Anomaly Detection and Data Recovery on Mini Batch Distillation Column based Cyber Physical System

Wedar Panji Mardyaningsih
Department of Electrical Engineering
Institut Teknologi Bandung
Bandung, Indonesia
wedar.panji@s.itb.ac.id

Pranoto Hidaya Rusmin
Department of Electrical Engineering
Institut Teknologi Bandung
Bandung,Indonesia
pranoto@lskk.ee.itb.ac.id

Budi Rahardjo
Department of Electrical Engineering
Institut Teknologi Bandung
Bandung Indonesia
rahard@gmail.com

***Abstract*— The development of industrial revolution 4.0 in industrial sector opened a cyber gap for outsiders to pose a threat to the system. Industrial control systems initially designed to meet SRA (Safety, Reliability, and Availability) priorities are now beginning to be pressed to consider security aspects related to the magnitude of the impact that can be caused due to external attacks. In making a safe Cyber Physical System (CPS) based automation, risk assessment will be used to determine the level risk of threat. Mini distillation column batch based CPS will be implemented as the approach of CPS in industrial sector. Anomaly detection based data-driven model and data recovery method is proposed to lower the impact of attack on this system.**

Keywords—CPS, anomaly detection, data recovery, risk assessment

I. INTRODUCTION

In recent years the industry has increasingly developed into the realm of connectivity between automatics and cyber, known as the industrial revolution 4.0. The industry begins to enter the stage where data from production machines can be collected and shared in real time [1]. This development makes the physical system and cyber becomes integrated as Cyber Physical System (CPS). As the physical system is integrated with cyber, security aspects must be considered especially when the system has a large impact on society [2].

However physical system that already integrated with cyber system will become more vulnerable because of the opportunity to be controlled remotely, the lack of a proper security mechanism, the exchange of data that is quite critical, and the vendor's low priority in taking into account security aspects when create the physical system [3]. Furthermore at the beginning, the control systems development in the industry was not designed to prioritize security aspects but more to availability aspects for increasing production [2].

The application of cyber technology opens a gap that is large enough to launch attacks on industrial processes. Based on the global annual IT security risks survey on 2017 conducted by Kaspersky Lab and B2B International, from over 900 surveyed industrial company, there are 28% of them admitted that they faced an attack in 2017, compared to 20% in 2016. It is also revealed that 48% of the company lack of insight and having difficulties on identifying cyberattacks [4]. Direct attacks occurred in 2000 when a former employee damaged the operation of a sewage treatment system in Queensland, Australia. This former employee changes the configuration at the pumping station so that raw waste material in large capacities meets the road and environment [5]. Increasing cases of cyberattack and threats make it is necessary to identify and implement the right solution [3].

One of the CPS applications in the industrial field is in the oil industry. Where one of the important processes in this stage of industry is distillation to obtain pure products from distilled liquids [6]. On a large scale, data related to production in the oil industry will be transmitted wirelessly to devices in the field. Wireless data transmission has a concept similar to the implementation of smart grid [7].

Recent studies [8] already implementing CPS on mini distillation column as the prototype of oil industry. Data from sensors in the plant, which are boiler temperature sensors, upper column temperatures, and concentration sensors will be transmitted through the Message Queuing Telemetry Transport (MQTT) protocol to the controller. The controller will do computation to create control signal based on data sensor that already transmitted. However this system has not considering the safety term yet. As the implementation is using MQTT protocol for the communication, several research already conduct to determine the threat and attack scenario that could be happen on system that used this protocol. Several attacks on data privacy, authentication, integrity, and port obscurity could happen on MQTT [9].

Data integrity of sensor is important to derived the control signal for physical system. Therefore, the sensor data value must be monitored during operations to prevent the occurring of failed data. Failed sensor data could lead into the enlarged of control signal. There are two categories to provide system working condition information, which are model-based method and data driven method [10]. Model-based method is appropriated on achieving precise and stable prediction result of the system condition. However data driven method is much easier to be implemented on most system because it is only depend on data monitoring of the system.

