






# Cyber-Immune Line Current Differential Relays

Ahmad Mohammad Saber , *Member, IEEE*, Amr Youssef , *Senior Member, IEEE*, Davor Svetinovic , *Senior Member, IEEE*, Hatem H. Zeineldin , *Senior Member, IEEE*, and Ehab F. El-Saadany , *Fellow, IEEE*

**Abstract**—Industrial advancements in information and communications technology facilitated the widespread use of line current differential relays (LCDRs) for protecting critical transmission lines due to their fast, sensitive, selective, and secure performance. Despite their advantages, LCDRs' reliance on vulnerable communication networks to swap current measurements makes them vulnerable to cyberattacks. In this article, a scheme is proposed to protect LCDRs from direct-false-tripping (DFT), fault-masking (FM), and sympathetic-tripping (ST) cyberattacks, which have not been studied together before for transmission-level LCDRs. The proposed scheme utilizes a deep neural network (DNN), trained offline on features extracted from only the measurements available for LCDRs. The trained DNN model can then be implemented within LCDRs. Unlike the previous solutions, which only differentiate between faults and DFT cyberattacks, the proposed scheme actively differentiates between authentic and manipulated LCDR measurements to detect and mitigate possible cyberattacks. The performance of the proposed scheme is evaluated using the IEEE 39-bus benchmark system. Our results show that the proposed scheme can accurately detect different forms of DFT, ST, and FM cyberattacks while maintaining the LCDR's protective characteristics. The proposed scheme is tested for real-time capability using an OPAL-RT simulator.

Manuscript received 7 December 2022; revised 16 May 2023; accepted 21 August 2023. This work was supported in part by the CIRA-013-2020, Khalifa University, UAE, and in part by the VRI20-07, ASPIRE Virtual Research Institute Program, Advanced Technology Research Council, UAE. Paper no. TII-22-5007. (Corresponding author: Ahmad Mohammad Saber.)

Ahmad Mohammad Saber is with the Advanced Power and Energy Center (APEC), Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi 127788, UAE (e-mail: ahmad.m.saber@ieee.org).

Amr Youssef is with the Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: youssef@ciise.concordia.ca).

Davor Svetinovic is with the Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi 127788, UAE, and also with the Department of Information Systems and Operations Management, Vienna University of Economics and Business, 1020 Wien, Austria (e-mail: davor.svetinovic@ku.ac.ae).

Hatem H. Zeineldin is with the Advanced Power and Energy Center (APEC), Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi 127788, UAE, and also with the Electric Power Engineering Department, Cairo University, Giza, Egypt (e-mail: hatem.zeineldin@ku.ac.ae).

Ehab F. El-Saadany is with the Advanced Power and Energy Center (APEC), Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi 127788, UAE, and also with the University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: ehab.elsaadany@ku.ac.ae).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TII.2023.3310769>.

Digital Object Identifier 10.1109/TII.2023.3310769

**Index Terms**—Cyber-physical security, false-tripping attacks, fault-masking attacks, line current differential relays (LCDRs), smart grid security, sympathetic-tripping attacks.

## I. INTRODUCTION

RECENTLY, the EU, USA, and NATO have led several initiatives for neutralizing cyber threats against national critical infrastructures [1]. Among critical infrastructures, power grids were targets of several nationwide cyber-launched attacks, some of which aimed to inflict societal harm by inducing blackouts [1]. Line current differential relays (LCDRs) are utilized by utilities when fast, sensitive, selective, and protectively secure protection is required, e.g., to protect critical transmission lines [2]. Even though LCDRs are excellent relays, they are highly vulnerable to cyberattacks, which makes them attractive targets for cyberattacks targeting smart power systems. Typically, an LCDR-protected line is equipped with two LCDRs; one LCDR is installed near one line end. Each LCDR collects time-stamped local current and voltage measurements, then both LCDRs interchange current measurements. LCDRs then continuously compare local and remote current measurements to detect faults.

For LCDRs, remote measurements

- 1) are stamped based on the GPS signal, which malicious entities may spoof;
- 2) are transmitted over a two-way communication network, which often involves vulnerable media, e.g., microwave and radio communications;
- 3) pass through the local area networks (LANs) of both the sending and receiving substations, while LANs are vulnerable to cyberattacks as demonstrated by the infamous cyberattacks on the Ukrainian grid.

Consequently, both the magnitudes and angles of these remote measurements are prone to manipulation by adversaries and can no longer be assumed authentic all the time.

From a power system's viewpoint, a cyberattack on an LCDR can have one of three goals. The first goal is to directly trip the line protected by the attacked LCDR under the healthy operation of the system, referred to in this article as a direct false-tripping cyberattack. This category of cyberattacks is the most studied in the literature due to its simplicity since no prior knowledge of the power system is required but the LCDR's working principle and how to attack it [3], [4]. An example of these attacks is when the remote measurements of the targeted LCDR are multiplied by a large number to resemble a fault current [3], [4]. The second

goal of cyberattacks is to mask a physical fault on the line from the LCDR, which is analogous to removing LCDRs from the system. This type of cyberattack is denoted as a fault-masking cyberattack [3]. The third possible goal, a more stealthy one, is to force the LCDR to trip due to a nearby external fault, e.g., on an adjacent line [5]. In this article, cyberattacks performed with this goal are denoted sympathetic-tripping cyberattacks since they resemble the sympathetic-tripping phenomena of unit protective relays where a differential protection scheme maloperates due to an external fault [6].

Being attractive targets of potential cyberattacks, LCDRs have recently been the focus of several papers that aimed to improve the LCDRs' cybersecurity. To avoid GPS-signal-spoofing cyberattacks on LCDRs, a GPS-free differential protection scheme was proposed in [7]. In [8], software-defined networking for operational technology was proposed, as a communication architecture of enhanced security, for communication-based relays. Nougain et al. [5] proposed resilient protection for LCDRs in medium-voltage dc microgrids against direct false-tripping and sympathetic-tripping cyberattacks. However, this solution cannot be applied to large ac transmission systems. To improve LCDRs' logic against cyberattacks, different techniques, e.g., [3], [4], and [9], were proposed to differentiate between internal faults and direct false-tripping cyberattacks on the LCDR protecting this line.

Inherently, these solutions fail to detect fault-masking cyberattacks since these solutions are only involved *after* the LCDR is triggered (by a fault or a direct false-tripping cyberattack) while fault-masking cyberattacks keep the attacked LCDR idle. In addition, most of these solutions cannot detect sympathetic-tripping cyberattacks, in which local measurements of the attacked LCDR are disrupted (by a nearby fault, similar to an internal fault), unlike the situation in direct false-tripping cyberattacks where the local measurements are normal. Therefore, there is a gap in developing a solution to secure LCDRs against cyberattacks of different goals, under system variations, and within the same time frame LCDRs would take to detect faults if there were no cyberattacks.

This article presents a unified scheme for detecting direct false-tripping, fault-masking, and sympathetic-tripping cyberattacks against LCDRs in transmission systems. The proposed scheme is a single-stage multiple-cyberattack detection scheme. In contrast with the previous approaches, the proposed scheme continuously monitors LCDR's measurements, which are the only required inputs, to detect possible cyberattacks, whether they falsely trigger the LCDR or keep it untriggered. In addition, the proposed scheme, based on deep learning, is trained on all the possible states of LCDRs. This includes the ability to differentiate between direct false-tripping and sympathetic-tripping cyberattacks, which have different effects on LCDR's local measurements and were not covered by the previous works on the transmission level. Deep neural networks (DNNs) are utilized in this article as the primary building block of the proposed scheme due to their superior speed and classification accuracy to serve as a guideline for protection engineers and LCDR manufacturers. The main contributions this article presents are as follows:

- 1) Developing a scheme for detecting cyberattacks on LCDRs, including fault-masking and sympathetic-tripping cyberattacks, for the first time.
- 2) The proposed scheme requires no additional measurements than those available for LCDRs.

In addition, the performance of the proposed scheme is evaluated as follows:

- 1) under cyberattacks and faults of different parameters;
- 2) under varying operating conditions and system dynamics;
- 3) under uncertainty stemming from measurement noise;
- 4) in real time using OPAL-RT real-time simulator.

The rest of this article is organized as follows. Section II explains LCDRs' vulnerability to cyberattacks. The threat model is demonstrated in Section III, followed by the proposed scheme in Section IV. Performance evaluation and real-time verification are conducted in Sections V and VI. Afterward, a discussion is performed in Section VII. Finally, Section VI concludes this article.

