Real-time anomaly intrusion detection for a clean water supply system, utilising machine learning with novel energy-based features.

1st Andres Robles-Durazno
School of Computing
Edinburgh Napier University
Edinburgh, UK
a.roblesdurazno@napier.ac.uk

2nd Naghmeh Moradpoor
School of Computing
Edinburgh Napier University
Edinburgh, UK
n.moradpoor@napier.ac.uk

3rd James McWhinnie
School of Engineering and Built Environment
Edinburgh Napier University
Edinburgh, UK
j.mcwhinnie@napier.ac.uk

4th Gordon Russell
School of Computing
Edinburgh Napier University
Edinburgh, UK
g.russell@napier.ac.uk

Abstract— ***Industrial Control Systems have become a priority domain for cybersecurity practitioners due to the number of cyber-attacks against those systems has increased over the past few years. This paper proposes a real-time anomaly intrusion detector for a model of a clean water supply system. A testbed of such system is implemented using the Festo MPA Control Process Rig. A set of attacks to the testbed is conducted during the control process operation. During the attacks, the energy of the components is monitored and recorded to build a novel dataset for training and testing a total of five traditional supervised machine learning algorithms: K-Nearest Neighbour, Support Vector Machine, Decision Tree, Naïve Bayes and Multilayer Perceptron. The trained machine learning algorithms were built and deployed online, during the control system operation, for further testing. The performance obtained from offline and online training and testing steps are compared. The captures results show that KNN and SVM outperformed the rest of the algorithms by achieving high accuracy scores and low false-positive, false-negative alerts.***

Keywords— Industrial Control System, Energy Monitoring, SCADA, KNN, SVM, Anomaly Detection, IDS.

I. INTRODUCTION

An industrial control system (ICS) is a wide class of automation system used to provide control and monitoring functionality in large industrial facilities and critical infrastructures such as chemical, transportation, nuclear, pharmaceutical, oil refineries, power generation plants and water/sewage treatment plants[1]. In early years, the components involved in an ICS, such as the Programming Logic Controller (PLC), did not have any computing or networking capability, however, the development of technology over the years and the introduction of Industry 4.0[2] has led to developing powerful ICS components such as remote terminal units, PLC's and sensors that have networking and wireless capabilities.

These advanced ICS components allow to monitor and operate a control process in almost any location across the globe, although, its online availability has also attracted the attention of different cyber threat actors that have various motivations such as hacktivists, state-sponsored actors, cybercriminals and cyber terrorists[3]. For instance, Stuxnet[4] is the first documented cyberweapon, discovered in 2010, that targeted the Iranian's nuclear program. It was designed to sabotage the centrifuges used to enrich the uranium. This attack raised awareness of cybersecurity issues for critical infrastructures around the world. Further, documented reports show an increment of 39% of attacks against ICS in 2018 when compared to 2017 [5].

This paper extends our previous work[6], which shows the feasibility of detecting anomalies on control systems using machine learning, and proposes a real-time energy-based anomaly intrusion

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

detection system (EBIDS) for a model of a clean water supply system, utilising machine learning. Current related work focuses on building and testing machine learning models with publicly available datasets [9][8] or from virtual testbeds [7], however, these models lack online validation. For instance, a machine learning model might obtain consistent metrics during the evaluation process, but it might not have the same behaviour when deployed online. Therefore, in this paper, we develop and evaluate the performance of our anomaly detection system in offline and online mode. Further, unlike other approaches that rely on the information collected from the control network packets, our EBIDS proposes a stronger ICS architecture by adding an extra layer of protection at the lowest control level.

This paper is organized as follows. Section II describes the related work in the field. Section III gives a brief overview of the design and implementation. Section IV refers to the EBIDS evaluation, while the conclusions are presented at section V followed by Acknowledgments.

II. RELATED WORK

In this section, existing work related to intrusion detection schemes for ICS's is discussed.

