm that penetrated an air-gapped network by exploiting a USB thumb drive autorun vulnerability [13].

Several supervised machine learning algorithms have shown good results in the detection of signature-based attacks which normally are detected by intrusion detection systems (IDS) but behavior-based attacks which can be termed anomalies or outliers have been difficult to detect or predict based on the dynamic attack strategies deployed by the attackers [14–17]. The choice of the machine learning algorithm to use is influenced by some key factors which include accuracy, computational capability, prediction speed, false alarm rates, and their application to real-time systems [8,18]. There is a need to identify and mitigate false data signals which may be introduced in the form of man-in-the-middle (MitM) attacks [19]. False data injection attacks (FDIA) which are deceptive can modify measured values thereby introducing errors which will result to system failures [20]. Disgruntled employees pose a huge threat to the OT as they can become insider threats with good knowledge of the production facility. Intentional malicious insider attacks usually have a huge impact with a high percentage of success [21]. The oil and gas industry in Nigeria has been faced with myriad of challenges ranging from pipeline vandalism, theft, illegal bunkering, and now intrusion attacks [22,23]. This work is focused on the detection and prevention of amorphous cyber-attacks on the networks of oil and gas facilities using machine learning and real-time SCADA dataset.

In this work, an ambitious attempt was made to reinforce the discussion on the detection and mitigation of MitM attacks using real-time datasets. The aim is to evaluate the performance of various machine learning candidates and propose a range of highly performing options while noting some trade-offs. This will help in developing a reliable security system capable of detecting amorphous network intrusions or attacks on a PCN using different machine-learning algorithms. To achieve this objective, the following measurable steps were taken:

1. We inject values to the collated real-time dataset at a specific date and time to make a distinction between normal operation and anomalous conditions which may represent plant shutdown, equipment failure, process upset conditions, or cyber-attacks.
2. We apply different machine learning algorithms on the dataset to determine the most effective algorithm in identifying anomalies.
3. We perform a comparative study of the results from the different machine learning algorithms, to determine their application to anomaly detection,
4. We perform a comparative review of the actualized results with results from other researchers to validate the achieved results.
5. In addition to the real-time dataset, we demonstrated the superior performance of the bagged tree and coarse tree algorithms using three public datasets namely: WUSTL-2018, ORNL PowerGrid, and TON_IoT

The paper is organized as follows: Section 1 is the introduction, Section 2 is the review of related works, Section 3 is the comparison of different machine learning algorithms, Section 4 is the results and discussion, and Section 5 is the conclusion and recommendation for future work. Acronyms used in this article are listed in the abbreviations section.

2. Related Works

The integration of standard open network technology has continuously exposed process control networks to malicious cyber-attacks. The need arises to ensure secured communication between the process sensors, the controllers, and the final control elements [6,24]. The connection of the PCN to the internet has also contributed to the growth of cyber-attack incidents with dangerous consequences [25]. The deployment of off-the-shelf IT equipment with its inherent vulnerabilities and associated failures has also contributed to the exposure of the PCN to cyberattacks [26]. Unstructured and unpredictable attacks are termed outliers to signature-based detections. These nonconforming patterns are termed anomalies, and detection of their kind of activities could be performed using unsupervised machine learning algorithms [27].

Ramotsoela et al. [14] noted that signature-based IDS are disadvantageous as they are unable to detect unknown attacks [14]. The constant dynamic modes of attacks used by the attackers are the major challenge of the work done by the authors in [28]; they used machine learning classifiers as an effective IDS where data were pre-processed to remove unrelated attributes from the dataset [28]. However, the public dataset used lacked sufficient details to detect recent attack types including zero-day attacks, advanced persistent threats (APTs), and their derivatives. Similarly, using the NSL-KDD dataset, the authors in [16] proposed unsupervised machine learning techniques based on a clustering approach to minimize false positives as a solution to unknown attacks including zero-day attacks [16]. Several IDS solutions exist but they cannot detect these unpatterned attacks which may be in the form of DoS, DDoS, MitM, or even zero-day attacks [29]. In [30], the authors reviewed different machine learning capabilities and concluded that the effectiveness and efficiency of a machine learning algorithm-based solution depend on the features and characteristics of the data as well as the performance of the algorithm [30].

Rosa et al. [17], in their work on intrusion detection using anomaly detection, observed that integration of different complex solutions for monitoring of networks will require prolonged network downtime, which will have limited application to a PCN. In their quantitative comparison of 17 unsupervised anomaly detection algorithms [15,27,31] adopted unsupervised anomaly detection algorithms becuase they are known to be the most suitable

way to deal with zero-day attacks and concluded that minimizing the misclassification of unsupervised anomaly detection algorithms is highly desirable and is a key challenge.

Melnick [19] explained the different forms of MitM which include session hijacking, IP spoofing, and replay attack in which any of the attack forms will lead to the attacker taking over the communication between the sensors and the controllers with the intention of disrupting the process control [32]. In [33], the authors explained that fairness, data trustworthiness, reliability, and availability are necessary for the actualization of cyber-physical systems, for example, smart cities with robust system architecture for secured high bandwidth systems and low-latency diffusion [33], whereas supervised machine learning is taught by example and uses labeled data to detect known attacks [34,35], unsupervised machine learning can analyze huge volumes of data to identify hidden patterns, clusters, and outliers, thereby can be very effective in detecting anomalies in datasets which include process upsets, shutdowns or faulty equipment as well as attacks [15,27,31,36,37]. Deep learning algorithms have shown great results in supervised and unsupervised machine learning applications using very large datasets, timely learning ability, produced great accuracy, and increased prediction speed with negligible false alarm rates [38–40]. Bierbrauer et al. [40] leveraged the NSL-KDD dataset to show the application of decision tree models in detecting APT attacks with high detection accuracy.

Al-Abassi et al. [39] used an ensemble deep learning-based cyber-attack detection method specifically designed for industrial control systems; the outcome of their work yielded good results, but they recommended the use of real-time datasets with properly defined attack types. In [41], the authors use a new approach to determine DDoS attack patterns on SCADA systems using machine learning applied three machine learning techniques using the KDDCup'99 dataset, but recommend using a real-time SCADA dataset for improved results. The characteristics, features, and operations of Stuxnet and APTs were reviewed by [42], in addition to highlighting their recent trends, attack features, and prediction of future attacks. Maynard and McLaughlin [43] investigated packet pilfering, injection attacks through HTTP, and command injection attacks. They were focused on DDoS flood attacks without the capability of modifying packets sent by other hosts as in the case of MitM attacks. Wilson et al. [44] in their work, Deep Learning-Aided Cyber-Attack Detection in Power Transmission Systems, noted the need to consider multiple sources of uncertainty and variations from subsystems renewable energy, smart grids, and so on. Furthermore, Husák et al. [45] in their Survey of Attack Projection, Prediction, and Forecasting in Cyber Security, emphasized the need for attack forecasting to minimize the impact on systems.

Researchers had used different machine learning approaches with datasets such as NSL-KDD and KDD CUP'99 to address the issue of cyber-attacks in the past. These datasets have existed for a long time and may not be a full representation of modern-day dynamic attack modes. Hence, the products of such research may not be relied upon for the detection of recent amorphous attacks. Consequently, recent works have either emphasized the use of facility-specific datasets, as argued in this paper, or the use of recent and updated datasets such as the WUSTL-2018 [35], ORNL PowerGrid [46], and the TON_IoT datasets [47]. Table 1 summarizes the recent related works and research gaps that motivated this work.

Table 1. Related Studies and identified research gaps.

Reference	Major Findings	Limitations
Pu et al. [16]	Unsupervised clustering-based anomaly detection method to minimize false positives	i. Identical and repeated NSL-KDD datasets which affect the learning ability of the algorithm and the final output ii. possibility of generating lots of false positives which could deceive real network traffics

Table 1. *Cont.*

Reference	Major Findings	Limitations
Rosa et al. [17]	Integration of different techniques and algorithms for networking monitoring	i. Integration of the complex solution will require prolonged network downtime.
Zoppi et al. [27]	Quantitative comparison of 17 Unsupervised anomaly detection algorithms	i. High rate of misclassification of unknown attacks ii. High computational complexity,
Abrar et al. [28]	Machine learning approach using different algorithms to solve intrusion detection problems (supervised learning)	i. The model could not detect zero-day attacks. ii. Not enough details in the public dataset, iii. Computational complex programs
Joloudari et al. [38]	Deep learning and decision tree algorithms for advanced persistent threat attack detection	i. Could not extract important features from the NSL-KDD dataset. ii. The dataset did not reflect the real scenario of the target idea.
Al-Abassi et al. [39]	Generalized ensemble deep learning for cyber-attack detection in industrial control system	i. Could not distinguish between system downtime and actual real-time attacks.

3. System Model, Threat Modeling Framework and Simulation Setup

3.1. Intrusion Detection Using Machine Learning Models

This study reveals the different forms of unpatterned attacks on the PCN with their resulting effects on the people, assets, and the environment as depicted in Figure 1. The compromise of the intercommunication between the sensors, controllers, and the final control elements could lead to devastating outcomes which may range from fatalities to environmental impact. This study reviewed the application of different machine learning algorithms in the modeling of these attacks using the 68,722 real-time SCADA datasets from the oil and gas industry. The performances of the different machine learning algorithms were assessed, which include isolation forest, k-nearest neighbor (kNN), Python Outlier detection (PyOD) which incorporates interquartile range (IQR), kNN, local outlier factor (LOF), long short-term memory, support vector machines (SVM) and decision tree algorithms. The 68,722 real-life SCADA data were extracted from an oil and gas facility.

