energies

MDPI

Article

Stealthy Cyberattacks Detection Based on Control Performance Assessment Methods for the Air Conditioning Industrial Installation

Jakub Filip Możaryn [1,*], Michał Frątczak [2], Krzysztof Stebel [2], Tomasz Kłopot [2], Witold Nocoń [2], Andrzej Ordys [1] and Stepan Ozana [3]

[1] Faculty of Mechatronics, Institute of Automatic Control and Robotics, Warsaw University of Technology, 00-661 Warsaw, Poland
[2] Department of Automatic Control and Robotics, Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, 44-100 Gliwice, Poland
[3] Department of Cybernetics and Biomedical Engineering, Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, 708 00 Ostrava, Czech Republic
* Correspondence: jakub.mozaryn@pw.edu.pl

Abstract: This paper aims to study the workflow of the detection centre of stealthy attacks on industrial installations that generate an increase in energy consumption. Such long-lasting, undetected attacks on industrial facilities make production more expensive and less competitive or damage the installation in the long term. We present the concept of the remote detection system of cyberattacks directed at maliciously changing the controlled variable in an industrial process air conditioning system. The monitored signals are gathered at the PLC-controlled installation and sent to the remote detection system, where the discrepancies of signals are analysed based on the Control Performance Assessment indices. The results of performed tests prove the legitimacy of the adopted approach.

Keywords: cyberattack; control variable; feedback system; cyberattack detection; process air conditioning station

Citation: Możaryn, J.F.; Frątczak, M.; Stebel, K.; Kłopot, T.; Nocoń, W.; Ordys, A.; Ozana, S. Stealthy Cyberattacks Detection Based on Control Performance Assessment Methods for the Air Conditioning Industrial Installation. *Energies* **2023**, *16*, 1290. https://doi.org/10.3390/en16031290

Academic Editor: Yun Liu

Received: 9 December 2022
Revised: 16 January 2023
Accepted: 19 January 2023
Published: 25 January 2023

1. Introduction

On 23 December 2015, the power grid of Ukraine was hacked, resulting in power outages for roughly 230,000 consumers in Ukraine for 1–6 h. Around 0.015% of the daily electricity consumption in Ukraine was not supplied (up to 73 MWh of electricity) [1]. The attack was distributed in an email via an infected Word document or PowerPoint attachment. Then the BlackEnergy 3 malware remotely compromised the information systems of three energy distribution companies in Ukraine and temporarily disrupted consumer electricity supply [2].

TXOne Networks, the OT zero trust and Industrial IoT (I-IoT) security company has published a 2021 cybersecurity report [3] which focuses on the vulnerabilities that can affect Industrial Control Systems (ICS). According to the report, the number of advisories dramatically increased in 2021, when there were 389 advisories published, compared with 249 a year earlier. The growing number of cyberattacks aiming at disrupting critical infrastructure (CI) clearly shows that hackers seek new attack vectors for their potentially dangerous activities.

The CI is the set of systems and related objects consisting of buildings, devices, installations, and services essential to the security of the state and its citizens that ensure the efficient functioning of public administration, institutions, and entrepreneurs [4]. CI consists of the following systems: (a) supply of energy and fuels, (b) communications, (c) ICT (Information and Communication Technologies) networks, (d) financial, (e) food

supply, (e) water supply, (f) health protection, (g) transport, (h) rescue, (i) public administration, (j) production, and (k) storage, warehousing and usage of chemical and radioactive substances, including pipelines of hazardous substances. CI plays a key role in the state's functioning and in citizens' lives. Because of events caused by forces of nature or human activities, CI may be destroyed or damaged and its operation may be disrupted, which may endanger the life and property of citizens. Such events negatively affect the economic development of the country. Therefore, protecting CI is the priority of every state. The primary objectives of CI tasks include safeguarding against threats and minimizing disruptions and damage, facilitating swift remediation, and minimizing economic and societal impacts. The protection of CI consists of all activities aimed at ensuring the functionality, continuity of operations and integrity of CI to prevent threats, risks or vulnerabilities, limit and neutralize their effects, and restore this infrastructure quickly in the event of failures, attacks and other events interfering with its proper functioning. In many states, cooperation with private enterprises is important because, in many cases, a substantial part of the CI of key importance for state security is privately owned.

In modern industrial companies, there exist overlapping technologies, i.e., Information Technologies (IT) regarding information, its flow, and administration, and Operating Technologies (OT) regarding the operation of physical processes and the machines (e.g., controllers, actuators, sensors) used to implement them. Such synergy is called IT/OT convergence and the two-way flow of information between these technologies brings the production process closer to the business world. For example, a visible trend has been observed in the monitoring and control of industrial plants based on the Industrial Internet of Things (I-IoT) devices and Computing Cloud (e.g., Control as a Service—CaaS) [5]. Despite significant improvements in cost, flexibility, and maintenance, it also introduces new problems that need to be addressed on the OT level, such as cybersecurity. Conventional ICSs are traditionally equipped with signal-induced fault detectors searching for anomalies in control and sensor signals concerning the behaviour of the ICS. They consist of estimating the state of the system and comparing the estimated states with the states measured by the sensors (i.e., residuals). Many works exist on defining faulty states based on the computed residuals (e.g., Chi-Square or CUSUM).

Until recently, the issues of detecting anomalies were carried out independently as Intrusion Detection Systems (IDS) in case of cyberattacks (security) or Advanced Diagnostic Systems (ADS) in case of technical faults (safety). However, cyberattacks in the ICS can currently be seen as an anomaly generator [6]. Considering the industrial process specificity, process model, and controller performance, the ADS should be equipped with the methods to detect and distinguish cyberattacks and process faults in OT infrastructure, thus working in parallel and exchanging information with IDS [7].

The three main cyberattack types on ICS can be distinguished:

- **Integrity attacks** that aim to degrade the control performance of the ICS (e.g., False Data Injection Attacks (FDIA), Man-In-The-Middle (MITM) attacks).
- **Availability attacks** that aim to disrupt the operations of some control equipment (e.g., DoS attacks),
- **Confidentiality attacks** that aim to collect information from the ICS (e.g., eavesdropping attacks).

Such attacks can be **stealthy attacks (covert attacks)** that generate anomalies while keeping fault detectors below their detection threshold and damaging or intruding into the system in the long term (e.g., Stuxnet), or **non-stealthy attacks** that are often quick-in-time attacks with huge impact.

Covert attacks involve access to sensor measurements, system controllers, and sufficient knowledge of system operations [8–11]. Some attacks aim at understanding the control architecture (e.g., control law implemented in controllers, the response of supervisions, fault detection threshold) or knowing the field equipment (e.g., sensors, actuators) to launch further integrity or availability attacks [12].

There are three main areas of possible cyberattacks on the ICS with a set of attack vectors each [13,14]:

- **Cyberattacks on software**, e.g., Buffer Overflow, SQL injection, Cross-Site Scripting (XSS).
- **Cyberattacks on hardware**, i.e., accessing the physical location of the ICS in an unauthorised way to damage and modify the operational procedure of the system, e.g., make changes on certain threshold values.
- **Cyberattacks on communication**, i.e., exploiting the communication channel and protocol vulnerabilities, exploiting unnecessary ports and services.

