A Deep Learning-Based Cyberattack Detection System for Transmission Protective Relays

Yew Meng Khaw, *Member, IEEE*, Amir Abiri Jahromi, *Member, IEEE*, Mohammadreza F. M. Arani, *Member, IEEE*, Scott Sanner, Deepa Kundur, *Fellow, IEEE*, and Marthe Kassouf

***Abstract*—The digitalization of power systems over the past decade has made the cybersecurity of substations a top priority for regulatory agencies and utilities. Proprietary communication protocols are being increasingly replaced by standardized and interoperable protocols providing utility operators with remote access and control capabilities at the expense of growing cyberattack risks. In particular, the potential of supply chain cyberattacks is on the rise in industrial control systems. In this environment, there is a pressing need for the development of cyberattack detection systems for substations and in particular protective relays, a critical component of substation operation. This article presents a deep learning-based cyberattack detection system for transmission line protective relays. The proposed cyberattack detection system is first trained with current and voltage measurements representing various types of faults on the transmission lines. The cyberattack detection system is then employed to detect current and voltage measurements that are maliciously injected by an attacker to trigger the transmission line protective relays. The proposed cyberattack detection system is evaluated under a variety of cyberattack scenarios. The results demonstrate that a universal architecture can be designed for the deep learning-based cyberattack detection systems in substations.**

***Index Terms*—Cyberphysical systems, transmission protective relays, cyberattack detection systems, deep learning, operational technology.**

Manuscript received April 29, 2020; revised September 15, 2020; accepted November 8, 2020. Date of publication November 25, 2020; date of current version April 21, 2021. This work was supported in part by the NSERC Discovery Grants Program; in part by the NSERC Strategic Partnerships for Projects Programs; and in part by the Fonds de Recherche du Québec—Nature et Technologies Postdoctoral Fellowship. Paper no. TSG-00658-2020. *(Corresponding author: Yew Meng Khaw.)*

Yew Meng Khaw and Deepa Kundur are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: ymkhaw@ece.utoronto.ca; dkundur@ece.utoronto.ca).

Amir Abiri Jahromi is with the School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, U.K. (e-mail: a.abirijahromi@leeds.ac.uk).

Mohammadreza F. M. Arani is with the Department of Electrical, Computer and Biomedical Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: marani@ryerson.ca).

Scott Sanner is with the Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: ssanner@mie.utoronto.ca).

Marthe Kassouf is with Hydro-Quebec Research Institute (IREQ), Varennes, QC J3X 1S1, Canada (e-mail: kassouf.marthe@ireq.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2020.3040361.

Digital Object Identifier 10.1109/TSG.2020.3040361

I. Introduction

CRITICAL infrastructures including electric power systems are undergoing a digital transformation and their dependence on information technology is expected to significantly increase in the coming years. The integration of information technology (IT) with operational technology (OT) in critical infrastructures improves efficiency, sustainability and consumer-centricity at the expense of increased cyberattack vulnerability [1], [2]. The high-profile cyberattacks against critical infrastructures in recent years like cyberattacks against the Ukrainian power grid illustrate the increasing exposure of these critical infrastructures to cyberattacks [3], [4]. These have promoted the detection and mitigation of cyberattacks to a top priority for governments and regulatory agencies as well as utilities [5].

Substations are at the forefront of digital transformation in electric power systems. The deployment of the IEC 61850 protocol in substations is expected to revolutionize the substation automation system by improving reliability, reducing costs and allowing interoperability between intelligent electronic devices (IEDs) while facilitating the realization of Internet of Things through remote access to substation assets and IEDs [6], [7]. Despite the unquestionable benefits of substation digitalization in automating and streamlining protection, control and asset management, it introduces complex cybersecurity concerns that need to be appropriately addressed [8]. This is mainly because the substation communication protocols are insecure as they must operate under the limited processing capability of intelligent electronic devices (IEDs) as well as various operational considerations such as speed, reliability, user-friendliness and openness [9]. Moreover, the security-by-obscurity philosophy that has traditionally been used as a defensive strategy for proprietary information and communication technologies (ICT) in substations no longer applies to emerging standards and interoperable communication protocols like IEC 61850 [10]. At the same time, the possibility of supply chain cyberattacks against industrial control systems (ICS), such as Stuxnet [11], [12], is a growing concern in the utilities and regulatory agencies.

In order to address the growing cybersecurity concerns in electric utilities, different standards and initiatives have been launched by standards organizations like the International Society of Automation (ISA) [13]–[15] and International Electrotechnical Commission (IEC) [16], research institutes

like Electric Power Research Institute (EPRI) [17] and government agencies including U.S. Department of Energy [18], [19] to develop cybersecurity measures and tools for cyber-assets in power systems. Moreover, the North American Electric Reliability Corporation (NERC) has established and enforced Critical Infrastructure Protection (CIP) standards to identify, categorize and protect cyber-assets that are essential to the reliable operation of the bulk electric system [20].

Transmission line protective relays are one of the most critical protection and control devices in substations. Coordinated cyberattacks targeting these relays have the potential to cause simultaneous tripping of multiple transmission lines and a widespread blackout [21]. As such, it is crucial to enhance the cybersecurity of transmission line protective relays. Existing research to address this problem can be classified as proposing either novel relay logic, anomaly detection or rule-based detection methods. Cyber-resilient logic designs have been proposed in [22] and [23] respectively for distance protection and line differential protective relays. A rule-based intrusion detection system has been presented in [24] for the IEC 61850 protocol. In [25], anomaly detection systems have been proposed for substation automation systems. An integrated host- and network-based anomaly detection system has been presented in [26] for substations. The semantics of sampled value (SV) and Generic Object Oriented Substation Event (GOOSE) messages have been employed in [27] to identify intrusions, anomalies, or abnormal behaviors in the IEC 61850 protocol. The aforementioned anomaly detection systems can successfully detect and mitigate some cyberattacks against IEC 61850 GOOSE and SV communication packets as well as IEDs by examining the logs of intruders' footprints. Yet, they are unable to detect new cyberattacks that continuously evolve. A cyberattack can target the payload of communication packets through a supply chain attack or a combined man-in-the-middle (MITM) and false data injection (FDI) attack that modify the sensor readings of current and voltage measurements to trigger unwanted relay action while aiming to maintain stealth. We assert that detection of such complex attacks are better addressed through advanced data analytics.

In recent years, there has been a growing focus on the application of machine learning for the detection and mitigation of cyberattacks against power systems [28]–[30]. Nevertheless, the application of machine learning for cybersecurity enhancement of protective relays has received little or no attention. Both misuse-based and anomaly-based techniques can be used for cyberattack detection. The misuse-based methods employ known signatures of cyberattacks; typically, such approaches have the advantage that they can detect such known cyberattacks with high recall rates, but demonstrate limitations in detecting previously unseen attacks. This is while anomaly-based approaches rely on learning and baselining the normal behavior of power systems. The main merit of anomaly-based techniques is their capability to detect zero-day attacks [31]. Moreover, it is possible to obtain training data for dynamic behaviors of power systems than the evolving and clandestine signatures of cyberattacks. A machine learning-based anomaly detection approach also removes the need to manually enumerate specifications and rules based on the communication protocol, as is required in specification-based detection techniques.

Support vector machine and principal component analysis have been used in [32] to detect stealthy attacks against state estimation. The compromised meters have been detected in [33] using an artificial intelligence-based method. In [34], conditional deep belief network is applied to recognize behavior patterns of FDI attacks using historical measurement data. False data injection attacks against phasor measurement units (PMU) have been detected in [35] using deep learning. A semi-supervised method has been employed in [36] for anomaly detection in an IEC 61850-based smart distribution substation. A non-nested generalized exemplar and state extraction method has been used in [37] for intrusion detection. Machine learning-based data analytics have been employed in [38] to identify the root causes of the transmission protection mal-operation such as cyberattacks. Nevertheless, the method presented in [38] has not been designed to detect or prevent cyberattacks against transmission line protection in real-time.

