Proceedings Paper:
Jahromi, MZ, Jahromi, AA, Sanner, S et al. (2 more authors) (2020) Cybersecurity Enhancement of Transformer Differential Protection Using Machine Learning. In: 2020 IEEE Power & Energy Society General Meeting (PESGM). 2020 IEEE Power & Energy Society General Meeting (PESGM), 02-06 Aug 2020, Montreal, QC, Canada. IEEE . ISBN 978-1-7281-5509-8

https://doi.org/10.1109/pesgm41954.2020.9282161

Cybersecurity Enhancement of Transformer Differential Protection Using Machine Learning

Martiya Zare Jahromi, Amir Abiri Jahromi, Scott Sanner, and Deepa Kundur
University of Toronto
Toronto, Canada
ssanner@mie.utoronto.ca, dkundur@ece.utoronto.ca

Marthe Kassouf
Hydro-Quebec Research Institute (IREQ)
Varennes, Canada
kassouf.marthe@ireq.ca

***Abstract*—The growing use of information and communication technologies (ICT) in power grid operational environments has been essential for operators to improve the monitoring, maintenance and control of power generation, transmission and distribution, however, at the expense of an increased grid exposure to cyber threats. This paper considers cyberattack scenarios targeting substation protective relays that can form the most critical ingredient for the protection of power systems against abnormal conditions. Disrupting the relays operations may yield major consequences on the overall power grid performance possibly leading to widespread blackouts. We investigate methods for the enhancement of substation cybersecurity by leveraging the potential of machine learning for the detection of transformer differential protective relays anomalous behavior. The proposed method analyses operational technology (OT) data obtained from the substation current transformers (CTs) in order to detect cyberattacks. Power systems simulation using OPAL-RT HYPERSIM is used to generate training data sets, to simulate the cyberattacks and to assess the cybersecurity enhancement capability of the proposed machine learning algorithms.**

***Index Terms*—Cyberphysical systems, operational technology, machine learning, differential protective relays, transformers.**

I. Introduction

RAPID deployment of standard and interoperable ICT in power systems has raised serious cybersecurity concerns in regulatory agencies and power utilities over the past decade [1]. This is mainly because the security-by-obscurity philosophy used as a defensive strategy for proprietary ICT in power systems has become obsolete in the emerging standard and interoperable ICT paradigm of smart grids [2].

In order to address the growing cybersecurity concerns, the North American Electric Reliability Corporation has established and enforced Critical Infrastructure Protection (CIP) standards. The CIP standards demand utilities to identify, categorize and protect cyber assets that are essential to the reliable operation of the bulk electric system [3]. Nevertheless, CIP standards do not provide any guideline about the cybersecuity measures/tools that should be developed to improve the cyber-resiliency of the assets. As such, different standards and initiatives have been launched by National Standard Institutes like ISA [4]–[6], research institutes like Electric Power Research Institute (EPRI) [7] and government agencies like Department of Energy (DOE) [8], [9] to develop cybersecurity measures/tools for cyber assets in power systems.

The digitalization of power grids over the past decade has increased the cyberattack surfaces across several grid components and promoted the cybersecurity enhancement of its assets like substation protective relays to a top priority for regulatory agencies and utilities, in particular following the successful cyberattacks against Ukrainian power infrastructures [10], [11]. In a substation, protective relays form the most critical defensive ingredient of power system against abnormal conditions [12], [13]. Thus, their maloperations initiated by cyberattacks may cause major consequences for power systems such as widespread blackouts.