In small and medium enterprises, SCADA (Supervisory Control And Data Acquisition) systems are vital to the ICS. The common practices of the SCADA system designers and operators with low-security levels cause them to be extremely vulnerable to various OT cyberattacks [15,16]. The broadly discussed and analysed virus Stuxnet is a typical example of a long-term covert attack damaging the system. It was revealed after it had caused over 1000 failures of the uranium enrichment centrifuges [17]. Another example is Triton malware targeting the SCADA/ICS system of the Saudi Arabian petrol company Petro Rabigh which went unnoticed for three years before being detected [18,19]. Covert cyberattacks present a significant challenge in terms of detection. Two commonly employed methods for detecting such attacks include analysing sensor measurements for deviations from expected correlations [11] and examining system dynamics for deviations from expected behaviour due to an attacker's imperfect knowledge of the system that can be detected by monitoring the residuals [20–22]. Recent studies show the significant role of artificial intelligence and machine learning methods in cyberattack detection in industrial installations. In such approaches, anomaly detection using deep learning models is used to identify and detect attacks on SCADA systems by learning the characteristics of malicious activity and differentiating them from normal features. The advantages of such systems over other methods indicated in the literature are minimal feature engineering or assumptions about the data distribution. A deep learning system for Automated Guided Vehicles (AGV) based on the Internet of Things (IoT) has been proposed in [23] and evaluated for its ability to identify and simulate the normal state of the AGV while detecting network instability caused by cyberattacks. In [24], a long short-term memory (LSTM)-based intrusion detection system (IDS) has been implemented and evaluated to detect cyber-physical attacks on a water treatment testbed. The research presented in [25] describes the method based on 1D convolutional neural networks and autoencoders with Primary Component Analysis (PCA) to improve cyberattack detection rates in industrial installations. In [26], the authors propose a stacked deep learning method to detect malicious attacks in SCADA systems and provide a comprehensive evaluation on several industrial benchmark datasets. Several software and hardware solutions have been proposed and implemented in the industry to increase the resiliency of the SCADA system to cyberattacks. One of them is zero-trust network architecture [27] with software-based approaches, such as the Elliptic Curve Digital Signature Algorithm (ECDSA) [28]. Hardware-based solutions have also been proposed, including using a Trusted Platform Module (TPM) to create a trusted chain for IoT devices and enhance the security of SCADA and automation systems [29,30].

In this paper, we propose to use the Control Performance Assessment (CPA), used to measure the quality of a control system, for cyberattack detection [31,32]. The CPA is based on the study of the chosen indexes [33] calculated for the signals gathered from the plant devices. Such indexes can be grouped into the following classes: (a) Step Response Indexes, (b) Data-Based Integral Measures, (c) Statistical Measures, (d) Model-Based Measures, (e) Frequency-Based Measures. The assessment requires methodologies and indexes (Key Performance Indicators) that enable measuring the system's quality and undertaking necessary improvement steps. CPA methods also allow benchmarking of different systems to prioritise maintenance actions. Furthermore, some of the measures may show a reason for the inappropriate operation, which is useful in detecting the deterioration of the system work. In this article, we discuss the use of data-based CPA statistical

measures, allowing the detection of possible anomalies in the system, and searching for the deterioration and possible statistically important changes within measured signals. The proposed method for detecting stealthy attacks was evaluated using a simulated attack on a controlled variable (CV) in the air conditioning system. The process model and detection system were implemented on two industrial workstations and PLCs connected remotely with secure tunnelling communication. The proposed method was tested under various operational scenarios.

2. Motivation

In [34], the authors presented an experimental evaluation of sensor attacks and defence mechanisms in feedback systems. Such attacks assume that the attacker can stealthily manipulate sensor readings in the control system, thus making the control system oblivious to the fact that the desired set points of process variables are not achieved. On the one hand, this will immediately affect the product quality, resulting in high costs of wasted raw materials and energy. In some cases, quality control in the plant should be able to detect the problem with deteriorating quality relatively quickly, and a proper investigation should lead to uncovering the stealthy sensor reading problem.

In this article, we evaluate the problem of stealthy manipulation of a selected control variable, especially in a feedback system requiring two independent control variables having opposite effects on the process variable. For example, when temperature control requires heating and cooling, the attacker may try to change the operating regime of the cooling process, thus forcing the heating part of the system to increase energy usage to compensate for the temperature drop caused by the attack-related cooling. In such a case, the feedback control system will correctly maintain the controlled temperature according to the desired setpoint, thus preventing product quality deterioration. This will obviously prevent costs associated with raw materials wasting but will increase the cost of consumed energy and will only be possible to detect using continuous or periodical inspection of the control system.

The authors conducted a quality audit of the operation of the control systems at the real plant where the installation similar to that described above is located. The audit showed the behaviour of the system described in the article. The situation was not the result of a cyberattack, but an improperly designed control system. Operation of the system in such a state generated significant losses for the plant. These losses went unnoticed for months of operation of the system. This inspired us to describe a possible scenario in which such a situation would be caused by deliberate action.

Therefore, in this paper, we demonstrate a cyberattack directed at maliciously manipulating a controlled variable (CV) in a feedback system and propose methods to detect such attacks. The process under consideration requires both cooling and heating to keep the desired temperature of the processed air. It is assumed that the heater's energy consumption (for example, the electric current) is monitored. It is very often fulfilled in practice, e.g., for diagnostics purposes. However, because the cooler in the system is assumed to operate independently and, in many cases, requires energy consumption for the preparation of the cooling agent in advance, a straightforward identification of concurrent cooler and heater operation is not sufficient for detecting malicious manipulation of the cooler.

3. Models and Methods

3.1. Feedback System under Attack

The feedback system under consideration is an air conditioning unit, in which fresh air of inlet temperature Tin = 20 °C passes through both a cooling heat exchanger (HEx-C) and a heating heat exchanger (HEx-H) (Figure 1). Such approaches are used, for example, in air conditioning systems for paint shops. A process variable (PV) in this feedback system is the measured temperature Tout of the conditioned air (TT—temperature transducer). A split range control algorithm (TC) uses two different control variables: a cooling unit controlled by the cooling control variable (CCV) when PV exceeds the set point (SP), or the

heating unit controlled by the heating control variable (HCV) when SP exceeds PV. The heating unit is supplied with hot water at 90 °C, the temperature is controlled by changing its flow of 0–20 L/min, and heating power consumption is measured (A). The cooling unit is supplied by a glycol at 1°C and temperature is controlled by manipulating its flow of 0–20 L/min. It is assumed that the feedback system is properly tuned and inadvertent fast switching between the cooler and the heater is avoided.

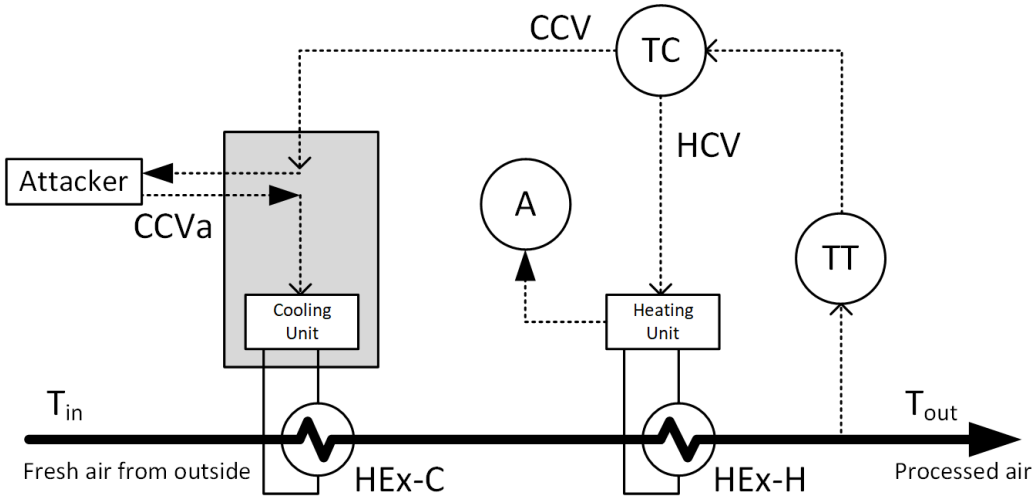

Figure 1. Feedback control of the air conditioning unit.