This article expands on the novel deep learning-based cyberattack detection system that we presented in [39] which was limited to distance protective relays and symmetrical three-phase faults. In this article, we present a novel deep learning-based cyberattack detection system for transmission line protective relays including distance protective relays, overcurrent protective relays and differential protective relays and for multiple fault scenarios. A 1-dimensional convolutional based autoencoder is used for cyberattack detection, leveraging the strength of unsupervised learning to detect previously unseen attacks. The proposed cyberattack detection system is trained with current and voltage datasets representing different types of faults occurring on the protected transmission line. The cyberattack detection system is then employed to detect current and voltage measurements that are tampered with by an attacker to trigger the transmission protective relays. The proposed cyberattack detection system is evaluated for various cyberattacks including combined MITM and FDI attack, attacks on instrument transformer tap settings and replay attack. It is demonstrated that a well-tuned deep learning-based cyberattack detection system performs well for different types of transmission protective relays which highlights the possibility of designing a universal architecture for the deep learning-based cyberattack detection systems in substations, eliminating the need for the costly and time-consuming process of tuning a model architecture for every combination of fault and relay element types.

The main contributions of this article are as follows:

- A novel deep learning-based cyberattack detection system with a universal architecture is proposed for detection and mitigation of false tripping cyberattacks against transmission line protective relays in substations.
- The performance and validity of the proposed cyberattack detection system is examined for the following:
  - Various transmission line protective relays including distance protective relays, overcurrent protective relays and differential protective relays.

- Different types of faults including three-phase-to-ground, two-phase-to-ground, single-phase-to-ground, and phase-to-phase faults.
- Different cyberattack scenarios including 1) combined MITM and FDI attack, 2) attacks on instrument transformer tap settings, and 3) replay attack.

It is worth noting that the proposed method is different from deep learning-based fault detection systems. The deep learning-based fault detection systems replace the protective relay logics for fault detection and isolation. This is while the proposed method is used in conjunction with protective relay elements to detect and mitigate false tripping cyberattacks against protective relays using operational technology data.

The remainder of this article is organized as follows. The modeling of cyberattacks against transmission line protective relays is described in Section II. Section III presents the proposed cyberattack detection system. The training, validation and testing steps of the proposed cyberattack detection system are presented in Section IV. The simulation results are provided in Section V. A brief discussion about the challenges facing the development of machine learning-based cyberattack detection systems for protective relays and directions for future research are provided in Section VI before concluding the paper in Section VII.

II. The Modeling of Cyberattacks Against Transmission Line Protective Relays

Transmission lines are normally protected by primary/main and back-up protections in power systems using the principles of distance, overcurrent and differential relaying. High-speed protection is an essential requirement for transmission lines because it preserves system stability, reduces damage to critical assets, improves power quality, and simplifies protective relay coordination. This has motivated the use of communication-assisted protection including current differential and pilot protection as the primary/main protection for transmission lines. This is while the step-distance and overcurrent protection remain as the widely used back-up protection for transmission lines.

The architecture of IEC 61850 substation automation system for transmission line protection is illustrated in Fig. 1. The merging units (MU) collect the analog measurements from the current transformers (CT) and voltage transformers (VT) and perform the analog-to-digital conversion. The MUs then transmit the measurements to the IEDs over the IEC 61850 substation LAN using SV messages. The transmission line protection logics for distance, overcurrent and differential relaying are implemented in the IEDs. The current differential and pilot protection logics receive the required information from the remote substation through the inter-substation communication network by GOOSE and SV messages.

It is worth noting that IEC 62351, a family of standards on data and communications security for power system management, was introduced to address the cybersecurity concerns associated with the IEC 61850 protocol [16]. Specifically, the

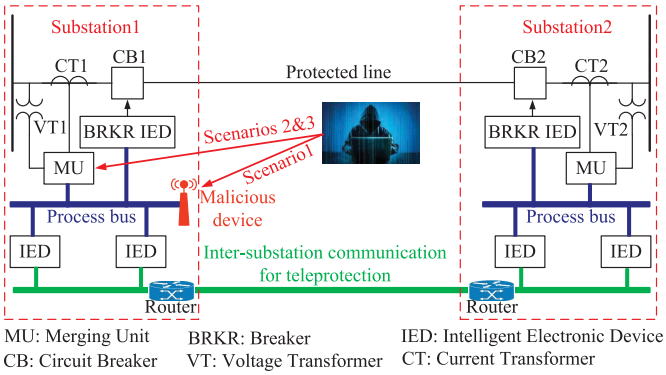

Fig. 1. Architecture of IEC61850 substation automation system for transmission line protection.

implementation of IEC 62351 will enhance the overall cybersecurity of the substation automation system by incorporating confidentiality and integrity measures like role-based access control that restricts unnecessary permissions, message level authentications and encryption mechanisms. Yet, no encryption mechanism was specified in the IEC 62351 standard for SV messages because of the time critical nature of these messages [40]. Instead, according to the IEC 62351-6 standard, the cybersecurity for information exchange of these time-critical messages relies on the supposition that SV messages are restricted to a logical substation LAN. Consequently, a breach on the substation LAN is sufficient to compromise power system applications that utilize SV messages. Moreover, the majority of the cyberattacks considered in our paper target operational technology data rather than information technology data. Authentication or other security measures proposed in the IEC 62351 standard would not prevent cyberattacks that are considered in the paper as discussed below.

The objective of the cyberattacker in this article is to cause the false tripping of transmission lines through falsifying the measurements from the instrument transformers to the transmission line protective relays. In other words, the particular type of cyberattack considered is one that aims to deceive protective relays into incorrectly assessing that a fault exists leading to unwanted breaker action. That is, no fault actually exists, but the attack induces the protection system to pick up as if there is. Hence, we aim to *distinguish* the presence of actual faults from these cyberattacks that attempt to mimic and fabricate the presence of faults that do not exist. Three scenarios are considered here to achieve this objective. The first scenario is executed through the process bus while the remaining two scenarios are executed by compromising a merging unit as illustrated in Fig. 1.

A. Attack Scenario 1

In the first scenario, we assume that a cyberattacker has remote access to the substation automation system through a malicious device which is connected to the process bus. The cyberattacker is assumed to recruit a substation employee who has authority to access communication devices in the substation to install the malicious device. The cyberattacker with

access to the process bus through the malicious device disrupts the flow of SV packets from the merging unit to the IEDs and forwards the SV packets with falsified payloads to the IEDs using a combination of MITM and FDI attack. Specifically, the attacker injects random false data with the appropriate magnitude, thus, coercing the transmission line protective relays to issue false tripping commands. When targeting an overcurrent relay, the attacker injects random current measurements with large magnitude to mimic a fault condition. Similarly, with a differential relay as a target, the attacker injects random current measurements with large magnitude while also ensuring the differential relay receives current measurements of different magnitude from both terminals of the transmission line. For the false tripping of distance relay, the attacker injects both current measurements of high magnitude and voltage measurements of low magnitude.

B. Attack Scenario 2

In the second scenario, we assume that the attacker has remote or physical access to the merging unit and modifies the settings of the CT/VT through the merging unit. The attack is assumed to be executed by an insider with access to the substation automation system or through a remote access to the process bus similar to the first scenario. The attacker could have recruited a disgruntled internal employee or may have obtained stolen or leaked legitimate operator credentials that allow remote access to the substation communication network. The tap settings of the instrument transformers allow users to change the voltage and current ratios between the primary and secondary windings of the transformers. For example, an attacker can change the tap settings of a current transformer such that a larger current is observed downstream to the current transformer. The attacker can also tamper with the tap settings of a voltage transformer such that the protective relays receive voltage measurements of lower magnitude, mimicking the voltage behavior in a fault condition.