This paper is organized as follows; in section II the author will explain about risk assessment conducted as a method to define the risk on the system. Section III will present the experiments and analysis of anomaly detection. Section IV will represent the recovery method when attack happen. Finally section V will present the conclusion and future work.

II. RISK ASSESSMENT

Risk assessment is intended to determine possible threats, risks, and vulnerabilities that might occur in the system. On this research, risk assessment conducted based on NIST SP-300 standards.

A. System Characterization

The CPS architecture was built using a plant in the form of a mini distillation column intended for alcohol distillation. In the system architecture built, there are a Mini PC as broker MQTT, a database server as the control signal computation unit, and several sensor and actuator transmitter [8]. This architecture can be seen on figure 1. Local network communication protocol is implemented by MQTT protocol that connects local controllers, sensors, local brokers and database servers. This paper will focused on analyzing the risk assessment on local network as it has direct impact on the physical system operations.

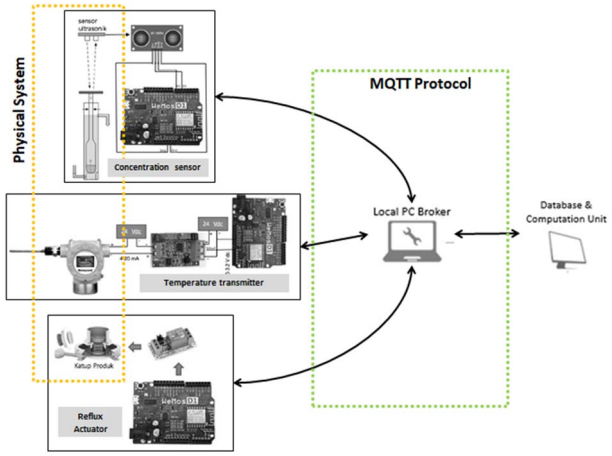

Fig. 1. CPS architecture implemented on mini distillation column batch

B. Threat Identification

Threats are the possibility of something from outside the system that can generate vulnerabilities in the system so that the function of the system cannot run as it should. Identifying threats is done by determining the source of the threat and the action of the threat based on examples of cases that have occurred and possible identification based on the current system. As mentioned in [3] the source of the threat to the MQTT protocol can be defined as follows external attacker, curious user, and malicious internal user. The types of threat from [3], [9], [11] is summarized on figure 2.

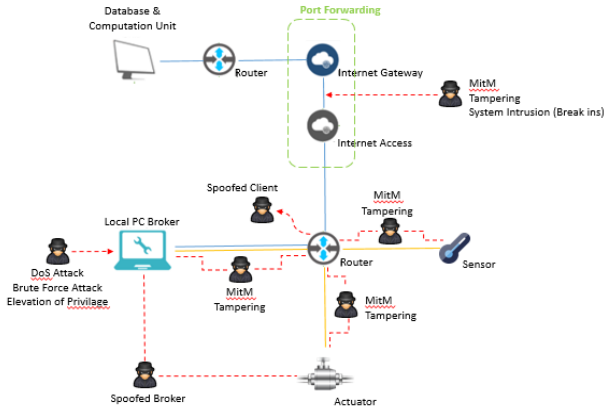

Fig. 2. Threat model analysis on mini distillation column batch based CPS

C. Risk Analysis

Risk analysis is done to determine the level of risk in the system. To determine the level of risk, the threat model has to be analyzed in the term of considering the likelihood of threat to exploit the vulnerabilities and the impact of threat to the system.