In [7] the authors propose a sequence-aware intrusion detection in Industrial Control Systems (S-IDS) which is capable of identify patterns of ICS network events, extract their semantic meaning and models known behaviours over time. They record network messages and log entries to define ICS device operations by employing discrete-time Markov chains. The S-IDS proposed by the authors is a layered structure that collects information from Modbus network traffic and log files. To evaluate their approach, the authors train the S-IDS with data obtained from water treatment and purification system that used Modbus protocol for network communication. To simulate the attacks the authors, inject malicious traces on the network traffic prior to sending the data to the S-IDS. Addressing their results, the rate of false/positive alarms generated by the S-IDS is reduced when they include information of the ICS infrastructure and physical process. The attacks injected on the network traffic is also detected. It can be argued whether the S-IDS can validate tampered log files or crafted network packets that contain malicious data.

In [8] the authors proposed a support vector machine (SVM) approach for cyber-attack detection on Industrial Control Systems Monitoring. Their proposed approach includes building a discriminant model, using SVM, between benign and malicious traffic by analysing intervals and length of control network packets. Their testbed involves two tanks of water equipped with real control devices and controlled automatically. They collected ten datasets under normal and attack conditions. Their attack scenarios involved a formal pen testing using the Metasploit Framework Rapid7 attack tool. Addressing their captured results, their SVM model achieved 95% of precision and 0.50 error rate on average for the datasets recorded. It is unclear whether the attacks executed aimed to disrupt the control network or the ICS operation. Further, is not discussed the importance of the features chosen among others available on the control network packets.

In [9] the authors proposed a novel one-class classification approach for Cyberattack detection in a water distribution system. The novelty of their approach relies on the use of the truncated Mahalanobis distance in the decision function of the classifier, which, improves the classification speed when compared to similar one-class classifiers. In order to test their approach, they recorded a dataset that corresponds to the final stage of a real water drinking distribution plant. Further, the dataset includes four simulated attacks to components such as pump, flowmeters and sensors that compose the ICS. Their captures results outperformed other approaches, for the four types of attacks included in the dataset, by achieving 100%, 88.8%, 91.3% and 82.3% of detection rate. However, is unclear how the authors obtained the detection rate or how it is evaluated. Moreover, the authors do not indicate whether the dataset includes information from the control process of network features.

In [10] the authors propose a Neural Network approach for anomaly detection in a water treatment

system. To conduct the research, they used a dataset obtained from the SWaT testbed, which is an operational scaled-down water treatment plant. The authors propose several techniques to improve the anomaly detection which include exponentially weighted smoothing, mean p-powered error measure, individual error weight for each variable and disjoint prediction window. Addressing their captured results, their machine learning models achieved 96.7%, 95.2% and 93.6% for MLP, CNN and RNN respectively. Although, it is argued whether this approach is applicable in an online environment since real-world applications demand high processing power.

In [11] the authors propose a multilayer data-driven approach for cyber-attack detection on Industrial Control Systems. Their proposed detection system is structured with a defence in depth concept that employs supervised and unsupervised machine learning models for intrusion detection. Their experimental setup includes a SCADA system and a testbed that simulates a two-loop nuclear system. The dataset contains malicious and benign network traffic and host system data collected by the Windows performance monitor. The malicious traffic includes packet sniffing using MITM, DoS, data exfiltration, false data injection and tampering, and simultaneous cyber-attacks which leads to a small loss of coolant accident. Addressing their captured results, they achieved a true positive rate of 98.84% for KNN followed by 98,27% for bagging, 97.69% for random forest and 94.80% for decision tree. Although their approach shows promising results, they still trust in the information collected from the network and their set of attacks is network-based.

In [12] the authors propose a one-class support vector machine (OCSVM) for intrusion detection in a SCADA system. They used datasets that contain malicious and benign traffic from a SCADA network that mainly involves MODBUS/TCP traffic for offline training. Their attack scenarios include man in the middle (MITM) by address resolution protocol (ARP), SYNC flooding and honeypot interaction. Addressing their captured results, the OCSVM intrusion detection was able to produce 98.42% and 99.12% accuracy for two online detection testing. It can be argued whether the evaluation of a machine learning model can be determined with by only one metric: accuracy. Further, their online detection process is unclear, and it does not provide a comparison between the result obtained during the offline validation and online testing.