Cybersecurity of protective relays has been analyzed in the literature from different perspectives. In [14], the impact of cyberattacks against protective relays has been examined from operational point of view. An analytical reliability assessment framework has been proposed in [15] for quantifying the impacts of cyberattacks against intelligent electronic devices (IEDs) and protection systems. The impact of cyber-physical attacks against communication-assisted protection schemes has been studied in [16]. A rule-based intrusion detection system has been presented in [17] for IEC 61850 protocol. Cyber-resilient designs have been proposed in [18] and [19] respectively for distance protection and line differential protection relays. Various anomaly detection systems (ADS) have been presented in [20], [21] for substations. A collaborative intrusion detection system (IDS) has been presented in [22] for generic object oriented substation event (GOOSE) and sampled value (SV) packets in IEC 61850 protocol. A machine learning based method has been used in [24] to identify cyberattacks against phasor measurement units (PMUs).

In this paper, a fully connected autoencoder is employed for detection and mitigation of cyberattacks against transformer differential protective relays. The fully connected autoencoder is trained with three-phase current measurements from current transformers (CTs) at both sides of a transformer. The fully connected autoencoder is then used to detect anomalies in three-phase currents. The simulation results demonstrate the capability of fully connected autoencoder to detect cyberattacks against transformer differential protective relays.

The main contributions of this paper are as follows.

- A method is presented for cybersecurity enhancement of transformer differential protective relays using machine learning.
- The performance of the proposed machine learning

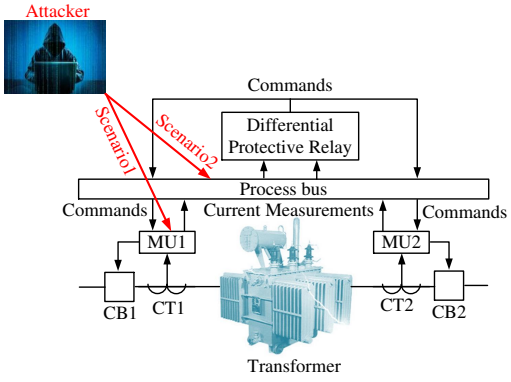

Fig. 1. Transformer differential protective relay.

method is investigated for two different cyberattack scenarios.

The remainder of this paper is organized as follows. Section II presents the basics of transformer differential protective relays and the potential cyberattack scenarios. The proposed machine learning method for cybersecurity enhancement of transformer differential protective relays is presented in Section III. Section IV provides the test system, training data sets and simulation results. The conclusions of the paper are given in Section V.

II. Transformer Differential Protective Relays and Cyberattack Scenarios

A. Transformer Differential Protective Relays

The primary objective of transformer protective relays is to detect internal transformer faults with high sensitivity and isolate the transformer as quickly as possible. Fast detection and de-energization of transformer faults minimizes the damages to the transformer as well as the need for subsequent repairs [12]. This task is normally performed by the transformer differential protection.

The differential relaying is a powerful relaying principle that can be used for any asset like transformers, lines, buses, and so forth. The differential protective relays are designed to measure the geometrical difference between electrical quantities in particular current measurements and operate when the difference goes beyond a certain threshold.

B. Cyberattack Scenarios

We consider a substation automation scheme that employs the IEC 61850 protocols GOOSE and SV for communication between protective relays and merging units. The analog measurements generated by the current transformers CT1 and CT2 in Fig. 1 are respectively converted by merging units MU1 and MU2 to SV packets. The GOOSE commands generated by the differential protective relay in Fig. 1 are conveyed to the merging units MU1 and MU2 in order to trigger actions respectively on the circuit breakers CB1 and CB2.

We consider two cyberattack scenarios against the transformer differential protective relay. In the first scenario (referred to as Scenario 1), we assume the implementation of a compromised electronic component inside the merging unit MU1, thus, corresponding to a supply chain attack such as in [25]. The embedded malicious electronic component is able to tamper the data received from CT1, for example, by changing the magnitude of current measurements.