## II. CYBERATTACKS ON LCDRS

Practically, two LCDRs protect each line. An LCDR, placed near one line terminal, receives a set of local and remote measurements. Local current and voltage measurements, denoted as  $i_1$  and  $v_1$ , respectively, are directly sent by the collocated current and potential transformers to the LCDR, so there is no room for manipulating these measurements remotely. On the contrary, the remote current measurement ( $i_2$ ) is communicated by the far-end LCDR after being time stamped, and therefore  $i_2$  is susceptible to manipulation [10].

### A. LCDRs' Characteristics

After receiving the time-stamped local and remote measurements, the LCDR determines the differential current ( $i_d$ ) using

$$i_d(t) = i_1(t) + i_2(t) \quad (1)$$

During normal operation or external faults,  $i_d$  is close to zero. Yet,  $i_d$  has a large magnitude during internal faults. To avoid maloperation and account for line capacitive currents, LCDRs employ restraining characteristics, so an LCDR trips only if

$$|i_d(t)| \geq i_{op}(t) \quad (2)$$

where  $|i_d|$  is the magnitude of the differential current, and  $i_{op}$  is the operating current determined as

$$i_{op}(t) = \begin{cases} i_{d_i} + s_1 \times i_r(t) & i_r(t) \leq i_{b_L} \\ i_{d_i} + s_1 \times i_{b_L} + s_2(i_r(t) - i_{b_L}) & i_r(t) > i_{b_L} \end{cases} \quad (3)$$

where  $i_{d_i}$  is an initial differential current setting,  $i_{b_L}$  is the bias current limit, and  $s_1$  and  $s_2$  are slopes of the LCDR characteristic lines, as illustrated in Fig. 1. Finally,  $i_r$  is the restraining current determined as

$$i_r(t) = |i_1(t)| + |i_2(t)| \quad (4)$$

Consequently, LCDRs are characterized by excellent speed, dependability, selectivity, and protection-security. In addition,

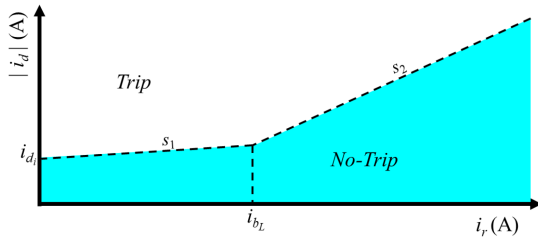


Fig. 1. Characteristic of LCDRs.

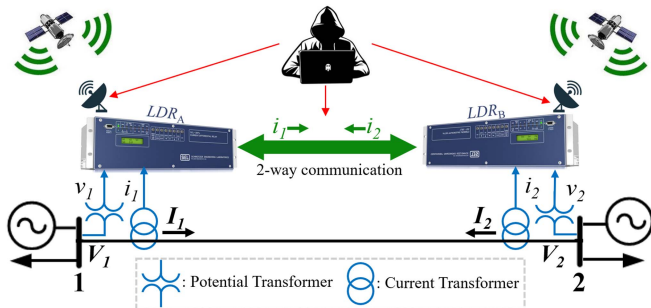


Fig. 2. Illustration of possible cyberattacks on LCDRs.

thanks to the developments in measurements and information technologies, modern LCDRs are immune to problems related to communication delay, noise, and externally induced saturation or transients in instrument devices. However, due to LCDRs' dependency on vulnerable infrastructures, e.g., the vulnerable communication media and GPS signal, LCDRs are vulnerable to cyberattacks.

### B. Vulnerability of LCDRs to Cyberattacks

Fig. 2 illustrates two main intrusion points through which malicious entities can intrude and modify the remote measurement  $i_2$ , which are the (1) GPS-signal receiver, and (2) the two-way communication network (TWCN). First, the received GPS signal can be overwhelmed by noise with a similar frequency to the true signal, which is equivalent to manipulating the phase angle of the remote measurement [7]. Second, the TWCN involves two vulnerabilities, which are (1) the LANs of the two substations hosting the LCDRs, and (2) the wide area network through which the remote measurements are exchanged. As demonstrated by the cyberattacks on the Ukrainian grid in 2015 and 2016, modern substations' LANs are vulnerable to intrusion, mainly due to the vulnerability of the IEC 61850 automation standard on which these substations are based [3]. Furthermore, many of the utilized media in TWCNs are vulnerable, e.g., wireless (radio/microwave), or involve optical–electrical interfacing devices, e.g., routers and switches, which can be exploited for cyberattacks [3], [7]. Consequently, malicious entities can exploit one or more of the vulnerable/weak points in the TWCN to manipulate the remote measurements of the LCDR.

By exploiting the aforementioned intrusion points, several mechanisms of cyberattacks can be performed on LCDRs. First,

time-synchronization attacks (TSAs) can be performed by attacking the GPS signal, which is the mechanism used by LCDRs to synchronize measurements [7]. For an LCDR under a TSA, (1) becomes

$$i_{d_{\text{TSA}}}(t) = i_1(t) + i_2(t) + \theta_{\text{TSA}}(t) \quad (5)$$

Meanwhile, the wide attack space of TWCN allows multiple cyberattack mechanisms to be performed to achieve either a direct false-tripping, a fault-masking, or a sympathetic-tripping cyberattack. For instance, malicious entities can modify the phase angle of the remote measurement or insert a delay in the form of a TSA. False-data injection attacks (FDIAs) on LCDRs are also possible, where an additional current measurement vector is injected into the remote measurements by the attackers, adding to their magnitude or angle. Under FDIAs, an LCDR's differential current appears as

$$i_{d_{\text{FDIA}}}(t) = i_1(t) + i_2(t) + i_{\text{FDIA}}(t) \quad (6)$$

In addition, LCDRs are vulnerable to the well-known man-in-the-middle attacks (MitMAs), in which the adversary intercepts the remote measurements to eavesdrop, modify or replace the communicated message contents. In this article, the effect of MitMAs on the LCDR is modeled as

$$i_{d_{\text{MitMA}}}(t) = i_1(t) + i_{2_{\text{MitMA}}}(t) \quad (7)$$

Like MitMAs, replay attacks on LCDRs can be performed by replaying a prerecorded signal instead of the actual remote signal, e.g., replaying a prerecorded fault current signal. Alternatively, the attacker can repeat a specific signal pattern, e.g., repeatedly replaying the normal (healthy) remote current to mask a fault. Herein,  $i_d$  determined by an LCDR under a replay attack can be modeled as

$$i_{d_{\text{replay}}}(t) = i_1(t) + i_{2_{\text{replay}}}(t) \quad (8)$$

### III. THREAT MODEL

In this article, following previous literature, *direct false-tripping cyberattacks* are defined as those explicitly performed when the system is fault free to fool the targeted LCDR into tripping its line unnecessarily [3]. On the other hand, *fault-masking* and *sympathetic-tripping* cyberattacks co-occur with a fault on the system, so both cyberattack categories are defined and studied separately from direct false-tripping attacks in this article. In fault-masking cyberattacks, malicious entities manipulate remote measurements to mask an internal fault, i.e., on the line protected by an LCDR, and the attacked LCDR never senses this fault. Therefore, this cyberattack is analogous to halting the primary protection in terms of the effect on the stability of the system and potential damage to sensitive equipment. This attack category is performed stealthily by modifying the remote measurements similar to the traditional MitMA model. In contrast, the objective of a sympathetic-tripping cyberattack is to, finally, trip the line protected the LCDR due to an out-of-the-line or external fault since this cyberattack can be more misleading and stealthy, i.e., than direct false-tripping cyberattacks on a healthy system, due to the overall system disturbance caused by the nearby fault on an adjacent line.

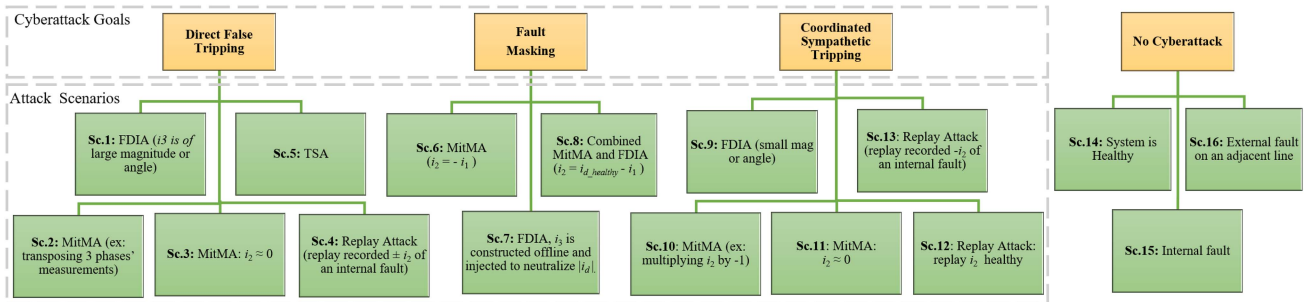


Fig. 3. Tree of LCDR's states under different cyberattacks and operating conditions.