C. Attack Scenario 3

In the third scenario, we assume that a malware installed on the merging unit is used to perform a replay attack by replacing measurements from the CT/VT with previously recorded fault measurements to cause false tripping of the transmission line protective relays. The malware can be installed on the merging unit through a supply chain attack or a threat agent with physical or remote access to the substation automation system. The malware can then eavesdrop and disrupt the information exchange between the instrument transformer and merging unit as well as between the merging unit and IED. This allows the attacker to record current and voltage measurements during fault scenario, which can be injected at a later time as a replay attack.

III. The Proposed Cyberattack Detection System

The objective of the proposed cyberattack detection system is to detect patterns in the measurements from instrument transformers, *i.e.*, CTs and VTs, that do not conform to the normal behavior of measurements. Note that the notion of normal behavior of measurements in this article includes both power system fault-free dynamics and dynamics during power system faults. One distinction of the proposed approach is that patterns in OT data are harnessed for the purpose of anomaly detection. Hence, in contrast to typical IT intrusion detection approaches that make use of communication packet semantics or logs of intruder footprints, we make use of data closer to the physical impacts of the attacks. Hence, time-series current and voltage measurements at the process bus level of substations are the inputs employed for data analytics.

A. Configuration of the Proposed Cyberattack Detection System in Substations

Anomaly detection systems using machine learning approaches have received considerable attention in recent years in various application domains including cybersecurity [31], [41], [42]. Several factors such as the nature of the input data, the availability of the labeled datasets as well as the constraints and requirements induced by the application domain determine the choice of the machine learning approaches for anomaly detection. As stated above, time-series current and voltage measurements at the process bus level of substations are the inputs employed for in the cyberattack detection system (CDS). Moreover, traditional IEDs and automation devices in substations are resource constrained devices with just enough memory and computational power to perform their tasks. This prevents the implementation of the power and resource demanding cyberattack detection systems that use machine learning-based methods within the IEDs and automation systems in substations. Yet, IEDs and automation devices with more powerful processors may emerge in the coming years with the ability to implement machine learning-based methods in order to respond to the growing need of power utilities to leverage machine learning techniques in their system operations. Finally, the evolving and clandestine nature of cyberattacks as well as their rarity against protective relays limit the possibility of obtaining and effectively modeling these anomalous behavior in contrast to normal behavior in substations for which there is significantly more data and more predictable characteristics. In this environment, semi-supervised and unsupervised machine learning approaches are in a superior position for cyberattack detection in contrast to supervised machine learning approaches.

Considering the aforementioned factors, we propose a centralized deep learning-based CDS for transmission line protective relays performed by additional physical devices with sufficient computational power separate from the IEDs as illustrated in Fig. 2. The cyberattack detection system is external to the IEDs and MUs within the substation and is connected to them via the process bus and inter-substation communication network. The proposed cyberattack detection system functions in two steps: 1) the offline training, validation and testing step and 2) the real-time operational step. In the offline training, validation and testing step, the proposed model learns the normal behavior of the current and voltage measurements during transmission line faults. The cyberattack

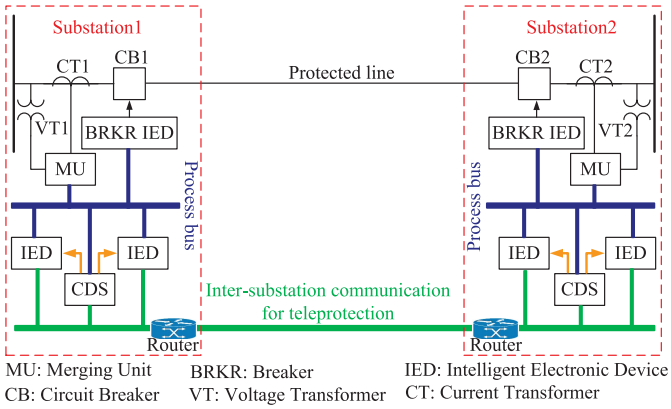

Fig. 2. The configuration of the proposed cyberattack detection system (CDS) in a substation.

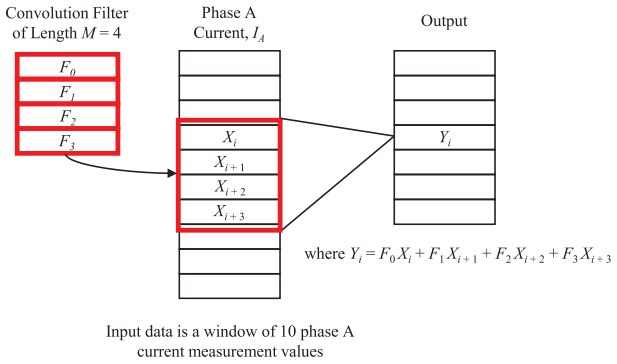

Fig. 3. A 1-dimensional convolution operation on 10 measurement samples with a convolution filter F of length $M = 4$.

detection system will go live within the substation when calibration through the offline training, validation and testing step is finalized. In the real-time operational step, the cyberattack detection system identifies anomalous measurements that do not conform to the normal behavior of measurements. The cyberattack detection system has two modes of operation in real-time: 1) cyberattack detection mode, and 2) cyberattack detection and mitigation mode. In the detection mode, the cyberattack detection system only generates an alarm after detecting anomalous measurements and does not intervene with the functionality of the protective relays in the IED. This is while, in the detection and mitigation mode, the cyberattack detection system sends commands to the IEDs to block the anomalous measurements in order to avoid transmission line false tripping.

It is worth noting that the proposed cyberattack detection system classifies any anomalous measurement as a cyberattack. This means that in detection mode, an alarm is generated, or in detection and mitigation mode, commands are sent to the IEDs to block anomalous measurements. In typical anomaly detection frameworks, the type of anomaly is not distinguished because they are not explicitly modeled. There are advantages to this treatment because new cyberattacks previously unknown can be accounted for as long as they involve anomalous measurements. If the source of anomalous measurements is to be distinguished, offline post forensic analysis like the one proposed in [38] is required.

B. A Deep Learning Autoencoder-Based Cyberattack Detection System

We now outline an unsupervised deep learning approach to anomaly detection using an autoencoder. Such an approach allows for the detection of zero-day attacks and removes the need to manually enumerate specifications and rules based on a specific communication protocol and cyberattack type. A deep learning approach also allows us to leverage the availability of a large volume of high-fidelity data that can be obtained for model training. The autoencoder consists of two parts; encoder and decoder. The encoder, f, compresses the input data, x, to a latent space, z, with dimensions typically smaller than the input data. The decoder, g, reconstructs an estimate of the input data from the latent space z. As the autoencoder is trained to be an identity system, the latent space z of smaller dimensionality must necessarily capture the most salient features of the input. Since the latent space is particular to the type of training data, inputs deviating from the training dataset will result in high reconstruction errors and flagged as anomalous data. The reconstruction error is computed from the mean squared error (MSE) between the reconstructed output and the input data to the autoencoder. Different types of models can be used in the autoencoder such as a fully-connected network, recurrent neural network and convolutional neural network. The reader should note that autoencoders are generally considered unsupervised methods because although labels of the normal training data are known, they are not explicitly incorporated during the training process as the original voltage and current input itself is also employed in the role of the labels. The objective of the autoencoder is to build a model of the normal data with the reasoning that data which is abnormal cannot be properly reproduced (i.e., autoencoded) by an autoencoder trained on only the normal data. As the autoencoder itself does not explicitly predict normal or abnormal labels, the autoencoder is generally considered unsupervised.

In this article, a 1-dimensional convolutional based autoencoder is used for the cyberattack detection system. Here, both the encoder and decoder make use of 1-dimensional convolution stages that consist of sliding a filter kernel over the data set and applying a dot product. The output of the convolution operation is given in (1).

$$\begin{aligned} Y(i) &= (X * F)(i) \\ &= \sum_{m=0}^{M-1} F_m X_{i+m} \end{aligned} \tag{1}$$

where Y denotes the output of the convolution operation, X denotes the 1-dimensional data input, F denotes the convolution filter of length M, '$*$' denotes the convolutional operator and i denotes the input data index. The convolution operation is illustrated in Fig. 3. The CNN model allows for parameter sharing in F which reduces the total number of trainable parameters, resulting in computational savings during model training with less memory requirements and higher statistical efficiency [42].