In this paper, the online and offline performance of the machine learning algorithms is discussed. Unlike the work provided by [9], the performance of the machine learning algorithms is shown in detail highlighting the strengths and weaknesses found in each one of them. Further, despite the work of [7], [9], in this research, the features used to build the machine learning models are obtained from hard-wired current sensors placed in the middle of the sensors that compose the ICS and the PLC. Moreover, this research provides results obtained from the execution of a novel set of attacks against a physical testbed which differs from the work presented by [10], [12] where the authors use virtual testbeds or common network attacks such as DoS and Man in the middle.

III. DESIGN AND IMPLEMENTATION

For testing purposes, we implemented a model of clean water supply system(CWSS) using the Festo MPA Compact workstation[13] in the configuration shown in Fig 1.

A. Normal Operation

The CWSS model aims to maintain the required tank water level setpoint in the tank B102. To achieve this, the water stored in the tank B101 is pumped via a variable speed drive so that the required tank water level can be maintained while the demand from the tank varies throughout the valve V106. We propose a water demand model for the seven day week, which is based on the real model of power consumption in the UK [14]. We keep this water demand model simplistic, so it could be reproduced in the future. More details about this model can be found in our previous work [15]. The water level of the tank is measured as the process variable (PV) for a proportional-integral-derivative (PID) closed-loop control of the delivery pump to maintain the required tank water level setpoint (SP). A detailed explanation of this testbed can be found in our previous work [16].

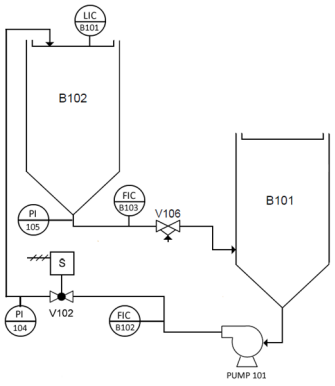

Fig. 1. Festo MPA Process Control Rig Diagram.

A. *Attack Scenarios*

Table I summarizes the set of attacks executed to the CWSS. Technical details regarding the attacks executed can be found in our previous work[15]. Malicious traffic includes crafted packets, executed by a malicious insider capable of accessing to the PLC over the control network, that overwrite the input/output memory of the PLC, as a result, the normal operation of the control process is compromised.

TABLE I SET OF ATTACKS EXECUTED TO THE CONTROL SYSTEM

Attack	Effect
Modifying Setpoint in the Working Memory	Water Level Increases/Decreases 2-2.5 litres. It depends on the value sent from attacker to the Input Memory of the PLC.
Attack on Ultrasonic Sensor	Water Level Increases/Decreases. It depends on the value sent from attacker to the Input Memory of the PLC.
Attack on Flow In	Affects Pump Operation, consequently the water level in the reservoir tank.
Attack on Pump	Water level decreases 0.5-1 Litres.
Attack on Flow Out	Affects the Control Operation when using feedforward Controller.
Attack on Pressure In	Slightly affects the normal operation of the control system. The water level increases/decreases 0.1 - 0.2 litres.
Attack on Pressure Out	Affects the control operation when using a PI controller that takes the Pressure Out as Input for calculating the water level, otherwise this does not affect the control operation.

B. *Energy-based IDS Architecture*

Historically, ICS devices such as PLC's and I/O were not networked and lacked computing and communication capabilities[1]. The emerge of Industry 4.0 [17] has led to developing ICS devices able to exchange data over the Internet. Further, the convergence of IT and ICS networks allow to manage, monitor and control industrial processes from remote locations. Fig. 2 shows a typical architecture of an IT and ICS network with security devices such as firewalls placed over the network [18]. When it comes to cybersecurity; defence in depth[19] is one of the well-known approaches, it comprises of a series of defensive mechanisms that are layered in the network in order to protect the assets. For instance, Fig 2 shows one firewall inspecting the incoming/outgoing traffic from the internet whereas the second firewall inspects the traffic between the corporate and control network.