### A. Tree of LCDR Cyberattacks and States

In most of the previous works, only three states of LCDRs were considered, which are (1) untriggered in a healthy system, (2) triggered by an internal fault, and (3) triggered by a direct false-tripping cyberattack. In contrast, a broader spectrum of possible states is considered in this article, including different ways/mechanisms to perform either direct false-tripping, fault-masking, or coordinated sympathetic-tripping cyberattacks. Fig. 3 illustrates the threat tree showing 16 possible LCDR states/scenarios, denoted Sc. 1–Sc. 16. Cyberattack scenarios branch from the three attack goals. Sc. 1–Sc. 5 are examples of possible direct false-tripping attacks on LCDRs. In Sc. 1, a large-magnitude current is injected, resulting in a significant difference between the local and the received remote current phasors and fooling the LCDR to trip. A similar effect can be obtained by injecting an additional phase angle of  $180^\circ$ . In Sc. 2, a form of MitMAs, the three remote current phasor measurements are transposed, resulting in an angle difference between local and perceived-remote phasors in the order of  $120^\circ$  and forcing the LCDR to trip. Sc. 3 is another form of MitMAs in which  $i_2$  is instead replaced by a current of zero magnitude, which satisfies (1)–(4), thus the LCDR trips. In Sc. 4, a prerecorded fault current is replayed instead of  $i_2$ . As an option, the replayed current can be multiplied by  $-1$ . Both cases are sufficient to trip the LCDR and bypass attack-detection mechanisms that allow tripping only if remote currents are like those expected under faults. Sc. 5 involves repeating a steady-state value of  $i_2$  during strong dynamics, e.g., switching a nearby capacitor bank, where, had there been no attack, both local and remote currents experience a strong dynamic behavior, i.e., increase in magnitude. This scenario's mismatch between local and manipulated-remote current measurement will likely fool the LCDR to trip.

Fault-masking cyberattacks are modeled in Scenarios 6–8. To keep  $i_d$  close to zero during internal faults, thus masking an internal fault, an FDIA, MitMA, or a combined attack can be performed. In Sc. 6, it is assumed that a MitMA is performed so that  $i_{2\text{MitMA}}$  equals  $-i_1$  during the fault, resulting in a near-zero value for  $i_d$ . Alternatively, in Sc. 7, it is assumed that the fault-imposed  $i_d$  is known to the adversaries, for example, in a coordinated cyber-physical attack context where the physical attacker initiates the fault. Knowing fault current allows attackers

to construct and inject  $i_3$  that diminishes  $|i_d|$  during a fault and keep it below  $i_{\text{op}}$ . In Sc. 8, a combined FDIA-MitMA is assumed to be performed so that, after manipulating  $i_2$ , the differential current as determined by the target LCDR remains equal to and in the form of  $i_{d\text{healthy}}$  (differential current under normal operation).

Sympathetic-tripping attacks are modeled for the first time in Sc. 9–13. In these scenarios, the attack starts just before and during an external fault, which could be intentional, and is performed by merely manipulating remote currents but in different ways. In Sc. 9, it is assumed that magnitudes of remote currents are diminished, e.g., by multiplying them with a factor of  $M$ , such that  $M \ll 1$ . In Sc. 10, and similar to 9, remote currents are multiplied by an  $M$ , where  $M = -1$ . Alternatively, in Sc. 11, as in an MitMA, authentic  $i_2$  is completely replaced by a current of zero magnitude. In Sc. 12, the usual healthy/normal  $i_2$  measurements are replayed instead of the actual  $i_2$  of the line. In Sc. 13, for further confusion, a prerecorded  $i_2$  of any fault is replayed after negating its angle by  $180^\circ$ . In Sc. 9–13, the LCDR trips due to the inconsistency between local and remote currents since local currents are disrupted due to the external fault.

Finally, under no cyberattack, the system could be either healthy (Sc. 14), experiencing an internal fault (Sc. 15), or an external fault (Sc. 16).

### B. Cyberattacks' Assumptions

1) *Required knowledge*: Since cyberattacks on LCDRs are mainly attacks on the specific line (and system) protected by the LCDR, it is assumed that the adversary has a basic understanding of power systems, line protection, and LCDRs' principle in specific. In addition, for fault-masking attack cases performed through FDIA only, it is assumed that fault currents are generally known to the attackers, for example, through knowledge of the system [11].

2) *Attackers' capabilities*: In addition, it is assumed that attackers are capable of remotely intruding into the LCDRs, through the cyber layer only, to spoof the GPS signal received by the LCDR (for TSAs) and manipulate the remote current measurements through the TWCN (for FDIA, MitMAs, and replay attacks). In detail, the adversary is assumed to be able to intercept, eavesdrop, modify, replace, and drop the remote measurements' packets only [3], [12]. For fault-masking and

sympathetic-tripping attacks, it is assumed that the remote-currents' manipulation is 1) initiated just before the fault occurs through a cyber MitMAs, or, possibly, 2) performed in a mutually coordinated cyber-physical attack with a physical initiator [13], [14]. Finally, similar to previous works, it is assumed that only remote measurements ( $i_2$ ) are attacked, local measurements ( $i_1$  and  $v_1$ ) are secure, and a few LCDRs can be attacked simultaneously for direct false-tripping, fault-masking, and sympathetic-tripping goals.

#### IV. DEVELOPING CYBER-IMMUNE LCDRS

This article examines three distinct categories of cyberattacks that had previously been overlooked or addressed in isolation. Following a similar approach would result in three solutions, one to detect each attack category. Moreover, previous solutions for detecting direct false-tripping attacks were developed to differentiate between this cyberattack category and internal faults. The primary rationale was that during faults, both local and remote measurements are disrupted but correlated, while under direct false-tripping cyberattacks, the power system and hence local measurements are healthy, and only remote measurements are perturbed. This previous approach has two main drawbacks. First, it fails to detect sympathetic-tripping cyberattacks where local measurements are not healthy but disrupted by an external fault in a way that is similar to how they are disrupted under an internal fault. Second, the previous approach facilitates performing fault-masking cyberattacks since, in that approach, tripping is allowed only if both local and remote measurements correspond to local and remote measurements of a fault. Therefore, a fault-masking attack can easily bypass such an approach by manipulating remote measurements, e.g., by replacing them with small-magnitude currents, during an internal fault.

##### A. Proposed Solution's Philosophy and Requirements

The proposed scheme immunizes LCDRs against direct false-tripping, fault-masking, and sympathetic-tripping cyberattacks. Therefore, the proposed scheme performs several functions simultaneously. These include, when the LCDR is idle, 1) actively looking for possible fault-masking cyberattacks, and 2) ensuring that these masked faults are internal. In addition, once the LCDR is triggered, the solution must 3) differentiate between faults and direct-false/sympathetic-tripping cyberattacks by correlating local and remote measurements, and 4) if a fault is determined, the proposed scheme must confirm the fault zone before tripping. The above has to be performed 5) accurately and 6) simultaneously, and 7) with a minimal time delay to preserve the LCDR's protective merits. Finally, 8) the proposed scheme uses only  $i_1$ ,  $i_2$ , and  $v_1$  measurements. DNNs have been widely used for complex online classification problems [15]. As universal function approximators, DNNs are known for their accuracy, lightweight, and ability to classify within a fraction of a millisecond. DNNs are chosen in this article as the main component of the proposed scheme for a cyber-immune LCDR against cyberattacks. Similar techniques could also be used, depending on each manufacturer's choice. Fig. 4 shows how the proposed DNN-based scheme works in conjunction with

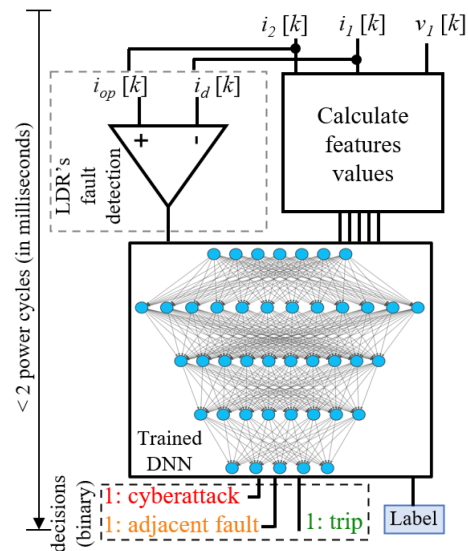


Fig. 4. Information flow in LCDRs secured by the proposed scheme.