Likelihood analysis is conduct with classify each threat into three levels, namely:

- ***High*** : easy method to conduct threat, control not effective
- ***Medium*** : medium knowledge to conduct threat, control can detain the threat
- ***Low*** : hard method, control can prevent or stops the threat

Impact analysis is conduct with classify each threat into three levels, namely:

- ***High*** : permanent impact, system operation stops
- ***Medium*** : temporary impact, system still can operate even the performance is decreasing
- ***Low*** : temporary impact, system performance still good

Determining the level of risk is done using the risk-level matrix in table 1.

TABLE I. RISK LEVEL MATRIX

		Impact		
		Low(10)	**Medium(50)**	**High(100)**
Threat Likelihood Determination	**High(1.0)**	**Low** 10x1.0=10	**Medium** 50x1.0=50	**High** 100x1.0=100
	Medium (0.5)	**Low** 10x0.5=5	**Medium** 50x0.5=25	**Medium** 100x0.5=50
	Low(0.1)	**Low** 10x0.1=1	**Low** 50x0.1=5	**Low** 100x0.1=10

From the analysis, the threat with high risk level is data tampering/modification because in the term of control, the modification of data could lead into high error when system is running. Besides that the current system has no mitigation action to prevent data tampering on system

D. Data Tampering Scenario

The scenario in the case of system modification is compiled to describe a number of scenarios that will be taken to be reviewed as a case so that an effective solution can be obtained. Data tampering is an action of inserting, deleting, or altering the information [3]. On this paper, inserting false data will be used as the scenario of data tampering. Attacker will insert invalid value of concentration sensor to broker. When database get the invalid data, the decision made by computation of control signal will have disturbed by the invalid data. The author make assumption that control signal transmission has safe connection.

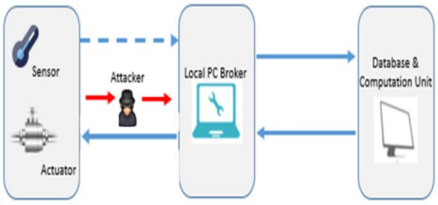

Fig. 3. Data tampering scenario using invalid data insertion

III. DETECTION OF ATTACK

There are two approaches on detecting the attack in the case of modification data attacks, namely using anomaly data approach and delay node approach. Anomaly is defined as the process of findings patterns in dataset whose behavior is normal or not [12]. On this paper the author will use data anomaly approach as the model of system. In anomaly data approach, the attack on the system will be modelled as disturbance. Because the scenario of modification happened on the transmission of concentration sensor data, the disturbance will be added at sensor data feedback as the figure 4.

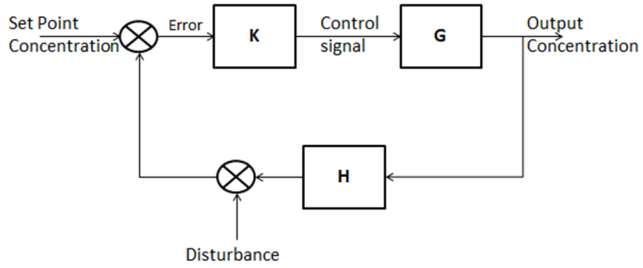

Fig. 4. Data modification as disturbance on closed loop system

The concentration value of the sensor in normal conditions will be modeled using a Neural Network (NN). NN consists of some computational units called neuron. An input received on neuron will be multiplied by corresponding edge weight. The sum of weight will be applied by activation function to produce output as equation (1). The weight and neuron bias will be updated recursively until error between output prediction and actual output is converged [13].

$$y_k = W.x_{k-N} + b \tag{1}$$

where :

- y_k : output from sensor reading
- W : weight matrix
- x_{k-N} : state matrix
- b : bias

In this implementation feed-forward neural network, loss function was used in the form of mean squared error (MSE) to predict the continue output as equation (2).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_{true} - y_{predict}\right)^2 \tag{2}$$

where N is the sum of data.

This modeling is done using open loop control on 70% and 80% reflux ratio. In this experiment, data at time t has relation with the data before at t-1. Because of that the model uses the first data as an input and the next data as an output. This method is then carried out sequentially from the first data until the last data as figure 5. The chosen model was the one which has lowest train loss (figure 6) and the prediction was fit enough with test data (figure 7).