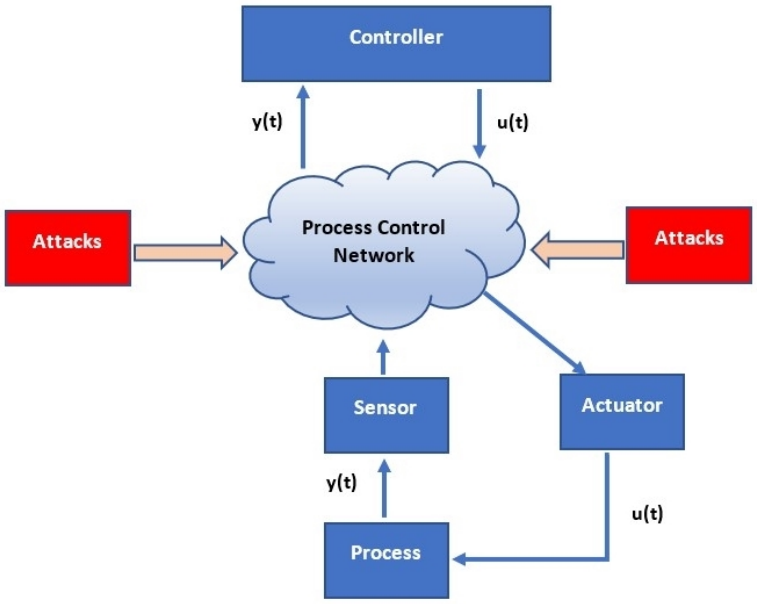


Figure 1. Interconnection of the PCN components under attack.

To simulate the impact of amorphous cyber-attacks on the oil and gas industry, a three-phase separator was selected as a case study (see Figure 2). Usually, the natural crude oil flowing from the wellbore which contains entrapped gas and water is fed into a vessel called a three-phase separator. This gravity vessel separates the crude into oil, water, and gas based on their densities [48–51]. In this study, a three-phase separator was used as a case study for ease of computation and simulation to showcase the effect of false data injection in SCADA.

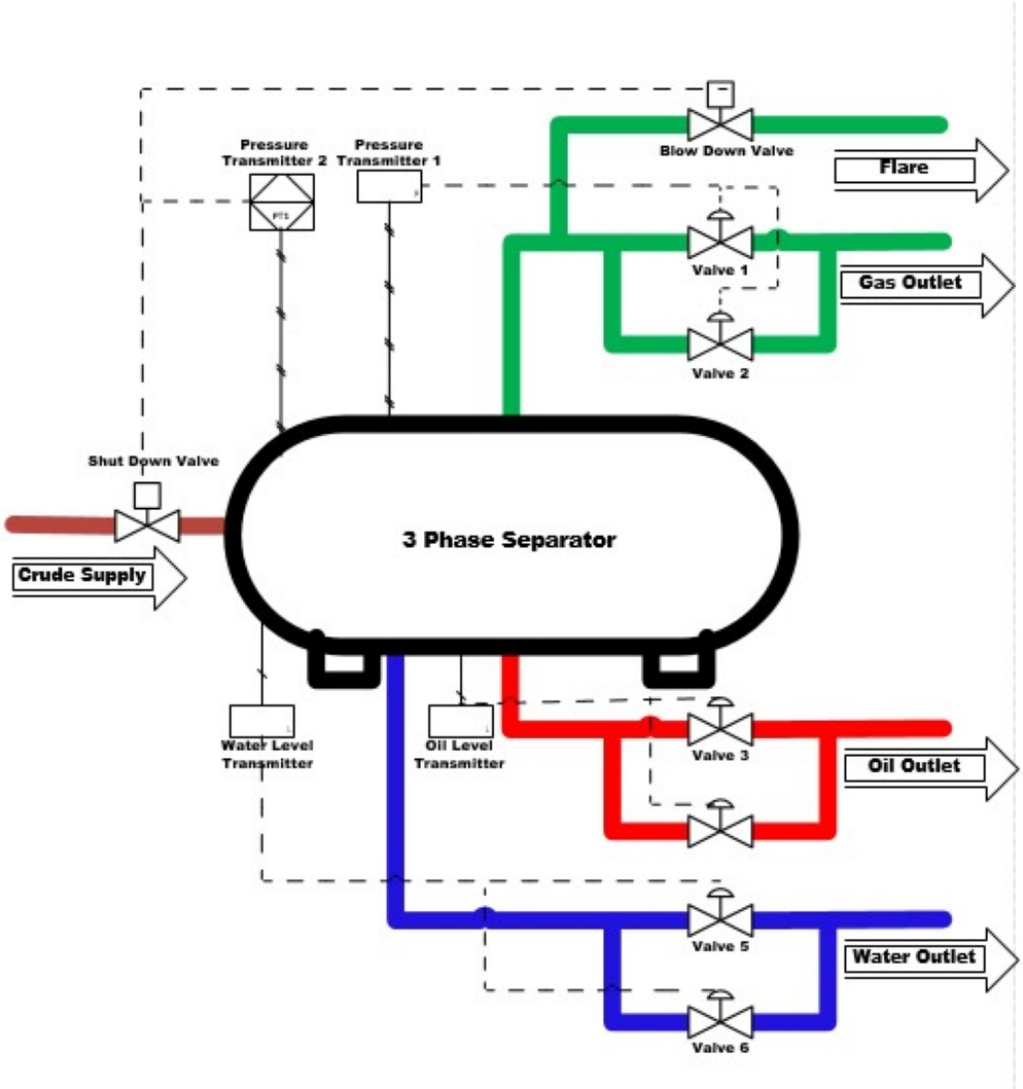


Figure 2. A three-phase separator.

Figure 2 shows a three-phase separator that receives crude oil from the well bore through the shutdown valve and separates the received crude oil into gas, oil, and water. The three-phase separator has three outlets: gas outlet, crude oil outlet, and water outlet. The process variables measured from the vessel include supply pressure, discharge pressure, pressure in the vessel, level of oil with water, level of oil, the temperature of the supplied fluid, vessel temperature, and temperature of the individual discharge lines, whereas the flow was measured on the respective outlet lines. To prevent process upset and its escalation, there is a need for the continuous monitoring of the multivariable inputs with consideration to their interactions in the vessel during the retention time. The 68,722 dataset used in this study's simulation are the three-phase separator vessel pressure data. The outcomes of the simulations using the different machine learning algorithms on the same

dataset are documented in the results session. A detailed overview of the system model of this research is shown in Figure 3.

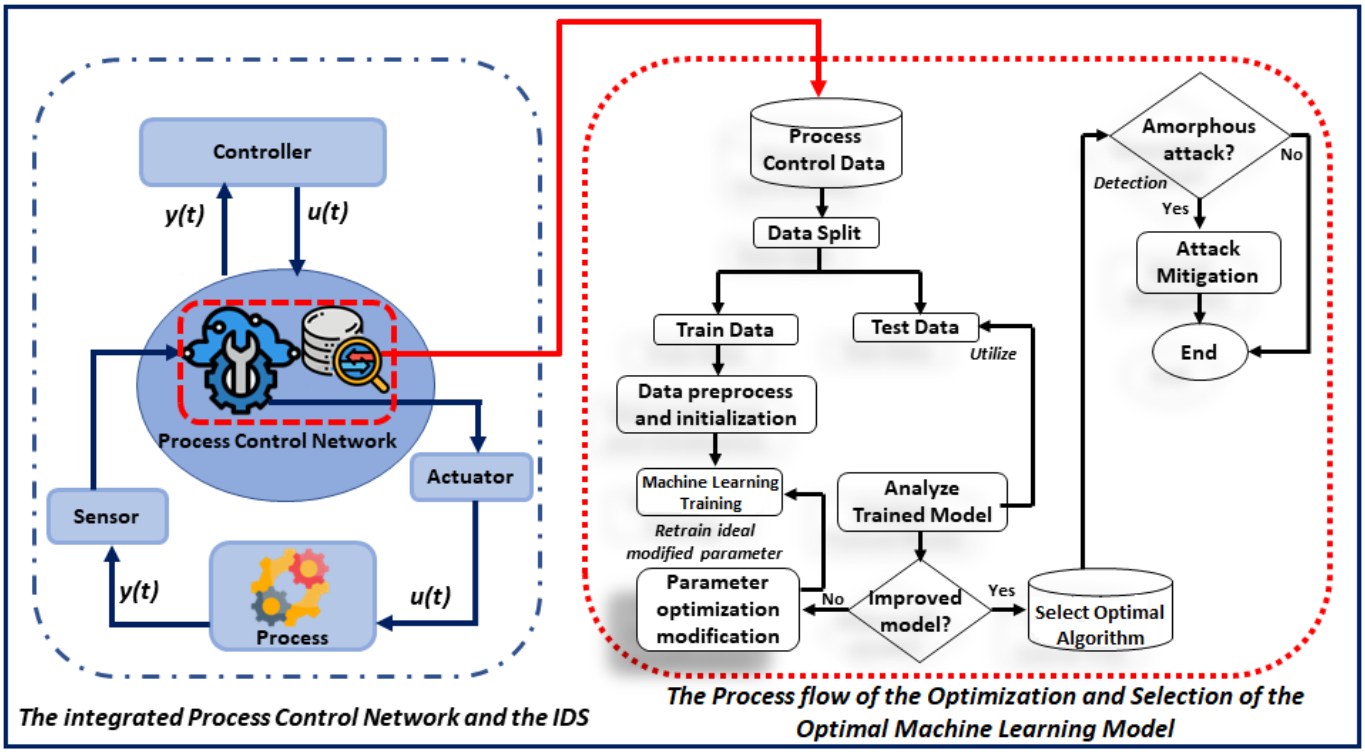


Figure 3. System Model showing the Experimentation and Selection of the Optimal Machine Learning Algorithm.