LCDRs. Only  $i_2$ ,  $i_1$ , and  $v_1$  are required as inputs. The features extracted from  $i_1$ ,  $v_1$ ,  $i_2$ , and the LCDR's output are continuously fed to the pretrained DNN to observe for possible cyberattacks and issue/confirm a trip order as needed. The main requirement of the proposed scheme is training DNNs on a dataset comprising the possible cyberattack scenarios, as highlighted in the previous section. This requirement can be easily met today, thanks to the recent developments in systems modeling and simulator automation software packages used in the power systems industry. Notably, a good training dataset covers all the planned operating scenarios of the power system where the LCDR to be secured is installed. These operating scenarios are often bounded on the transmission level and, therefore, can be simulated.

##### B. DNNs for a Unified Cyberattacks' Detector

A DNN, as illustrated in Fig. 4, consists of a number of successive layers ( $L$ ), which can be generally divided into three groups; 1) the input layer, 2)  $(L - 2)$  fully connected layers, and 3) the classification layer. Each layer ( $l$ ) is comprised of a number of neurons ( $n$ ). Initially, the input data vector ( $X_0$ ) of  $N_0$  dimensions is fed to the input layer, which has the same width as  $X_0$ , as depicted in Fig. 4. Similarly, the data are then forward fed through the  $(L - 2)$  hidden layer and the output layer. On the neuron level,  $X_{n,l}$ , defined as the output of neuron  $n$  in layer  $l$  is determined as

$$X_{n,l} = f_{n,l} (w_{n,l,i}^T X_{l-1} + b_{n,l}) \quad (9)$$

where  $X_{l-1}$  is the vector of input data to layer  $l$ .  $f_{n,l}$  and  $b_{n,l}$  are the activation function and bias for the given neuron, respectively.  $w_{n,l,i}$  is the weight-linking neuron  $i$  in layer  $l - 1$  to the neuron  $n$  in layer  $l$ . In general,  $b_{n,l} \in \mathbb{R}$  and  $w_{n,l} \in \mathbb{R}^{N_{l-1}}$ .

1) *Hyperparameters of the utilized DNN and its training process*: For DNNs, configuring their hyperparameters is the most critical step [16], which is the goal of the training phase. The DNN's hyperparameters include: total number of layers:  $L$ ,

number of neurons/nodes in each layer,  $f$  for each node, which can be, for example, sigmoidal, hyperbolic, rectified linear unit (ReLU) or none [16], [17],  $b$  term for every node, and  $w$  linking every two connected nodes. In addition, these parameters need to be identified: number of training episodes/epochs, and split ratio (SR) of the dataset, i.e., what percentage of the data will be used for training and what percentage is left for validation.

Typically, the training process is performed over several epochs. This process is often started with random hyperparameter values. Afterward, in each epoch, a DNN model with different—more optimal—hyperparameter values is trained and validated, i.e., on the training-and-validation dataset. For all the epochs, the accuracy of each model is recorded versus the model hyperparameters. The model's performance is commonly evaluated in terms of classification error. Herein, the loss function, denoted  $e$ , is computed as

$$e = \frac{\text{Number of misclassified cases}}{\text{Total number of cases}} \quad (10)$$

Finally, the optimal hyperparameters of a DNN are considered those of the model that resulted in the highest accuracy in the training-validation phase.

In this article, we train the DNN model using Bayesian optimization (BO) in conjunction with k-fold cross-validation, as explained next. The BO optimizes the hyperparameter values over the epochs. Within each epoch, we evaluate the objective function (i.e., the validation performance) on each fold separately and then average the results across the k to obtain a more reliable estimate of the model's performance. This can help to reduce the variance in the estimated performance and improve the robustness of the hyperparameter tuning process, as explored previously in references, such as [18].

*2) Bayesian optimizer for hyperparameters optimization:* Several techniques were developed to facilitate the training and validation phase, e.g., random search, grid search, and BO. These techniques mainly aid in hyperparameter optimization, reducing error, and training time. This article selects BO as the governing technique for the training and validation epochs. In BO, the problem of hyperparameter tuning is solved sequentially, i.e., in epochs, as an optimization problem with model hyperparameters being the optimization variables and the objective function (obj) described as

$$\text{obj}(h) = \min(e) = y \quad (11)$$

where  $h$  is the set of hyperparameter values. In other words, the goal of this process is to find the point (DNN model) which minimizes the training-validation error, which is equivalent to maximizing the model's accuracy. Initially, the hyperparameter values are selected randomly. Afterward, the BO utilizes an acquisition function to determine the next set of hyperparameter values to try. Different acquisition functions can be used. In this article, the expected improvement (EI), whose details can be found in [19], is utilized as the acquisition function of the BO. With EI acquisition function, the BO evaluates the EI in obj, and hyperparameter values that may increase obj are ignored.

During training, the BO seeks to maximize the DNN model's accuracy. Herein, the hyperparameter optimization ends when

the maximum number of epochs is reached. After stopping, the DNN model with the lowest error in all the training-validation epochs is selected as the best model. The hyperparameter values associated with this model are considered optimal.

*3) Cross-validation:* After the BO selects a set of hyperparameter values for the DNN model to try in a specific epoch, the accuracy of this model is obtained in a k-fold cross-validation manner [3], [18]. The training-and-validation dataset, which must contain samples of all operating conditions of the LCDR, is shuffled and then split into k-equal folds, numbered 1 to k. (The value of k is to be consistent with the desired SR.) The DNN model is then trained and validated k times. In the kth iteration, the kth fold is held for validation, the model is trained on the remaining folds, the kth model's accuracy is recorded, and so forth. Finally, the accuracy of the k models is averaged to obtain the accuracy of the DNN model in this epoch.

### C. Features

In this research, the DNNs are trained on features extracted only from the measurements today's LCDRs' have, i.e.,  $i_1$ ,  $i_2$ , and  $v_1$ , therefore eliminating the need for additional power systems measurements, parameters, distribution function or probability. Table I summarizes the classification features. In this table,  $P$  denotes the line phases (A, B, and C), while  $S$  denotes the three sequences (positive, negative, and zero).  $Z$  is the apparent impedance defined in this article as the ratio between the same phase's local voltage and local current. In addition, operators  $|\cdot|$ ,  $\text{Re}(\cdot)$ , and  $\text{Im}(\cdot)$  denote the magnitude, real, and imaginary parts, respectively. Finally,  $\theta$  is the phase angle, e.g.,  $\theta_{i_2, i_1}^P$  is the difference between phase angles of  $i_2$  and  $i_1$ . The chosen features are selected to maximize extracted information from the available measurements, hence features are diversified between magnitude and angle-based features, differential features, and features related to the power factor and  $Z$ . Most of the utilized features are already calculated by modern LCDRs, e.g., [10], and, in general, all features are obtained via simple additions and multiplications, thus optimizing the proposed scheme's complexity.

## V. PERFORMANCE EVALUATION

To validate the performance of the proposed scheme in enforcing the cybersecurity of LCDRs, extensive cases are simulated, which represent all the prescribed scenarios in Section III, including direct false-tripping, fault-masking, and sympathetic-tripping cyberattacks, internal and external faults, and healthy-system dynamics. An extensive search space is established by generating comprehensive cases. Next, the proposed solution is trained, and its performance is evaluated. Afterward, the proposed scheme is tested under measurement noise.