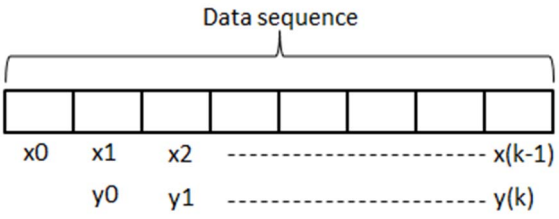

Fig. 5. Data sequence on system modelling

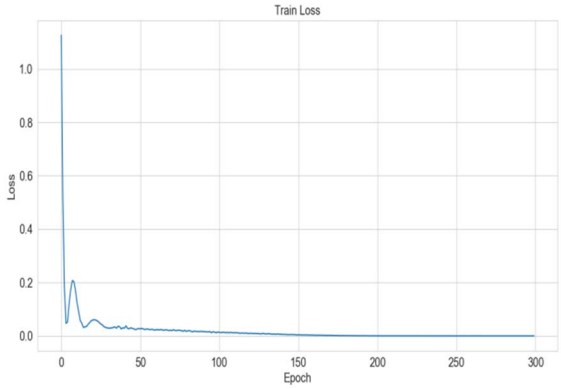

Fig. 6. Train loss

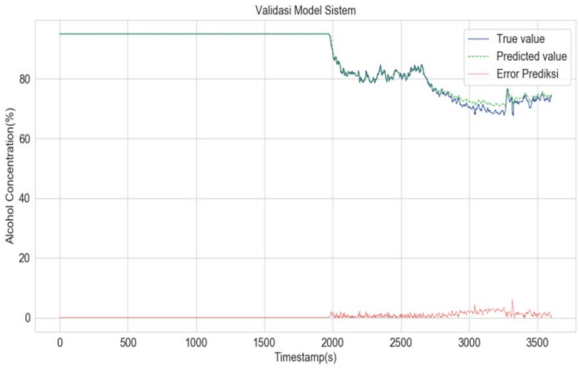

Fig. 7. Prediction and actual data in normal condition

The model that has been obtained is then used to predict the concentration value when the system operates. The sensor value prediction is done in real time with the input being the actual sensor value at that time. Predicted values are then compared with actual values. The difference between these values is called prediction error as in equation (3).

$$error_{prediction} = y_{true} - y_{predict} \tag{3}$$

The decision rule is formulated as if the prediction error is less or equal to the threshold level then the algorithm indicates normal behaviour, otherwise it is anomaly (abnormal) [14]. This decision rule then implemented on this paper as the difference in value above the threshold will indicate the modification of data. While the error value below the threshold will be considered as noise.

The threshold is determined by approaching the error prediction as noise of sensor reading. It can be seen from table II that the maximum error prediction from several experiment conditions is 13.56. Therefore the threshold to differentiate between anomaly and process noise will be 13.56.

TABLE II. ERROR PREDICTION FROM OPEN LOOP AND CLOSED LOOP CONDITION

System Condition	Error Prediction		
	Min	Max	Range
Open loop 70% reflux	3.23E-04	13.5164	13.52
Open loop 80% reflux	9.16E-05	13.5643	13.56
Closed loop Kp 1	7.63E-06	12.2996	12.30

The detection experiment was carried out by injecting false concentration data sensor. This false data will destruct the mean average of sensor data detected by database computer. The injections were held several times like it can be seen in figure 8.