3.2. Threat Modeling Description

Amorphous forms of threats exist for PCN ranging from external threats to internal threats as shown in Figure 4. External threats through the internet may exploit the login details of authorized personnel which include the remote employee or the vendor technical support personnel who may be connecting through the internet to the business network and then to the PCN. Adequate controls may have been applied on the firewalls to prevent access of unauthorized persons but there is a limitation of exploited accounts. All other external threats such as internet threats which may come as a result of the interface between the PCN and the business network can be detected and prevented using the developed system. Insider threats are internal users within the organization, such as employees, former employees, business associates, and contractors who have malicious intents, with correct authentication to the network and knowledge of the company information. Insider attacks have been identified as one of the most dangerous cyber threats for critical infrastructures (CIs), as the attacker continues sending real and legitimate control commands to other network devices, it can lead to catastrophic damage to CIs. Insider attacks usually have a huge impact and greater success rate because it is difficult to predict when they want to attack and how they want to attack thereby making it unpreventable [21,52–54]

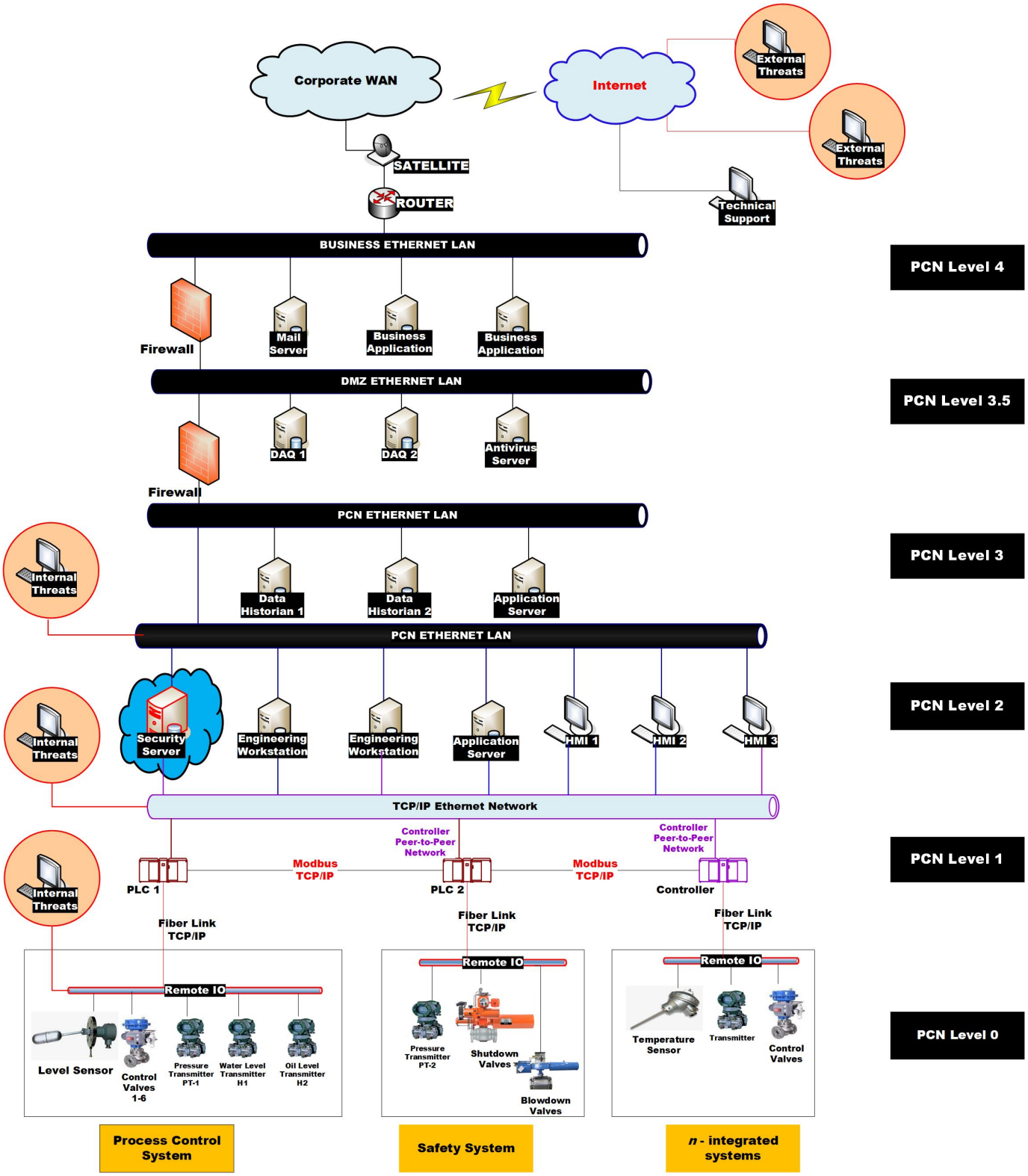


Figure 4. PCN Levels showing the vulnerability point at each level. This was leveraged in the threat modeling.

Other forms of threats are denial-of-service (DoS) attacks, which floods the systems, servers, or networks with unsolicited traffic to exhaust and interrupt network communication resources and bandwidth [20]; distributed-denial-of-service (DDoS) attacks, where the attackers use multiple compromised devices to launch DoS attack and can bombard a central server with simultaneous data requests thereby disrupting the services; man-in-the-middle (MitM) attacks, where the attackers deliberately insert themselves into a two-party communication interrupting the traffic flow, with the intention of data theft and false data injection attacks [20,55]; man-on-the-side (MotS) attacks allow an adversary to read and inject packets, but not modify packets sent by other hosts as in the case of MitM attacks [43]; and zero-day exploit attacks, which occur after a network vulnerability has been announced but before a patch or solution is implemented. Advanced persistent threats (APTs) are highly sophisticated threats that focus on the critical industrial sector. APTs have the capability to remain undetected for an extended period after they gain unauthorized access to a computer network, and usually do not need the internet for spreading; a typical example is 2010 Stuxnet [42].

From Figure 4, both the external threats and the insider threats are identified.

Level 0—Production process equipment, sensors and actuators;

Level 1—Controllers and real-time control of the production process;

Level 2—Human machine interface (HMI), engineering work stations and servers;

Level 3—Data historians, advanced control;

Level 3.5—Demilitarized zone (DMZ), interface between PCN and business network;

Level 4—Business or enterprise network.

Levels 0 to 3.5 are termed the process control network which may consist of different technologies, topologies, protocols, and communication mediums. The communication medium for levels 0 to 2 could be standard or proprietary depending on the integration systems and vendor equipment used. Level 3 and 3.5 utilizes standard open systems Ethernet technology whereas Level 4 utilizes open systems Ethernet LAN technology. The newly introduced security server which houses the models for cyber-attack detection and with the capability of monitoring the real-time data exchange between the input/output devices and the controllers is integrated at the layer 2 of the PCN alongside the engineering workstations with direct access to the data exchange between the controllers and the field devices.

3.3. Simulation and Experimental Setup

The 68,722 pressure datasets collated from a three-phase separator were applied such that 60% was used for model training, 25% for model testing, and 15% for model validation in other to achieve optimal performance and avoid overfitting of the models. Python 3.0 libraries, deep learning toolkit, and MATLAB were used for the simulations on a CPU with the optimal configuration of Intel(R) Core (TM) i7 CPU @ 3.00 GHz, 16 GB RAM, and GPU Tesla K80. Other hyperparameters such as learning rates, batch size, number of epochs, time steps, and contamination parameters were used during the simulations to improve the model results and optimize the learning abilities of the different algorithms as seen in Table 2. To validate the tree algorithm performance, additional simulations were performed using three public datasets - WUSTL-2018 [35], ORNL [46], and TON_IoT [47]. This is critical to reinforce the performance results.

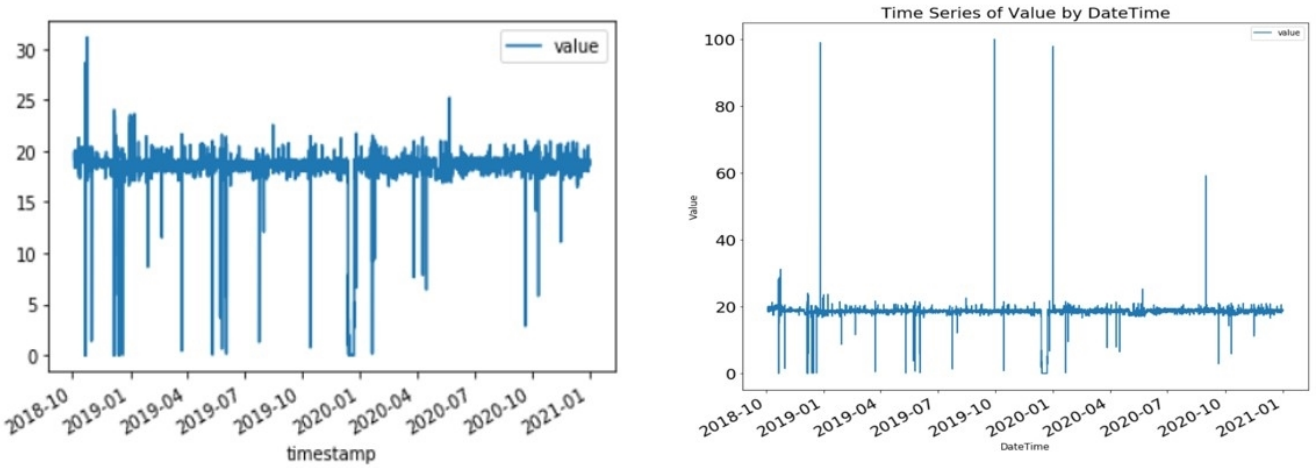
Table 2. Behavior of SCADA pressure dataset using different machine learning algorithms.