The assumed mode of cyberattack is through the cooling unit. If the attacker gets access to the internal data processing of the cooling system, the CCV value can be read and changed by the attacker to a new cooling control variable (CCVa). Moreover, it is assumed that our control system cannot monitor the inner variables of the cooling unit and the malicious manipulation of the cooling unit will not be directly detected. This assumption seems justified since many cooling units are sold as single and closed systems, with only a limited number of process variables exposed to the plantwide control system. Therefore, the attacker can force the cooling unit to operate and decrease the air temperature, even if cooling is not required. The feedback control system will react accordingly by increasing the power consumed by the heating unit, and the temperature of the conditioned air will be maintained. However, the operating costs of the air conditioning unit will be significantly increased. Because the inner parameters of the cooling unit are not monitored, such a situation may last for prolonged periods. The heating unit's power consumption is measured using electric current measurement. For instance, thyristor power controllers often enable easy reading of output power.

Therefore, the following assumptions are made in the presented demonstration of cyberattack detection. The measured variables are the temperature of the fresh air Tin, the temperature of the processed air Tout with its set point SP, and power consumption based on the electrical current measurement A. The unknown or unmeasurable parameters are the power consumption of the cooling unit and the cooling control variable CCVa, manipulated by cyberattack. Additionally, it is impossible to prevent the cooler from working simultaneously as the heating unit and vice versa since the closed cooling unit needs to prepare ice water in advance.

In this article, we assume only a limited scope of cyberattack. First, it is assumed that the setpoint temperature SP is greater or equal to Tout. Hence, only the heater unit is being used by the split range controller. Secondly, it is assumed that the attacker maliciously manipulates the cooling controlled variable by increasing it and cooling the fresh air, thus forcing the controller to increase power consumption.

3.2. Proposed Attack Detection Approach

In our research, we use the standard control performance assessment method based on the Minimum-Variance (MV) benchmark to reveal the possible cyber threats. The proposed MV benchmark (as a reference performance bound) can be estimated from data monitored online (e.g., process value, control value). The only assumption is that the system delay estimate is known.

In CPA, the reference best feedback control used to benchmark is the Minimum-Variance Control (MVC, i.e., optimal H2 control) [35]. MVC produces the smallest possible closed-loop output variance and it is worse for any other linear controllers. The MVC-based assessment compares the actual system-output variance σ_y^2 to the output variance σ_{MV}^2 as obtained using an MVC applied to an estimated time-series model from measured output data. The so-called Harris index (HI) is defined as [36].

$$\eta_{\mathrm{MV}} = \frac{\sigma_{\mathrm{MV}}^2}{\sigma_y^2} \tag{1}$$

The Harris index is calculated from the measured data and is given in the interval, where a value close to 1 indicates the best possible control concerning the theoretically achieved output variance, while 0 means the worst performance, including unstable control. The Harris index is typically calculated for the process value; however, it can be used as the measure to assess any signal variance and in our case, can be adapted to the course of the control signal, allowing for the detection of potential anomalies (changes in variance) caused, among others, by cyberattacks. There are two advantages to using η_{MV} over σ_y^2: (a) it is independent of the underlying disturbances, and (b) η_{MV} is bounded between 0 and 1; thus we can set the threshold value that will indicate the deterioration of the signal, which can be due to the possible cyberattack.

We calculate the Harris index as follows [37].

$$\widehat{\eta_{\mathrm{MV}}} = \frac{(n-b-m+1)\widehat{\sigma_{\mathrm{MV}}^2}}{\widetilde{u}^T\widetilde{u} + \overline{u}^2} \tag{2}$$

where n is the sample length, b is the estimated delay, m is a model rank.

The estimate of the residual mean square error is given by

$$\widehat{\sigma_{\mathrm{MV}}^2} = \frac{\left(\widetilde{u} - \widetilde{X}\hat{\alpha}\right)^T\left(\widetilde{u} - \widetilde{X}\hat{\alpha}\right)}{(n-b-2m+1)} \tag{3}$$

To calculate the estimate $\hat{\sigma_{\mathrm{MV}}^2}$, we solve the set of linear equations

$$\left(\widetilde{X}^T\widetilde{X}\right)\hat{\alpha} = \widetilde{X}^T\widetilde{u} \tag{4}$$

where

$$\widetilde{u} = \begin{bmatrix} \widetilde{u}_n \\ \widetilde{u}_{n-1} \\ \vdots \\ \widetilde{u}_{b+m} \end{bmatrix}, \widetilde{X} = \begin{bmatrix} \widetilde{u}_{n-b} & \widetilde{u}_{n-b-1} & \cdots & \widetilde{u}_{n-b-m+1} \\ \widetilde{u}_{n-b-1} & \widetilde{u}_{n-b-2} & \cdots & \widetilde{u}_{n-b-m} \\ \vdots & \vdots & \vdots & \vdots \\ \widetilde{u}_m & \widetilde{u}_{m-1} & \cdots & \widetilde{u}_1 \end{bmatrix}, \underline{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} \tag{5}$$

and

$$\widetilde{u}_n = u_n - \overline{u} \tag{6}$$

is the corrected deviation of the control value u_n from its mean value $\overline{u}$.

3.3. System Architecture

In Figure 2, the experimental set-up used in the research is presented. The proposed architecture generally assumes that identifying cyberattacks is outsourced and performed by a remotely connected data centre, as outsourcing practice is becoming common nowadays. The presented cyberattack detection methods could also be realised using locally implemented systems, for example, edge computing [38]. Such an approach, however, needs more scalability and closer integration of the attack detection system with the hardware infrastructure of the control system. Therefore, it was decided to prepare a distributed system which fulfils the industrial requirements, considering the security of the data.

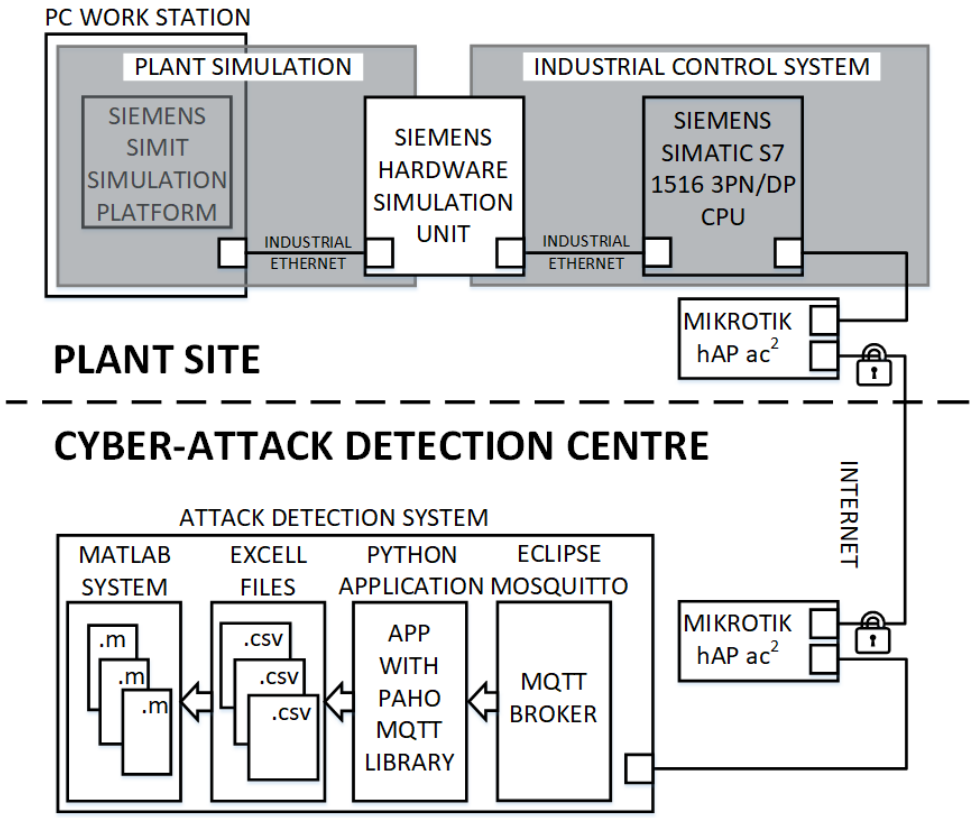

Figure 2. Distributed laboratory setup based on outsourcing idea.