For the encoder section of the cyberattack detection system that embeds the input into a low-dimensional latent space

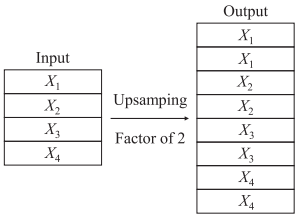

Fig. 4. A 1-dimensional upsampling with a factor of 2 applied to an input.

where similar inputs should embed near each other, we use a neural network consisting of interleaved layers of convolutional operations followed by a nonlinear activation and max pooling [42]. An example of a common nonlinear activation function is the rectified linear unit (ReLU) which is a piecewise linear function as defined in (2).

$$\sigma(x) = \max(x, 0) \tag{2}$$

The pooling layer has the effect of reducing the input dimension by downsampling the input. Common approaches include average pooling and max pooling layers which slide a small window at a given stride, taking the average and maximum value respectively within the window to produce a downsized dataset. In the decoder section of the autoencoder, the data in the latent space is expanded back to the original input dimensions. Our decoder consists of convolution operations interleaved with 1-dimensional upsampling layers [43]. We provide an example of a 1-dimensional upsampling operation with an upsampling factor of 2 in Fig. 4.

IV. Training, Validation and Testing of the Cyberattack Detection System

Different types of faults including three-phase-to-ground faults, two-phase-to-ground faults, single-phase-to-ground faults and phase-to-phase faults may occur on transmission lines. Naturally, the signatures of each of these faults are distinct. Moreover, the types of input data used by different protective relays such as distance, overcurrent and differential protective relays are different. For instance, distance relays make use of both current and voltage measurements while overcurrent and differential relays rely solely on current measurements.

The differences between the types of faults and inputs to the protective relays render it impossible to train a single deep learning model for all types of faults and protective relays. Yet, we posit that a universal architecture can be designed for the deep-learning model in the cyberattack detection system. The implementation of a universal architecture in the cyberattack detection system eliminates the cumbersome need for optimizing the architecture for every variety and combination of faults and protective relays. Hence, in this article, we consider a universal architecture for the deep-learning model and train it for each type of fault and protective relay separately. This approach results in a deep-learning model with the universal architecture, with different model weights for each combination of faults and protective relays. Each of the deep learning models becomes active by the activation of the corresponding protective relay element and remains inactive for the activation of all other protective relay elements. Note that the training, validation and testing steps are conducted offline. Therefore, the computational complexity and execution times are not limiting factors.

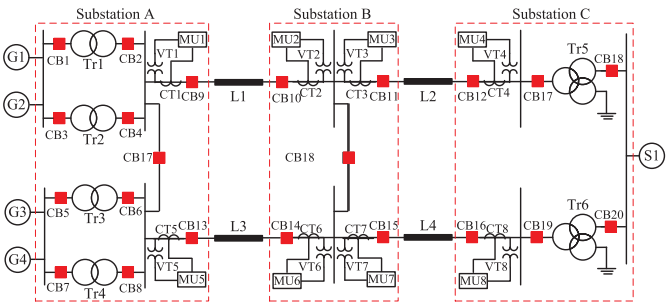

Fig. 5. The IEEE PSRC D6 benchmark test system.

A. Transmission Test System

Fig. 5 illustrates the IEEE power system relaying committee (PSRC) D6 benchmark test system [44]. The test system connects a power plant with four 250 MVA generator units to a 230 kV transmission network through two parallel 500 kV transmission lines. The test system is comprised of three substations. Substation A connects the power plant to the 500 kV transmission lines. Substation B is a switching substation and is located 280 km from substation A. Substation C is located 220 km from Substation B and models the connection to a 230 kV transmission system that is modeled as an infinite bus. The transmission lines are protected by the principles of distance, overcurrent and differential protection.

B. Training Dataset

The transmission test system in Fig. 5 is simulated in OPAL-RT HYPERSIM to generate training datasets. The simulations are performed for a duration of 200 milliseconds with the fault initiating randomly between t=100 ms to t=120 ms. The starting time of the fault is varied between t=100 ms to t=120 ms in the simulations to ensure fault occurs at different parts of the current and voltage waveforms. Note that the period of one cycle is approximately 16.7 ms in a 60 Hz power system. Moreover, the generation levels and fault locations on the transmission line L1 are changed in each simulation to generate datasets under different operating conditions and fault location scenarios. The generation levels of G1-G2 and G3-G4 are varied in unison between 300 MW to 400 MW with a step size of 10 MW. The fault location is changed along the transmission line L1 with a step size of 10 km. The simulations are performed for three-phase-to-ground, two-phase-to-ground, single-phase-to-ground, and phase-to-phase faults. The fault impedance is assumed to be zero. In total, 50,820 simulations are performed to generate training datasets for each type of fault.

The measurements are collected for all three-phases. The current measurements are collected from CT1 and CT2 in Fig. 5 and the voltage measurements are collected from VT1. The measurements are collected at the sampling rate of 4800

samples per second to comply with IEC 61850-9-2 standard for SV packet specifications [45]. As such, each simulation run contains 960 samples per measurement per phase.

C. Training and Optimization of the Autoencoder Architecture

The 1-dimensional convolutional based autoencoder described in Section III-B is trained with three-phase measurements corresponding to the inputs of the associated protective relay. The autoencoder associated with the overcurrent protective relay is trained with three-phase current measurements from CT1. The autoencoder associated with the distance relay is trained with three-phase current and voltage measurements from CT1 and VT1 and the autoencoder associated with the differential relay is trained with three-phase current measurements from CT1 and CT2.

The autoencoder is trained with 70% of the 50,820 simulations. The validation and test datasets each comprises 15% of the 50,820 simulations. An important parameter for autoencoder training is the input data length, *i.e.*, the number of input samples fed to the autoencoder. In this article, a sliding window of 50 ms, *i.e.*, 240 samples of current/voltage measurements for each phase, is fed to the autoencoder as input. As such, each window consists of 3 cycles of measurements. Thus, the 200 ms simulation data is split into sliding windows of 50 ms data. As the sliding window slides over the entire simulation sample, the autoencoder is trained. Note that data standardization is performed before the data is fed to the autoencoder. Consider a training dataset D that contains the measurement points, $x_1, x_2, \ldots, x_N$. In data standardization, the data for each measurement type is scaled to unit variance and zero mean as given in (3)-(5) where the mean, μ, and standard deviation, σ, is calculated across the entire training dataset.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{n} \tag{3}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{n}} \tag{4}$$

$$x_{stand} = \frac{x - \mu}{\sigma} \tag{5}$$

The cyberattack detection system is further trained to reduce the loss function, which is the MSE between the input and the reconstructed output of the autoencoder as given in (6).

$$L = \|g(f(x)) - x\|_2^2 \tag{6}$$

ADAM, a state-of-the-art stochastic gradient-based optimization algorithm [46], is used for model training to minimize the loss function. ADAM employs an adaptive learning rate and momentum via a moving average of the gradients and squared gradients, for faster convergence over a straightforward gradient descent algorithm.

A universal architecture is used in this article for the cyberattack detection system as discussed in Section IV. The architecture is optimized via grid search for different number of layers, number of convolution filters and pooling size. The architecture is optimized in this article to obtain the highest recall rate for the replay attack scenario instead of the lowest loss value. This is because the lowest loss value does not necessarily result in the best cyberattack detection performance. The final architecture is chosen based on the highest recall rate observed in the validation dataset. Using labelled replay attack measurement samples during the validation step allows the selection of a better tuned model architecture at the expense of slightly biasing the performance of the cyberattack detection system towards the replay attack scenario. Nevertheless, the cyberattack detection system is observed to perform well in all other attack scenarios considered in this article.