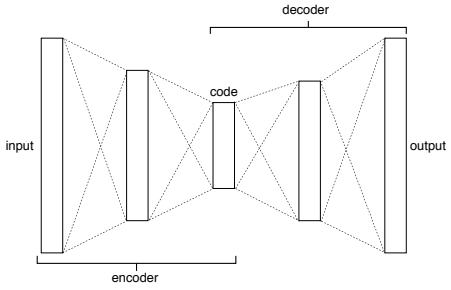

Fig. 2. Autoencoder structure

In the second scenario (referred to as Scenario 2), an attacker gains remote access to the substation process bus through the use of stolen legitimate operator credentials and a remote connection to the substation communication network. The attacker then performs false data injection attack by injecting falsified SV packets on the process bus forcing the differential protective relay to misoperate.

In both scenarios, falsified current measurements trigger the differential protective relay and de-energize the transformer in the absence of physical faults. Although substation and control center operators observe the transformer tripping, they would not be able to re-energize the transformer before performing a comprehensive investigation about the reason behind transformer tripping based on utility guidelines. Having machine learning algorithms for anomaly detection would provide a mitigation strategy to prevent differential protective relay misoperation caused by cyberattacks.

III. Proposed Machine Learning Method

In this paper, we leverage a fully connected autoencoder for cybersecurity enhancement of differential protective relays by using it to detect anomalies.

From a technical perspective, a fully connected autoencoder consists of two parts: an encoder and a decoder as illustrated in Fig. 2. An encoder consists of an input layer, a variable number of hidden layers and a code (embedding) layer. The code layer connects the encoder and decoder. The decoder consists of the same number of hidden layers as the encoder and an output layer. An autoencoder is trained with inputs as output target labels. When provided with a new input post-training, it will nonlinearly embed the input into a code and then nonlinearly decompress this code to approximately reconstruct the input at the output [26].

In this work, we train the autoencoder on current measurements that are free of cyberattacks with the aim that it will be able to accurately compress and reconstruct such attack-free measurements. It is noteworthy that the attack-free measurements are easy to collect from power systems since power systems are not under attack normally and cyberattacks

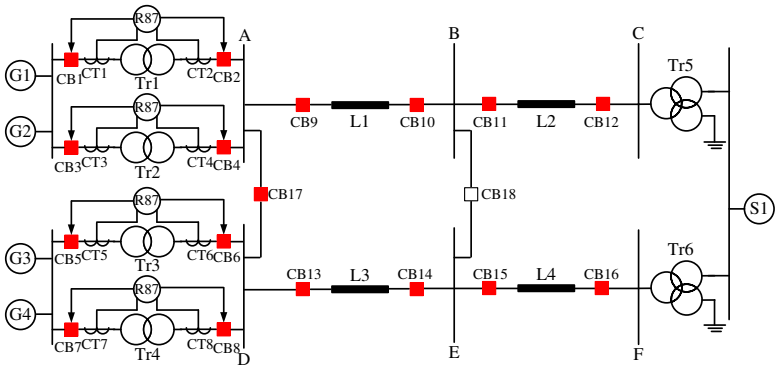

Fig. 3. The IEEE PSRC D6 benchmark test system.

are very rare incidents in power systems. This is in line with assumption in anomaly detection systems in the literature [27]. Since the autoencoder has not been trained on data containing cyberattacks, we hypothesize that reconstructions of anomalous measurements occurring during attacks may not be reconstructed well [27], [28]. Thus, we aim to use a threshold of the autoencoder reconstruction error as a means of detecting anomalous measurements that may indicate cyberattacks.

IV. Simulation Results

A. Test System

Fig. 3 illustrates the IEEE power system relaying committee (PSRC) D6 benchmark test system [29], [30]. The benchmark test system represents a power plant with four generators G1-G4 that are connected to the main grid through a 500kV transmission system. The 500kV transmission system consists of four transmission lines L1-L4 and the main grid is modeled as an infinite bus S1. The power plant transformers Tr1-Tr4 are protected by differential protective relays.