The system consists of a plant site and a cyberattack detection centre. The two parts communicate using a secure, tunnelled connection based on the Mikrotik hAP ac2 device. Physically, the plant site was located at the Silesian University of Technology in Gliwice, Poland, and the cyberattack detection centre was at the Warsaw University of Technology in Warsaw, Poland.

The plant site consists of a PC workstation on which the air conditioning system (Figure 1) is simulated. The simulation is implemented in the Siemens Simit Simulation Platform (Figure 3), which is commonly used in industrial practice for the virtual commissioning of control systems [39,40]. This module simulates a ProfiNet process interface based on industrial Ethernet and connects the control system and process simulation. The industrial control system was implemented using Siemens Simatic S7-1516-3 PN/DP PLC. This PLC implements the control algorithm and provides the capability of using the MQTT protocol, which enables safe communication with the distant cyberattack detection centre. MQTT was chosen as it provides security and ease of implementation. In other applications, OPC UA may also be considered as it provides similar functionality.

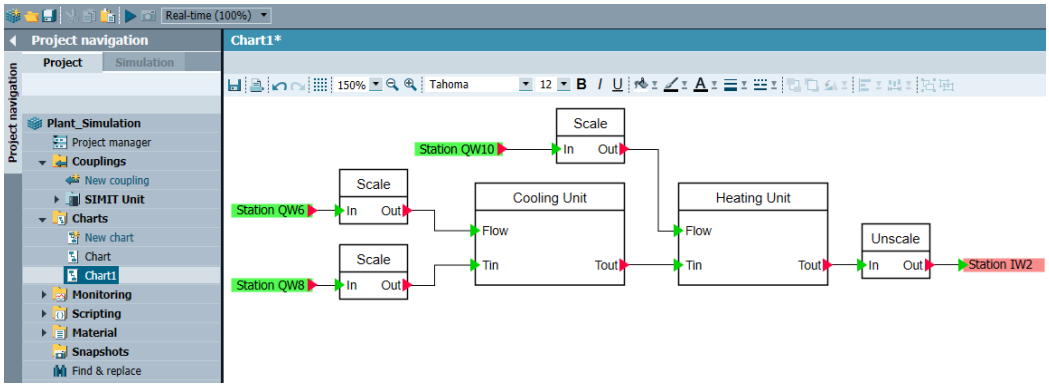

Figure 3. Simulation of the air conditioning unit in Siemens SIMIT.

The distant cyberattack detection centre is based on a set of applications, including a data acquisition module that retrieves all necessary data from the plant side PLC using the MQTT protocol. Eclipse Mosquitto is used as the MQTT broker and mediates communication by both clients. From the client's point of view, communication is done only with the broker and direct communication between clients is not possible. This principle facilitates the scalability of the MQTT network and enables easy expansion of the data set exchanged between clients. Additionally, all data is encrypted using TLS and user authentication based on login, and a password is provided. Data acquisition and storage are implemented in Python, acts as a MQTT client and uses the paho.mqtt.python library. Data is stored using csv files that are, in turn, imported into MATLAB for cyberattack detection analysis.

4. Experimental Results

The proposed cyberattack detection method has been verified for periodic signals maliciously added to the control variable of the cooling unit. Two different attacks were analysed: a triangular and a sinusoidal signal (Figures 4 and 5) added to the cooling control signal. The amplitude and frequency of the attack signals have been chosen so that the influence of the attack signal on process temperature is well within the noise range of the signal and is not clearly visible, but to maximise the consumed energy by both cooling and heating units. Therefore, although process operators pay close attention to process variables (temperature in this case), such an attack would not have been easily detected. Potentially, this attack may be seen by observing the control signal of the heating unit; therefore, the Harris index is computed, which detects changes in the analysed signal variance. The Harris index was calculated for N = 1000 samples of the measured heating control signal, for a model of rank m = 30 and for a time delay tau = 1 sample with a moving window of n = 200 samples. This initial set of parameters was suggested to detect significant variations in control signal variance by comparing the calculated Harris index with variance calculated directly from the control signal (initial values were chosen based on [32]).

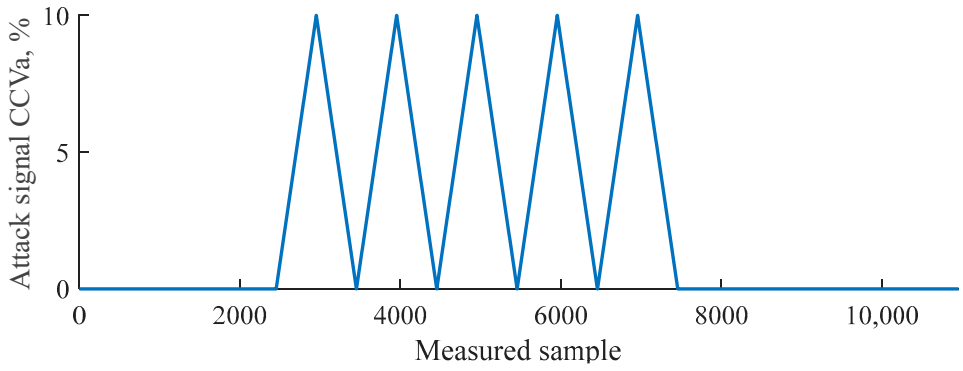

Figure 4. Triangular attack signal.

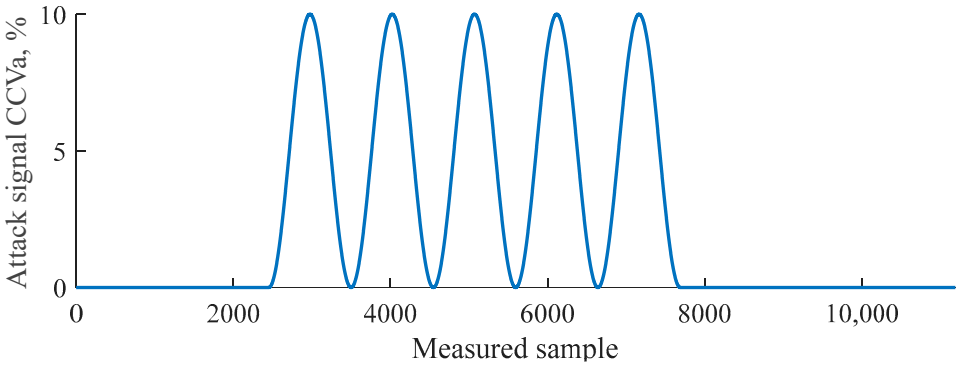

Figure 5. Sinusoidal attack signal.

Figure 6 presents results for a triangle attack signal added as the CCVa signal, particularly the effect on the measured heating unit current HU [%]. Figure 7 presents results of adding a sinusoidal attack signal as the CCVa signal. As can be seen, HU [%] is a good basis for detecting the attack. Since the variance of the HU signal increases, the Harris index decreases and can be thresholded to generate the attack detection signal. The threshold was selected as 0.2 based on historical data in this case. A slight delay in the detection signal concerning the actual attack is visible.