TABLE I
AUTOENCODER ARCHITECTURE

Encoder			Decoder		
1.	Convolution	32 filters	13.	Convolution	256 filters
2.	Convolution	32 filters	14.	Convolution	256 filters
3.	Max Pooling	Pool Size 2	15.	Upsampling	Factor 6
4.	Convolution	64 filters	16.	Convolution	128 filters
5.	Convolution	64 filters	17.	Convolution	128 filters
6.	Max Pooling	Pool Size 4	18.	Upsampling	Factor 5
7.	Convolution	128 filters	19.	Convolution	64 filters
8.	Convolution	128 filters	20.	Convolution	64 filters
9.	Max Pooling	Pool Size 5	21.	Upsampling	Factor 4
10.	Convolution	256 filters	22.	Convolution	32 filters
11.	Convolution	256 filters	23.	Convolution	32 filters
12.	Max Pooling	Pool Size 6	24.	Upsampling	Factor 2
			25.	Convolution	3 filters
			26.	Convolution	3 filters

The final architecture is summarized in Table I and illustrated in Fig. 6. In all convolutional layers, we used a convolution filter size of 10, convolutional stride length of 1 and used ReLU as the activation function. When choosing the final model weights, we chose the weights at the epoch that results in the highest recall rate within 100 epochs. For example, in the three-phase-to-ground fault scenario, we used 80 epochs for the overcurrent relay, 70 epochs for the distance relay and 60 epochs for the differential relay. This is commonly known as early stopping. Again, this choice of the final model weights is done based on the validation dataset. The deep learning model is implemented with Keras with a Tensorflow backend [47].

V. SIMULATION RESULTS

In this section, we examine the performance of the proposed deep learning-based cyberattack detection system. Three cyberattack scenarios including 1) combined MITM and FDI attack, 2) attack on instrument transformer tap settings, and 3) replay attack are considered. In each scenario, we investigate the performance of the cyberattack detection system for different types of faults and different protective relay principles.

Anomalous or attack data are data that deviates from normal behavior as recognized by the cyberattack detection system during model training. These attack cases represent rare occurrences resulting in an imbalanced dataset with very small number of positive cases. Using accuracy as our performance metric is therefore inapt. Consider a dataset with 1000 measurement samples with only 1 attack sample. A naive cyberattack detection system that always classifies an input as negative or normal will achieve an accuracy of 99.9%. As

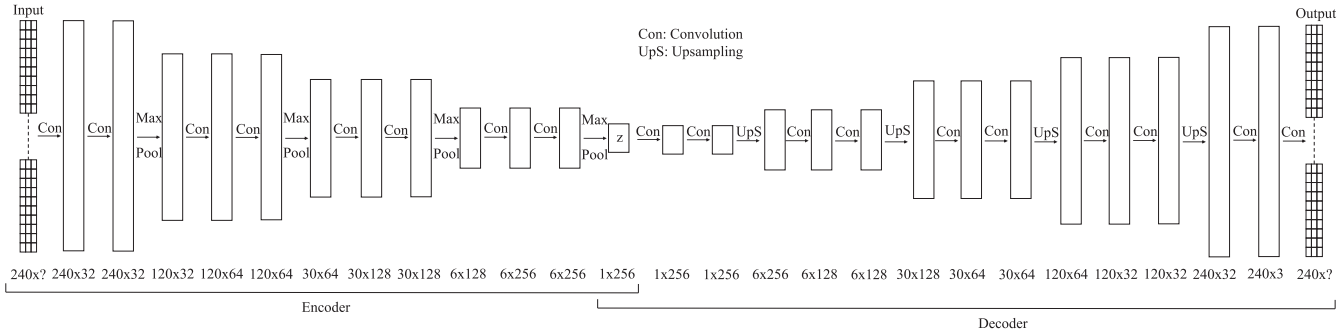

Fig. 6. The architecture of the proposed cyberattack detection system where the input/output dimensions depend on the type of the protective relay.

such, the precision and recall metrics are employed to measure the performance of the proposed cyberattack detection system.

$$precision = \frac{\#\ True\ Positive}{\#\ True\ Positive + \#\ False\ Positive} \tag{7}$$

$$recall = \frac{\#\ True\ Positive}{\#\ True\ Positive + \#\ False\ Negative} \tag{8}$$

True Positive represents cyberattacks that are correctly detected by the cyberattack detection system. False Positive represents measurements with normal behavior that are incorrectly classified as a cyberattack. False Negative represents cyberattacks that are not detected by the cyberattack detection system. True Negative represents measurements with normal behavior that are correctly classified as legitimate measurements. # represents the count of each event. Therefore, precision is the fraction of attack classifications made by the cyberattack detection model that is correct. Recall is the fraction of actual attacks that are "recalled", *i.e.*, correctly classified as attacks by the cyberattack detection system.

As discussed in Section III-B, the deep learning-based cyberattack detection system is capable of reconstructing measurements with low reconstruction error when applied to data exhibiting normal characteristics. This is while reconstruction error is high for anomalous measurements that deviate from the training data. Hence, a threshold for the reconstruction error can be set for cyberattack detection. The threshold for cyberattack detection is set at 1.5 times of the maximum MSE between the input and the reconstructed output observed with the training dataset. This conservatively high threshold ensures low false positive rates.

TABLE II
PERFORMANCE OF THE CDS: RANDOM FDI ATTACK

Fault Scenario	Relay Type	Precision	Recall
Single-Phase-to-Ground (A-G)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Two-Phase-to-Ground (A-B-G)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Three-Phase-to-Ground (A-B-C-G)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Phase-to-Phase (A-B)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Three-Phase (A-B-C)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%

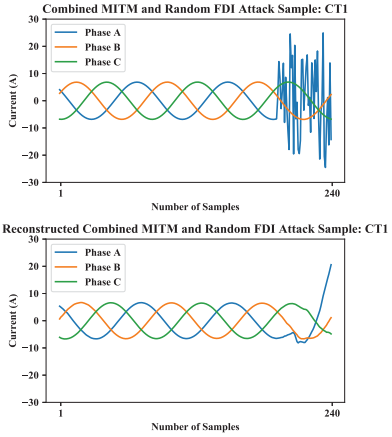

Fig. 7. Reconstruction of the measurements during a combined MITM and random FDI attack on the overcurrent relay.

A. Combined Man-In-The-Middle and Random False Data Injection Attack

In this scenario, we assume that a cyberattacker has remote access to the substation automation system through a malicious device which is connected to the process bus. We further assume that the cyberattacker understands the principles of transmission line protective relays but does not have knowledge about the dynamics of the transmission network under attack. Thus, the cyberattacker injects random measurements to the process bus to trigger the transmission line protective relays. In the case of the overcurrent relay, the cyberattacker injects current measurements with large magnitudes to the process bus. In the case of the distance relay, the cyberattacker injects current and voltage measurements with high and low magnitudes respectively to the process bus to represent a fault. In the case of the differential relay, the cyberattacker injects different current measurements with high magnitudes to the process bus. The performance of the cyberattack detection system considering different types of faults and protective relay principles for the combined MITM and FDI attack is summarized in Table II. A sample of measurements during a combined MITM and FDI attack on the overcurrent relay is illustrated in Fig. 7. As illustrated in Fig. 7, the autoencoder reconstructs the injected false data with high error.