### A. Generation of Extensive Case Studies

To evaluate the performance of LCDRs augmented with the proposed scheme, time-domain simulations are carried out on a power transmission system. The IEEE 39-bus transmission-level

TABLE I  
SUMMARY OF CLASSIFICATION FEATURES

Features	Number
$ V_1^p ,  i_1^p ,  V_1^s ,  i_1^s ,  i_2^p ,  i_2^s ,$	1–6
$\theta_{v_1}^p, \theta_{i_1}^p, \theta_{i_2}^p, \theta_{v_1}^s, \theta_{i_1}^s, \theta_{i_2}^s,$	7–12
$ i_d^p ,  i_d^s ,  i_2^p  -  i_1^p ,  i_2^s  -  i_1^s , \theta_{i_d}^p, \theta_{i_d}^s, \theta_{i_2, i_1}^p, \theta_{i_2, i_1}^s,$	13–20
$\text{Cos}(\theta_{v_1, i_1}^p), \text{Cos}(\theta_{v_1, i_1}^s), \text{Sin}(\theta_{v_1, i_1}^p), \text{Sin}(\theta_{v_1, i_1}^s),$	21–24
$ Z_1^p ,  Z_1^s , \theta_{z_1}^p, \theta_{z_1}^s, \text{Re}(Z_1^p), \text{Re}(Z_1^s), \text{Im}(Z_1^p), \text{Im}(Z_1^s)$	25–32

Each feature is inputted 3 times, one time for each phase/sequence.

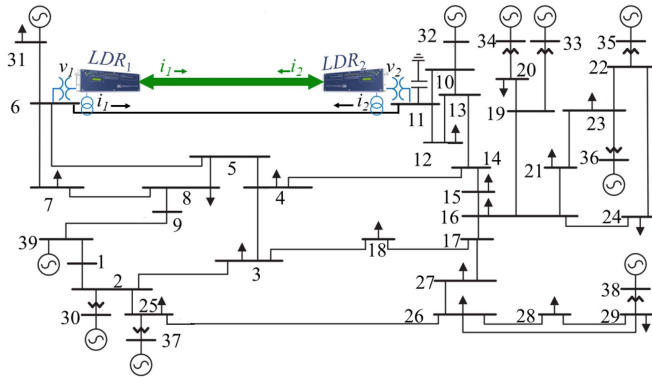


Fig. 5. Line 6–11 in the IEEE 39-bus system protected by LCDRs.

benchmark system, which is representative of a part of the U.S. grid, is used in this section. The system, whose details can be found in [20], is simulated in PSCAD/EMTDC environment. In this test system, line 6–11, being critical, is protected by  $\text{LCDR}_1$  and  $\text{LCDR}_2$  near bus 6 and 11, respectively, as shown in Fig. 5. For  $\text{LCDR}_1$ , which is focused on in this case study,  $i_1 = 0.273 \angle 97.8^\circ$  kA while  $i_2 = 0.29 \angle -73.51^\circ$ , and  $v_1 = 141.993 \angle -100.36$  kV under normal operation. The settings of the LCDR are configured following [3], [10] so that  $i_{d_i}$ ,  $i_{b_L}$ ,  $s_1$ , and  $s_2$  equal 0.05 kA, 0.585 kA, 0.2, and 0.4, respectively. Therefore, normally,  $i_r = 0.302$  kA,  $i_d = 0.046 \angle -9.53^\circ$  kA, and  $i_{op} = 0.1104$  kA. A wide spectrum of possible internal faults is simulated on this power system to ensure that the proposed scheme does not confuse them with the investigated cyberattacks. Internal faults on line 6–11 are simulated by varying all the following fault-current-affecting parameters [3], [17], which are the following.

- 1) *Fault type*: All the possible types, i.e., A-G, B-G, C-G, AB, BC, CA, AB-G, BC-G, CA-G, ABC, and ABC-G, are simulated.
- 2) *Fault resistance*: A wide range of fault resistance values is simulated covering 0.001, 1, 2, 3, 4, 6, 8, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, and 300  $\Omega$ .
- 3) *Fault location*: Faults are simulated at 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% of the line measured from bus 6.

In total, 2376 fault cases are generated, which will be used for both fault scenarios and to model fault-masking cyberattack

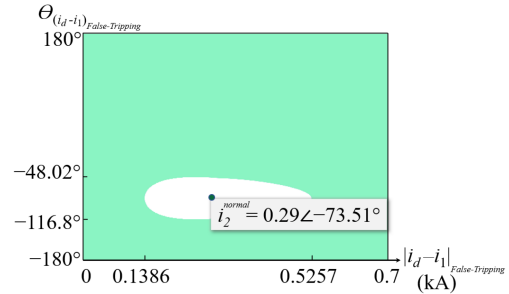


Fig. 6. Loci of remote current for false tripping of  $\text{LCDR}_1$ .

scenarios. Similarly, several external faults are simulated on adjacent lines, near bus 6 and bus 11, with varying types, locations, and resistances, which will also be used to model sympathetic-tripping cyberattack cases as detailed in the following paragraphs.

1) *Simulating direct false-tripping cyberattack scenarios*: Under the healthy operation, (1)–(4) can be numerically solved to obtain the graph illustrated in Fig. 6. In the figure, light green space represents the locus of remote currents, as received by  $\text{LCDR}_1$ , that would force the LCDR to trip line 6–11 under normal operation. Sc.1–Sc.4 explained in Section III are simulated. In Sc. 1, 800 FDIA are simulated using currents of large magnitudes and angles. To model Sc. 2, 5 cases are simulated where measurements of two and three phases are switched, e.g., switching A with B. Sc. 3 is simulated by replacing  $i_2$  with 0.0001 kA during a system dynamic event, which is switching the capacitor on bus 11. For Sc. 4, 800 replay attack cases are simulated. Authentic remote measurements are replaced with remote measurements from a prerecorded fault scenario, i.e., from the generated fault dataset. Chosen faults are of different types, resistances, and locations. In total, 800 TSA are simulated, in Sc. 5, by delaying the time stamp of remote measurements in increments of 1/16 cycle with a minimum of 1/25 cycle. Table II depicts examples of direct false-tripping cyberattack cases.

Herein, manipulated  $i_2$  values fall in the trip region of Fig. 6 so  $\text{LCDR}_1$  is fooled to trip.

2) *Simulating fault-masking cyberattack scenarios*: To simulate fault-masking cyberattacks, in Sc. 6 and Sc. 7,  $i_2$  is manipulated—during internal faults—so that  $|i_d|$  remains close to zero, which can be achieved by FDIA and MitMAs as explained above. For faults to be masked, internal faults of

**TABLE II**  
EXAMPLES OF DIRECT FALSE-TRIPPING CYBERATTACK SCENARIOS

Scenario	Manipulated $i_2$ (kA)
No attack ( <i>normal operation</i> )	0.29 $\angle -73.51^\circ$
1 ( <i>large magnitude</i> )	1.16 $\angle -74^\circ$
1 ( <i>increasing angle</i> )	0.29 $\angle 106.49^\circ$
1 ( <i>manipulating mag. and angle</i> )	0.29 $\angle -246.31^\circ$
2 ( <i>switching phase A with B</i> )	0.29 $\angle 47^\circ$
2 ( <i>switching phase A with C</i> )	0.29 $\angle -194^\circ$
3 ( $ i_2 $ kept close to zero)	0.0001 $\angle -73.51^\circ$
4 ( <i>3-ph fault mid-line at 1 <math>\Omega</math></i> )	3.6 $\angle 176.47^\circ$
5 ( <i>1/4 a cycle delay</i> )	0.29 $\angle -16.49^\circ$

different resistances, locations, and types are randomly selected from the faults dataset. To simulate the worst-case scenario, masked internal faults are simulated simultaneously with external faults on adjacent lines in Sc. 8. Overall, 2400 fault-masking cyberattack cases are simulated.

3) *Simulating sympathetic-tripping cyberattack scenarios:* In Sc. 9–Sc.13, remote currents of LCCR<sub>1</sub> are manipulated simultaneously with faults on adjacent lines to force-trip LCCR<sub>1</sub>. For these 500 external fault cases, all scenarios, i.e., Sc. 9–Sc. 13, are simulated, as explained next, resulting in 2500 cases. In Sc. 9, additional current is injected to keep  $i_2$  close to zero or decrease its angle by approximately  $180^\circ$ , therefore introducing a large  $i_d$  and forcing LCCR<sub>1</sub> to trip. To achieve a similar result, in Sc. 10 and Sc. 11, the remote current is multiplied by  $-1\angle 1^\circ$  and  $0\angle 0^\circ$ , respectively, in two forms of MitMAs. In Sc. 12,  $i_2$  is replaced by a constant value of  $0.29\angle -73.51^\circ$  during the external faults, forcing the LCCR to trip. Similar to previous scenarios, in Sc. 13, prerecorded internal fault remote currents are replayed, i.e., instead of real-time  $i_2$ . Attacks are performed simultaneously with faults on adjacent lines to confuse LCCR<sub>1</sub> further, thus fooling it into maloperation and tripping of line 6–11.