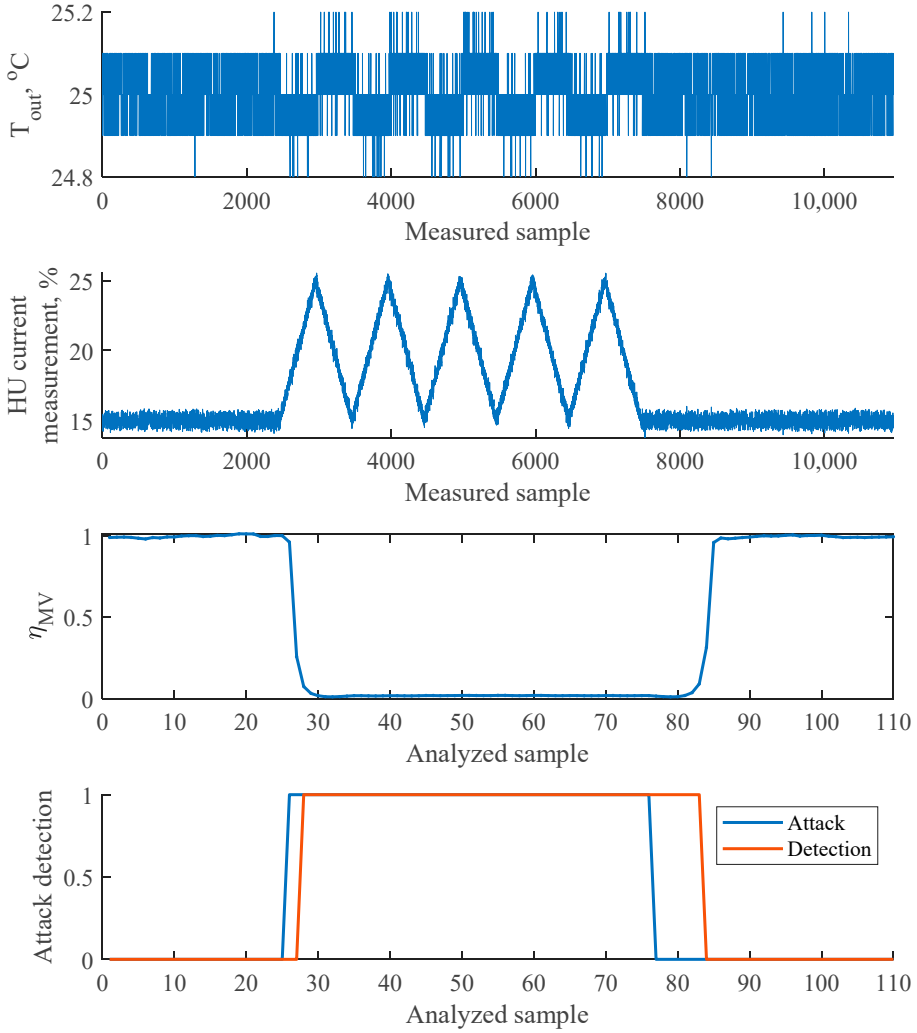

Figure 6. Effect of the triangular attack signal on process control.

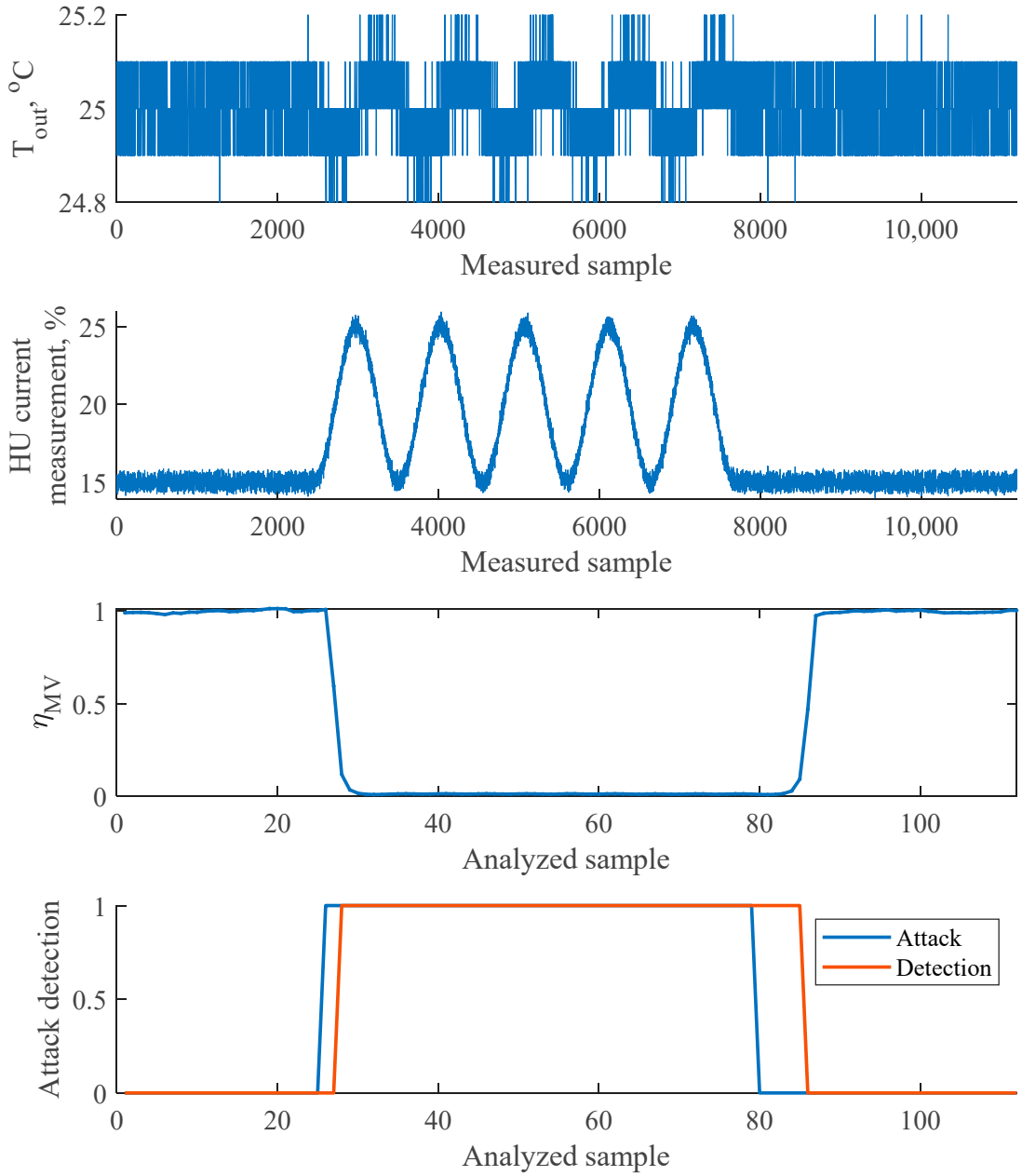

Figure 7. Effect of the sinusoidal attack signal on process control.

The efficiency of the presented approach was verified for triangular and sinusoidal attack signals with decreased amplitudes, thus with reduced impact of the cyberattack on cooling unit. A "less effective", from the attacker's perspective, cyberattack results in a decreased variation of the analysed control signal and increased calculated values of the Harris index, which sometimes have not exceeded the pre-set threshold. However, if the impact of a cyberattack on the variance of the control signal is insignificant, then the extra energy consumption caused by a cyberattack is negligible.

Results presented in Figures 6 and 7 have been generated assuming that no natural disturbances caused by the process itself are present (for example, varying demand for processed air) or from varying parameters of fresh air from the outside (for example, varying temperature and/or humidity). Figure 8 presents a natural disturbance added into the process, representing changes in air demand for the air conditioning system. Figures 9

and 10 present results for an additional sinusoidal process disturbance having a lower frequency concerning the attack signal itself.