TABLE III
PERFORMANCE OF THE CDS: ATTACKS AGAINST INSTRUMENT TRANSFORMER TAP SETTINGS

Fault Scenario	Relay Type	Precision	Recall
Single-Phase-to-Ground (A-G)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Two-Phase-to-Ground (A-B-G)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Three-Phase-to-Ground (A-B-C-G)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Phase-to-Phase (A-B)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%
Three-Phase (A-B-C)	Overcurrent	100%	100%
	Differential	100%	100%
	Distance	100%	100%

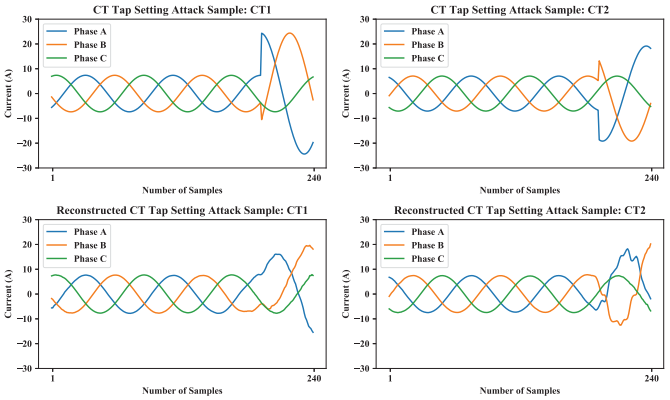

Fig. 8. Reconstruction of the measurements during an attack against the instrument transformer tap settings to trigger the differential protective relay.

TABLE IV
PERFORMANCE OF THE CDS: REPLAY ATTACK

Fault Scenario	Relay Type	Precision	Recall
Single-Phase-to-Ground (A-G)	Overcurrent	100%	95.2%
	Differential	100%	91.0%
	Distance	100%	96.1%
Two-Phase-to-Ground (A-B-G)	Overcurrent	100%	92.4%
	Differential	100%	82.1%
	Distance	100%	95.4%
Three-Phase-to-Ground (A-B-C-G)	Overcurrent	100%	91.7%
	Differential	100%	88.8%
	Distance	100%	94.5%
Phase-to-Phase (A-B)	Overcurrent	100%	93.3%
	Differential	100%	88.5%
	Distance	100%	87.4%
Three-Phase (A-B-C)	Overcurrent	100%	93.8%
	Differential	100%	88.4%
	Distance	100%	95.7%

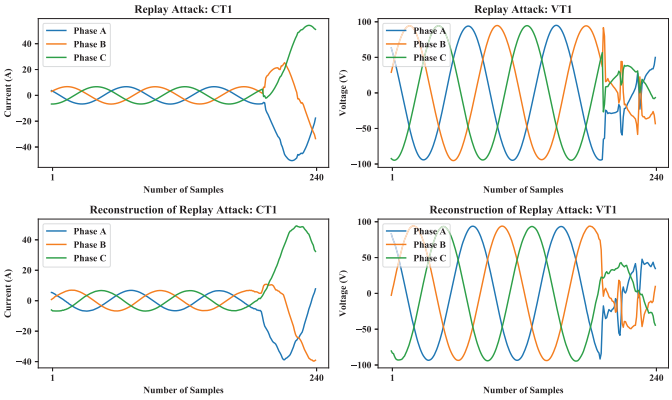

Fig. 9. Reconstruction of the measurements during a replay attack on the distance relay.

B. Tampering of Instrument Transformer Tap Settings

In the second scenario, we assume that the attacker has remote or physical access to the merging unit and modifies the settings of the CT/VT through the merging unit. In the case of overcurrent relay, the attacker changes the tap setting of the current transformer CT1 such that a large current magnitude is seen by the overcurrent relay. In the case of the distance relay, the attacker changes the tap settings of current transformer CT1 and voltage transformer VT1 such that it triggers the distance relay. In the case of differential relay, the cyberattacker changes the tap settings of the current transformers CT1 and CT2 to trigger the differential relay. The performance of the cyberattack detection system considering different types of faults and protective relay principles for the attacks against the instrument transformer tap settings is summarized in Table III. A sample of measurements during a cyberattack on the instrument transformer tap settings to trigger a differential protective relay is illustrated in Fig. 8. As illustrated in Fig. 8, the autoencoder poorly reconstructs the measurements resulting in successful detection of the cyberattack due to significant deviation of the attack data from normal behavior.

C. Replay Attack

In the third scenario, we assume that a malware inside the merging unit performs a replay attack by replacing the measurements from CT/VT with previously recorded measurements to cause false trippings of the protective relays. We considered various scenarios ranging from unsynchronized to fully synchronized injection of the actual fault measurements to cause false line tripping. Note that the replay attack assumes a very strong capability on the part of the attacker. The performance of the cyberattack detection system considering different types of faults and protective relay principles for the replay attack is summarized in Table IV. A sample of the measurements during a replay attack against the distance relay is illustrated in Fig. 9. As illustrated in Fig. 9, the autoencoder poorly reconstructs the measurements resulting in successful detection of the cyberattack. It should be noted that in the case of a fully-synchronized replay attack, the proposed cyberattack detection system was not able to detect the attacks. In such attack scenarios, the measurements received is essentially the same as measurements received in real-fault conditions.

D. Computational Complexity of the Proposed Cyberattack Detection System

The proposed cyberattack detection system was able to detect the cyberattacks approximately 25 ms after the starting point of the cyberattack, *i.e.*, after receiving 120 samples of falsified current/voltage measurements. Moreover, it takes the autoencoder slightly under 4 ms to reconstruct the measurements using i7-9700K CPU with RTX2080 GPU. This

sums up to a minimum real-time delay of 29 ms in processing the data, slightly less than 2 cycles. It is worth noting that the operating time of the cyberattack detection system may become larger than the operating time of commercial protective relays [48]. Thus, further investigation is needed to ensure that the sensitivity of protective relays will not be compromised by the proposed cyberattack detection system.

As discussed previously in Section IV-C, a sliding window of 50 ms (or 3 cycles) equivalent to 240 samples of current/voltage measurements for each phase, is consecutively fed in real-time to the autoencoder as input for possible cyberattack detection after the system becomes active. Moreover, the proposed cyberattack detection system needs less than two cycles of processing time to distinguish a cyberattack from a legitimate fault. Therefore, a buffer with a capacity to capture 5 cycles (400 samples per phase) of current and voltage measurements, which is practical and reasonable, would be sufficient to enable real-time operation.

VI. Discussion

The development of deep learning-based cyberattack detection systems for improving the cybersecurity of protective relays in substations is at its embryonic stage. Despite the promising results obtained in this article, several open challenges should be addressed before it can be applied to real systems. The first is related to the scarcity of fault data in substations that is required to train the cyberattack detection system. One can overcome this challenge by developing dynamical models that represent the real substations in time-domain simulators and validating the dynamical models with data from the fault recorders in substations. We emphasize that it is impossible to develop machine learning-based cyberattack detection systems for protective relays without access to high fidelity training datasets. Thus, the development of accurate time-domain dynamical models of substations is the essential first step for the advancement of machine learning-based cyberattack detection systems.

There are numerous scenarios and practical considerations that should be taken into account before implementing the proposed model in practice. For instance, one needs to investigate the impact of scenarios such as current transformer saturation, the existence of short lines, in-feeds, out-feeds, different fault impedances, varying penetration of distributed energy resources to see how these scenarios would impact the performance of the proposed model. Moreover, several considerations such as different transmission network topologies, communication packet loss and noise should be taken into account. For digital substations, an IEC 61850-9-2 merging unit publishes 80 SV packets per cycle in a 60 Hz power system, which means that measurements are transmitted every 208.3 microseconds. Protective relays are designed such that they transition to an offline status if more than two consecutive SV packets are missed [49]. This means packet loss will not be the limiting factor as the protective relays will no longer be in service under these conditions. Moreover, digital communication systems are designed such that noise is minimized with various filters commonly employed in industrial devices like IEDs and MUs to further reduce any effect of noise on the data [50]. Hence, the impact of noise and packet loss are neglected in this article.