Algorithm	Accuracy (%)	Training Time (ms)	MCE	Prediction Speed (obs/s)
Decision Trees				
Fine Tree (FT)	100	1.1708	0	1,200,000
Medium Tree (MT)	100	1.0781	0	1,300,000
Coarse Tree (CT)	100	0.45488	0	1,000,000
Optimizable Tree	100	21.323	0	1,300,000
Discriminant Analysis				
Linear Discriminant (LDR)	100	1.843	24	1,100,000
Quadratic Discriminant (QDR)	99.2	1.1597	518	1,600,000
Optimizable Discriminant	100	25.029	24	1,600,000
Logistic Regression (LR)	100	3.205	N/A	1,100,000
Naive Bayes				
Gaussian Naive Bayes (GNB)	99.2	1.4947	518	1,400,000
Kernel Naive Bayes (KNB)	100	65.633	8	4500
Optimizable NB	100	918.96	8	3800
Support Vector Machines (SVM)				
Linear SVM	100	7.3065	25	780,000
Quadratic SVM	100	383.79	17	1,500,000
Cubic SVM	80.2	1657.3	13,588	930,000
Fine Gaussian SVM	100	7.433	5	610,000
Medium Gaussian SVM	100	5.3155	1	760,000
Coarse Gaussian SVM	100	5.1452	20	1,100,000
Optimized SVM	100	7490.9	25	1,100,000
Nearest Neighbors				
Fine KNN	100	3.6447	0	820,000
Medium KNN	100	2.0989	5	460,000
Coarse KNN	99.9	3.5228	35	130,000
Cosine KNN	99.9	17.422	35	17,000
Cubic KNN	100	2.3157	5	380,000
Weighted KNN	100	2.1524	0	450,000
Ensemble Learning (EL)				
Boosted Trees	99.9	5.0025	35	1,200,000
Bagged Tree	100	8.5874	0	320,000
Subspace Discriminant	100	4.5421	24	260,000
Subspace KNN	100	12.777	0	93,000
RUS-Boosted Tree	100	2.4396	20	960,000
Optimized Ensemble	100	232.87	0	530,000

4. Result Discussion and Performance Evaluation

The extracted real-time 68,722 pressure values which is an essential process variable from the SCADA system were plotted against the date and time. Pressure is a critical process variable in this process as over-pressurization could lead to explosion and under-pressurization could lead to the implosion of the process vessel, either with catastrophic results which will impact adversely the people, assets, and the environment. The features of

the extracted real-time data plotted in Figure 5a show that the data do not contain extremely high or extremely low values of pressure for the period under review. For the purpose of simulating the man-in-the-middle (MitM) attack, extreme values of pressure were injected into the dataset on specific dates and times. Figure 5b shows the plot of SCADA pressure against the date and time with the anomalies injected.



(a) Plot of Pressure (value) against Date and Time with 68,722 raw data samples without anomaly *(y-axis is the pressure while the x-axis is the Date and Time)*

(b) Plot of Pressure (value) against Date and Time, 68,722 data samples with anomalies injected *(y-axis is the pressure while the x-axis is the Date and Time)*

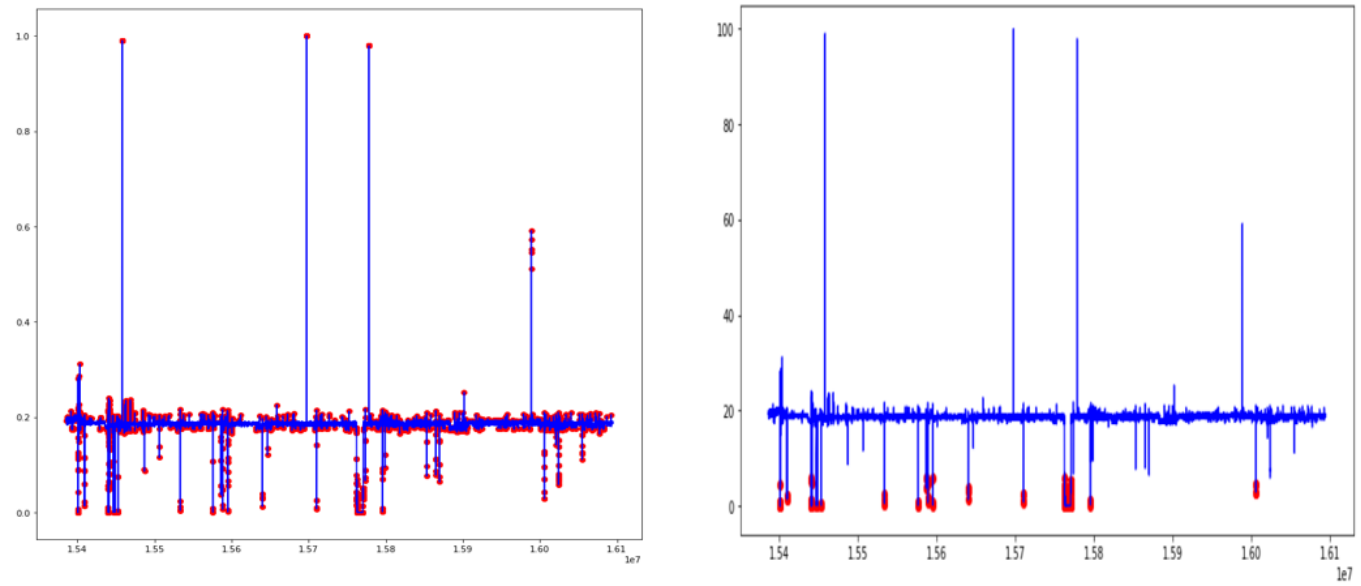
Figure 5. Visualization of the extracted pressure values from the SCADA with and without anomalies.

In Figure 6a, with the contamination parameter set to 0.1, the isolation forest algorithm showed high sensitivity in detecting changes in the pressure values for the period under review including the extreme high-pressure values, and detected all as anomalies. This can be termed high False Alarm Rates (FAR). With the contamination parameter set to 0.01, the Isolation Forest was able to detect as anomalies the extreme low-pressure values only with reduced FAR, but it was unable to identify the extremely high anomalies in the dataset and this makes this algorithm for real-time detecting MitM attacks as shown in Figure 6b.

In Figure 7a, with step set to 34361, batch size of 32 and 20 epochs, the long short-term memory (LSTM) algorithm detected some of the extreme pressure values for the period under review. Changing the batch size to 128 as in Figure 7b, the algorithm detected all the extreme high-pressure values as anomalies though with FAR. The algorithm was unable to identify the extremely low anomalies in the dataset, which makes it unreliable for real-time detection of MitM attacks.

Figure 8a–c show the plot of Python Outlier Detection (PyOD) incorporating interquartile range (IQR), k-nearest neighbor (kNN), and local outlier factor (LOF). The results of this algorithm show high sensitivity in detecting pressure value changes by all three algorithms. Although the IQR could detect extreme high-pressure and low-pressure with high FAR, kNN, and LOF failed to detect extreme high-pressure values correctly. KNN and LOF accuracy of about 70% and a high FAR makes them unsuitable for detecting and mitigating MitM attacks.

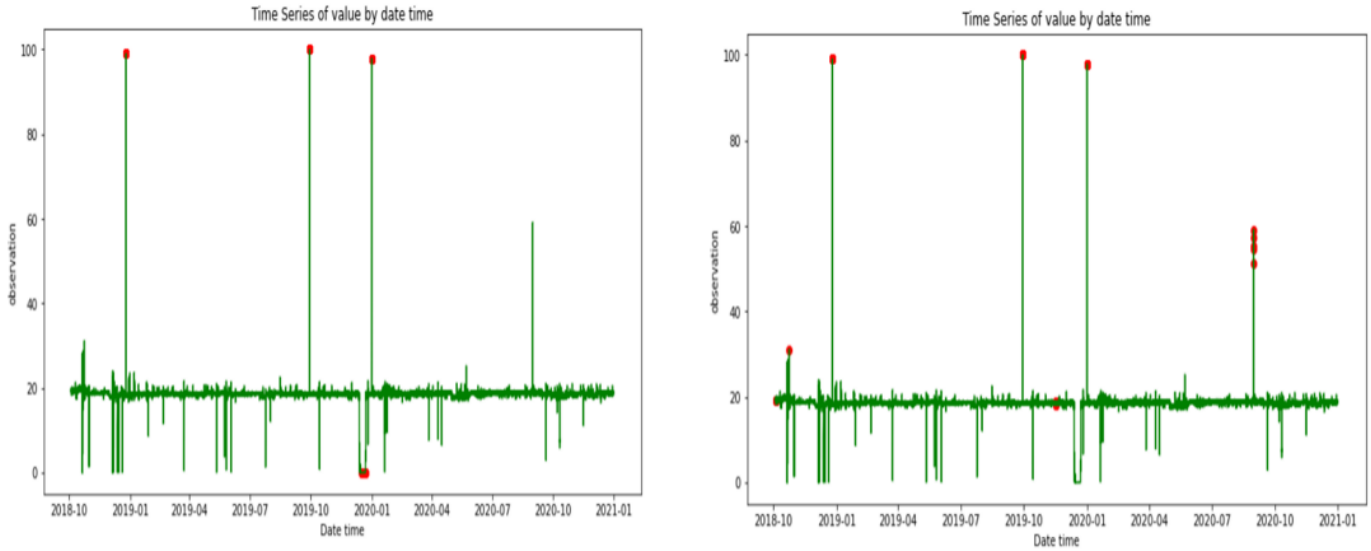
We applied the same 68,722 real-time SCADA pressure dataset to several other machine learning algorithms and compared their performance metrics, which are accuracy, Receiver operator characteristics (ROC), confusion matrix, training time, misclassification error (MCE), and prediction speed; the outcome is shown in Table 2. Based on these combined machine learning metrics as shown in Table 2, it was concluded that the coarse tree algorithm has significant performance and can detect MitM attacks effectively with negligible FAR.