B. Training Data Set

We employed OPAL-RT HYPERSIM to implement and simulate PSRC D6 test system and generate the data set for machine learning. The operating points of the generators G1-G4 are changed between 200 MW and 400 MW in 50 MW step size. The three-phase internal fault of the transformer Tr1 is simulated for 16 randomly selected starting times to ensure faults occur at different parts of the current waveforms. This amounts to 10000 simulations. In order to be consistent with SV packet specifications in IEC 61850 standard, we capture 4800 samples per second for current measurements per phase. The 10000 simulations are performed for 1.5 seconds and the transformer faults are initiated randomly between t=1 s to t=1.02 s. It is noteworthy that the period of one cycle is approximately 0.0167 s in a 60Hz power system.

An important parameter for generating training data set is the input data length, *i.e.*, number of samples in each training data. In this paper, a sliding window of 10-ms, *i.e.*, 48 samples of current for each phase is considered. Thus, each training data contains 288 current samples, *i.e.*, 144 three-phase current samples from each side of the transformer. In order to generate a 10-ms sliding window, we first extract a 20-ms window from the 1.5 s simulation results containing 48 samples before the starting point of the fault and 47 samples after the starting point of the fault. Next, a sliding window of 10-ms slides sample by sample till it covers the 20-ms window. This amounts to 48 10-ms windows per simulation. Considering 10000 simulations, 2000 simulations data are used for testing. Hence, the training data set contains 48*8000=384000 10-ms current measurement data.

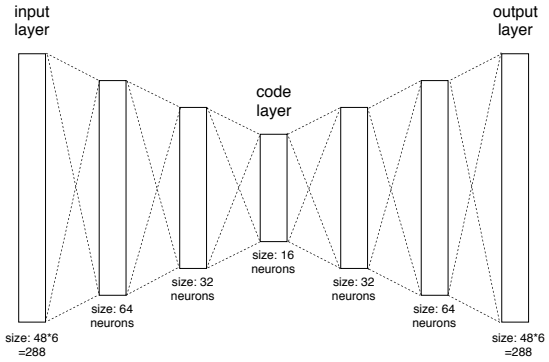

Fig. 4. Proposed Fully Connected Autoencoder Structure

C. Cyberattack Test Sets

In a real world, the anomaly detection data is heavily imbalanced and attacks are rare events. As we mentioned, we have 2000 simulation data sequences for test. To make an imbalanced data set, for each scenario, we replace 100 of the sequences with malicious data sequences. As discussed in Section II. B, we consider two cyberattack scenarios against transformer differential protective relay. Malicious data sequences are generated based on the aforementioned cyberattack scenarios. In Scenario 1, the current measurements from the current transformer CT1 are scaled to trigger the transformer differential protective relay. In Scenario 2, the current measurements from current transformer CT1 are replaced by false data.

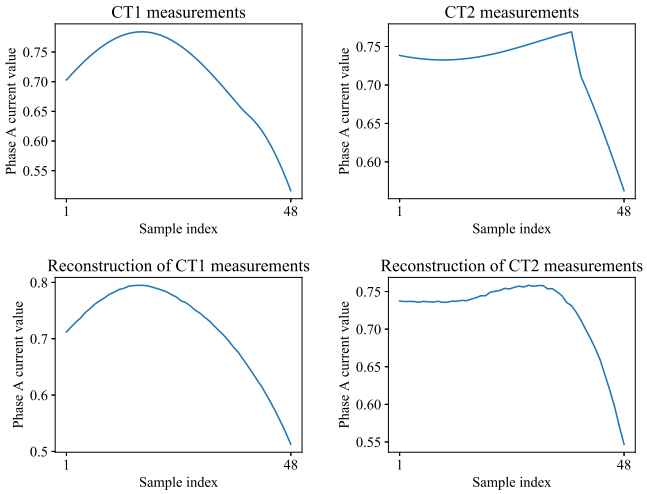

Fig. 5. Reconstruction of phase A current measurements from CT1 and CT2 during transformer physical fault.