The energy-based IDS (EBIDS) proposed in this paper adds an extra layer of protection to the control system, because it is placed at the lowest in the control process level and it is hard-wired to the PLC/Sensors. For instance, our previous work [15] demonstrates that the Input/Output memory of Siemens PLC's can be overwritten by crafting and sending packets to the PLC over the network [20], hence the values obtained from those spaces of memory are vulnerable to cybersecurity risks and exploits, which is unlikely to the values obtained by the EBIDS because its architecture makes it not accessible from the IT/ICS network.

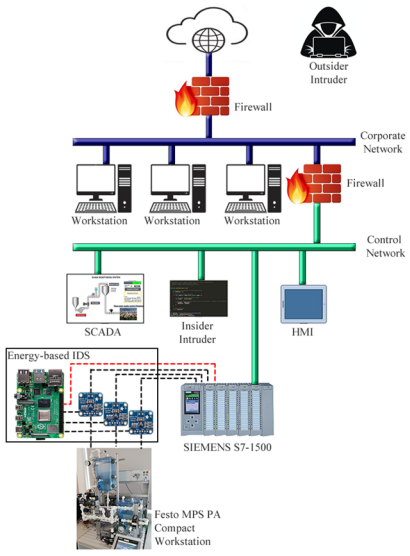

Fig. 2. EBIDS architecture.

C. Dataset

The dataset contains malicious and benign traffic that is recorded during a one-day operation. The EBIDS is tested using the dataset collected from the CWSS implemented for this research. Fig. 1 shows the sensors/actuators that are monitored: ultrasonic sensor B101, Pump 101, Flowmeter_in B102, Pressure_in 104, Pressure_out 105 and Flow_out B103. Each one of the sensors/actuators is hard-wired to the INA 219 sensor[21] and Input/output memory of the Siemens S7-1500 PLC[22]. The INA 219 sensor provides four energy features: voltage in the shunt resistor, the voltage in the INA 219 board, current and power. Thus, the dataset used in the pre-processing phase of the machine learning process contains 24 features in total. Fig 3 shows the original dataset obtained from the testbed and the balanced dataset after applying SMOTE oversampling technique[23]. SMOTE has been successfully applied and widely used in similar researches.

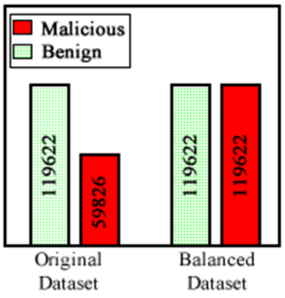

Fig. 3. Energy-based dataset.

D. Machine learning algorithms

The discussion of ML algorithms is beyond the scope of this paper, however, we provide literature with further information and technical discussions regarding those algorithms[25]. In this paper, we use traditional ML algorithms that were applied to similar researches discussed in section II. The following supervised ML algorithms are chosen for training and testing.

- K-Nearest Neighbour (KNN).
- Support Vector Machine (SVM).
- Decision tree (DT).
- Multilayer Perceptron (MLP).
- Naïve Bayer (NB).

E. Machine learning evaluation metrics

Choosing the right metrics for evaluating a machine learning algorithm influences how its performance is measured and compared with other approaches[26]. The metrics are usually derived from the confusion matrix, which is a summary of prediction results on a classification problem. Table II shows a confusion matrix, True Negative (TN) represents the number of benign samples correctly classified as benign, True Positive (TP) represents the number of malicious samples correctly classified as malicious, False Negative (FN) represents the number of benign samples incorrectly classified as malicious and finally, False Positive (FP) represents the number of malicious samples incorrectly classified as benign[27].