4) *Simulating no-cyberattack scenarios:* In this part, scenarios 14–16 are simulated. In Sc. 14, different healthy current and voltage fluctuations and dynamics are simulated. Normal fluctuations are recorded at varying line impedances at varying X/R ratios. Both parameters are varied in steps of 0.1% within the tolerance of  $\pm 5\%$ . These fluctuations are also simulated under generation connection/disconnection on bus 31 starting at different time instances, from 0.5 to 0.5167 s in steps of 1 ms. Internal faults, i.e., Sc. 15, are those generated above for line 6–11. For Sc. 16, 1000 external faults are simulated in addition to those generated for sympathetic-tripping cyberattacks.

## B. Training and Validation of the Utilized DNN

The training of the DNN is performed in MATLAB/Simulink environment. Each sample in the generated datasets is represented by the features summarized in Table I. The labeled datasets are then randomly split into 80% for training and validation and 20% for testing. The training-validation process is performed in consecutive epochs that the BO governs. In each

**TABLE III**  
DETAILED CONFUSION MATRIX

True Class	(A)	(B)	(C)	(D)	(E)	(F)
(A)	100%	-	-	-	-	-
(B)	-	98%	-	-	2%	-
(C)	-	-	99.8%	-	-	0.2%
(D)	-	-	-	100%	-	-
(E)	-	0.3%	-	-	99.7%	-
(F)	-	-	1%	-	-	99%

(A) Direct False-Trip Cyberattacks. (B) Sympathetic-Trip Cyberattacks. (C) Fault-Masking Cyberattacks. (D) Healthy System and no cyberattack. (E) Internal Fault and no cyberattack. (F) External Fault and no cyberattack.

epoch, the DNN model is also cross-validated  $k$  times. Initial values of the hyperparameters are determined randomly. Next, starting the first epoch, the BO optimizer determines the set of DNN hyperparameter values to try next, i.e., the values most likely to improve the accuracy, based on the EI acquisition function, and given the current model and any previously observed data points. Following this, in the same epoch, the DNN model—with hyperparameters determined by the BO—is evaluated using fivefold cross-validation. That is, the training-validation dataset is divided into five equal folds, the DNN model is trained, and the objective function is evaluated separately five times. Each time, one different fold is left out for validation, and the DNN model is trained on the four other folds and then validated on the left-out fold. Finally, the accuracy of the DNN model in this epoch is the average of the five accuracies obtained in the fivefold cross-validation, which helps obtain a more reliable estimate of the model's performance. The training epochs are continued in the same manner as the first epoch. In the end, optimal hyperparameter values are those of the model associated with the lowest classification error  $e$  among all the epochs. In this research, the optimized DNN comprises three fully connected layers, with 211, 117, and 115 neurons, respectively, ReLU-activation, and  $\lambda$  of  $3.2102 \times 10^{-5}$ . The training concludes in 37 313 s on a PC with a 1.8-GHz i7-8565 CPU, and 16 GB of RAM.

## C. Results Discussion

The trained model is now tested on 20% of the dataset held out for testing. Our results, summarized in Table III, indicate that the overall classification error of the proposed scheme is 0.4%.

In addition to the accuracy, the proposed solution's performance can be evaluated using the precision, recall, and F1-score evaluation metrics, which can be calculated for each class separately [21]. These metrics can be determined using the following:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (12)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (13)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$



**TABLE IV**  
PRECISION, RECALL, AND F1-SCORE RESULTS

Metric	(A)	(B)	(C)	(D)	(E)	(F)
Precision	100%	99.69%	99.01%	100%	98.03%	99.8%
Recall	100%	98%	99.8%	100%	99.7%	99%
F1-score	1.00	0.9868	0.9939	1.00	0.9884	0.9922

(A) Direct False-Trip Cyberattacks. (B) Sympathetic-Trip Cyberattacks. (C) Fault-Masking Cyberattacks. (D) Healthy System and no cyberattack. (E) Internal Fault and no cyberattack. (F) External Fault and no cyberattack.

Using these metrics, the performance of the proposed scheme on the extensive case studies can be analyzed for each class depicted in Table IV. Overall, the proposed scheme exhibits good performance on the three metrics, with averaged *precision*, *recall*, and *F1-score* values of 99.4217%, 99.4167%, and 0.9936, respectively. Detailed analysis and interpretation of the results are presented in the upcoming paragraphs.

On the one side, for cyberattacks on the LCDR, initially, direct false-tripping cyberattacks are correctly classified, mainly because they do not manipulate local measurements, so it is easy to detect the inconsistency between local and remote measurements. For sympathetic-tripping cyberattacks, which are harder to detect than direct false-tripping cyberattacks, 98% are correctly classified. The remaining 2% is confused for internal faults. This confusion can be explained in terms of the common disruption in both local and remote measurements in both classes. In this research, the misclassified sympathetic-tripping cyberattacks are simultaneous with nearby and strong three-phase faults on the adjacent lines. Moreover, 99.8% of fault-masking cyberattacks are detected since, even though the remote measurements are manipulated to keep the  $i_d$  below the tripping threshold, local voltage and current measurements are not manipulated, which facilitates the detection of these attacks. Only 0.2% of fault-masking cyberattacks in this research are confused for external fault. These external faults have mainly low impedance, resulting in a slight increase in  $i_2$ . This increase is similar to that caused by a masked internal fault near the protected line's far end.

On the other side, for the noncyberattack cases, all the healthy cases are correctly classified, which means that the proposed scheme is not likely to maloperate for normal power system dynamics. Furthermore, 99.7% of internal faults are correctly detected, ensuring the protective sensitivity of the LCDR is not affected. Meanwhile, the remaining percentage of internal faults is confused for sympathetic tripping cyberattacks. For the same reason, some of these attacks are confused for faults. The minor percentage of (masked) faults that are not cleared on time by the LCDR will be cleared by the local or remote backup protection, but a few milliseconds after the LCDR would normally do if there were no cyberattacks [3]. Besides, 99% of external faults are correctly identified, ensuring that the proposed scheme does not highly impact protective selectivity. It is noteworthy that in the remaining 1% of external fault cases where the proposed solution suspects there might be a fault-masking cyberattack and, hence, prioritizes protecting the power system from possible sustained faults, the line autoreclosers will operate and restore

**TABLE V**  
SUMMARIZED RESULTS FROM A CYBER PERSPECTIVE

True	A cyberattack	Not a cyberattack
A cyberattack	99.3%	0.7%
Not a cyberattack	0.43%	99.57%

**TABLE VI**  
SUMMARIZED RESULTS FROM A POWER SYSTEMS VIEWPOINT

True Action	Trip	Block
Trip	99.75%	0.25%
Block	0.75%	99.25%

the healthy line a few milliseconds post the circuit breaker's operation.

Table V recapitulates the obtained results from a different angle. From a cyber viewpoint, on the one hand, 99.3% of all cyberattacks targeting the LCDR are correctly detected as cyber intrusions. Therefore, the proposed scheme helps in defending the protection system against cyberattacks. On the other hand, 99.57% of noncyberattack states of the LCDR are correctly classified as no-intrusion cases, i.e., the false intrusion alarm rate is less than 0.43%.

It is important to evaluate the proposed scheme from a power systems viewpoint, where the main concern is protecting the system from faults as they threaten its stability. Undetected faults are physically damaging to the power system. As Table VI depicts, overall, LCDRs augmented with the proposed scheme correctly generate the required tripping decisions in 99.75% of the cases. Hence, the fault-detection sensitivity of LCDRs, previously compromised by cyberattacks, is now enforced after employing the proposed scheme. The minor percentage of missed tripping orders should be handled by backup protection, as explained earlier.

#### D. Performance Under Uncertainty in Measurements

On top of the previous results, this section evaluates the performance of the proposed scheme under uncertainty, which could stem from the noise produced by inaccurate measurement devices. Measurement error is modeled as an additive, white, and Gaussian noise, with a signal-to-noise ratio of 35 dB [3]. The proposed scheme is tested under the aforementioned noise without retraining, and the performance results are depicted in Table VII.

The overall accuracy dropped to 98.98%. Nevertheless, better and more accurate measurement devices can compensate for this drop.