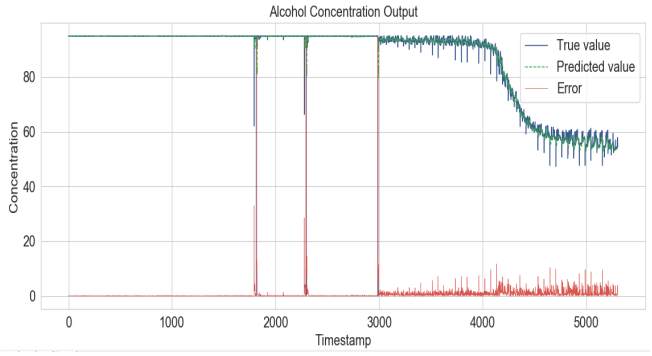

Fig. 8. Anomaly detection on 70% reflux setting

Anomaly detection mechanism will give notification when it detects the attack. The notification can be seen in figure 9.

```
====================Anomaly report=========================
Anomaly detected at 1921.5 s. Error prediction: 32.897255
Anomaly detected at 1945.04 s. Error prediction: 94.69729
Anomaly detected at 2440.15 s. Error prediction: 28.588753
Anomaly detected at 2458.49 s. Error prediction: 94.71648
Anomaly detected at 3207.08 s. Error prediction: 94.94364
```

Fig. 9. Anomaly report on attack detection

IV. DATA RECOVERY

The CPS-based mini distillation system has a higher level of complexity compared to other IT systems. Moreover the system requires a continuous operation so that if the system is turned off to take an incident response action, it will cause a loss. Thus a new approach is needed. This recovery approach includes switching control functions to system failure. In this situation the process is still ongoing even the system optimality is reduced [15].

Data recovery is conduct to normalize the anomaly data, so that the control signal will not be exploding because of the anomaly data. Data recovery is done using anomaly threshold filter to eliminate the anomaly data and kalman filter to eliminate noise.

A. Kalman Filter

Kalman filter is based on minimizing the mean square error recursively. Kalman filter is often used for target tracking, noise cancellation and other dynamics system [16]. There are five steps on kalman filter, which are process update, measurement update, covariance update, kalman gain computation, and error estimates update. On this paper kalman filter is conducted by using the actual raw value of concentration sensor data as the system equation. Process noise covariance matrix and measurement noise covariance matrix are made to be constant in this paper.

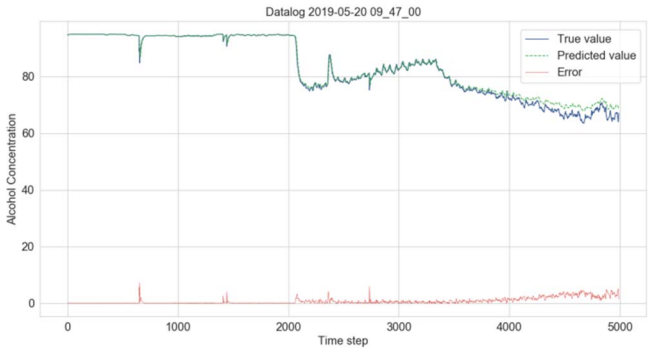

Fig. 10. Kalman filter output on 70% reflux ratio

Fig. 11. Kalman filter output on 80% reflux ratio

As it can be seen from figure 10 and 11, the noise was reduced so that the image becomes clearer. However some anomaly point still can be seen like a glitch on several points. It is because the filter system input still used actual raw data that the anomaly is still considered as normal data in this section.

B. Anomaly Threshold

Several approaches has been created on recover anomalous sensor data. As in [17], mutual information level between sensors is used to determine the other available sensors data that can be used to recover the anomalous sensor data. The highest mutual information sensor will be chosen to predicting the anomalous sensor value.