TABLE II CONFUSION MATRIX

Class	Classified as Benign	Classified as Malicious
Benign	True Negative (TN)	False Positive (FP)
Malicious	False Negative (FN)	True Positive (TP)

Our research focuses on critical infrastructure such as a clean water supply system, for that reason, we emphasise in maximizing the detection rate and minimizing as much as possible the number of false alarms reflected in the EBIDS. The metrics used to evaluate the results obtained from this research are explained as follows. Accuracy, shown in equation (1), is the ratio of correct predictions over the total number of predictions.

$$Accuracy = \frac{TP + TN}{TN + FN + FP + TP} \quad (1)$$

False Negative Rate (FNR) represented in equation (2) indicates the ratio of malicious traffic classified as benign.

$$FNR = \frac{FN}{TP + FN} \quad (2)$$

False Positive Rate (FPR), shown in equation (3) indicates the ratio of benign samples classified as malicious.

(3)

$$FPR = \frac{FP}{TN + FP}$$

F. *Energy-based IDS Operation*

The EBIDS has two components which are shown in Fig 4. An EBDIS detection classifier, which is built offline using a free software machine learning library for python[28]: scikit-learn and the real-time detection application. These two components are explained as follows:

1) Offline. The EBIDS detection classifier is trained offline with a dataset collected from the testbed. The dataset contains newly engineered energy-based traces of malicious and benign traffic obtained from the sensors/actuators that compose the testbed. The pre-processing step in machine learning improves the quality of the raw data collected from the testbed converting it into a clean set of usable information. The steps involved in data preprocessing includes *a*) removing the noise from the energy-based dataset by applying a low pass digital filter[29] on it, due to the data collected includes external factors such as noise. *b*) Using feature selection techniques such as Chi-Square and Information Gain to remove features that do not contribute to the energy-based ML model.

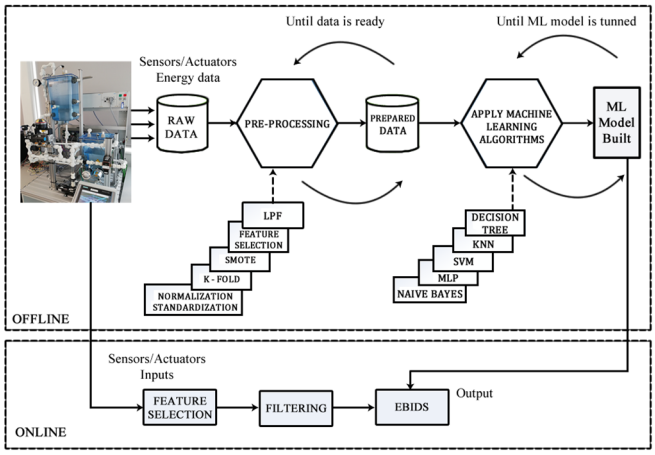

Fig. 4. Online Detection Energy-based IDS.

c) Using smooth techniques to adjust the class distribution of the dataset. *d*) Testing the effectiveness of the machine learning models by splitting the dataset into k consecutive folds for cross-validation. *e*) Scaling the dataset by applying Standardization/Normalization techniques. Following the dataset was split into training and testing datasets. The training dataset is composed of 80% of the entire data and it was used to train our ML model. The remaining 20% of the data is used to evaluate the performance of the trained ML model. A detailed explanation of the offline machine learning process can be found in our previous work[6]. Finally, we use the joblib library [30] available on Python to build the ML model and save it as a file for online evaluation.

2) Online. In the online phase of the process, the EBIDS uses the classifier built in the offline phase to detect the set attacks executed to the Input/Output memory of the Siemens PLC. We use the same joblib library described before to recover the machine learning model of each one of the ML algorithms. The ML model is deployed online in a Raspberry PI that collects, filters and selects the newly engineered energy-based features chosen during the feature selection process. The EBIDS raises an alarm to the operator when an anomaly is present in the control process. The EBIDS analyses 12 features, 20 times per second on average.

IV. EBIDS EVALUATION

This section summarises the results obtained from the evaluation of the proposed system during the experimentation phase. In the filtering process we applied a cutting edge and complex low pass filter at the pre-processing step, however, the same filter could not be applied during the online evaluation because the complex filter calculates its parameters based on the entire dataset provided. This scenario could not be replicated during the online testing, for that reason we opted for implementing our own low pass filter. Fig. 5 shows the results in terms of accuracy for online and offline evaluation. KNN achieved the highest accuracy during the offline evaluation followed closely by MLP. DT and SVM achieved above 98% of accuracy, whereas, NB shows the worst performance achieving 95.5% only. KNN and SVM showed a similar performance during the online and offline evaluation. The difference of accuracy among DT, NB and MLP during the online and offline training is more significative.