D. Metrics for Measuring the Cyberattack Detection Performance of Fully Connected Autoencoder

Two metrics including precision and recall are considered to measure the performance of the fully connected autoencoder as follows.

$$Precision = \frac{(TruePositive)}{(TruePositive) + (FalsePositive)} \quad (1)$$

$$Recall = \frac{(TruePositive)}{(TruePositive) + (FalseNegative)} \quad (2)$$

In (1)-(2), True Positive represents anomalous data that are correctly identified by fully connected autoencoder. False Positive represents three-phase transformer fault data that are incorrectly identified as anomalous data. False Negative represents anomalous data that are identified as transformer three-phase fault. True Negative represents transformer three-phase faults that are correctly identified as transformer three-phase fault.

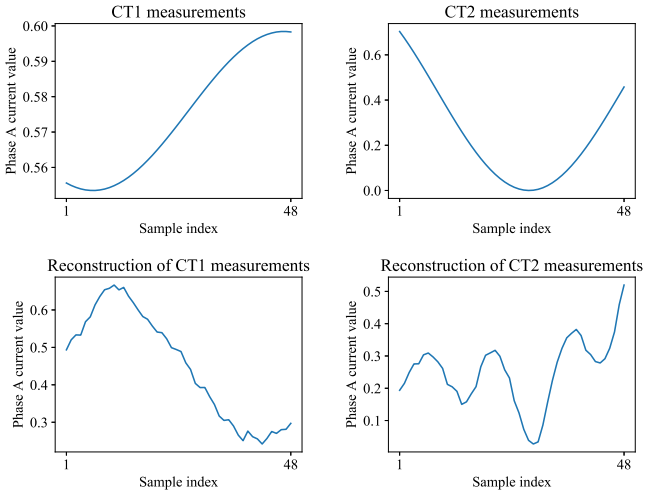

Fig. 6. Reconstruction of phase A current measurements from CT1 and CT2 after a supply chain cyberattack.

E. Fully Connected Autoencoder Architecture

As illustrated in Fig. 4 for the input layer, data has been flattened to a vector size of $(window\ \ size) * (count\ of\ \ features) = 48 * 6$. Hence, we have input layer of size 288, and output layer with the same size. The code size has been set to size 16. The autoencoder has two hidden layers for encoder/decoder. The hidden layers in the encoder have 64 and 32 neurons respectively, and the hidden layers in the decoder have 32 and 64 neurons.

We examined different values for each parameter in order to tune hyper-parameters of the autoencoder as summarized in Table I. The parameters that produced the lowest validation errors are selected. It is noteworthy that the validation errors are not reported here for the sake of conciseness. The Relu activation function is used for all layers except the last layer which uses linear activation function. The Adam Optimizer is further used for optimization. The learning rate is set to 0.01. It is noteworthy that we used Tensorflow and Keras libraries for the implementation of the autoencoder.

TABLE I
PARAMETER VALUES EXAMINED FOR HYPER-PARAMETER SELECTION

Parameter	Values
Learning rate	{0.01, 0.001}
Hidden layers in encoder/decoder	{1, 2, 3, 4, 5}
Neurons in first hidden layer	{32, 64, 128}

Fig. 5 illustrates the capability of the fully connected autoencoder to reconstruct phase A current measurements from CT1 and CT2 during transformer physical fault.

F. Case Studies

The fully connected autoencoder has been tested for the two cyberattack scenarios defined in Section II.B.

1) Scenario 1: The autoencoder performance is examined for different changes in current measurement magnitudes. The scaling of current measurements ranging between 1.1 to 5 are considered and tested. The fully connected autoencoder was able to identify the attacks with 100% precision and 100% recall. It is noteworthy that the autoencoder becomes active when the pick up element of the differential protective relay detects a physical fault on the transformer and becomes active. Thus, the autoencoder is capable of detecting and blocking the anomalous current measurements. This mitigation strategy detects and prevents cyberattacks before causing differential protective relay misoperation and transformer false tripping but would not be able to identify the source of the cyberattack.