As mentioned in [14], data is decided into anomaly when it has passed the threshold value of error prediction. From the previous part, the threshold which obtained from several experiments is 20% (in the scale of alcohol concentration). Therefore when the error prediction is above the threshold, the actual value will not be used to predict the next output. Instead the predicted value of the previous actual value will be the input to predict the next output. After anomaly value was eliminated, the results from the previous step will be the input for kalman filter to eliminate the noise.

$$y_k = \begin{cases} Cx_k & ,without\ attack \\ Cx_{k-N} & ,with\ attack \end{cases} \quad (4)$$

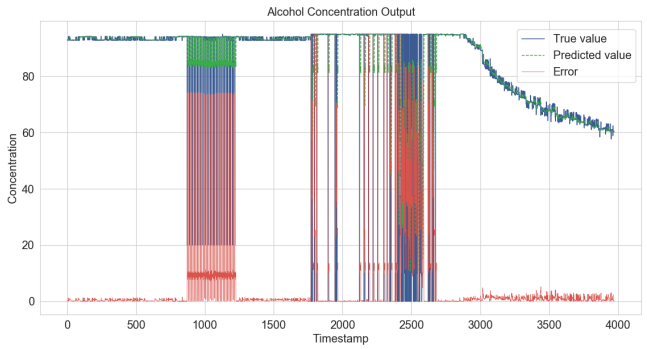

Fig. 12. Anomaly attack scenario on closed loop proporsional controller

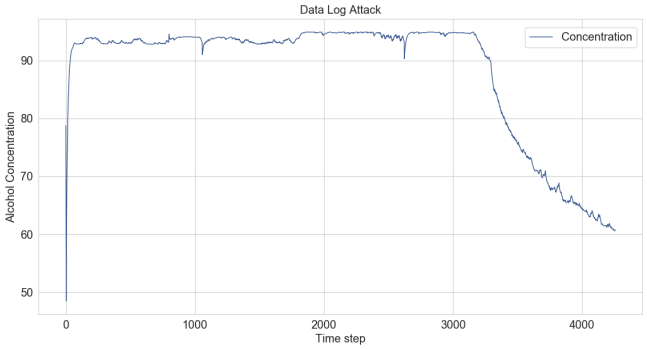

Fig. 13. Concentration output after recovery scenario and kalman filter

After implementing anomaly threshold elimination, some points that have error prediction above 13.56% were eliminated. It results on normalized data that resemble the actual normal data. In figure 12 the noise is reduced and the anomaly point is disappeared. The same results also occurred on the experiment represented by figure 13.

The experiment is also hold by comparing the distillate mixture ratio (DMR) between the attacked data recovery and non-recovery. The equation for DMR can be seen below

$$DMR = \frac{\psi_{distillate} \cdot V_{distillate}}{\psi_{mixture} \cdot V_{mixture}} \quad (5)$$

where ψ is the alcohol concentration(%) and V is the volume. As it can be seen form table III the recovery mechanism give improvement of system performance as the DMR of the system with recovery is higher than without recovery.

TABLE III. DMR PERFORMANCE INDICATOR ON RECOVERY TESTING

Condition	Initial Condition	Distillate Condition	DMR
Attack	30%, 3L	83%, 0.75L	69.17%
Recovery	30%, 3L	88%, 0.8L	78.22%

V. CONCLUSIONS AND FUTURE WORKS

Risk analysis is conduct to determine that data tampering has the highest risk level on the case of mini distillation column based CPS. Data driven framework for anomaly detection and data recovery based on neural network is proposed in this study. The main contributions of this article can be concluded that: (1) Anomaly of concentration sensor data can be detected in real time using neural network model as linear approach of the system. (2)Anomaly elimination based on threshold and kalman filter can be used as data recovery method for invalid data insertion in the case of data tampering attacks. In future, research works will be carried out to detect anomaly and recover data for other types of attacks.

ACKNOWLEDGMENT

This work is partly supported by Honeywell Indonesia.

REFERENCES

[1] F. Shrouf, J. Ordieres, and G. Miragliotta, “Smart Factories in Industry 4 . 0 : A Review of the Concept and of Energy Management Approached in Production Based on the Internet of Things Paradigm,” *IEEE Int. Conf. Ind. Eng. Eng. Manag.*, pp. 697–701, 2014.