(a) Plot of Isolation Forest Algorithm anomaly detection with 68,722 dataset, contamination parameter set to 0.1 *(y-axis is the pressure while the x-axis is the Date and Time)*

(b) Plot of Isolation Forest Algorithm anomaly detection with 68,722 dataset, contamination parameter set to 0.01 *(y-axis is the pressure while the x-axis is the Date and Time)*

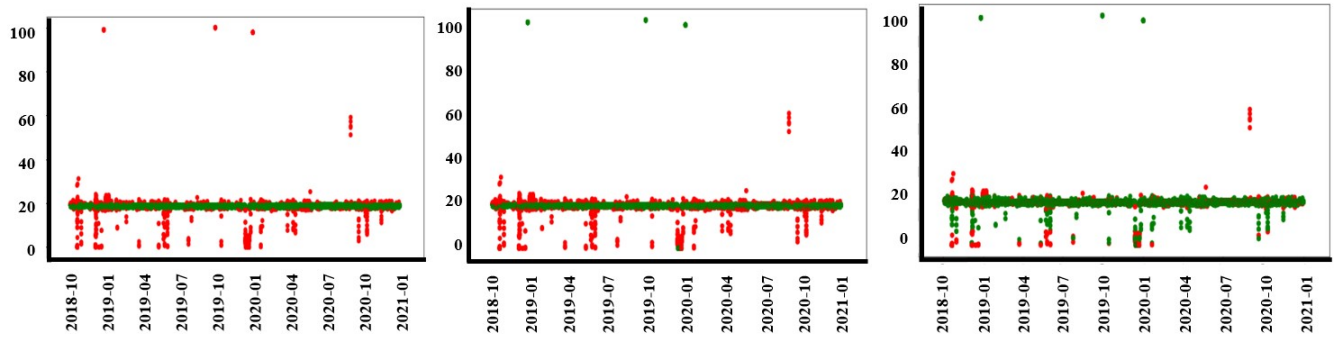
Figure 6. Effect of contamination parameter on the isolation forest algorithm.



(a) Plot of LSTM Algorithm anomaly detection with a dataset of 68,722, time step of 34361, batch size of 32 and 20 epochs *(y-axis is the pressure while the x-axis is the Date and Time)*

(b) Plot of LSTM Algorithm anomaly detection with a dataset of 68,722, time step of 34361, batch size of 128 and 20 epochs *(y-axis is the pressure while the x-axis is the Date and Time)*

Figure 7. Effect of batch size variation on the LSTM algorithm.



(a) Plot of Inter Quartile Range (IQR) Algorithm anomaly detection with a dataset of 68,722 *(y-axis is the pressure while the x-axis is the Date and Time)*

(b) Plot of k Nearest Neighbor (kNN) Algorithm anomaly detection with a dataset of 68,722 *(y-axis is the pressure while the x-axis is the Date and Time)*

(c) Plot of Local Outlier Factor (LOF) Algorithm anomaly detection with a dataset of 68,722 *(y-axis is the pressure while the x-axis is the Date and Time)*

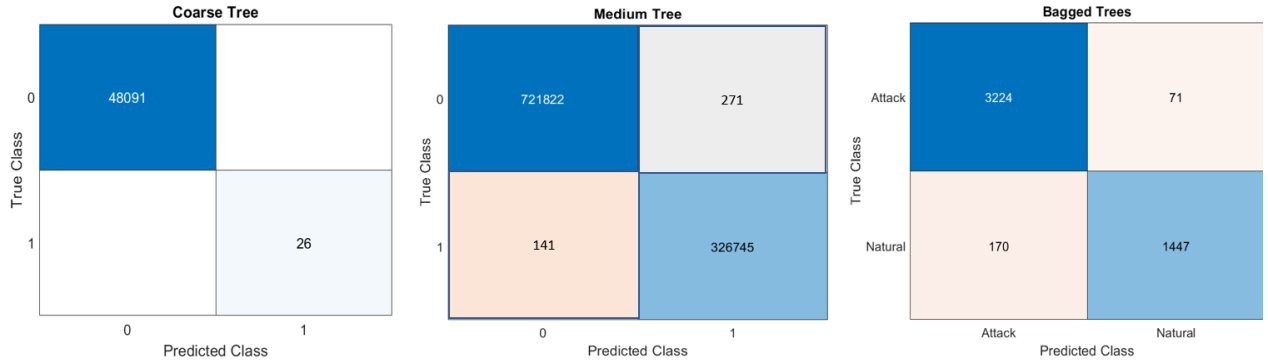
Figure 8. Plot of local outlier factor, KNN performance, and interquartile range result.

In addition, a thorough comparison was made between the results achieved with the real-time dataset and three public datasets, namely, the WUSTL-2018 [35], ORNL Power Grid [46], and TON_IoT datasets [47]. It is important to state that the FAR recorded with the SCADA was zero as compared to other datasets used by other researchers. The result is shown in Table 3, confirming the superior performance of tree algorithms such as the coarse tree and ensemble tree such as the bagged tree. In addition, although the bagged tree was consistent in most of the datasets in terms of high accuracy, it came with the challenge of high training time justifying the use of the coarse tree where an accuracy-time trade-off is needed.

Table 3. Best and worst-performing machine learning algorithms on various public datasets.

Datasets /Algorithm	Accuracy (%)	Training Time (ms)	FAR	Prediction Speed (obs/s)
SCADA Pressure Dataset				
Coarse Tree	100	0.4549	0	1,000,000
Cubic SVM	80.2	1657.3	13,588	930,000
WUSTL-SCADA-2018 Dataset [35]				
Medium Tree	100	5.6605	412	4,100,000
Subspace Discriminant	93.1	101.64	72,009	110,000
ORNL POWER GRID Dataset [46]				
Bagged Tree	95.1	4.8021	241	2500
Quadratic Discriminant	52.4	1.6364	2339	120,000
TON_IoT DATASET [47]				
Bagged Tree	100	1789.5	9	61,000
Coarse Tree	82.4	94.643	81,043	1,000,000

Figure 9a–c show the plot of the Confusion matrix of the tree algorithm with the best performance using the 68,722 real-time SCADA pressure dataset, which shows zero false positives as compared with other WUSTL and ORNL datasets used by other researchers which produced 141 and 170 false positives, respectively.



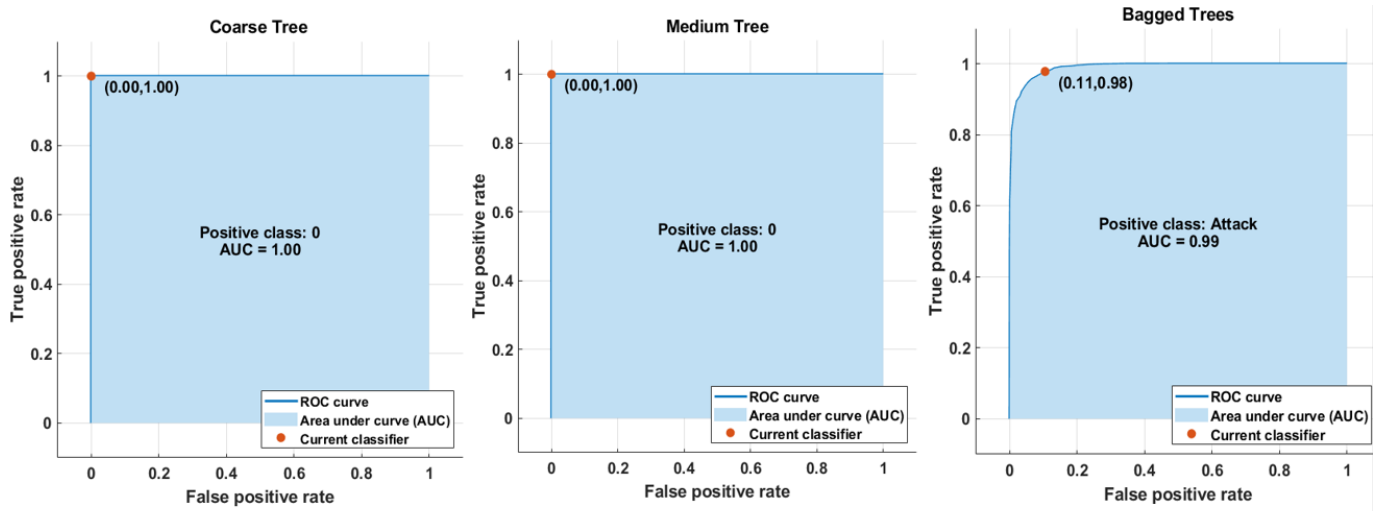
(a) Confusion Matrix of the Best Performed Tree Algorithm in SCADA Pressure Dataset

(b) Confusion Matrix of the Best Performed Tree Algorithm in WUSTL-SCADA-2018 Dataset

(c) Confusion Matrix of the Best Performed Tree Algorithm in ORNL (Power Grid) SCADA Dataset

Figure 9. Plot of confusion matrix of the tree algorithms using the facility dataset and public datasets.