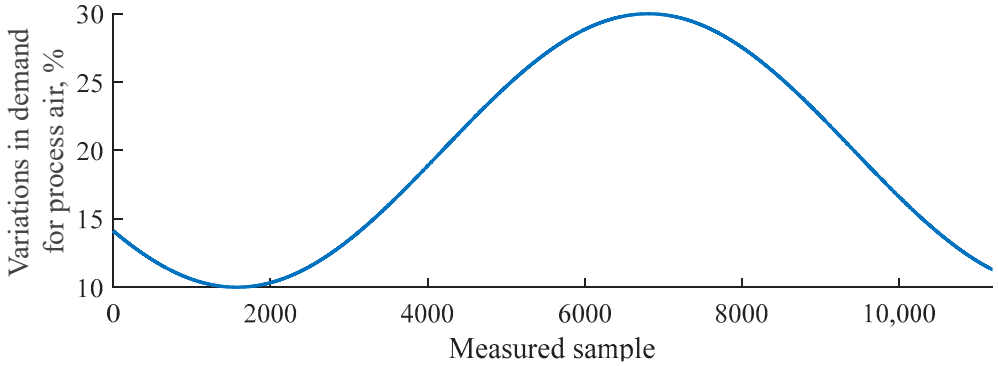

Figure 8. Changes in process air demand that represent natural disturbances in the process.

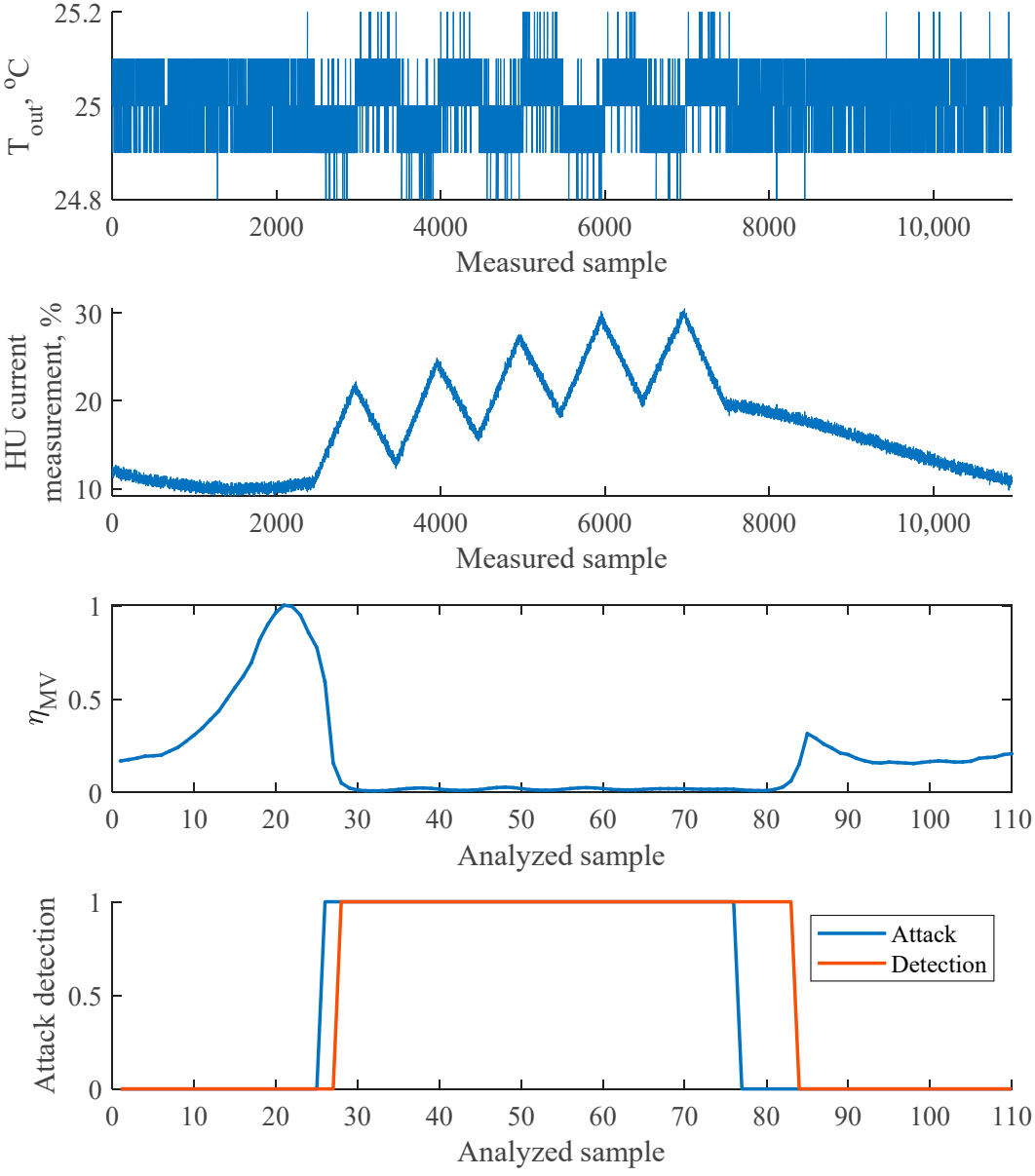

Figure 9. Triangular attack signal on top of a low-frequency natural sinusoidal disturbance.

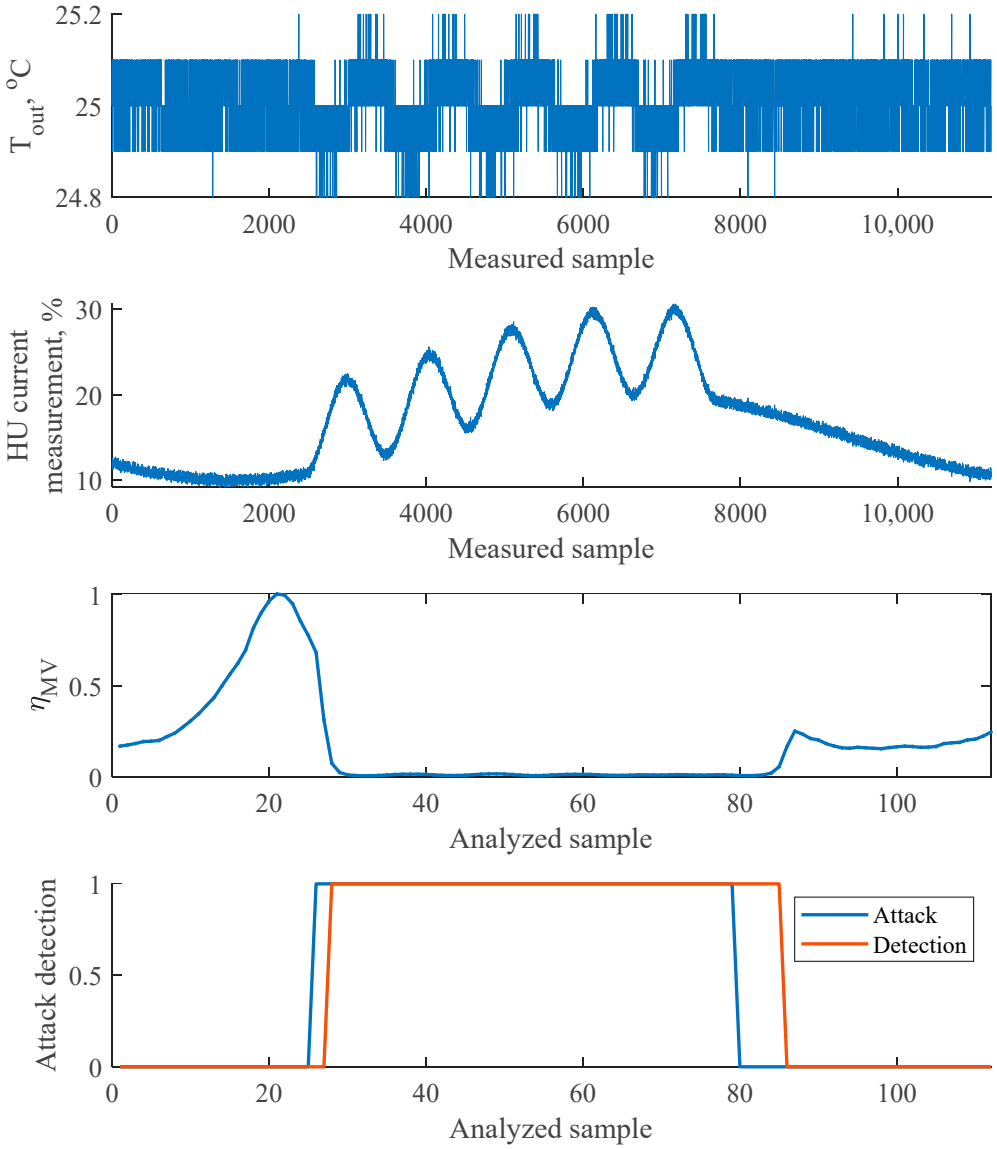

Figure 10. Sinusoidal attack signal on top of a low-frequency natural sinusoidal disturbance.