[2] D. Syed, T.-H. Chang, D. Svetinovic, T. Rahwan, and Z. Aung, “Security for Complex Cyber-Physical and Industrial Control Systems : Current Trends , Limitations , and Challenges Security for Complex Cyber-Physical and Industrial Control Systems : Current Trends , Limitations , and Challenges,” in *Pacific Asia Conference on Information Systems (PACIS)*, 2017, p. 180.

[3] S. N. Firdous, Z. Baig, C. Valli, and A. Ibrahim, “Modelling and Evaluation of Malicious Attacks against the IoT MQTT Protocol,” *2017 IEEE Int. Conf. Internet Things IEEE Green Comput. Commun. IEEE Cyber, Phys. Soc. Comput. IEEE Smart Data*, pp. 748–755, 2017.

[4] D. Berard, “Kaspersky Lab Survey: Targeted Attacks Increase for Industrial Organizations,” *Kaspersky Lab Survey*, 2018. [Online]. Available: https://usa.kaspersky.com/about/press-releases/2018_targeted-attacks-increase-for-industrial-organizations#_ftn1. [Accessed: 26-May-2019].

[5] A. Humayed, J. Lin, F. Li, and B. Luo, “Cyber-Physical Systems Security — A Survey,” *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1802–1831, 2017.

[6] V. T. Minh, “Modeling and control of distillation column in a petroleum process,” *Proc. 2010 5th IEEE Conf. Ind. Electron. Appl. ICIEA 2010*, pp. 259–263, 2010.

[7] H. Zhou *et al.*, “Analysis of production data manipulation attacks in petroleum cyber-physical systems,” *IEEE/ACM Int. Conf. Comput. Des.*, pp. 1–7, 2016.

[8] I. Budiawan, “Cyber Physical System-Based Automation : Studi Kasus Kendali PD Berbasis Kejadian pada Kolom Distilasi Mini,” 2018.

[9] S. Andy, B. Rahardjo, and B. Hanindhito, “Attack Scenarios and Security Analysis of MQTT Communication Protocol in IoT System,” *2017 4th Int. Conf. Electr. Eng. Comput. Sci. Informatics*, pp. 1–6, 2017.

[10] L. Liu, Q. Guo, D. Liu, and Y. Peng, “Data-driven Remaining Useful Life Prediction Considering Sensor Anomaly Detection and Data Recovery,” *IEEE Access*, vol. 7, pp. 1–1, 2019.

[11] A. W. Atamli and A. Martin, “Threat-based Security Analysis for the Internet of Things,” *2014 Int. Work. Secur. Internet Things Threat.*, pp. 35–43, 2014.

[12] J. D. Parmar and J. T. Patel, “Anomaly Detection in Data Mining : A Review,” *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 7, no. 4, pp. 32–40, 2017.

[13] A. Singh, “Anomaly Detection for Temporal Data using Long Short-Term Memory (LSTM),” KTH ROYAL INSTITUTE OF TECHNOLOGY, 2017.

[14] P. Filonov, A. Lavrentyev, and A. Vorontsov, “Multivariate Industrial Time Series with Cyber-Attack Simulation: Fault Detection Using an LSTM-based Predictive Data Model,” pp. 1–8, 2016.

[15] L.-E. Schneller, “Developing an Industrial Control Systems Cybersecurity Incident Response Capability,” Jul. 2009.

[16] G. Badhwar, S. Badhwar, and A. S. Sappal, “Noise reduction using kalman filter,” *Int. J. Latest Trends Eng. Technol.*, vol. 7, no. 1, pp. 639–643, 2016.

[17] L. Liu, D. Liu, Q. Guo, Y. Peng, and J. Liang, “SDR: Sensor data recovery for system condition monitoring,” *I2MTC 2018 - 2018 IEEE Int. Instrum. Meas. Technol. Conf. Discov. New Horizons Instrum. Meas. Proc.*, no. 1, pp. 1–6, 2018.