Figure 10a–c shows the plot of the receiver operator characteristics (ROC) curve of the best-performing tree algorithm using the 68,722 real-time SCADA pressure dataset, which shows coarse tree produced the best result with zero false positives and better area under the curve (AUC). In contrast, WUSTL and ORNL showed in the medium tree and bagged tree, respectively, with lesser AUC. Besides the SCADA pressure dataset, the ensemble bagged tree is generally desirable across public datasets when accuracy is considered over another trade-off such as training time.



(a) ROC of the Best Performed Tree Algorithm in SCADA Pressure Dataset

(b) ROC of the Best Performed Tree Algorithm in WUSTL-SCADA-2018 Dataset

(c) ROC of the Best Performed Tree Algorithm in ORNL (Power Grid) SCADA Dataset

Figure 10. Plot of receiver operator characteristics (ROC) curves.

The TON_IoT dataset [47] was used for result validation. Figures 11 and 12 reveal the true positive rate (TPR) and false discovery rate (FDR) of the bagged tree and coarse tree, respectively. Although the coarse tree performed best in our field data, bagged tree remains the best ensemble learning method capable of making up for the weaknesses and leveraging the strengths of tree algorithms, as it detected the MiTM attacks.

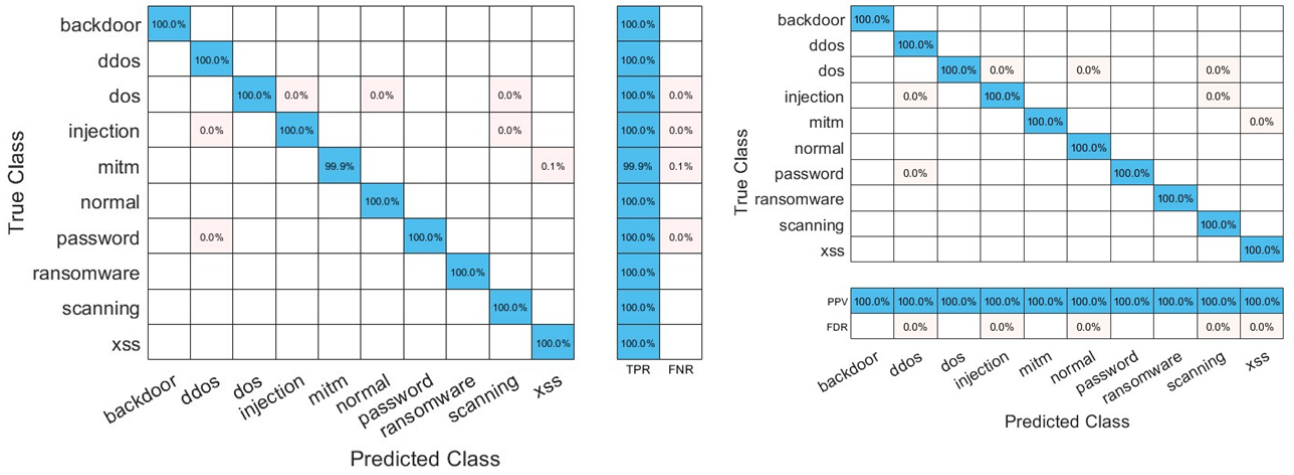


Figure 11. Bagged trees graph showing the (**a**) probability of the positive classification of all attacks (TPR/sensitivity) and (**b**) the proportion of all the identified attacks that may not be attacked (false discovery rates (FDR)) on the TON_IoT datasets.

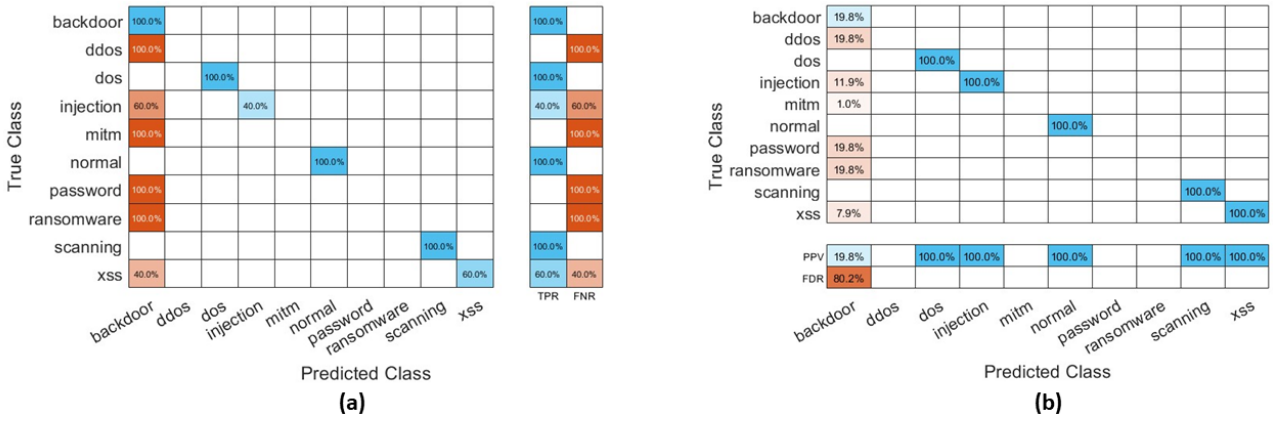


Figure 12. Coarse trees graph illustrating the (**a**) probability of the positive classification of all attacks (TPR/sensitivity) and (**b**) the proportion of all the identified attacks that may not be attacked (false discovery rates (FDR)) on the TON_IoT datasets.

The TPR and FDR performance of the coarse tree and ensemble tree on other datasets (SCADA pressure, WUSTL-2018, and ORNL Power Grid) are shown in Figures 13 and 14. The results are consistent with the confusion matrix and ROC results.

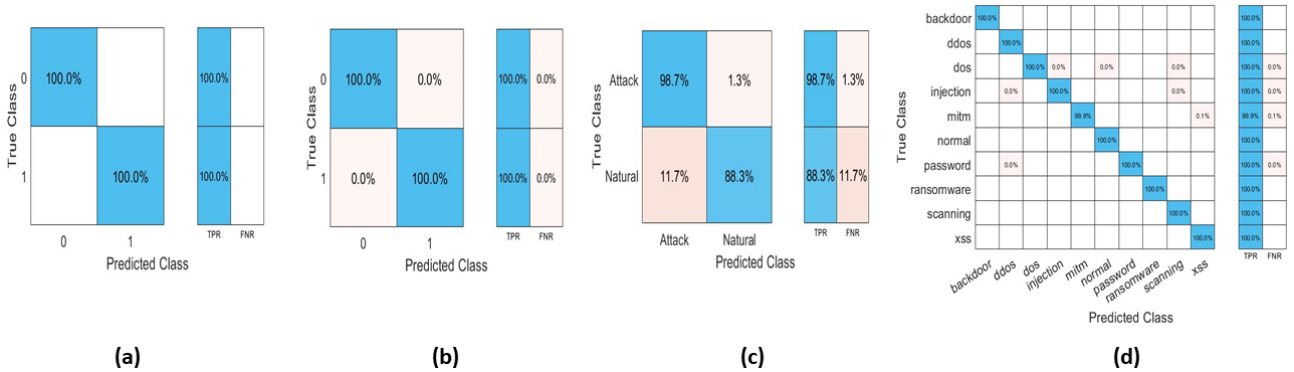


Figure 13. Graph illustrating the probability of the positive attack classification (TPR/sensitivity) across all evaluated datasets ((**a**) SCADA Pressure, (**b**) WUSTL-2018, (**c**) ORNL PowerGrid (**d**) TON_IoT).

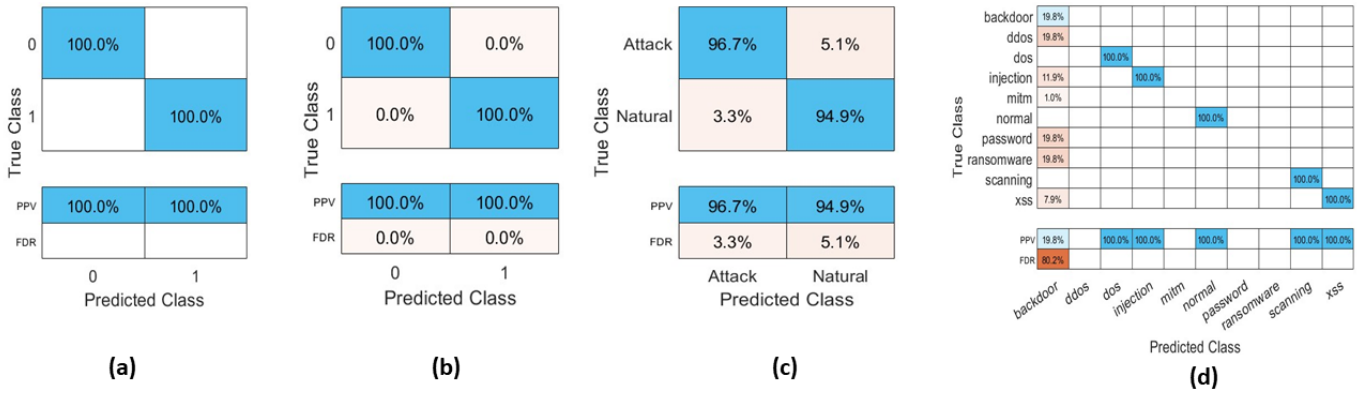


Figure 14. Graph demonstrating the proportion of all the identified attacks that may not be attacked (false discovery rates (FDR)) across all evaluated datasets (**a**) SCADA Pressure, (**b**) WUSTL-2018, (**c**) ORNL PowerGrid (**d**) TON_IoT.