In this case, the variance of the HU signal is considerably larger, even when no attack is currently active, leading to increased changes in the Harris index. Based on historical data, a different threshold value has been selected as 0.1 and the attack is assumed active if the Harris index is lower than 0.1.

The ability to differentiate between cyber-faults and process faults is a crucial issue that warrants further examination. The CPA method presented in this article can be compared with a methodology and example described in [6], where cyberattack detection was conducted utilizing algorithms designed to detect and isolate process faults based on the analysis of residuals (Fault Isolation System—FIS). In instances of cyberattacks, the identification of diagnostic signal patterns for individual scenarios can prove to be a challenging task. Therefore, various extensions have been proposed to address this: multiple symptom assessment, symptom onset timing, and symptom sequence [6]. In such cases, the main advantage of the CPA method is its simplicity, as it requires signal analysis only without prior identification of the cause–effect relationships, complicated modelling, and residual analysis, as in the case of FIS extensions.

5. Conclusions

The method of stealthy attack detection on the industrial installation based on the data-driven statistical control performance measure (Harris index) was presented. As an example, we used the simulation of the air conditioning installation where the problem of stealthy manipulation of a selected control variable was evaluated, especially in a feedback system requiring two independent control variables having opposite effects on the process variable. The proposed monitoring and cyberattack detection system has been implemented on the two industrial-type workstations and PLC controllers (one for the process workstation and the second for the anomaly detection centre), connected remotely using secure tunnelling communication.

Precise detection of a cyberattack requires additional analysis of the situation, for example, by observation of network traffic [41]. There are a several ways to distinguish between process faults and cyberattacks in industrial installations:

- **Monitoring for unusual patterns or anomalies in system behaviour:** Process faults will often manifest as abnormal behaviour or unexpected output from the system, whereas cyberattacks may involve unusual network traffic, system resource usage, or other anomalies.
- **Looking for signs of tampering or unauthorized access:** Process faults typically do not result from intentional tampering, whereas cyberattacks may involve unauthorized access to the system or manipulation of its controls.
- **Reviewing system logs and event histories:** These can provide important clues about the root cause of an issue, such as when it occurred, what triggered it, and what actions were taken in response.
- **Robust security measures implementation:** Ensuring that the industrial installation has robust security measures, e.g., firewalls, intrusion detection and prevention systems, and secure authentication and access controls, can help to prevent or mitigate the effects of cyberattacks.
- **Expert consultations:** If one is unsure whether an issue results from a process fault or a cyberattack, it may be helpful to consult with experts in the field who can guide and assist in identifying and addressing the issue.

Regarding the Harris index as a potential measure for cyberattack detection, it should be emphasised that it requires proper tuning of the parameters, i.e., sample length, estimated delay, and model rank. Moreover, further experimental research should be performed to choose the detector thresholds for different types of attacks. The presented results suggest that the Harris index may potentially be used to detect periodic attack signals being added into one of the control variables. In reality, process operators rarely regularly focus on the control signal; therefore, such a tool would support the operator and technology crews in detecting process cyberattacks. Control signal variance may change due to other reasons, for example, because of control units wearing out. The increased variance, however, unequivocally points to the problem with the control performance.

The proposed solution can be easily generalised for different signals gathered from the plant and can be used in other industrial domains. However, it is important to carefully evaluate the performance of the proposed anomaly detection method to determine its suitability for a particular application. The threshold value and the range of noise values in the system significantly impact the effectiveness of the diagnostic system. A highly sensitive system can quickly detect an attack but may also be prone to false positives due to larger signal values. False positives can be a concern in any anomaly detection system, and it is important to consider the potential sources of variability in the system and how they might affect the performance of the method. This should involve collecting data from the system under various operating conditions to evaluate the method's performance for different scenarios. It may also be necessary to fine-tune the method's parameters (e.g., the threshold for detecting an anomaly) to achieve the desired level of performance. The optimisation of the diagnostic block is a topic worthy of further consideration.

Author Contributions: Conceptualization, J.F.M., T.K. and A.O.; Methodology, J.F.M., M.F. and K.S.; Software, J.F.M., M.F. and K.S.; Validation, J.F.M., K.S., T.K. and S.O.; Formal analysis, W.N.; Investigation, J.F.M., M.F. and W.N.; Resources, T.K. and M.F.; Data curation, K.S.; Writing—original draft preparation, J.F.M., W.N. and T.K.; Writing—review and editing, W.N., T.K. and S.O.; Visualization, W.N. and K.S.; Supervision, A.O.; Project administration, A.O.; Funding acquisition, A.O. and T.K. All authors have read and agreed to the published version of the manuscript.

Funding: Andrew Ordys and Jakub Możaryn acknowledge support from the National Agency of Academic Exchange (NAWA), "Polish Returns," grant no: PPN/PPO/2018/1/00063/U/00001; and from the POB Research Centre Cybersecurity and Data Science of Warsaw University of Technology within the Excellence Initiative Program-Research University (ID-UB, grant no: 1820/38/Z01/POB3/2020). This work was partly financed by a grant from the Silesian University of Technology—a subsidy for maintaining and developing the research potential in 2022 (grant no: 02/060/BK_23/0043). The APC was co-funded by the Warsaw University of Technology and the Silesian University of Technology.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gilbert, D. Black Energy Cyber Attacks Against Ukrainian Government Linked to Russia, International Business Times: Zugegriffen 2014. Available online: http://www.ibtimes.co.uk/blackenergy-cyber-attacks-against-ukrainian-government-linked-russia-1467401 (accessed on 1 December 2022).
2. Paganini, P. *Black-Energy Used as a Cyber Weapon Against Ukrainian Critical Infrastructure*; Infosec Institute: Madison, WI, USA, 2017.
3. TXOne. 2021. Cybersecurity Report. 2022. Available online: https://www.txone.com/security-reports/2021-cybersecurity-report/ (accessed on 1 December 2022).
4. Hokstad, P.; Utne, I.B.; Vatn, J. *Risk and Interdependencies in Critical Infrastructures*; Springer: London, UK, 2012.
5. Możaryn, J.; Ordys, A.; Stec, A.; Bogusz, K.; Al-Jarrah, O.Y.; Maple, C. Design and Development of Industrial Cyber-Physical System Testbed. In *Advanced, Contemporary Control*; Springer: Cham, Switzerland, 2020; pp. 725–735.
6. Kościelny, J.; Syfer, M.; Ordys, A.; Wnuk, P.; Możaryn, J.; Fajdek, B.; Puig, V.; Kukiełka, K. Towards a unified approach to detection of faults and cyber-attacks in industrial installations. In Proceedings of the 2021 European Control Conference (ECC), Rotterdam, The Netherlands, 29 June–2 July 2021; pp. 1839–1844.
7. Syfert, M.; Ordys, A.; Kościelny, J.M.; Wnuk, P.; Możaryn, J.; Kukiełka, K. Integrated Approach to Diagnostics of Failures and Cyber-Attacks in Industrial Control Systems. *Energies* **2022**, *15*, 6212. [CrossRef]
8. Pasqualetti, F.; Dorfler, F.; Bullo, F. Attack detection and identification in cyber-physical systems. *IEEE Trans. Autom. Control.* **2013**, *58*, 2715–2729. [CrossRef]
9. Teixeira, A.; Amin, S.; Sandberg, H.; Johansson, K.H.; Sastry, S.S. Cyber security analysis of state estimators in electric power systems. In Proceedings of the 49th IEEE Conference on Decision and Control (CDC), Atlanta, GA, USA, 15–17 December 2010; IEEE: Piscataway, NJ, USA; pp. 5991–5998.
10. De Sá, A.O.; Da Costa Carmo, L.F.R.; Machado, R.C. Covert attacks in cyber-physical control systems. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1641–1651. [CrossRef]
11. Van Long, D.O.; Fillatre, L.; Nikiforov, I. Sequential monitoring of SCADA systems against cyber/physical attacks. *IFAC-PapersOnLine* **2015**, *48*, 746–753.
12. Syfert, M.; Kościelny, J.M.; Możaryn, J.; Ordys, A.; Wnuk, P. Simulation Model and Scenarios for Testing Detectability of Cyberattacks in Industrial Control Systems. In *Intelligent and Safe Computer Systems in Control and Diagnostics, Proceedings of the International Conference on Diagnostics of Processes and Systems Chmielno, Kashubia, Poland, 5–6 September 2022*; Springer: Cham, Switzerland, 2023; pp. 73–84.
13. Smith, R.S. A decoupled feedback structure for covertly appropriating networked control systems. *IFAC Proc. Vol.* **2011**, *44*, 90–95. [CrossRef]
14. Zhu, B.; Joseph, A.; Sastry, S. A taxonomy of cyber attacks on SCADA systems. In Proceedings of the 2011 International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing, Dalian, China, 19–22 October 2011; pp. 380–388.
15. 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, 22–15 March 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–5.
16. Buchanan, S.S. Cyber-Attacks to Industrial Control Systems since Stuxnet: A Systematic Review. Ph.D. Thesis, Capitol Technology University, Laurel, MD, USA, 2022.