Another interesting research direction is to examine the cybersecurity of the proposed cyberattack detection system. While the proposed cyberattack detection system addresses the cybersecurity vulnerabilities of protective relays, its addition may present an additional attack surface which needs further investigation. An attacker may target the proposed cyberattack detection system through the process bus to perform attacks against protective relays and IEDs. Yet, cyberattackers with access to the process bus can directly target the protective relays without the need to compromise the proposed cyberattack detection system. Moreover, authentication can be used for the output signals of the proposed cyberattack detection system to improve cybersecurity [51]. This is while it is impossible to use authentication for SV packets considering the large number of SV packets that should be processed by an IED in each cycle, *i.e.*, 80 packets per cycle in a 60 Hz system.

Last but not least, it is important to highlight that there is no one-size-fits-all solution to the cybersecurity challenges of industrial control systems like substation protection and control. The cybersecurity challenges in these systems can only be overcome by considering a holistic approach and implementing layered protective measures and defence-in-depth models.

VII. Conclusion

This article presented a deep-learning based cyberattack detection system for transmission line protective relays. The proposed cyberattack detection system is trained with measurements representing different types of faults. Moreover, the cyberattack detection system is trained with different sets of inputs depending on the principle of the protective relay under study such as distance, overcurrent or differential protective relays. The simulation results verified the capability of the proposed cyberattack detection system in identifying different types of cyberattacks including 1) combined MITM and FDI attack, 2) tampering of instrument transformer tap settings, and 3) replay attack. The simulation results further highlighted that a universal architecture can be designed for the deep-learning model in the cyberattack detection system. The implementation of such a universal architecture eliminates the cumbersome need for optimizing the architecture for each type of fault and protective relay and significantly facilitates the development of the cyberattack detection system for the protective relays in substations. The challenges facing the development of machine learning-based cyberattack detection systems for protective relays and directions for future research have been further discussed.

References

[1] D. R. Harp and B. Gregory-Brown, "IT/OT convergence: Bridging the divide," Atlanta, GA, USA, NexDefense, White Paper, 2016.

[2] E. D. Knapp and R. Samani, *Applied Cyber Security and the Smart Grid: Implementing Security Controls Into the Modern Power Infrastructure*. Amsterdam, The Netherlands: Syngress, 2013.

[3] R. M. Lee, M. J. Assante, and T. Conway, *Analysis of the Cyber Attack on the Ukrainian Power Grid Defense Use Case*, Elect. Inf. Sharing Anal. Center, Washington, DC, USA, Mar. 2016.

[4] J. Slowik, "Anatomy of an attack: Detecting and defeating CRASHOVERRIDE," Hanover, MD, USA, Dragos Inc., White Paper, Oct. 2018.

[5] "Cyber threat and vulnerability analysis of the U.S. electric sector," Idaho Nat. Lab., Idaho Falls, ID, USA, Rep. INL/EXT-16-40692, Jun. 2017.

[6] K. P. Brand, V. Lohmann, and W. Wimmer, *Substation Automation Handbook*. Bremgarten, Switzerland: Utility Autom. Consulting Lohmann, 2003.

[7] "Communication networks and systems for power utility automation—Part 90-4: Network engineering guidelines," Int. Electrotechn. Commission, Geneva, Switzerland, Rep. TR 61850–90–4:2013, Aug. 2013. [Online]. Available: http://webstore.iec.ch/

[8] *Cyber Security Requirements for Substation Automation, Protection and Control Systems*, IEEE Standard C37.240-2014, 2014.

[9] P. Ackerman, *Industrial Cybersecurity: Efficiently Secure Critical Infrastructure Systems*. Birmingham, U.K.: Packt, 2017.

[10] 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.*, Jun. 2007, pp. 1–8.

[11] Z. Basnight, J. Butts, Jr., J. Lopez, and T. Dube, "Firmware modification attacks on programmable logic controllers," *Int. J. Critical Infrastruct. Protect.*, vol. 6, no. 2, pp. 76–84, 2013.

[12] R. Langner, "Stuxnet: Dissecting a cyberwarfare weapon," *IEEE Security Privacy*, vol. 9, no. 3, pp. 49–51, May/Jun. 2011.

[13] K. Stouffer, J. Falco, and K. Scarfone, *Guide to Industrial Control Systems (ICS) Security*, document SP 800-82, NIST, Gaithersburg, MD, USA, 2015.

[14] E. Smith, S. Corzine, D. Racey, D. Patrick, H. Colin, and J. Weiss, "Going beyond cybersecurity compliance: What power and utility companies really need to consider," *IEEE Power Energy Mag.*, vol. 14, no. 5, pp. 48–56, Sep./Oct. 2016.

[15] *ISA99, Industrial Automation and Control Systems Security*. Accessed: Aug. 2016. [Online]. Available: http://isa99.isa.org/ISA99%5CWiki/Home.aspx

[16] International Electrotechnical Commission, *Power Systems Management and Associated Information Exchange—Data and Communications Security. Part 1: Communication Network and System Security—Introduction to Security Issues*, IEC Standard TS 62351-1, 2007.

[17] *Creating Security Metrics for the Electric Sector*, Elect. Power Res. Inst., Washington, DC, USA, Dec. 2015.

[18] *Thr Energy Sector Control Systems Working Group (ESCSWG), Roadmap to Achieve Energy Delivery Systems Cybersecurity*, U.S. Dept. Energy, Washington, DC, USA, 2011.

[19] *U.S. Department of Energy Office of Electricity Delivery & Energy Reliability, Multiyear Plan for Energy Sector Cybersecurity*, U.S. Dept. Energy, Washington, DC, USA, 2018.

[20] *Critical Infrastructure Protection (CIP) Reliability Standards*, North Amer. Elect. Rel. Corporat., Atlanta, GA, USA, 2017. [Online]. Available: http://www.nerc.com

[21] A. Abiri-Jahromi, A. Kemmeugne, D. Kundur, and A. Haddadi, "Cyber-physical attacks targeting communication-assisted protection schemes," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 440–450, Jan. 2020.

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

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

[24] U. K. Premaratne *et al.*, "An intrusion detection system for IEC61850 automated substations," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2376–2383, Oct. 2010.

[25] 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.

[26] 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, Apr. 2014.

[27] 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.

[28] H. Karimipour, A. Dehghantanha, R. M. Parizi, K. R. Choo, and H. Leung, "A deep and scalable unsupervised machine learning system for cyber-attack detection in large-scale smart grids" *IEEE Access*, vol. 7, pp. 80778–80788, 2019.

[29] S. Ahmed, Y. Lee, S.-H. Hyun, and I. Koo, "Feature selection–based detection of covert cyber deception assaults in smart grid communications networks using machine learning," *IEEE Access*, vol. 6, pp. 27518–27529, 2018.

[30] Y. Wang, M. M. Amin, J. Fu, and H. B. Moussa, "A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids," *IEEE Access*, vol. 5, pp. 26022–26033, 2017.

[31] Y. Xin *et al.*, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35365–35381, 2018.

[32] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1644–1652, Sep. 2017.

[33] K. Khanna, B. K. Panigrahi, and A. Joshi, "AI-based approach to identify compromised meters in data integrity attacks on smart grid," *IET Gen. Trans. Distrib.*, vol. 12, no. 5, pp. 1052–1066, 2018.

[34] Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2505–2516, Sep. 2017.

[35] J. Wei and G. J. Mendis, "A deep learning-based cyber-physical strategy to mitigate false data injection attack in smart grids," in *Proc. Joint Workshop Cyber Phys. Security Resilience Smart Grids (CPSR-SG)*, Vienna, Austria, 2016, pp. 1–6.

[36] A. Valdes, R. Macwan, and M. Backes, "Anomaly detection in electrical substation circuits via unsupervised machine learning," in *Proc. IEEE 17th Int. Conf. Inf. Reuse Integr. (IRI)*, Pittsburgh, PA, USA, 2016, pp. 500–505.

[37] U. Adhikari, T. H. Morris, and S. Pan, "Applying non-nested generalized exemplars classification for cyber-power event and intrusion detection," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 3928–3941, Sep. 2018.