5. Conclusions

The outcome of this study is the evaluation of different machine learning algorithms on the 68,722 SCADA real-time datasets using the following combined machine learning performance metrics: high accuracy, earliest training time, fastest prediction speed, negligible MCE, and less computation power requirement. Based on these combined machine learning performance metrics using the 68,722 datasets, it was concluded that the coarse tree algorithm showed the best performance, and is regarded as the most suitable for the detection of MitM attacks in a process control network of an oil and gas installation. This study can be improved upon by evaluating more machine learning algorithms as well as the use of more real-time SCADA datasets which improve the detection of other forms of cyber-attacks. More real-time SCADA data samples are required in order to develop very accurate and reliable anomaly detection models. Because the tree algorithms performed best, it is recommended to perform more hyperparameter tuning of the ensemble tree, which will aid the trade-off between accuracy and training time. The use of different datasets is a helpful approach to validation. However, such datasets should have a comparable relationship to avoid misjudgment of performance. In this work, the coarse tree had a limitation of confusing the MiTM attack with a backdoor attack, and more work will be required to handle multiclass detection and classification. In our approach, we solved this problem by employing an ensemble tree but it came with a training time burden.

Author Contributions: Conceptualization, U.O.O., C.I.N. and F.K.O.; methodology, all authors contributed equally; software, U.O.O., C.I.N. and M.M.A.; validation, U.O.O., C.I.N., F.K.O., C.C.M., J.-K.C.O. and I.O.A.; formal analysis, U.O.O. and C.I.N.; investigation, all authors contributed equally; resources, U.O.O. and C.I.N.; data curation, U.O.O., C.I.N. and M.M.A.; writing—original draft preparation, all authors contributed; writing—review and editing, U.O.O. and C.I.N.; visualization, U.O.O., C.I.N. and M.M.A.; supervision, F.K.O., J.-K.C.O., I.O.A. and C.I.N.; project administration, U.O.O. and C.I.N.; funding acquisition, U.O.O. and C.I.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research work received no external funding

Data Availability Statement: Not applicable.

Acknowledgments: The authors wish to thank Love Allen Chijioke Ahakonye of the Networked Systems Laboratory, IT-Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea for assisting in the simulation during the preliminary stage of the project.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

APT	Advance Persistent Threats
AUC	Area Under Curve
CI	Critical Infrastructure
DDoS	Distributed Denial-of-Service
DoS	Denial-of-Service
DMZ	Demilitarized Zone
FAR	False Alarm Rates
FDIA	False Data Injection Attacks
FDR	False Discovery Rates
FNR	False Negative Rates
FPR	False Positive Rates
HMI	Human Machine Interface
ICS	Industrial Control Systems
ICT	Information and Communication Technology
IDS	Intrusion detection systems
IIoT	Industrial Internet of Things
IoT	Internet of Things
IP	Internet Protocol
IQR	Interquartile Range
IT	Information Technology
kNN	k-Nearest Neighbors
LDR	Linear Discriminant Regression
LOF	Local Outlier Factor
LSTM	Long Short-Term Memory
MATLAB	Matrix Laboratory
MCE	Misclassification error
MitM	Man-in-the-Middle
MotS	Man-on-the-Side
NSL-KDD	National Security Laboratory Knowledge Discovery in Databases
ORNL	Oak Ridge National Laboratories
OT	Operation Technology
PCN	process control network
PLC	Programmable Logic Controller
PPV	Positive Predictive Values
PyOD	Python Outlier detection
ROC	Receiver Operator Characteristics
SCADA	Supervisory Control and Data Acquisition
SVM	Support Vector Machines
TPR	True Positive Rates
USB	Universal Serial Bus
WUSTL	Washington University in St. Louis
Xss	Cross-site Scripting Attack

References

1. Smurthwaite, M.; Bhattacharya, M. Convergence of IT and SCADA: Associated Security Threats and Vulnerabilities. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *790*, 012041. [CrossRef]
2. CISA; FBI. DarkSide Ransomware: Best Practices for Preventing Business Disruption from Ransomware Attacks. *Cybersecur. Advis.* **2021**. Available online: https://www.cisa.gov/news-events/cybersecurity-advisories/aa21-131a (accessed on 29 May 2023).
3. Marchetti, M.; Pierazzi, F.; Guido, A.; Colajanni, M. Countering Advanced Persistent Threats through Security intelligence and big data analytics. In Proceedings of the 2016 8th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 31 May–3 June 2016; pp. 243–261. [CrossRef]
4. Kaspersky. APT trends report Q2 2022. Available online: https://securelist.com/apt-trends-report-q2-2022/106995/ (accessed on 29 May 2023)
5. Kaspersky. APT trends report Q2 2019. Available online: https://securelist.com/apt-trends-report-q2-2019/91897/ (accessed on 29 May 2023)

6. Irmak, E.; Erkek, İ. An overview of cyber-attack vectors on SCADA systems. In Proceedings of the 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, 22–25 March 2018; pp. 1–5. [CrossRef]
7. Stergiopoulos, G.; Gritzalis, D.A.; Limnaios, E. Cyber-Attacks on the Oil & Gas Sector: A Survey on Incident Assessment and Attack Patterns. *IEEE Access* **2020**, *8*, 128440–128475. [CrossRef]
8. Nwakanma, C.I.; Ahakonye, L.A.C.; Njoku, J.N.; Eze, J.; Kim, D.S. Effective Industrial Internet of Things Vulnerability Detection Using Machine Learning. In Proceedings of the 2022 5th Information Technology for Education and Development (ITED), Abuja, Nigeria, 1–3 November 2022; pp. 1–8. [CrossRef]
9. Ahakonye, L.A.C.; Nwakanma, C.I.; Lee, J.M.; Kim, D.S. Agnostic CH-DT Technique for SCADA Network High-Dimensional Data-Aware Intrusion Detection System. *IEEE Internet Things J.* **2023**, *10*, 10344–10356. [CrossRef]
10. Ogu, R.E.; Achumba, I.E.; Okoronkwo, C.D.; Chukwudebe, G.A.; Chukwuchekwa, N. An IoT Solution for Air Quality Monitoring and Hazard Identification for Smart City Development. In Proceedings of the 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON), Lagos, Nigeria, 5–7 April 2022; pp. 1–5. [CrossRef]
11. Ogu, R.E.; Chukwudebe, G.A.; Achumba, I.E.; Chukwuchekwa, N.; Ezenugu, I.A. A Robust IoT-based Air Quality Monitoring Node for Multi-Location Deployment. *Int. J. Eng. Res. Technol.* **2022**, *11*, 146–151.
12. Ahakonye, L.A.C.; Nwakanma, C.I.; Lee, J.M.; Kim, D.S. SCADA intrusion detection scheme exploiting the fusion of modified decision tree and Chi-square feature selection. *Internet Things* **2023**, *21*, 100676. [CrossRef]
13. Alves, T.; Das, R.; Morris, T. Embedding Encryption and Machine Learning Intrusion Prevention Systems on Programmable Logic Controllers. *IEEE Embed. Syst. Lett.* **2018**, *10*, 99–102. [CrossRef]
14. Ramotsoela, D.; Hancke, G.; Abu-Mahfouz, A. Attack detection in water distribution systems using machine learning. *Hum. Cent. Comput. Inf. Sci.* **2019**, *9*, 1–22. [CrossRef]
15. Zoppi, T.; Ceccarelli, A.; Bondavalli, A. Unsupervised Algorithms to Detect Zero-Day Attacks: Strategy and Application. *IEEE Access* **2021**, *9*, 90603–90615. [CrossRef]
16. Pu, G.; Wang, L.; Shen, J.; Dong, F. A hybrid unsupervised clustering-based anomaly detection method. *Tsinghua Sci. Technol.* **2021**, *26*, 146–153. [CrossRef]
17. Rosa, L.; Cruz, T.; de Freitas, M.B.; Quitério, P.; Henriques, J.; Caldeira, F.; Monteiro, E.; Simões, P. Intrusion and anomaly detection for the next-generation of industrial automation and control systems. *Futur. Gener. Comput. Syst.* **2021**, *119*, 50–67. [CrossRef]
18. Ahakonye, L.A.C.; Nwakanma, C.I.; M., L.J.; Kim, D.S. Efficient Classification of Enciphered SCADA Network Traffic in Smart Factory Using Decision Tree Algorithm. *IEEE Access* **2021**, *9*, 154892–154901. [CrossRef]
19. Melnick, J. Top 10 Most Common Types of Cyber Attacks. Available online: https://blog.netwrix.com/2018/05/15/top-10-most-common-types-of-cyber-attacks/ (accessed on 29 May 2023).
20. Mahrukh, M.;Thomas, M.S. Load Altering Attacks- a Review of Impact and Mitigation Strategies. In Proceedings of the 2023 International Conference on Recent Advances in Electrical, Electronics and Digital Healthcare Technologies (REEDCON), New Delhi, India, 1–3 May 2023; pp. 397–402.
21. Tang, S.; Liu, Z.; Wang, L. Power System Reliability Analysis Considering External and Insider Attacks on the SCADA System. In Proceedings of the 2020 IEEE/PES Transmission and Distribution Conference and Exposition, Chicago, IL, USA, 12–15 October 2020; pp. 1–4. [CrossRef]
22. Hunga, M.O.; Adishi, E. Oil Theft, Illegal Bunkering and Pipeline Vandalism: It's Impact on Nigeria Economy, 2015–2016. *IIARD Int. J. Econ. Bus. Manag.* **2017**, *3*, 47–65.
23. Wilson, G. The Nigerian State and Oil Theft in the Niger Delta Region of Nigeria. *J. Sustain. Dev. Afr.* **2014**, *16*, 69–81.
24. Mohammed, A.S.; Saxena, N.; Rana, O. Wheels on the Modbus - Attacking ModbusTCP Communications. In Proceedings of the WiSec '22: Proceedings of the 15th ACM Conference on Security and Privacy in Wireless and Mobile Networks, Miami, FL, USA, 15–17 May 2022; pp. 288–289. [CrossRef]
25. Jang-jaccard, J.; Nepal, S. A survey of emerging threats in cybersecurity. *J. Comput. Syst. Sci.* **2014**, *80*, 973–993. [CrossRef]
26. Amin, S.; Litrico, X.; Sastry, S.; Bayen, A. Cyber security of water SCADA systems- part II: Attack detection using enhanced hydrodynamic models. *IEEE Trans. Control Syst. Technol.* **2013**, *21*, 1679–1693. [CrossRef]
27. Zoppi, T.; Ceccarelli, A.; Capecchi, T.; Bondavalli, A. Unsupervised Anomaly Detectors to Detect Intrusions in the Current Threat Landscape. *ACM/IMS Trans. Data Sci.* **2021**, *2*, 1–26. [CrossRef]
28. Abrar, I.; Ayub, Z.; Masoodi, F.; Bamhdi, A.M. A Machine Learning Approach for Intrusion Detection System on NSL-KDD Dataset. In Proceedings of the 2020 International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 18–20 September 2020; pp. 919–924. [CrossRef]
29. Kulugh, V.E.; Mbanaso, U.M.; Chukwudebe, G.A. Cybersecurity Resilience Maturity Assessment Model for Critical National Information Infrastructure. *SN Comput. Sci.* **2022**, *3*, 217. [CrossRef]
30. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 1–21. [CrossRef]
31. Zoppi, T.; Ceccarelli, A.; Salani, L.; Bondavalli, A. On the educated selection of unsupervised algorithms via attacks and anomaly classes. *J. Inf. Secur. Appl.* **2020**, *52*, 102474. [CrossRef]
32. Khraisat, A.; Alazab, A.A. A critical review of intrusion detection systems in the internet of things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. *Cybersecurity* **2021**, *4*, 1–31. [CrossRef]