## VI. VERIFICATION THROUGH REAL-TIME SIMULATIONS

In this section, the proposed scheme is evaluated in real time using the setup shown in Fig. 7. This setup mainly involves 1) a real-time digital simulator (RTDS), 2) an oscilloscope, and 3) a PC. The utilized RTDS belongs to the family of OP5700 RCP/Hardware-In-the-Loop field programmable gate array (FPGA)-based real-time simulator [22]. This RTDS employs a high-end reconfigurable FPGA and a group of Intel Xeon

TABLE VII  
PERFORMANCE UNDER MEASUREMENT NOISE

True	(A)	(B)	(C)	(D)	(E)	(F)
(A)	98.7%	-	-	1.3%	-	-
(B)	-	98.1%	-	-	1.4%	0.5%
(C)	-	-	99.6%	-	-	0.4%
(D)	-	-	-	98.9%	1.1%	-
(E)	-	0.5%	-	-	99.5%	-
(F)	-	-	0.9%	-	-	99.1%

(A) Direct False-Trip Cyberattacks. (B) Sympathetic-Trip Cyberattacks. (C) Fault-Masking Cyberattacks. (D) Healthy System and no cyberattack. (E) Internal Fault and no cyberattack. (F) External Fault and no cyberattack.

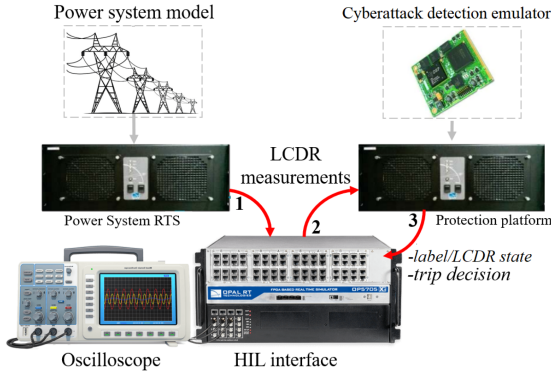


Fig. 7. Real-time simulation setup.

E5 quad-core processors running 2.3–3 GHz. In this experiment, the DNN model, trained in the previous section, is imported to RT-LAB software (2019 release), which converts the power system and solution models into C language for running on the OPAL-RT machine. In the OPAL-RT simulator, the DNN model is loaded to one of the real-time simulator’s CPU cores, which is then connected in the loop with another CPU core that runs the remainder of the power system. The sampling frequency is chosen as 1 kHz. For monitoring, the trip/no-trip signal is sent from OPAL-RT, through one of its input/output ports, to a Tektronix-DPO4054B oscilloscope [23].

Using this setup, several cases corresponding to all the scenarios investigated in this article are simulated to obtain the average prediction time of LCDRs combined with the proposed scheme. Fig. 8 illustrates the total time,  $T$ , taken by the proposed solution to operate and detect a case of a fault-masking cyberattack. As the figure depicts, the DNN detects the fault-masking cyberattack and generates the adequate trip order in less than 1 ms. (For better visualization, the time scale of the oscilloscope’s display is set to 400  $\mu$ s.) Our results show that the total DNN operating time is approximately 0.76 ms in real time. This result ensures that the overall proposed scheme operates within the time limit for LCDRs on transmission level, which is 2 power cycles, given modern LCDRs operate within 1.5 cycles [10]. Therefore, the complexity of the proposed scheme is accepted.

## VII. DISCUSSION

While the main focus of this article is to protect the highly vulnerable LCDRs from false-tripping, fault-masking,

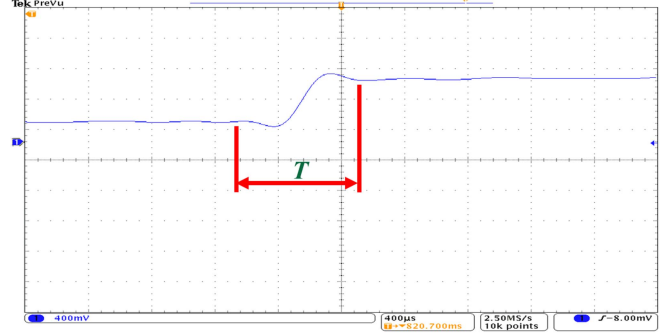


Fig. 8. Screenshot of the oscilloscope showing the time  $T$  taken by the DNN to detect a cyberattack and generate the necessary trip command.

and sympathetic-tripping cyberattacks, it will be interesting to investigate cyberattacks that may target the proposed scheme itself, e.g., using adversarial samples. In our problem, all the features used by the utilized DNN come from only the Lcdr measurements  $i_1$ ,  $v_1$ , and  $i_2$ . Since both  $i_1$  and  $v_1$  are local measurements that the remote cyberattackers cannot compromise, the attack surface available for intruders to launch adversarial attacks is relatively small compared with when the Lcdr is left unsecured and can be easily attacked by false-tripping, fault-masking, or sympathetic-tripping cyberattacks.

Several previous works have proposed mechanisms that can be used to defend DNNs against adversarial attacks, e.g. adversarial training [24], [25], robust optimization [26], using ensemble methods [27], and/or explainable techniques [28]. In addition, it may be useful to include domain knowledge, i.e., the power system protection system has specific constraints and requirements that may not be present in other applications. By incorporating domain knowledge into the design of the DNN, the system can be more robust to adversarial attacks. For example, the DNN can be designed to consider the power system’s physical constraints of the power system, such as voltage and current limits. Furthermore, since cyberattackers usually want to minimize the attack time, i.e., to maintain stealth and avoid getting caught, keeping the trained DNN model hidden/restricted, from outsiders, by the power system operators shall further decrease the chances of a successful adversarial attack.

## VIII. CONCLUSION

This article proposed a scheme for securing LCDRs against direct false-tripping, fault-masking, and sympathetic-tripping cyberattacks. The proposed scheme runs in conjunction with the Lcdr and is based only on the same measurements available for LCDRs, which are  $i_1$ ,  $v_1$ , and  $v_2$ . The proposed scheme employs a DNN to detect and classify the aforementioned cyberattacks on LCDRs. Offline, the DNN is trained on a synthetically generated dataset consisting of Lcdr measurements for the three main cyberattack categories: internal faults, external faults, and healthy fluctuations. The investigated cyberattack categories were modeled using a broad spectrum of attack mechanisms, including FDIAs, TSAs, MitMAs, and replay attacks. Our results show that the proposed scheme 1) can accurately detect and classify direct false-tripping, fault-masking, and sympathetic-tripping

cyberattacks, 2) is protectively secure, selective, and sensitive, 3) is not affected by normal system dynamics or noise, and 4) increases the LCDR's fault-detection time by only a fraction of a millisecond. Our experimental results using OPAL-RT's simulator showed that the proposed scheme can detect cyberattacks in real time. Future work directions have also been discussed.