[38] A. Ahmed *et al.*, "Cyber physical security analytics for anomalies in transmission protection systems," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 6313–6323, Nov. 2019.

[39] Y. M. Khaw, A. A. Jahromi, M. F. M. Arani, D. Kundur, S. Sanner, and M. Kassouf, "Preventing false tripping cyberattacks against distance relays: A deep learning approach," in *Proc. IEEE Inter. Conf. Comm. Control Comput. Technol. Smart Grids (SmartGridComm)*, Beijing, China, 2019, pp. 1–6.

[40] *Power Systems Management and Associated Information Exchange—Data and Communications Security. Part 6: Security for IEC 61850*, IEC Standard TS 62351-6:2007, 2007.

[41] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surveys*, vol. 41, no. 3, pp. 1–58, Jul. 2009.

[42] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[43] *Keras Documentation Convolutional Layers*. Accessed: Apr. 2020. [Online]. Available: https://keras.io/layers/convolutional/

[44] H. Gras *et al.*, "A new hierarchical approach for modeling protection systems in EMT–type software," in *Proc. Intern. Conf. Power Syst. Transients*, Seoul, South Korea, Jun. 2017, pp. 1–6.

[45] *Implementation Guideline for Digital Interface to Instrument Transformers Using IEC 61850-9-2*, UCA Int. Users Group, Switzerland, 2004.

[46] D. P. Kingma and J. L. Ba, "ADAM: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, 2015, pp. 1–15.

[47] *Keras: The Python Deep Learning Library*. Accessed: Apr. 2020. [Online]. Available: https://keras.io/

[48] *A Technical Paper: Protection System Reliability*, North Amer. Elect. Rel. Corporat. Syst. Protect. Control Task Force, Atlanta, GA, USA, Nov. 2008.

[49] D. J. Dolezilek, "Using software-defined network technology to precisely and reliably transport process bus Ethernet messages," in *Proc. 14th Int. Conf. Develop. Power Syst. Protect.*, Belfast, U.K., Mar. 2018, pp. 1–6.

[50] D. Hou, A. Guzman, and J. Roberts, "Innovative solutions improve transmission line protection," in *Proc. Southern African Conf. Power Syst. Protect.*, Midrand, South Africa, Nov. 1998, pp. 1–25.

[51] T. Cui, D. Ishchenko, and R. Nuqui, *Security Filter: Secure Communication of Protection and Control Devices in IEC 61850 Substations*, Protect. Autom. Control World Americas, Raleigh, NC, USA, 2015.

Yew Meng Khaw (Member, IEEE) received the B.A.Sc degree in engineering science from the University of Toronto with a specialization in Energy Systems Engineering in 2018, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto in 2020, where his research focused on the intersection of deep learning and smart grid security. He is currently a Senior Consultant with the Data, Analytics and AI team, Ernst and Young, Canada. He is also a Visiting Researcher with the University of Toronto. His research interests include deep learning and its applications in cyber-physical security. More recently, his interests also include the broader applications of data science to create business values for his clients.

Amir Abiri Jahromi (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from McGill University, Montréal, QC, Canada, in 2016.

From January 2018 to December 2019, he was a Postdoctoral Fellow with the University of Toronto. In 2020, he was a Research Associate with the University of Toronto. He is currently a Lecturer with the School of Electronic and Electrical Engineering, University of Leeds. His research interests are in the fields of power system modeling, cyberphysical security, reliability, economics, and optimization of power systems.

Mohammadreza F. M. Arani (Member, IEEE) received the M.Sc. degree in electrical engineering from the University of Waterloo, Waterloo, Canada, in 2012, and the Ph.D. degree in energy systems from the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada, in 2017. From 2012 to 2013, he worked as a Research Associate with the University of Waterloo. He was an NSERC Postdoctoral Fellow with the University of Toronto, from 2017 to 2019. He joined Ryerson University, Toronto, Canada, as an Assistant Professor in July 2019. His research interests include cyber-physical security of smart grids, renewable and distributed generation, plug-in hybrid electric vehicles, microgrids dynamics and control, and power system stability.

Scott Sanner received the double B.S. degrees in computer science and electrical and computer engineering from Carnegie Mellon University in 1999, the M.S. degree in computer science from Stanford University in 2002, and the Ph.D. degree in computer science from the University of Toronto in 2008. He is currently an Assistant Professor in Industrial Engineering and was Cross-appointed in Computer Science with the University of Toronto. Previously, he was an Assistant Professor with Oregon State University and before that he was a Principal Researcher with National ICT Australia and an Adjunct Faculty with the Australian National University.

Deepa Kundur (Fellow, IEEE) received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical and computer engineering from the University of Toronto, Toronto, Canada, in 1993, 1995, and 1999, respectively.

She is currently a Professor and the Chair of the Edward S. Rogers Sr. Department of Electrical and Computer Engineering with the University of Toronto. She has authored over 200 journal and conference papers and is also a recognized authority on cyber security issues. Her research interests lie at the interface of cybersecurity, signal processing, and complex dynamical networks.

Prof. Kundur's research has received Best Paper recognitions at numerous venues, including the 2015 IEEE Smart Grid Communications Conference, the 2015 IEEE Electrical Power and Energy Conference, the 2012 IEEE Canadian Conference on Electrical and Computer Engineering, the 2011 Cyber Security and Information Intelligence Research Workshop, and the 2008 IEEE INFOCOM Workshop on Mission Critical Networks. She has also been the recipient of teaching awards at the University of Toronto and Texas A&M University. She has served in numerous conference executive organization roles including as the Publicity Chair for ICASSP 2021, the Track Chair for the 2020 IEEE International Conference on Autonomous Systems, the General Chair of the 2018 GlobalSIP Symposium on Information Processing, Learning and Optimization for Smart Energy Infrastructures, the TPC Co-Chair for IEEE SmartGridComm 2018. the Symposium Co-Chair for the Communications for the Smart Grid Track of ICC 2017, the General Chair for the Workshop on Communications, Computation and Control for Resilient Smart Energy Systems at ACM e-Energy 2016, the General Chair for the Workshop on Cyber-Physical Smart Grid Security and Resilience at Globecom 2016, the General Chair for the Symposium on Signal and Information Processing for Smart Grid Infrastructures at GlobalSIP 2016, the General Chair for the 2015 International Conference on Smart Grids for Smart Cities, the General Chair for the 2015 Smart Grid Resilience Workshop at IEEE GLOBECOM 2015, and the General Chair for the IEEE GlobalSIP 2015 Symposium on Signal and Information Processing for Optimizing Future Energy Systems. She currently serves on the Advisory Board of IEEE Spectrum. She is a Fellow of the Canadian Academy of Engineering, and a Senior Fellow of Massey College.

Marthe Kassouf received the B.Sc. degree in computer engineering from the École Supérieure des Ingénieurs de Beyrouth, Lebanon, in 1997, the M.Sc. degree in computer engineering from the École Polytechnique de Montréal, Canada, in 1999, and the Ph.D. degree in electrical engineering from McGill University, Canada, in 2008. Since 2008, she has been working as a Researcher with the Hydro Quebec Research Institute (IREQ), where she has been contributing to the implementation of different projects aiming at the enhancement of the information and telecommunications infrastructure supporting the power grid, mainly in the areas of wireless communication systems, time synchronization, and cybersecurity. She has been the Project Manager for the Cybersecurity Research Project with IREQ, since 2018. She was also an Adjunct Professor with the Department of Electrical and Computer Engineering, McGill University from 2013 to 2020. Her research interests are in telecommunication networks, time synchronization systems, power grid automation, and cybersecurity for smart grids. She is also an Active Member in the Working Group 15 of the International Electrotechnical Commission (IEC) Technical Committee 57, since 2015, she has been contributing to the development of IEC 62351 standards for the cybersecurity of power system information infrastructure.