33. Okafor, K.C.; Ndinechi, M.C.; Misra, S. Cyber-physical network architecture for data stream provisioning in complex ecosystems. *Trans. Emerg. Telecommun. Technol.* **2022**, *33*, 1–31. [CrossRef]
34. Wakefield, K. *A Guide to the Types of Machine Learning Algorithms: SAS UK*; SAS Institute: London, UK, 2021.
35. Teixeira, M.A.; Salman, T.; Zolanvari, M.; Jain, R.; Meskin, N.; Samaka, M. SCADA System Testbed for Cybersecurity Research Using Machine Learning Approach. *Future Internet* **2018**, *10*, 76. [CrossRef]
36. El Naqa, I.; Murphy, M.J. *What Is Machine Learning?* Springer: Berlin/Heidelberg, Germany, 2015; pp. 3–11.
37. Ndubuaku, M.U.; Anjum, A.; Liotta, A. Unsupervised anomaly thresholding from reconstruction errors. *Lect. Notes Comput. Sci.* **2019**, *11874LNCS*, 123–129. [CrossRef]
38. Joloudari, J.H.; Haderbadi, M.; Mashmool, A.; Ghasemigol, M.; Band, S.S.; Mosavi, A. Early detection of the advanced persistent threat attack using performance analysis of deep learning. *IEEE Access* **2020**, *8*, 186125–186137. [CrossRef]
39. Al-Abassi, A.; Karimipour, H.; Dehghantanha, A.; Parizi, R.M. An Ensemble deep learning-based cyber-attack detection in industrial control system. *IEEE Access* **2020**, *8*, 83965–83973. [CrossRef]
40. Bierbrauer, D.A.; Chang, A.; Kritzer, W.; Bastian, N.D. Cybersecurity Anomaly Detection in Adversarial Environments. *arXiv* **2021**, arXiv:2105.06742.
41. Alhaidari, F.A.; AL-Dahasi, E.M. New Approach to Determine DDoS Attack Patterns on SCADA System Using Machine Learning. In Proceedings of the 2019 International Conference on Computer and Information Sciences (ICCIS), Aljouf, Saudi Arabia, 3–4 April 2019; pp. 1–6. [CrossRef]
42. Al-Rabiaah, S. The "Stuxnet" Virus of 2010 As an Example of A "APT" and Its "Recent" Variances. In Proceedings of the 2018 21st Saudi Computer Society National Computer Conference (NCC), Riyadh, Saudi Arabia, 25–26 April 2018; pp. 1–5. [CrossRef]
43. Maynard, P.; McLaughlin, K. *Towards Understanding Man-on-the-Side Attacks (MotS) in SCADA Networks, Proceedings of the 17th International Joint Conference on e-Business and Telecommunications, ICETE 2020—Volume 2: SECRYPT, Lieusaint, Paris, France, 8–10 July 2020*; Samarati, P., di Vimercati, S.D.C., Obaidat, M.S., Ben-Othman, J., Eds.; ScitePress: Setubal, Portugal, 2020; pp. 287–294. [CrossRef]
44. Wilson, D.; Tang, Y.; Yan, J.; Lu, Z. Deep Learning-Aided Cyber-Attack Detection in Power Transmission Systems. In Proceedings of the 2018 IEEE Power & Energy Society General Meeting (PESGM), Portland, OR, USA, 5–10 August 2018; pp. 1–5. [CrossRef]
45. Husák, M.; Komárková, J.; Bou-Harb, E.; Čeleda, P. Survey of Attack Projection, Prediction, and Forecasting in Cyber Security. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 640–660. [CrossRef]
46. Morris, T. Industrial Control System (ICS) Cyber Attack Datasets. In *Proceedings of the Center for Cybersecurity Research and Engineering (CCRE)*; The University of Alabama in Huntsville: Huntsville, AL, USA. Available online: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets (accessed on 29 May 2023).
47. Moustafa, N. ToN_IoT datasets. *IEEE Dataport* **2019**. [CrossRef]
48. Song, S.; Liu, X.; Li, C.; Li, Z.; Zhang, S.; Wu, W.; Shi, B.; Kang, Q.; Wu, H.; Gong, J. Dynamic Simulator for Three-Phase Gravity Separators in Oil Production Facilities. *ACS Omega* **2023**, *8*, 6078–6089. [CrossRef]
49. Abdu Sabir, B.M.; Elamin, I.H.; Sadiq, H.R. Dynamic Modelling and Simulation of A Three-Phase Gravity Separator. *J. Karary Univ. Eng. Sci.* **2022**, *11*, 1–19. [CrossRef]
50. Wu, F.; Huang, K.; Li, H.; Huang, C. Analysis and Research on the Automatic Control Systems of Oil-Water Baffles in Horizontal Three-Phase Separators. *Processes* **2022**, *10*, 1102. [CrossRef]
51. Jonach, T.; Jordan, C.; Haddadi, B.; Harasek, M. Modelling and Simulation of 3-Phase Separators in the Oil and Gas Industry with Emphasis on Water Quality. *Chem. Eng. Trans.* **2022**, *94*, 1009–1014. [CrossRef]
52. Nasr, P.M.; Varjani, A.Y. Petri net model of insider attacks in SCADA system. In Proceedings of the 2014 11th International ISC Conference on Information Security and Cryptology, Tehran, Iran, 3–4 September 2014; pp. 55–60. [CrossRef]
53. Nasr, P.M.; Varjani, A.Y. Alarm based anomaly detection of insider attacks in SCADA system. In Proceedings of the 2014 Smart Grid Conference (SGC), Tehran, Iran, 9–10 December 2014; pp. 1–6. [CrossRef]
54. Gönen, S.; Sayan, H.H.; Yılmaz, E.N.; Üstünsoy, F.; Karacayılmaz, G. False data injection attacks and the insider threat in smart systems. *Comput. Secur.* **2020**, *97*, 101955. [CrossRef]
55. Radoglou-Grammatikis, P.; Dalamagkas, C.; Lagkas, T.; Zafeiropoulou, M.; Atanasova, M.; Zlatev, P.; Boulogeorgos, A.A.A.; Argyriou, V.; Markakis, E.K.; Moscholios, I.; et al. False Data Injection Attacks against Low Voltage Distribution Systems. In Proceedings of the GLOBECOM 2022–2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 4–8 December 2022; pp. 1856–1861. [CrossRef]