## REFERENCES

- [1] J. Przetacznik and S. Tarpova, "Russia's war on Ukraine: Timeline of cyber-attacks," in *Eur. Parliamentary Res. Serv.*, Jun. 2022, pp. 1–7. [Online]. Available: [europarl.europa.eu/thinktank/en/document/EPRS\\_BRI\(2022\)733549](https://eurparl.europa.eu/thinktank/en/document/EPRS_BRI(2022)733549)
- [2] H. Miller, J. Burger, N. Fischer, and B. Kasztenny, "Modern line current differential protection solutions," in *Proc. 63rd Annu. Conf. Protective Relay Eng.*, 2010, pp. 1–25.
- [3] A. Mohammad Saber, A. Youssef, D. Svetinovic, H. H. Zeineldin, and E. F. El-Saadany, "Anomaly-based detection of cyberattacks on line current differential relays," *IEEE Trans. Smart Grid*, vol. 13, no. 6, pp. 4787–4800, Nov. 2022.
- [4] Y. M. Khaw, A. Abiri Jahromi, M. F. M. Arani, S. Sanner, D. Kundur, and M. Kassouf, "A deep learning-based cyberattack detection system for transmission protective relays," *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2554–2565, May 2021.
- [5] V. Nougain, S. Mishra, and S. S. Jena, "Resilient protection of medium voltage DC microgrids against cyber intrusion," *IEEE Trans. Power Del.*, vol. 37, no. 2, pp. 960–971, Apr. 2022.
- [6] A. Sharma and B. K. Panigrahi, "Interphase fault relaying scheme to mitigate sympathetic tripping in meshed distribution system," *IEEE Trans. Ind. Appl.*, vol. 55, no. 1, pp. 850–857, Jan./Feb. 2019.
- [7] L. Chen et al., "Remedial pilot main protection scheme for transmission line independent of data synchronism," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 681–690, Jan. 2019.
- [8] A. Kemmeugne, A. Abiri Jahromi, and D. Kundur, "Resilience enhancement of pilot protection in power systems," *IEEE Trans. Power Del.*, vol. 37, no. 6, pp. 5255–5266, Dec. 2022, doi: [10.1109/TPWRD.2022.3175148](https://doi.org/10.1109/TPWRD.2022.3175148).
- [9] A. Ameli, A. Hooshyar, and E. F. El-Saadany, "Development of a cyber-resilient line current differential relay," *IEEE Trans. Ind. Inf.*, vol. 15, no. 1, pp. 305–318, Jan. 2019.
- [10] *Multilin L90 Line Current Differential System – GE Grid Solutions*. [Online]. Available: <https://www.gegridsolutions.com/multilin/catalog/l90.htm>
- [11] Y. Liu, M. K. Reiter, and P. Ning, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Secur.*, 2009, pp. 21–32.
- [12] D. Dolev and A. C. Yao, "On the security of public key protocols," in *Proc. 22nd Annu. Symp. Found. Comput. Sci.* 1981, pp. 350–357.
- [13] Z. Li, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "Bilevel model for analyzing coordinated cyber-physical attacks on power systems," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2260–2272, Sep. 2016.
- [14] V. Y. Pillitteri and T. L. Brewer, "Guidelines for smart grid cybersecurity," in *Proc. Nat. Inst. Standards Technol. Interagency Rep. 7628, Rev. 1*, vol. 1, Sep. 2014, pp. 12–61.
- [15] Y. Chen, M. G. Fadda, and A. Benigni, "Decentralized load estimation for distribution systems using artificial neural networks," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 5, pp. 1333–1342, May 2019.
- [16] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, AI-based, or both?," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7331–7376, Oct. 2019.
- [17] F. Martin and J. A. Aguado, "Wavelet-based ANN approach for transmission line protection," *IEEE Trans. Power Del.*, vol. 18, no. 4, pp. 1572–1574, Oct. 2003.
- [18] M. Feurer, A. Klein, K. Eggenberger, J. T. Springenberg, M. Blum, and F. Hutter, "Efficient and robust automated machine learning," in *Proc. Adv. Neural. Inf. Process. Syst.*, Dec. 2015, pp. 2962–2970.
- [19] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural. Inf. Process. Syst.*, vol. 25, 2012, pp. 1–9.
- [20] M. Pai, *Energy Function Analysis for Power System Stability*. Berlin, Germany: Springer, 2012.
- [21] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. J. Snead, I. A. Cree, and N. M. Rajpoot, "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1196–1206, May 2016.
- [22] *OP5700*, OPAL-RT Technologies, Inc. [Online]. Available: <https://wiki.opal-rt.com/display/HDGD/OP5700>
- [23] *DPO4000 Series Digital Phosphor Oscilloscopes Datasheet*, TEKTRONIX, Inc. [Online]. Available: [tek.com/en/datasheet/dpo4000-series-digital-phosphor-oscilloscopes-datasheet](https://www.tek.com/en/datasheet/dpo4000-series-digital-phosphor-oscilloscopes-datasheet)
- [24] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. 3rd Int. Conf. Learn. Representations*, 2015.
- [25] X. Yuan, P. He, Q. Zhu, and X. Li, "Adversarial examples: Attacks and defenses for deep learning," *IEEE Trans. Neural. Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2805–2824, Sep. 2019.
- [26] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," in *Proc. Int. Conf. Learn. Representations*, 2018.
- [27] F. Tramer, A. Kurakin, N. Papernot, D. Boneh, and P. McDaniel, "Ensemble adversarial training: Attacks and defenses," in *Proc. Int. Conf. Learn. Representations*, 2018.
- [28] Reference: M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 1135–1144.



**Ahmad Mohammad Saber** (Member, IEEE) received the B.Sc. degree in electrical engineering from Ain Shams University, Cairo, Egypt, in 2016, and the M.Sc. degree in electrical engineering from Cairo University, Giza, Egypt, in 2019. He is currently working toward the Ph.D. degree in electrical engineering and computer science with the Advanced Power and Energy Center, Khalifa University, Abu Dhabi, UAE.

Since 2016, he has held technical and commercial roles in several industries including power transformers manufacturing, water and wastewater treatment, and access control and security systems. He is currently an International Visiting Graduate Student with the University of Toronto, Toronto, ON, Canada. His current research interests include cyber-physical security, machine learning applications, power systems protection, distributed generation, renewable power planning, and integration.

Mr. Saber was a Reviewer for multiple IEEE transactions journals.



**Amr Youssef** (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electronics and communications engineering from Cairo University, Giza, Egypt, in 1990 and 1993, respectively, and the Ph.D. degree in computer engineering from Queens University, Kingston, ON, Canada, in 1997.

He was with Nortel Networks, Center for Applied Cryptographic Research, University of Waterloo, Waterloo, ON, Canada, IBM, Armonk, NY, USA, and Cairo University. He is currently a Professor with the Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, Canada. He has authored more than 240 referred journal and conference publications in areas related to his research interests. His current research interests include information security and cyber-physical systems security.

Dr. Youssef was on about 75 technical program committees of cryptography and data security conferences. He was the Co-/Chair for Africrypt 2010, Africrypt 2020, and the conference Selected Areas in Cryptography (SAC 2014, SAC 2006, and SAC 2001).



**Davor Svetinovic** (Senior Member, IEEE) received the doctorate degree in computer science from the University of Waterloo, Waterloo, ON, Canada, in 2006.

Previously, he worked with TU Wien, Vienna, Austria, and Lero—the Irish Software Engineering Center—Dublin, Ireland. He was a Visiting Professor and a Research Affiliate with MIT and MIT Media Lab, Cambridge, MA, USA. He is currently a Professor of computer science with the Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi, UAE, and the Department of Information Systems and Operations Management, Vienna University of Economics and Business, Vienna, (on leave), where he is the Head of the Institute for Distributed Ledgers and Token Economy, and the Research Institute for Cryptoeconomics. He has extensive experience working on complex multidisciplinary research projects. He has authored or coauthored more than 100 papers in leading journals and conferences and is a Highly Cited Researcher in blockchain technology. His research interests include cybersecurity, blockchain technology, cryptoeconomics, trust, software engineering, and developing advanced research capabilities and institutions in emerging economies.

Dr. Svetinovic is a Senior Member of ACM (Lifetime) and a Mohammed Bin Rashid Academy of Scientists Affiliate.



**Hatem H. Zeineldin** (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from Cairo University, Giza, Egypt, in 1999 and 2002, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2006.

He was with Smith and Andersen Electrical Engineering Inc., Toronto, ON, Canada, where he was involved in projects involving distribution system designs, protection, and distributed generation. He was a Visiting Professor with the Massachusetts Institute of Technology, Cambridge, MA, USA. He is currently with the Khalifa University of Science and Technology, Abu Dhabi, UAE, and on leave from Faculty of Engineering, Cairo University. His current research interests include distribution system protection, distributed generation, and micro grids.

Dr. Zeineldin is currently an Editor of IEEE TRANSACTIONS ON ENERGY CONVERSION.



**Ehab F. El-Saadany** (Fellow, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from Ain Shams University, Cairo, Egypt, in 1986 and 1990, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 1998.

He was with ECE Department, University of Waterloo, between 2000 and 2019, where he was the Director of the Power M.Eng. Program between 2010 and 2015. He is currently a Professor with EECS Department, Khalifa University, Abu Dhabi, UAE, and an Adjunct Professor with the ECE Department, University of Waterloo. He is currently the Director of the Advanced Power and Energy Center, Khalifa University. He is an Internationally Recognized Expert in the area of sustainable energy integration and smart distribution systems. His research interests include smart grid operation and control, microgrids, self-healing, cyber-physical security of smart grids, protection, power quality, distributed generation, and power electronics interfacing.

Dr. El-Saadany is an Editor of the IEEE TRANSACTIONS ON SMART GRID, IEEE TRANSACTIONS ON POWER SYSTEMS, and IEEE POWER ENGINEERING LETTERS. He is a Registered Professional Engineer in the Province of Ontario.