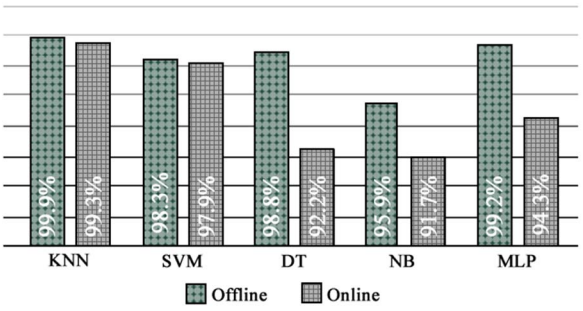

Fig. 5. Accuracy evaluation.

The metric accuracy shows the performance obtained by the classifiers but is not the only metric capable of evaluating the performance of the ML classifiers. Fig. 6 shows the false positive rate (FPR) achieved by the classifiers. This metric indicates the amount of benign traffic classified as malicious. KNN presents the best performance for both scenarios achieving 0.1% and 0.11% for offline and online evaluation. NB achieves 2.5% of FPR during the offline evaluation but increases to 6.8% in the online evaluation.

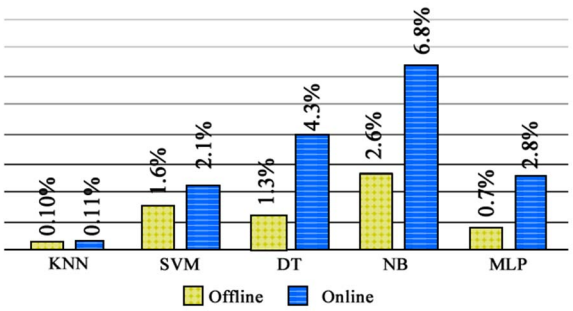

Fig. 6. False Positive Rate Evaluation.

The false-negative rate (FNR) represents the amount of malicious traffic classified as benign. In critical infrastructures, FNR alerts are more dangerous than FPR, because it indicates that the security system fails in detecting an attack that is occurring in the control application. Fig 7. Shows the results of the FPR metric. KNN shows the best performance for both scenarios achieving the lowest scores among the other classifiers. DT and MLP present considerable differences between offline and online evaluation. SVM shows a small difference between both evaluations but the score achieved is twice the score achieved by KNN.

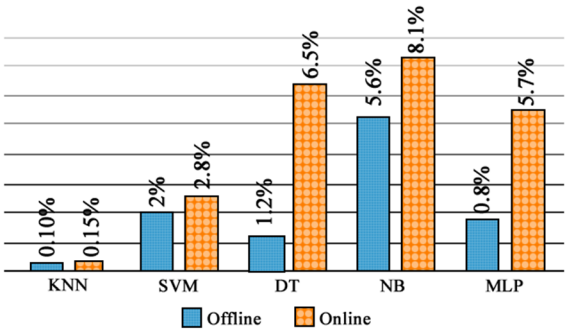

Fig. 7. False Negative Rate Evaluation.

V. Conclusions

This paper proposes a real-time anomaly intrusion detection for a clean water supply system, utilising machine learning with novel energy-based features. A model of a clean water supply system implemented in the Festo Rig was used to analyse the performance of the proposed detection system to cyber-attacks to the input memory of the PLC. The evaluation of the ML models showed a solid performance during the offline testing but only KNN and SVM showed the same consistency during the online evaluation. During the dataset collection, some parts of the normal operation were missed because attacks were executed at that time. This increased the number of false-positive alarms because the EBIDS was not able to recognize those missed parts.