Fig. 6 illustrates an example of phase A current measurements reconstruction for CT1 and CT2 during cyberattacks on MU1 using a supply chain attack. As it can be seen from Fig. 6, the autoencoder reconstructs the phase A current measurements from CT1 and CT2 with high error.

2) Scenario 2: The autoencoder was able to identify the false data injection attacks with 100% precision and 100% recall. A similar mitigation strategy to what has been considered for Scenario 1 is applied here.

Fig. 7 illustrates an example of phase A current measurements reconstruction for CT1 and CT2 during false data injection cyberattacks. As it can be seen from Fig. 7, the autoencoder reconstructs the phase A current measurements from CT1 and CT2 with high error.

V. Conclusion

This paper presented a method for cybersecurity enhancement of transformer differential protective relays using machine learning. A fully connected autoencoder is trained using sliding windows of 10-ms composed of the current measurements at each side of the transformer. The sliding windows contain 48 SV single-phase current measurements *i.e.*, 144 SV three-phase current measurements at each side of the transformer. The data sets used for the autoencoder training are generated and archived using OPAL-RT HYPERSIM for different operating points of the test system under study. Afterwards, the proposed autoencoder is employed for detection and mitigation of two different cyberattack scenarios. The simulation results demonstrate the capability and high performance of the proposed machine learning algorithm for detection and mitigation of cyberattacks against transformer differential protective relays. While these results are very encouraging, further research should investigate a larger range of scenarios to better understand the range of conditions where autoencoders perform well.

References

[1] G. N. Ericsson, "Cyber security and power system communication: Essential parts of a smart grid infrastructure," *IEEE Trans. Power Del.*, vol. 25, no. 3, pp. 1501-1507, Jul. 2010.

[2] S. Ward *et al.*, "Cyber security issues for protective relays; c1 working group members of power system relaying committee," *In Proc. IEEE Power Eng. Soc. Gen. Meet.*, June 2007, pp. 1-8.

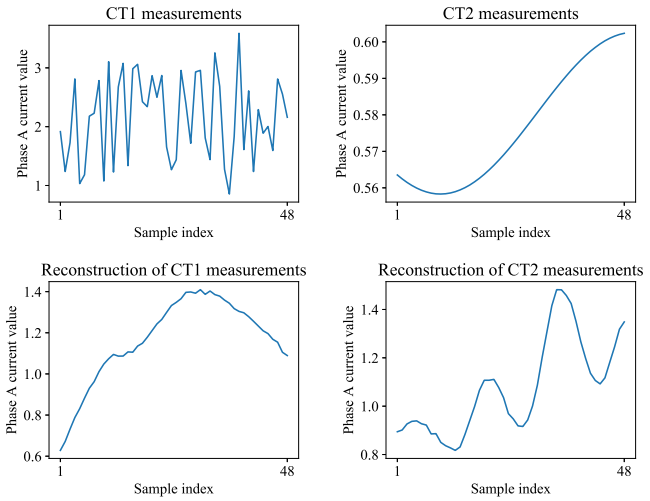

Fig. 7. Reconstruction of phase A current measurements from CT1 and CT2 after FDI cyberattacks.

[3] North American Electric Reliability Corporation (NERC) Critical Infrastructure Protection (CIP) Reliability Standards, http://www.nerc.com

[4] K. Stouffer, J. Falco, and K. Scarfone, "Guide to Industrial Control Systems (ICS) Security," NIST SP 800-82, 2015.

[5] E. Smith, S. Corzine, D. Racey, D. Patrick, H. Colin, and J. Weiss, "Going beyond cybersecurity compliance," *IEEE Power Energy Mag.*, vol. 14, no. 5, pp. 48-56, Sep. 2016.

[6] ISA99, Industrial Automation and Control Systems Security. Accessed on Aug. 2016.

[7] Electric Power Research Institute "Creating Security Metrics for the Electric Sector", Dec. 2015.