17. Alanazi, M.; Mahmood, A.; Chowdhury, M.J.M. SCADA Vulnerabilities and Attacks: A Review of the State-of-the-Art and Open Issues. *Comput. Secur.* **2022**, *125*, 103028. [CrossRef]
18. Albright, D.; Brannan, P.; Walrond, C. Did Stuxnet Take Out 1000 Centrifuges at the Natanz enrichment Plant? Institute for Science and International Security: Washington, DC, USA, 2010.
19. Myung, J.W.; Hong, S. ICS malware Triton attack and countermeasures. *Int. J. Emerg. Multidiscip. Res.* **2019**, *3*, 13–17. [CrossRef]
20. Di Pinto, A.; Dragoni, Y.; Carcano, A. TRITON: The first ICS cyber attack on safety instrument systems. *Proc. Black Hat USA* **2018**, *2018*, 1–26.
21. Schellenberger, C.; Zhang, P. Detection of covert attacks on cyber-physical systems by extending the system dynamics with an auxiliary system. In Proceedings of the 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, Australia, 12–15 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1374–1379.
22. Hoehn, A.; Zhang, P. Detection of covert attacks and zero dynamics attacks in cyber-physical systems. In Proceedings of the 2016 American Control Conference (ACC), Boston, MA, USA, 6–8 July 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 302–307.
23. Elsisi, M.; Tran, M.Q. Development of an IoT architecture based on a deep neural network against cyber-attacks for automated guided vehicles. *Sensors* **2021**, *21*, 8467. [CrossRef] [PubMed]
24. Zizzo, G.; Hankin, C.; Maffeis, S.; Jones, K. Intrusion Detection for Industrial Control Systems: Evaluation Analysis and Adversarial Attacks. *arXiv* **2019**, arXiv:1911.04278.
25. Kravchik, M.; Shabtai, A. Efficient Cyber Attack Detection in Industrial Control Systems Using Lightweight Neural Networks and PCA. *IEEE Trans. Dependable Secur. Comput.* **2019**, *19*, 2179–2197. [CrossRef]
26. Wang, W.; Harrou, F.; Bouyeddou, B.; Senouci, S.-M.; Sun, Y. Cyber-attacks detection in industrial systems using artificial intelligence-driven methods. *Int. J. Crit. Infrastruct. Prot.* **2022**, *38*, 100542. [CrossRef]
27. Alagappan, A.; Venkatachary, S.K.; Andrews, L.J.B. Augmenting Zero Trust Network Architecture to enhance security in virtual power plants. *Energy Rep.* **2022**, *8*, 1309–1320. [CrossRef]
28. Ullah, S.; Zheng, J.; Din, N.; Hussain, M.T.; Ullah, F.; Yousaf, M. Elliptic Curve Cryptography; Applications, challenges, recent advances, and future trends: A comprehensive survey. *Comput. Sci. Rev.* **2023**, *47*, 100530. [CrossRef]
29. Tidrea, A.; Korodi, A.; Silea, I. Cryptographic Considerations for Automation and SCADA Systems Using Trusted Platform Modules. *Sensors* **2019**, *19*, 4191. [CrossRef]
30. Gilles, O.; Pérez, D.G.; Brameret, P.-A.; Lacroix, V. Securing IIoT communications using OPC UA PubSub and Trusted Platform Modules. *J. Syst. Archit.* **2023**, *134*, 102797. [CrossRef]
31. Li, D.; Paynabar, K.; Gebraeel, N. A degradation-based detection framework against covert cyberattacks on SCADA systems. *IISE Trans.* **2021**, *53*, 812–829. [CrossRef]
32. Jelali, M. *Control Performance Management in Industrial Automation: Assessment, Diagnosis and Improvement of Control Loop Performance*; Springer: Berlin/Heidelberg, Germany, 2012. [CrossRef]
33. Domański, P.D. *Control Performance Assessment: Theoretical Analyses and Industrial Practice*; Springer: Cham, Switzerland, 2020; Volume 245.
34. Umsonst, D.; Sandberg, H. Experimental evaluation of sensor attacks and defense mechanisms in feedback systems. *Control. Eng. Pract.* **2020**, *124*, 105178. [CrossRef]
35. Astrom, K.J. *Introduction to Stochastic Control Theory*; Elsevier: Amsterdam, The Netherlands, 1971.
36. Harris, T.J. Assessment of control loop performance. *Can. J. Chem. Eng.* **1989**, *67*, 856–861. [CrossRef]
37. Desborough, L.; Harris, T. Performance assessment measures for univariate feedback control. *Can. J. Chem. Eng.* **1992**, *70*, 1186–1197. [CrossRef]
38. Georgakopoulos, D.; Jayaraman, P.P.; Fazia, M.; Villari, M.; Ranjan, R. Internet of Things and edge cloud computing roadmap for manufacturing. *IEEE Cloud Comput.* **2016**, *3*, 66–73. [CrossRef]
39. Bysko, S.; Bysko, S.; Fratczak, M.; Nowak, P.; Klopot, T.; Czeczot, J.; Stebel, K.; Laszczyk, P. PID Controller tuning by Virtual Commissioning-a step to Industry 4.0. *J. Phys. Conf. Ser.* **2022**, *2198*, 012010. [CrossRef]
40. Frątczak, M.; Nowak, P.; Kłopot, T.; Czeczot, J.; Bysko, S.; Bysko, S. Component-based simulation tool for virtual commissioning of control systems for heat exchange and distribution processes. In Proceedings of the International Conference on Automation, Online, 20–21 August 2020; Springer: Cham, Switzerland, 2020; pp. 67–77.
41. Nazir, S.; Patel, S.; Patel, D. Assessing and augmenting SCADA cyber security: A survey of techniques. *Comput. Secur.* **2017**, *70*, 436–454. [CrossRef]