The EBIDS proposed in this paper shows a different approach for cyber-attack detection than traditional network IDS. Its features are collected directly from the actuators/sensors that compose the control system instead of extracting the values from the ICS network traffic as most of the current approaches do. The main concern in using values collected from the ICS network traffic is trusting its integrity because it has been extensively probed that attackers can easily tamper network packets. It makes even worst for ICS network traffic because it lacks encryption.

Acknowledgement

This research is supported by the School of Computing and the School of Engineering and the Built Environment of Edinburgh Napier University.

REFERENCES

[1] K. Kamel and E. Kamel, "Introduction to PLC Control Systems and Automation," in *Programmable Logic Controllers: Industrial Control*, McGraw-Hill Education, 2014, pp. 1–31.

[2] M. Rüßmann *et al.*, "Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries," *Bus. Inf. Syst. Eng.*, vol. 6, no. 4, pp. 239–242, 2015, doi: 10.1007/s12599-014-0334-4.

[3] Kaspersky, "What happened to the Internet: attack on Cisco switches," 2018. [Online]. Available: https://www.kaspersky.com/blog/cisco-apocalypse/21966/. [Accessed: 29-Nov-2018].

[4] R. Langner, "Stuxnet: Dissecting a Cyberwarfare Weapon," *IEEE Secur. Priv.*, vol. 9, no. 3, pp. 49–51, May 2011, doi: 10.1109/MSP.2011.67.

[5] Cybersecurity Insiders, "Insider Threat 2018 Report," p. 41, 2018.

[6] A. Robles-Durazno, N. Moradpoor, J. Mcwhinnie, and G. Russell, "A supervised energy monitoring-based machine learning approach for anomaly detection in a clean water supply system," in *In Proceedings of the IEEE International Conference on Cyber Security and Protection of Digital Services (Cyber Security 2018)*, 2018, pp. 1–8, doi: 10.1109/CyberSecPODS.2018.8560683.

[7] M. Caselli, E. Zambon, and F. Kargl, "Sequence-aware intrusion detection in industrial control systems," *CPSS 2015 - Proc. 1st ACM Work. Cyber-Physical Syst. Secur. Part ASIACCS 2015*, pp. 13–24, 2015, doi: 10.1145/2732198.2732200.

[8] A. Terai, S. Abe, S. Kojima, Y. Takano, and I. Koshijima, "Cyber-attack detection for industrial control system monitoring with support vector machine based on communication profile," *Proc. - 2nd IEEE Eur. Symp. Secur. Priv. Work. EuroS PW 2017*, pp. 132–138, 2017, doi: 10.1109/EuroSPW.2017.62.

[9] P. Nader, P. Honeine, and P. Beauseroy, "Detection of cyberattacks in a water distribution system using machine learning techniques," *2016 6th Int. Conf. Digit. Inf. Process. Commun. ICDIPC 2016*, pp. 25–30, 2016, doi: 10.1109/ICDIPC.2016.7470786.

[10] D. Shalyga, P. Filonov, and A. Lavrentyev, "Anomaly Detection for Water Treatment System based on Neural Network with Automatic Architecture Optimization," 2018.

[11] F. Zhang, H. A. D. E. Kodituwakku, J. W. Hines, J. B. Coble, W. Hines, and J. B. Coble, "Multi-Layer Data-Driven Cyber-Attack Detection System for Industrial Control Systems Based on Network, System and Process Data," *IEEE Trans. Ind. Informatics*, vol. 3203, no. c, pp. 1–1, 2019, doi: 10.1109/tii.2019.2891261.

[12] L. A. Maglaras and J. Jiang, "Intrusion detection in SCADA systems using machine learning techniques," *Proc. 2014 Sci. Inf. Conf. SAI 2014*, pp. 626–631, 2014, doi: 10.1109/SAI.2014.6918252.

[13] FESTO, "MPS PA Compact Workstation with level, flow rate, pressure and temperature controlled systems." [Online]. Available: https://www.festo-didactic.co.uk/gb-en/learning- systems/process-automation/compact-workstation/mps-pa- compact-workstation-with-level,flow-rate,pressure-and- temperature-controlled-systems.htm?fbid=Z2IuZW4uNTUwLjE3LjE4Ljg4Mi40Mzc2. [Accessed: 07-Jul-2018].