[8] Department of Energy, Roadmap to Achieve Energy Delivery Systems Cybersecurity, 2011.

[9] Department of Energy, Multiyear Plan for Energy Sector Cybersecurity, 2018.

[10] R. M. Lee, M. J. Assante, and T. Conway, "Analysis of the Cyber Attack on the Ukrainian Power Grid Defense Use Case", Electricity-Information Sharing and Analysis Center (E–ISAC), March 2016.

[11] J. Slowik, Anatomy of an attack: Detecting and Defeating CRASHOVERRIDE, Dragos Inc. White Paper, Oct. 2018.

[12] T. D. J. Blackburn, Protective Relaying: Principles and Applications, 4th ed. CRC Press, 2014.

[13] M. Kezunovic, J. Ren, and S. Lotfifard, Design, *Modeling and Evaluation of Protective Relays for Power Systems*, Springer International Publishing, 2006.

[14] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Li, "Power system risk assessment in cyber attacks considering the role of protection systems," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 572–580 March 2017.

[15] M. Bahrami, M. Fotuhi-Firuzabad and H. Farzin, "Reliability evaluation of power grids considering integrity attacks against substation protective IEDs," *IEEE Transactions on Industrial Informatics*, *early access*.

[16] A. Abiri-Jahromi, A. Kemmeugne, D. Kundur and A. Haddadi, "Cyber-physical attacks targeting communication-assisted protection schemes," *IEEE Trans. Power Syst.*, *early access* 2019.

[17] U. K. Premaratne, J. Samarabandu, and T. S. Sidhu, "An intrusion detection system for IEC61850 automated substations," *IEEE Trans. Power Del.*, vol. 25, pp. 2376–2383, Oct. 2010.

[18] J. Hong *et al.*, "Cyber attack resilient distance protection and circuit breaker control for digital substations," *IEEE Trans. Indust. Inform.*, vol. 15, no. 7, pp. 4332–4341, July 2019.

[19] A. Ameli, A. Hooshyar, and E. F. El-Saadany, "Development of a cyber-resilient line current differential relay," *IEEE Trans. Indust. Inform.*, vol. 15, no. 1, pp. 305–318, Jan. 2019.

[20] C. W. Ten, J. Hong, and C. C. Liu, "Anomaly detection for cybersecurity of the substations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 865-873, Dec. 2011.

[21] J. Hong, C. C. Liu, and M. Govindarasu, "Integrated anomaly detection for cyber security of the substations," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1643-1653, April 2014.

[22] J. Hong, and C. C. Liu,"Intelligent electronic devices with collaborative intrusion detection systems" *IEEE Trans Smart Grid*, vol. 10, no. 1, pp. 271–281, Jan. 2019.

[23] Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, and C. Wang, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35365–35381, 2018.

[24] A. Ahmed *et al.*, "Cyber Physical Security Analytics for Anomalies in Transmission Protection Systems," *IEEE Trans. Indust. App.*, *early access*.

[25] Z. Basnight, J. Butts, Jr. J. Lopez and T. Dube, Firmware modification attacks on programmable logic controllers, International Journal of Critical Infrastructure Protection, vol. 6, pp. 76–84, 2013.

[26] I. Goodfellow, Y. Bengio, and A. Courville. Deep learning. *MIT press*, 2016.

[27] V. Chandola, A. Banerjee, V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, 41.3, 2009.

[28] M. Sakurada, and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, ACM, 2014.

[29] IEEE PSRC WG D6, "Power swing and out-of-step considerations on transmission lines," Jul. 2005.

[30] H. Gras, J. Mahseredjian, E. Rutovic, U. Karaagac, A. Haddadi, O. Saad, I. Kocar, and A. El-Akoum, "A new hierarchical approach for modeling protection systems in EMT–type software," *Intern. Conf. Power Syst. Transients*, Seoul, Republic of Korea, June 2017.