[14] NORDPOOL, "Consumption." [Online]. Available: https://www.nordpoolgroup.com/Market-data1/Power-system-data/Consumption1/Consumption/ALL/Hourly1/?view=table. [Accessed: 30-Apr-2018].

[15] A. Robles-Durazno, N. Moradpoor, J. McWhinnie, G. Russell, and I. Maneru-Marin, "Implementation and detection of novel attacks to the PLC memory of a clean water supply system," in *Communications in Computer and Information Science*, 2019, vol. 895, pp. 91–103, doi: 10.1007/978-3-030-05532-5_7.

[16] A. Robles-Durazno, N. Moradpoor, J. McWhinnie, G. Russell, and I. Maneru-Marin, "PLC Memory Attack Detection and Response in a Clean Water Supply System," *Int. J. Crit. Infrastruct. Prot.*, May 2019, doi: 10.1016/j.ijcip.2019.05.003.

[17] J. Schlechtendahl, M. Keinert, F. Kretschmer, A. Lechler, and A. Verl, "Making existing production systems Industry 4.0-ready," *Prod. Eng. Res. Dev.*, vol. 9, no. 1, pp. 143–148, 2015, doi: https://doi.org/10.1007/s11740-014-0586-3.

[18] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, "Guide to Industrial Control Systems (ICS) Security," 2015.

[19] B. Pretorius and B. van Niekerk, "Cyber-Security for ICS/SCADA," *Int. J. Cyber Warf. Terror.*, vol. 6, no. 3, pp. 1–16, Jul. 2016, doi: 10.4018/IJCWT.2016070101.

[20] A. Robles, "Packet Crafting, Scapy and S7-1500 PLC." [Online]. Available: https://github.com/andrex17/ics. [Accessed: 06-Apr-2019].

[21] Adafruit, "INA219 HIGH SIDE DC CURRENT SENSOR BREAKOUT - 26V ±3.2A MAX." [Online]. Available: https://www.adafruit.com/product/904. [Accessed: 15-Jan-2019].

[22] Siemens, "Our fastest controller for automation." [Online]. Available: https://www.siemens.com/global/en/home/products/automation/systems/industrial/plc/simatic-s7-1500.html. [Accessed: 09-Nov-2015].

[23] K. W. P. Chawla, N.V., Bowyer, K.W., Hall, L.O., "SMOTE: Synthetic Minority Over-Sampling Technique. Journal of Artificial Intelligence Research," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[24] M. C. Belavagi and B. Muniyal, "Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection," in *Procedia Computer Science*, 2016, doi: 10.1016/j.procs.2016.06.016.

[25] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: a survey," *Digit. Commun. Networks*, vol. 4, no. 3, pp. 161–175, 2018, doi: 10.1016/j.dcan.2017.10.002.

[26] I. Technology and I. Technology, "A R EVIEW ON E VALUATION M ETRICS F OR D ATA C LASSIFICATION E VALUATIONS," vol. 5, no. 2, pp. 1–11, 2015.

[27] K. M. Ting, "Confusion Matrix," in *Encyclopedia of Machine Learning and Data Mining*, C. Sammut and G. I. Webb, Eds. Boston, MA: Springer US, 2017, p. 260.

[28] G. Hackeling, *Mastering Machine Learning With Scikit-learn*. Packt Publishing, 2014.

[29] M. Hansen, M. Haugland, T. Sinkjær, and N. Donaldson, "Real time foot drop correction using machine learning and natural sensors," *Neuromodulation*, vol. 5, no. 1, pp. 41–53, 2002, doi: 10.1046/j.1525-1403.2002._2008.x.

[30] A. Malakhov, "Composable Multi-Threading for Python Libraries," *Proc. 15th Python Sci. Conf.*, no. Scipy, pp. 15–19, 2016, doi: 10.25080/majora-629e541a-002.